alternative procedure is more difficult, and hence was not used earlier. In this approach the constant matrix K_{0i} rather than the matrix $R(s)$ is permuted at each iteration i, where K_{0i} is the first term in the expansion

$$R(s)M_{i-1}(s)=G_i(s)[K_{0i}+K_{1i}s^{-1}+\cdots]. \tag{7}$$

The permutation of K_{0i} is denoted by K_0^i and at each iteration the rank of K_{0i} is checked by counting the number of nonzero elements along the diagonal of K_0^i.

The procedure is presented in an algorithmic form and for purposes of clarity is developed using an example of a (5×5) system. Let a (5×5) matrix $R(s)$ be given by

$$R(s)=\begin{bmatrix} k_1s^{-2} & k_2s^{-4} & k_3s^{-1} & k_4s^{-4} & k_5s^{-7} \\ k_6s^{-1} & k_7s^{-3} & k_8s^{-2} & k_9s^{-4} & k_{10}s^{-6} \\ k_{11}s^{-1} & k_{12}s^{-8} & k_{13}s^{-9} & k_{14}s^{-7} & k_{15}s^{-5} \\ k_{16}s^{-1} & k_{17}s^{-2} & k_{18}s^{-2} & k_{19}s^{-2} & k_{20}s^{-3} \\ k_{21}s^{-1} & k_{22}s^{-3} & k_{23}s^{-3} & k_{24}s^{-4} & k_{25}s^{-6} \end{bmatrix}.$$

For each iteration i, a suitable matrix $L(s)$, which multiplies $M_{i-2}(s)$ to yield $M_{i-1}(s)$, the matrix $G_i(s)$ as well as the rank of K_{0i} or K_0^i, r_i is specified. The most important step is the choice of $L(s)$, the procedure for obtaining which is explained below in iteration 1. It should be noted that in this method the matrices $L(s)$, $M_{i-1}(s)$, $G_i(s)$ and K_{0i} are defined with respect to the original matrix $R(s)$ and not its permutation.

Algorithm 3.1—i) Initialization: iteration $i = 1$; $L(s) = M_0(s) = I$, $G_1(s) = \text{diag}\ (s^{-1}, s^{-1}, s^{-1}, s^{-1}, s^{-1})$. The matrices K_{01} and K_0^1 are shown below.

$$K_{01}=\begin{bmatrix} 0 & 0 & k_3 & 0 & 0 \\ k_6 & 0 & 0 & 0 & 0 \\ k_{11} & 0 & 0 & 0 & 0 \\ k_{16} & 0 & 0 & 0 & 0 \\ k_{21} & 0 & 0 & 0 & 0 \end{bmatrix} \quad K_0^1=\begin{bmatrix} k_3 & 0 & 0 & 0 & 0 \\ 0 & k_6 & 0 & 0 & 0 \\ 0 & k_{11} & 0 & 0 & 0 \\ 0 & k_{16} & 0 & 0 & 0 \\ 0 & k_{21} & 0 & 0 & 0 \end{bmatrix}, \ r_1=2.$$

Evidently K_0^1 has been obtained by shifting the third column of K_{01} to the first column so that the first two diagonal elements of K_0^1 are nonzero, and hence the rank $r_1 = 2$.

First determine the linearly dependent rows of K_0^1 (or K_{01}) and then determine the columns of K_0^1 (or correspondingly, those of K_{01}) with nonzero elements which pertain to the linearly dependent rows of K_0^1. Here the last four rows of K_0^1 or K_{01} are linearly dependent and the second column of K_0^1 (equivalently the first one of K_{01}) has nonzero element corresponding to the linearly dependent rows of K_0^1. Hence, choose $L(s) = \text{diag}\ (s^{-1}, 1, 1, 1, 1)$, where s^{-1} appears only in the first column corresponding to the first column of K_{01}. Note that $L(s)$ is defined with respect to K_{01}, rather than K_0^1.

ii) Iteration $i = 2$; with $L(s)$ of i), define $M_1(s) = M_0(s) \cdot L(s)$. Express $R(s)M_1(s)$ as in (7) with $G_2(s) = \text{diag}\ (s^{-1}, s^{-2}, s^{-2}, s^{-2}, s^{-2})$. Permuting K_{02} to obtain K_0^2 shows that rank $r_2 = 3$. Following the procedure of i), it can be seen that $L(s) = \text{diag}\ [s^{-1}, 1, s^{-1}, 1, 1]$ is chosen.

iii) Iteration $i = 3$; with $L(s)$ of ii), define $M_2(s) = M_1(s)L(s)$. Express $R(s)M_2(s)$ as in (7) and permute K_{03} to obtain $G_3(s) = \text{diag}\ [s^{-2}, s^{-3}, s^{-3}, s^{-2}, s^{-3}]$ and $r_3 = 4$. The choice of $L(s)$ is $L(s) = \text{diag}\ [s^{-1}, s^{-1}, s^{-1}, 1, 1]$.

