p_s = compressive drag pressure, N/m^2
q = superficial flow rate of liquid, $m^3/m^2 \cdot s$
q_L = rate-thickness product, m^2/s
r = superficial flow rate of solids, $m^3/m^2 \cdot s$
x = distance from medium, m
w = mass of dry solid per unit area, kg/m^2

Greek Letters

α = filtration or flow resistance, m/kg
ϵ = porosity or volume fraction of voids, —
μ = viscosity, $N \cdot s/m^2$
ρ_s = true solid density, kg/m^3

LITERATURE CITED

Grace, H. P., "Resistance and compressibility of filter cakes," *Chem. Eng. Progr.*, **49**, 303 (1953).

Hutto, F. B., Distribution of porosity in filter cake, *ibid.*, **53**, 328 (1957).

Lu, Wei-Ming, "Theoretical and experimental analysis of variable pressure filtration," Ph.D. dissertation, Univ. Houston, Texas (1968).

Nieto, Shayler, "Studies in variable pressure filtration," M.S. thesis, Univ. Houston, Texas (1967).

Schlicter, Charles, U.S. Geol. Survey, 19th Ann. Rpt., Part II, 295-385 (1897-8).

Shirato, M., and S. Okamura, "Behavior of Gairome clay slurries under constant pressure filtration," *Kagaku Kogaku*, **20**, 678 (1956).

Shirato, M., M. Sambuichi, H. Kato, and T. Aragaki, "Internal flow mechanism in filter cakes," *AIChE J.*, **15**, 405 (1969).

Shirato, M., T. Murase, M. Negawa, and T. Senda, "Fundamental studies of expression under variable pressure," *J. Chem. Eng. Japan*, **3**, 105 (1970).

Tiller, F. M., and H. R. Cooper, "The role of porosity in filtration: part 5, porosity variation in filter cakes," *AIChE J.*, **8**, 495 (1962).

Tiller, F. M., and P. J. Lloyd, (eds.), *Theory and practice of solid-liquid separation*, Ch. 4, Dept. of Chem. Eng., Univ. Houston, Texas (1972).

Tiller, F. M., and Wei-Ming Lu, "The role of porosity in filtration 8; Cake non-uniformity in compression permeability cells," *AIChE J.*, **18**, 56 (1973).

Tiller, F. M., and H. Risbud, "Hydraulic deliquoring of filter cakes, First Pacific Chemical Engineering Congress," Part 1, p. 80, *Soc. Chem. Engrs. Japan*, Amer. Inst. Chem. Engrs. (1972).

Manuscript received August 1 and accepted September 12, 1973.

Overcoming Deficiencies of the Two-Level Method for Systems Optimization

GEORGE STEPHANOPOULOS and ARTHUR W. WESTERBERG

Department of Chemical Engineering
University of Florida, Gainesville, Florida 32611

Avery and Foss (1971) have demonstrated that the method of two-level optimization (Brosilow and Lasdon, 1965; Brosilow and Nunez, 1968) may not be generally applicable to chemical process design problems due to the mathematical character of commonly encountered objective functions. The purpose of this communication is to demonstrate that this shortcoming can in fact be overcome by using a new algorithm proposed by the authors (1973).

Consider the following sequential unconstrained problem. (Constraints and recycles do not change the following results, and we want to keep the presentation here as simple as possible.)

$$\text{Min} \left\{ F|F = \sum_{i=1}^{N} \phi_i(x_i, u_i),\ x_{i+1} = f_i(x_i, u_i), \quad i = 1, \ldots, N,\ x_1 = x_1^0 \text{ (given)} \right\} \tag{P1}$$

The Lagrangian for this problem is

$$L = \sum_{i=1}^{N} \{\phi_i(x_i, u_i) - \lambda_i^T x_i + \lambda_{i+1}^T f_i(x_i, u_i)\} + \lambda_1^T x_1^0 - \lambda_{N+1}^T x_{N+1}$$

The natural boundary condition on λ_{N+1} is $\lambda_{N+1} = 0$ which yields

$$L = \sum_{i=1}^{N} l_i(x_i, u_i, \lambda_i, \lambda_{i+1}) + \lambda_1^T x_1^0$$