iv) Iteration $i = 4$; define $M_3(s) = M_2(s)L(s) = \text{diag}\ [s^{-3}, s^{-1}, s^{-2}, 1, 1]$, where $L(s)$ has been obtained in iii). $G_4(s) = \text{diag}\ [s^{-3}, s^{-4}, s^{-4}, s^{-2}, s^{-4}]$ and $r_4 = 4$. $L(s) = \text{diag}\ [s^{-1}, s^{-1}, s^{-1}, s^{-1}, 1]$ is chosen.

v) Iteration $i = 5$; $M_4(s) = M_3(s) \cdot L(s)$. $G_5(s) = \text{diag}\ (s^{-4}, s^{-5}, s^{-5}, s^{-3}, s^{-5})$ and

$$K_{05}=\begin{bmatrix} 0 & 0 & k_3 & 0 & 0 \\ k_6 & k_7 & k_8 & k_9 & 0 \\ k_{11} & 0 & 0 & 0 & k_{15} \\ 0 & 0 & 0 & k_{19} & k_{20} \\ k_{21} & k_{22} & 0 & k_{24} & 0 \end{bmatrix}, \quad K_0^5=\begin{bmatrix} k_3 & 0 & 0 & 0 & 0 \\ k_8 & k_6 & k_7 & k_9 & 0 \\ 0 & k_{21} & k_{22} & k_{24} & 0 \\ 0 & 0 & 0 & k_{19} & k_{20} \\ 0 & k_{11} & 0 & 0 & k_{15} \end{bmatrix}, \ r_5=5.$$

Since, $r_5 = 5$, $D(s) = M_4(s) = \text{diag}\ (s^{-4}, s^{-2}, s^{-3}, s^{-1}, 1)$ and $E(R(s)D(s)) = K_{05}$ is nonsingular and is given as denoted above. From our earlier discussion, it also follows that the Hermite form $\bar{H}(s)$ of $W_p(s)W_c(s)$ for $\pi(s) = s$ is given by $G_5(s)$, and the high-frequency gain matrix $\bar{K}_p = E(W_p(s)W_c(s)) = K_{05}$. Any other stable $\bar{H}(s)$ can be obtained by substituting a stable monic polynomial, $(s + a)$ with $a > 0$, for s in $G_5(s)$ [10].

IV. Conclusions

When the matrix $E(W_p(s))$ is nonsingular, $H(s)$ is diagonal and can be easily specified with the knowledge of relative degree of each element of $W_p(s)$. When it is triangular, the principal result of this note demonstrates that with the same prior information a diagonal precompensator $W_c(s)$ can be designed such that $W_p(s)W_c(s)$ has a diagonal Hermite form $\bar{H}(s)$. With the scheme presented here, almost all MIMO plants, satisfying (2), can be adaptively controlled, enlarging the class of plants considered in [4], [5], and using less prior information than that assumed in [7]–[9].

References

[1] K. S. Narendra and L. S. Valavani, "Stable adaptive controller design—Direct control," *IEEE Trans. Automat. Contr.*, vol. AC-23, Aug. 1978.

[2] K. S. Narendra, Yuan-Hao Lin, and L. S. Valavani, "Stable adaptive controller design, Part II: Proof of stability," *IEEE Trans. Automat. Contr.*, vol. AC-25, June 1980.

[3] A. S. Morse, "Global stability of parameter-adaptive control systems," *IEEE Trans. Automat. Contr.*, vol. AC-25, June 1980.

[4] R. V. Monopoli and C. C. Hsing, "Parameter adaptive control of multivariable systems," *Int. J. Contr.*, vol. 22, no. 3, pp. 313–327, 1975.

[5] G. C. Goodwin, P. J. Ramadge, and P. E. Caines, "Discrete time multivariable adaptive control," *IEEE Automat. Contr.*, vol. AC-25, June 1980.

[6] R. P. Singh and K. S. Narendra, "A globally stable adaptive controller for multivariable systems," in *Proc. 20th Conf. Decision. Contr.*, San Diego, CA, Dec. 1981; also in Center for Syst. Sci., Yale Univ., New Haven, CT, Tech. Rep. 8111, Apr. 1982.

[7] H. Elliott and W. A. Wolovich, "A parameter adaptive control structure for linear multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-27, Apr. 1982.

[8] G. C. Goodwin and R. S. Long, "Generalization of results on multivariable adaptive control," *IEEE Trans. Automat. Contr.*, vol. AC-25, Dec. 1980.

[9] A. S. Morse, "Parametrizations for multivariable adaptive control," in *Proc. 20th Conf. Decision Contr.*, San Diego, CA, Dec. 1981.

[10] R. P. Singh and K. S. Narendra, "Prior information in the design of multivariable adaptive controllers," Center for Syst. Sci., Yale Univ., New Haven, CT, Tech. Rep. 8304, 1983.

[11] A. S. Morse, "System invariants under feedback and cascade control," in *Math. System Theory, Vol. 131.* New York: Springer Verlag.

A Solution Method for the Static Constrained Stackelberg Problem Via Penalty Method

EITARO AIYOSHI AND KIYOTAKA SHIMIZU

Abstract—**This note presents a new solution method for the static constrained Stackelberg problem. Through our approach, the Stackelberg problem is completely transformed into a one-level unconstrained problem such that the newly introduced overall augmented objective function is minimized with repect to the leader's and the follower's variables jointly. It can be proved that a sequence of solutions to the transformed problems converges to the solution of the original problem, when the penalty parameters are updated.**

I. Introduction and Motivation

This note presents a new solution method for the static constrained Stackelberg problem by adopting the ideas of a penalty method that transforms it into the one-level unconstrained problem. The Stackelberg problem [1], [2] is described as a two-level optimization problem, in

Manuscript received February 6, 1984; revised April 3, 1984. This paper is based on a prior submission of June 24, 1983.

The authors are with the Faculty of Science and Technology, Keio University, Kohoku-ku, Yokohama, Japan.

0018-9286/84/1200-1111$01.00 ©1984 IEEE

which the follower's problem is subordinated to the leader's problem as a part of its constraints

$$\min_{x} F(x, \check{y}(x)) \tag{1a}$$

$$\text{subject to } G(x, \check{y}(x)) \leqq \mathbf{0} \tag{1b}$$

$$f(x, \check{y}(x)) = \min_{y} f(x, y) \tag{1c}$$

$$\text{subject to } g(x, y) \leqq \mathbf{0} \tag{1d}$$

where $x \in R^N$ and $y \in R^n$ are the leader's and the follower's variables, respectively, and F: $R^N \times R^n \rightarrow R^1$ and G: $R^N \times R^n \rightarrow R^M$ are the objective and the constraint functions related to the leader, while f: $R^N \times R^n \rightarrow R^1$ and g: $R^N \times R^n \rightarrow R^m$ are those concerned with the follower. Here, let $\check{y}(x)$ denote an optimal solution to the follower's problem corresponding to the leader's variable x, and impose the following assumptions.

1) For any fixed $x \in R^N$, optimal solution $\check{y}(x)$ is uniquely determined.

Such a problem cannot be solved by ordinary mathematical programming, and it is very hard to develop its computational method. In previous works [3], [4], we proposed a penalty approach such that the follower's part was replaced by an unconstrained problem with an augmented objective function

$$p(x, y; r) = f(x, y) + r\phi(g(x, y)), \qquad r > 0 \tag{2}$$

where ϕ was a continuous interior penalty function on the negative domain of R^m, and

$$\phi(g(x, y)) > 0, \qquad \text{as } y \in \text{int } \mathcal{Y}(x), \tag{3a}$$

$$\phi(g(x, y)) \rightarrow +\infty, \qquad \text{as } y \rightarrow \text{bd}\,\mathcal{Y}(x). \tag{3b}$$

Here, let $\mathcal{Y}(x)$ denote $\{y | g(x, y) \leqq \mathbf{0}\}$. Then the original problem (1) was transformed into

$$\min_{x} F(x, \check{y}(x; r)) \tag{4a}$$

$$\text{subject to } G(x, \check{y}(x; r)) \leqq \mathbf{0} \tag{4b}$$

$$P(x, \check{y}(x; r), r) = \min_{y} p(x, y; r). \tag{4c}$$

In [3], it was proved that the sequence $\{(x^k, \check{y}(x^k; r^k))\}$ of optimal solutions to problem (4), in response to a parameter sequence of r converging to zero, converges to the solution to problem (1) when one solves problem (4) for a fixed value of the parameter r, one should solve the follower's penalized unconstrained problem every time the trial value of the leader's variable was updated. Therefore, this approach was less efficient.

In this note, we use the same augmented objective function (2) and the penalty function (3) for the follower's problem, but the difference is that the follower's penalized problem is replaced by its stationary condition $\nabla_y p(x, y; r) = \mathbf{0}$, so that the original two-level problem is transformed into a one-level problem

$$\min_{(x,y)} F(x, y) \tag{5a}$$

$$\text{subject to } G(x, y) \leqq \mathbf{0} \tag{5b}$$

$$\nabla_y p(x, y; r) = \mathbf{0} \tag{5c}$$

$$g(x, y) < \mathbf{0} \tag{5d}$$

where the constraint (5d) prescribes the domain of the function p.