Next we define problem (P2) as

$$h(\lambda_1, \ldots, \lambda_N) = \left\{ \sum_{i=1}^{N} \underset{x_i, u_i}{\text{Min}}\ l_i(x_i, u_i, \lambda_i, \lambda_{i+1}) \right\} + \lambda_1^T x_1^0 \tag{P2}$$

and finally the general Lagrange problem (P3)

$$\underset{\lambda_1, \ldots, \lambda_N}{\text{Max}}\ h(\lambda_1, \ldots, \lambda_N) \tag{P3}$$

The two-level method involves solving (P3) by first guessing the multipliers $\lambda_1, \ldots, \lambda_N$, solving the subproblem minimizations in (P2), and then adjusting the multipliers until $h(\lambda_1, \ldots, \lambda_N)$ is maximized. If the point $(\hat{x}_1, \ldots, \hat{x}_N, \hat{u}_1, \ldots, \hat{u}_N; \lambda_1, \ldots, \lambda_N)$ is a saddle point of the Lagrangian, then the point $(\hat{x}_1, \ldots, \hat{x}_N, \hat{u}_1, \ldots, \hat{u}_N)$ solves problem (P1) and $(\lambda_1, \ldots, \lambda_N)$ solves problem (P3).

The two-level optimization procedure fails if no saddle point exists for the Lagrangian. Avery and Foss demonstrated that in a simple heat recovery network no such saddle point can exist. Their article shows that for their problem the stationary points for the sub-Lagrangians l_i in (P2) are never at a minimum with respect to the x_i and u_i. Under these conditions the actual minimum found in (P3) will typically be at extreme values, that is, on the constraints, of the x_i and u_i variables. Usually this result precludes any chance that the equality constraints will be satisfied for (P1), and thus solving (P3) will fail to solve (P1).

Let us define a new problem, which we shall call problem (P4), as follows:

$$\text{Min}\left\{ F^*|F^* = F + K\left\{ \sum_{i=1}^{N} [x_{i+1} - f_i]^T[x_{i+1} - f_i] \right\} + K(x_1 - x_1^0)^T(x_1 - x_1^0); x_{i+1} - f_i(x_i, u_i) = 0, \quad i = 1, \ldots, N; x_1 - x_1^0 = 0; K > 0 \right\} \text{(P4)}$$

The Lagrangian of this problem is

$$L^* = L + K \sum_{i=1}^{N} [x_{i+1} - f_i]^T[x_{i+1} - f_i] + K(x_1 - x_1^0)^T(x_1 - x_1^0)$$

Now,

$$\Delta L^* = \Delta L + K\left(\frac{\partial g}{\partial q^T}\Delta q\right)^T\left(\frac{\partial g}{\partial q^T}\Delta q\right)$$

where

$$q^T = [x_1^T, \ldots, x_N^T; u_1^T, \ldots, u_N^T]$$

and

$$g^T = [(x_1 - x_1^0)^T, (x_2 - f_1)^T, \ldots, (x_{N+1} - f_N)^T]$$

For permissible variations Δq (that is, satisfying the constraints) $\Delta L^* > 0$ always, and for nonpermissible variations where one might have $\Delta L < 0$, there is a $K > 0$ such that $\Delta L^* > 0$, and at a stationary point, L^* is always at a minimum. Therefore since a solution of (P4) is a solution to (P1), we conclude that the two-level optimization method can be applied to the chemical process design problems by solving this modified problem. The introduction of the quadratic terms destroys the separability of the system since when expanded they produce crossproduct terms $x_{i+1}{}^T f_i(x_i, u_i)$. However, one can recover a computational separability in trade for an iteration by expanding this term in a Taylor Series up to the first order terms (Stephanopoulos and Westerberg, 1973).