Our next step is to solve the one-level constrained problem (5) under a given follower's penalty parameter. To do so, we adopt the interior-exterior mixed penalty approach, and introduce a new overall (leader's) augmented objective function

$$Q(x, y; t, s, r) = F(x, y) + t\Phi(G(x, y)) + r\phi(g(x, y)) + s\Psi(\|\nabla_y p(x, y; r)\|. \tag{6}$$

Here Φ is a continuous interior penalty function on the negative domain of R^M, and

$$\Phi(G(x, y)) > 0, \qquad \text{if } (x, y) \in \text{int } \{(x, y) | G(x, y) \leqq \mathbf{0}\}, \tag{7a}$$

$$\Phi(G(x, y)) \rightarrow +\infty, \qquad \text{if } (x, y) \rightarrow \text{bd } \{(x, y) | G(x, y) \leqq \mathbf{0}\}. \tag{7b}$$

Ψ is a continuous and monotone increasing exterior penalty function on $[0, +\infty)$, and

$$\Psi(\|\nabla_y p(x, y; r)\|) = 0, \quad \text{if } (x, y) \in \left\{(x, y) \in \bigcup_{x \in R^N} x \times \text{int } \mathcal{Y}(x) | \nabla_y p(x, y; r) = \mathbf{0}\right\}, \tag{8a}$$

$$\Psi(\|\nabla_y p(x, y; r)\|) > 0, \quad \text{if } (x, y) \in \left\{(x, y) \in \bigcup_{x \in R^N} x \times \text{int } \mathcal{Y}(x) | \nabla_y p(x, y; r) \neq \mathbf{0}\right\}. \tag{8b}$$

Hence, we can completely transform the original problem (1) into the following one-level unconstrained problem such that the overall augmented objective function (6) is minimized with respect to the leader's and the follower's variables jointly:

$$\min_{(x,y)} Q(x, y; t, s, r). \tag{9}$$

In the orthodox way of updating the penalty parameters, in order to solve problem (5) using the function Q, one has to let the penalty parameters associated with Q go to their limits under a fixed penalty parameter in p, and then one must repeat this process every time r in problem (5) is updated to solve the original problem. In the next section, however, the main theorem states that all the penalty parameters can be updated simultaneously toward their limits in our doubly penalized approach. This result makes the penalty approach more efficient than the methods of previous works [3], [4].

Approaches through replacement of the follower's part by the associated optimality condition can be found in [5], [6]. In the case with inequality constraints, the complementary slackness condition stands in the way of this approach. In [5], an attempt to overcome the difficulty can be seen only in the case where all functions are linear, while [6] does not consider the complementary slackness condition. Our penalty approach intends to avoid the above difficulty by introducing the augmented objective function in the follower's part.

II. Theoretical Results

Impose the following assumptions.

ii) int $\mathcal{Y}(x)$ which is given by $\{y | g(x, y) < \mathbf{0}\}$, is not empty and its closure is $\mathcal{Y}(x)$.

iii) int $\{(x, y) | G(x, y) \leqq \mathbf{0}\}$, which is given by $\{(x, y) | G(x, y) < \mathbf{0}\}$, is not empty and its closure becomes $\{(x, y) | G(x, y) \leqq \mathbf{0}\}$.

iv) The functions F, G, f, and g are continuous with respect to their arguments, and particularly f and g are differentiable in y and ϕ is differentiable in g. Furthermore, $\nabla_y f(x, y)$, $\nabla_y g(x, y)$ and $\nabla\phi$ are continuous with respect to their arguments.

v) The set $\{(x, y) | G(x, y) \leqq \mathbf{0}, g(x, y) \leqq \mathbf{0}\}$ is compact.

It can be easily proved that under assumptions ii)–v) there exists a solution to problem (9) for every positive penalty parameter on $\{(x, y) | G(x, y) < \mathbf{0}, g(x, y) < \mathbf{0}\}$, and the following lemma holds, which follows from some standard results of the penalty method.

Lemma 1: Let x be given arbitrarily such that $G(x, \check{y}(x)) < \mathbf{0}$ for the corresponding $\check{y}(x)$, and $\{y^k\}$ be a sequence of optimal solutions to problem $\min_y p(x, y; r^k)$ corresponding to a positive sequence $\{r^k\}$ converging to zero. Then, for any positive sequence $\{s^k\}$ diverging to infinity and any positive sequence $\{t^k\}$ converging to zero

$$\lim_{k \rightarrow \infty} Q(x, y^k; t^k, s^k, r^k) = F(x,, \check{y}(x)). \tag{10}$$ ■

Here, we add the following four further assumptions.

vi) The set $\{x | G(x, \check{y}(x)) < \mathbf{0}\}$ is not empty and its closure is $\{x | G(x, \check{y}(x)) \leqq \mathbf{0}\}$.

vii) The function ϕ is additively separable as $\phi(g) = \Sigma_{i=1}^m \varphi(g_i)$, where φ is a monotone increasing and continuously differentiable function on $(-\infty, 0)$.

viii) Let $I(x, y) = \{i | g_i(x, y) = 0\}$ for any fixed (x, y). Then $\nabla_y g_i(x, y)$, $i \in I(x, y)$ are linearly independent.

ix) The function f and g are convex in y for every fixed x.

Then the following lemma holds.

Lemma 2: Let $\{(x^k, y^k)\}$ be a sequence of optimal solutions to problem (9) corresponding to positive sequences $\{r^k\}$ and $\{t^k\}$ converging to zero and a positive sequence $\{s^k\}$ diverging to infinity. If assumptions i)–ix) are satisfied and there exists an accumulation point $(\bar{x}, \bar{y})$ of $\{(x^k, y^k)\}$, then $\bar{y}$ solves the follower's problem in response to $\bar{x}$.

Proof: First, in order to show

$$\lim_{k\to\infty} \nabla_y p(x^k, y^k; r^k) = \mathbf{0} \tag{11}$$

suppose that $\nabla_y p(x^k, y^k; r^k)$ does not converge to zero. Then, nor does $\Psi(\|\nabla_y p(x^k, y^k; r^k)\|)$ by continuity and monotonicity of Ψ. That is, there exist a positive M and a positive integer K such that $\Psi(\|\nabla_y p(x^k, y^k; r^k)\|) > M$ for $\forall k < K$. Therefore, as $s^k \to \Psi + \infty$, $s^k\Psi(\|\nabla_y p(x^k, y^k; r^k)\|) \to +\infty$. Thus, for any positive ϵ', there exists a positive integer K' such that

$$Q(x^k, y^k; t^k, s^k, r^k) > \epsilon' \qquad \text{for } \forall k > K'. \tag{12}$$

On the other hand, choosing $\bar{x}$ such that $G(\bar{x}, \hat{y}(\bar{x})) < \mathbf{0}$, and considering the corresponding follower's solution $\hat{y}(\bar{x})$ as well as the solution $\bar{y}^k$ $(= \hat{y}(\bar{x}; r^k)) \in \operatorname{int} \mathcal{Y}(\bar{x})$ to the problem $\min_y p(\bar{x}, y; r^k)$, we have, from Lemma 1,

$$\lim_{k\to\infty} Q(\bar{x}, \bar{y}^k; t^k, s^k, r^k) = F(\bar{x}, \hat{y}(\bar{x})).$$

That is, for any positive ϵ, there exists a positive integer $\bar{K}$ such that

$$|Q(\bar{x}, \bar{y}^k; t^k, s^k, r^k) - F(\bar{x}, \hat{y}(\bar{x}))| < \epsilon \qquad \text{for } \forall k > \bar{K}. \tag{13}$$

Now, setting

$$\epsilon' = F(\bar{x}, \hat{y}(\bar{x})) + \epsilon \tag{14}$$

and using (13), (14), (12) in turn, we have

$$Q(\bar{x}, \bar{y}^k; t^k, s^k, r^k) < F(\bar{x}, \hat{y}(\bar{x})) + \epsilon = \epsilon' < Q(x^k, y^k; t^k, s^k, r^k) \tag{15}$$

which contradicts the fact that (x^k, y^k) solves problem (9). Thus, we have (11).

Next, let us set $\lambda_i^k = r^k \partial\varphi(g_i(x^k, y^k))/\partial g_i$ and show that sequences $\{\lambda_i^k\}$, $i = 1, \cdots, m$ have accumulation points, respectively. First, $g(\bar{x}, \bar{y}) \leqq \mathbf{0}$, which follows from the fact that, for any positive integer k, $g(x^k, y^k) < \mathbf{0}$. Now, denote a newly convergent subsequence of $\{(x^k, y^k)\}$ to $(\bar{x}, \bar{y})$ by $\{(x^k, y^k)\}$, and the corresponding subsequence of $\{\lambda_i^k\}$ by $\{\lambda_i^k\}$. Since $g_i(\bar{x}, \bar{y}) < 0$, $i \notin I(\bar{x}, \bar{y})$, $\partial\varphi(g_i(\bar{x}, \bar{y}))/\partial g_i$ is finite by assumption vii). From this and the continuity of g_i, for $i \notin I(\bar{x}, \bar{y})$,

$$\lim_{k\to\infty} \lambda_i^k = 0 \times (\partial\varphi(g_i(\bar{x}, \bar{y}))/\partial g_i) = 0. \tag{16}$$

To show boundedness of $\{\lambda_i^k\}$, $i \in I(\bar{x}, \bar{y})$, suppose that there exists an unbounded sequence among them. Setting $\alpha_i^k = \lambda_i^k/\Sigma_{i=1}^m \lambda_i^k$, the sequences $\{\alpha_i^k\}$, $i \in I(\bar{x}, \bar{y})$ have accumulation points $\tilde{\alpha}_i$, $i \in I(\bar{x}, \bar{y})$, respectively, since $0 \leqq \alpha_i^k < 1$. Denote a newly convergent subsequence of $\{\alpha_i^k\}$ to $\tilde{\alpha}_i$ by $\{\alpha_i^k\}$, and the corresponding subsequence of $\{\lambda_i^k\}$ and $\{(x^k, y^k)\}$ by the same symbols, respectively. Dividing $\nabla_y p(x^k, y^k; r^k)$ by $\Sigma_{i=1}^m \lambda_i^k$, taking limits and using (11), we have

$$\begin{aligned} Q &= \lim_{k\to\infty} \nabla_y p(x^k, y^k; r^k) \Big/ \sum_{i=1}^m \lambda_i^k \\ &= \lim_{k\to\infty} \nabla_y f(x^k, y^k) \Big/ \sum_{i=1}^m \lambda_i + \sum_{i=1}^m \lim_{k\to\infty} \alpha_i^k \nabla_y g_i(x^k, y^k) \\ &= \mathbf{0} + \tilde{\alpha}^T \nabla_y g(\bar{x}, \bar{y}) \end{aligned} \tag{17}$$

where $\tilde{\alpha}^T = (\tilde{\alpha}_1, \cdots, \tilde{\alpha}_m)$. Since $\tilde{\alpha}_i = 0$, $i \notin I(\bar{x}, \bar{y})$ from (16), we have

$$\sum_{i \in I(\bar{x}, \bar{y})} \tilde{\alpha}_i \nabla_y g_i(\bar{x}, \bar{y}) = \mathbf{0}. \tag{18}$$

Since not all $\tilde{\alpha}_i$, $i \in I(\bar{x}, \bar{y})$ are equal to zero, (18) contradicts the independence of assumption (viii). Thus, the sequences $\{\lambda_i^k\}$, $i = 1, \cdots, m$, are all bounded, and they have finite accumulation points.

Incidentally, since λ_i^k are nonnegative under the monotonicity assumption vii), we get $\bar{\lambda}_i \geqq 0$, $i \in I(\bar{x}, \bar{y})$. From (16), we have also $\bar{\lambda}_i = 0$, $i \notin I(\bar{x}, \bar{y})$. That is, for $\bar{\lambda} = (\bar{\lambda}_1, \cdots, \bar{\lambda}_m)^T$

$$g(\bar{x}, \bar{y}) \leqq \mathbf{0}, \quad \bar{\lambda}^T g(\bar{x}, \bar{y}) = 0, \qquad \bar{\lambda} \geqq \mathbf{0}. \tag{19a}$$

And also, from (11), we have

$$\nabla_y f(\bar{x}, \bar{y}) + \bar{\lambda}^T \nabla_y g(\bar{x}, \bar{y}) = \mathbf{0}. \tag{19b}$$

Under the convexity assumption ix), (19) provides the necessary and sufficient condition of $\bar{y}$ to be an optimal solution to the follower's problem in response to $\bar{x}$. ∎

We now show the main theorem for convergence, by utilizing the above lemmas.

Theorem: Let $\{(x^k, y^k)\}$ be a sequence of optimal solutions to problem (9) corresponding to positive sequences $\{t^k\}$ and $\{r^k\}$ converging to zero, and a positive sequence $\{s^k\}$ diverging to infinity. If assumptions i)–ix) are satisfied, then the sequence $\{(x^k, y^k)\}$ has accumulation points, any one of which solves problem (1).

Proof: Since $\{(x^k, y^k)\}$ belongs to the compact set of the assumption v), it has an accumulation point. Denote any one of the accumulation points by $(\bar{x}, \bar{y})$ and a newly convergent subsequence of $\{(x^k, y^k)\}$ to $(\bar{x}, \bar{y})$ by $\{(x^k, y^k)\}$.

Feasibility of the Accumulation Points: Lemma 2 says that $\bar{y}$ solves the follower's problem with $\bar{x}$. Since $\hat{y}(\bar{x})$ is unique under the assumption i), we have $\bar{y} = \hat{y}(\bar{x})$. Furthermore, since $G(x^k, y^k) < \mathbf{0}$, the continuity of G implies that $G(\bar{x}, \bar{y}) \leqq \mathbf{0}$. Thus, $(\bar{x}, \bar{y})$ is feasible for problem (1).

Optimality of the Accumulation Points: Suppose that $(\bar{x}, \bar{y})$ does not solve problem (1), then there exists x' such that

$$G(x', \hat{y}(x')) \leqq \mathbf{0}, \tag{20}$$

$$F(x', \hat{y}(x')) < F(\bar{x}, \bar{y}) = F(\bar{x}, \hat{y}(\bar{x})). \tag{21}$$

Since $\hat{y}(x)$ is continuous at any x under assumptions i), ii), iv), and v), as proved by Hogan [7]. $F(x, \hat{y}(x))$ also becomes continuous at any x under the continuity of F. Therefore, consider an open ball $B(x'; \delta)$ around x' with radius δ; then there exists a number $\delta > 0$ such that $F(x, \hat{y}(x)) < F(\bar{x}, \bar{y})$ for $\forall x \in B(x'; \delta)$. Furthermore, (20) and assumption vi) imply that there exists another point $x'' \in \{x | G(x, \hat{y}(x)) < \mathbf{0}\} \cap B(x'; \delta)$. That is, there exists x'' such that

$$G(x'', \hat{y}(x'')) < \mathbf{0}, \tag{22}$$

$$F(x'', \hat{y}(x'')) < F(\bar{x}, \bar{y}). \tag{23}$$

Here, let

$$F(\bar{x}, \bar{y}) - F(x'', \hat{y}(x'')) = 2\epsilon, \qquad \epsilon > 0. \tag{24}$$

On the other hand, considering the solution $\hat{y}(x'')$ to the follower's problem with x'' and the solution $y''^k (= \hat{y}(x''; r^k)) \in \operatorname{int} \mathcal{Y}(x'')$ to the problem $\min_y p(x'', y; r^k)$, we have $y''^k \to \hat{y}(x'')$ as $k \to \infty$, based on the standard penalty method. This convergence, the continuity of G, and (22) imply the existence of a positive integer K such that

$$G(x'', y''^k) < \mathbf{0} \qquad \text{for } \forall k > K. \tag{25}$$

Together with the fact that $y''^k \in \operatorname{int} \mathcal{Y}(x'')$, (25) implies that (x'', y''^k) belongs to the domain of definition of Q for all $k > K$.

Meanwhile, Lemma 1 implies that $\lim_{k\to\infty} Q(x'', y''^k; t^k, s^k, r^k) = F(x'', \hat{y}(x''))$. That is, there exists a positive integer K'' such that

$$|Q(x'', y''^k; t^k, s^k, r^k) - F(x'', \hat{y}(x''))| < \epsilon \qquad \text{for } \forall k > K'' \tag{26}$$

for ϵ in (24). Besides, from $(x^k, y^k) \to (\bar{x}, \bar{y})$ and continuity of F, we have the existence of a positive integer $\tilde{K}$ such that

$$|F(x^k, y^k) - F(\bar{x}, \bar{y})| < \epsilon \qquad \forall k > \tilde{K} \tag{27}$$

for the same ϵ. Using (26), (24), (27), and positiveness of the penalty term, in turn, we have the following relations for all values of the integer $k > \max(K'', \tilde{K})$

TABLE I
COMPUTATIONAL RESULTS FOR EXAMPLE 1

r^k	s^k	t^k	x^k	y^k	Q^k	F^k	f^k
100.	0.01	100.	8.650	5.708	212.1	93.2	91.8
10.	0.1	10.	9.348	8.676	125.4	89.1	10.90
1.0	1.0	1.0	9.875	9.587	103.6	97.7	0.903
0.1	10.	0.1	9.929	9.826	100.0	98.6	0.1751
0.01	100.	0.01	9.977	9.912	99.8	99.5	0.0399
0.001	1000.	0.001	9.980	9.917	517.6	99.6	0.0343
True Values			10.000	10.000	-----	100.0	0.0000

$Q^k=Q(x^k,y^k;t^k,s^k,r^k)$, $F^k=F(x^k,y^k)$, $f^k=f(x^k,y^k)$

TABLE II
COMPUTATIONAL RESULTS FOR EXAMPLE 2

r^k	s^k	t^k	x_1^k	x_2^k	y_1^k	y_2^k	Q^k	F^k	f^k
100.	0.01	100.	12.44	17.84	-2.650	0.622	96.31	6.651	31.81
10.	0.1	10.	19.38	25.58	1.524	6.047	20.75	7.206	4.826
1.0	1.0	1.0	22.94	28.28	3.115	8.147	10.47	8.654	0.0495
0.1	10.	0.1	23.06	28.20	3.028	8.274	9.36	8.613	7.44×10^{-3}
0.01	100.	0.01	24.52	29.73	4.518	9.667	5.99	5.932	3.77×10^{-3}
0.001	1000.	0.001	24.52	29.71	4.515	9.693	5.91	5.825	1.80×10^{-4}
0.0001	10000.	0.0001	24.52	29.70	4.516	9.698	5.82	5.786	2.0×10^{-6}
True Values			25.00	30.00	5.000	10.000	-----	5.000	0.000

$Q^k=Q(\mathbf{x}^k,\mathbf{y}^k;t^k,s^k,r^k)$, $F^k=F(\mathbf{x}^k,\mathbf{y}^k)$, $f^k=f(\mathbf{x}^k,\mathbf{y}^k)$

$$Q(x'', y''^k; t^k, s^k, r^k)<F(x'', \check{y}(x''))+\epsilon$$
$$=F(\tilde{x}, \tilde{y})-\epsilon<F(x^k, y^k)\leqq Q(x^k, y^k; t^k, s^k, r^k). \tag{28}$$

Then this contradicts that (x^k, y^k) is an optimal solution to problem (9) for all positive integer $k > \max(K, K'', \bar{K})$. This completes the proof. ■

The theorem says that a limit point of $\{(x^k, y^k)\}$, which is generated by a series of problems (9) corresponding to a sequence $\{(r^k, s^k, t^k)\}$ such that $r^k \to 0, s^k \to +\infty, t^k \to 0$, if it exists, becomes a solution to the two-level problem (1). Consequently, if positive numbers r^k, s^k, and t^k are given, the problem can be solved easily by means of unconstrained optimization techniques.

III. NUMERICAL RESULTS

In order to verify the convergence shown by the theorem, we present illustrative examples as follows.

Example 1 [3]: We present the following simple problem:

$$\min_x x^2+(\check{y}(x)-10)^2$$
$$\text{subject to } -x+\check{y}(x)\leqq 0, \quad 0\leqq x\leqq 15$$
$$(x+2\check{y}(x)-30)^2=\min_y (x+2y-30)^2$$
$$\text{subject to } x+y\leqq 20$$
$$0\leqq y\leqq 20.$$

For this problem, we used SUMT type penalty functions in the subsidiary and the overall augmented objective function, and solved $\min_{(x,y)} Q(x, y; t^k, s^k, r^k)$ by means of Fletcher-Reeves' conjugate gradient method. The computational results are summarized in Table I.

Example 2: This is an example with two-dimensional variables $\mathbf{x}$ and $\mathbf{y}$.

$$\min_{\mathbf{x}} 2x_1+2x_2-3\check{y}_1(\mathbf{x})-3\check{y}_2(\mathbf{x})-60$$
$$\text{subject to } x_1+x_2+\check{y}_1(\mathbf{x})-2\check{y}_2(\mathbf{x})-40\leqq 0$$
$$0\leqq x_1\leqq 50, \quad 0\leqq x_2\leqq 50$$
$$(\check{y}_1(\mathbf{x})-x_1+20)^2+(\check{y}_2(\mathbf{x})-x_2+20)^2$$
$$=\min_{\mathbf{y}} (y_1-x_1+20)^2+(y_2-x_2+20)^2$$
$$\text{subject to } 2y_1-x_1+10\leqq 0, \quad 2y_2-x_2+10\leqq 0$$
$$-10\leqq y_1\leqq 20, \quad -10\leqq y_2\leqq 20.$$

The computational results are shown in Table II. These computational experiments confirm the convergence property and availability of our method. However, we faced computational difficulties of ill-conditioning in the vicinity of the boundary of the feasible region, when we chose r and t too small and s too large. It is hoped that the trouble can be overcome by various existing techniques for improvement of the penalty method. An unsettled issue is that the original problem (1) is a nonconvex program in general and so is the transformed problem (9). Therefore, it is desirable that problem (9) is solved by a solution technique producing a global optimum. Whereas, if one applies a standard gradient method to it, it is advisable to resolve problem (9) for several different initial points.

REFERENCES

[1] M. Simaan and J. B. Cruz, "On the Stackelberg strategy in nonzero-sum games," *J. Optimiz. Theory Appl.*, vol. 11, no. 5, pp. 533-555, 1973.
[2] M. Simaan, "Stackelberg optimization of two-level systems," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-7, no. 4, pp. 554-556, 1977.
[3] K. Shimizu and E. Aiyoshi, "New computational method for Stackelberg and min-max problems by use of a penalty method," *IEEE Trans. Automat. Contr.*, vol. AC-26, no. 2, pp. 460-466, 1981.
[4] E. Aiyoshi and K. Shimizu, "Hierarchical decentralized systems and its new solution by a barrier method," *IEEE Trans. Syst. Man, Cybern.*, vol. SMC-11, no. 6, pp. 444-449, 1981.
[5] J. F. Bard and J. E. Falk, "An explicit solution to the multilevel programming problem," *Comput. Operat. Res.*, vol. 9, no. 1, pp. 77-100, 1982.
[6] G. P. Papavassilopoulos, "Algorithms for leader-followers games," in *Proc. 18th Allerton Conf. Commun., Contr. Comput.*, 1980.
[7] W. W. Hogan, "Point-to-set maps in mathematical programming," *SIAM Rev.*, vol. 15, pp. 591-603, 1973.

Multidimensional Aperture Control

JORGE L. ARAVENA AND WILLIAM A. PORTER

***Abstract*—Industrial control applications often must cope with sensor information which is obscured by physical obstructions or is not synchronized with other sensors. This correspondence formulates a generalized information aperture model in a multidimensional setting. A basic optimal control problem is solved using aperture constraints.**

I. INTRODUCTION

In this correspondence, we consider a class of problems whose motivation is found in certain industrial control applications. In earlier

Manuscript received July 21, 1983; revised February 2, 1984. This paper is based on a prior submission of January 28, 1983. Paper recommended by P. R. Kumar, Past Chairman of the Stochastic Control Committee. This work was supported in part by the State Board of Regents under Grant 127-20-4101.

J. L. Aravena is with the Departmento de Ingenieria Electrica, Universidad de Santiago, Santiago, Chile.

W. A. Porter is with the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803-5901.

0018-9286/84/1200-1114$01.00 ©1984 IEEE