To provide a further insight in the two-level optimization procedure and its shortcomings, consider the following class of problems:

$$\text{Min}\left\{ F|F = \sum_{i=1}^{N} \phi_i(x_i, u_i), x_{i+1} - f_i(x_i, u_i) = z_{i+1}, \quad i = 1, \ldots, N, x_1 - x_1^0 = z_1 \right\} \quad \text{(P5)}$$

Let

$$w(z) = \underset{x_i, u_i, i=1,\ldots,N}{\text{Min}} \{F|x_{i+1} - f_i = z_{i+1}, \quad i = 1, \ldots, N, x_1 - x_1^0 = z_1\}$$

It has been shown (Everett, 1963) that if $\hat{x}_i, \hat{u}_i, i = 1, \ldots, N$ minimize the Lagrangian of a problem from the class of problems (P5), for given multipliers, then there exists a supporting hyperplane of the set

$$R = \{(z_0, z)|z_0 \geq w(z)\}$$

at the point $(w(z), z)$, where

$$w(z) = F(\hat{x}_1, \ldots, \hat{x}_N, \hat{u}_1, \ldots, \hat{u}_N) \quad \text{and}$$

$$\hat{x}_{i+1} - f_i(\hat{x}_i, \hat{u}_i) = z_{i+1}, i = 1, \ldots, N \quad \text{and} \quad \hat{x}_1 - x_1^0 = z_1$$

Figure 1 shows the failure of the two-level approach since there is no supporting hyperplane for the set R at the point $(w(0), 0)$. The introduction of the quadratic terms [see problem (P4)] causes $w(z)$ to move upwards and, under certain conditions (Stephanopoulos and Westerberg, 1973), the existence of a supporting hyperplane at the point $(w(0), 0)$ is guaranteed for finite K.

Let us now consider the heat recovery network described by Avery and Foss (1971). We shall use their nomenclature throughout the rest of this work. The cold stream D of their Figure 1 is to be heated from temperature T_0 to T_p. Three hot streams having flow rates a, b, c and temperatures t_a, t_b, t_c may be used for heating; here they are considered to have sufficient heat capacity and availability to accomplish the task. The distribution of the total heat load among the three exchangers is to be accomplished at the minimum cost of equipment which is related to the total heat transfer area by

$$C = \sum_{i=1}^{3} c_i = \sum_{i=1}^{3} \gamma_i A_i^{\alpha_i}$$

with γ and α positive and typically $\alpha = 0.6$. Let us decompose the process into three stages, shown in their Figure 2, and consider the following problem:

$$\text{Min } F^* = c_1(x_2) + c_2(y_1, y_2) + c_3(z_1) + K(x_2 - y_1)^2 + K(y_2 - z_1)^2$$

subject to

$$x_2 - y_1 = 0 \quad \text{and} \quad y_2 - z_1 = 0$$

The coupled terms x_2y_1 and y_2z_1 are expanded in a Taylor Series about the point $\hat{x}_2, \hat{y}_1, \hat{y}_2, \hat{z}_1$ and are approximated only by the linear terms in the expansion. The Lagrangian of the above problem becomes

$$\begin{aligned} L^* \simeq & \{c_1(x_2) - (\lambda_1 + 2K\hat{y}_1)x_2 + Kx_2^2 + K\hat{x}_2\hat{y}_1\} \\ & + \{c_2(y_1, y_2) + (\lambda_1 - 2K\hat{x}_2)y_1 - (\lambda_2 + 2K\hat{z}_1)y_2 \\ & + Ky_1^2 + Ky_2^2 + K\hat{x}_2\hat{y}_1 + K\hat{y}_2\hat{z}_1\} \\ & + \{c_3(z_1) + (\lambda_2 - 2K\hat{y}_2)z_1 + Kz_1^2 + K\hat{y}_2\hat{z}_1\} \\ & = l_1 + l_2 + l_3 \end{aligned}$$

The necessary conditions for the minimum of C are

$$\nabla_v L^* = 0 \quad \text{and} \quad \nabla_\lambda L^* = 0$$

where

$$v = [x_2, y_1, y_2, z_1]^T \quad \text{and} \quad \lambda = (\lambda_1, \lambda_2)^T$$

This results in the following equations:

$$\frac{\partial l_1}{\partial x_2} = 0, \quad \frac{\partial l_2}{\partial y_1} = \frac{\partial l_2}{\partial y_2} = 0, \quad \frac{\partial l_3}{\partial z_1} = 0, \quad x_2 = y_1, \quad \text{and} \quad y_2 = z_1$$

that is, each subsystem is at a stationary point with respect to the variables associated with it. The sufficient condition that a stationary point is minimum is that the Hessian matrix be positive definite at that point. Put $Q = y_2 - y_1$, the heat load in unit 2, and take

$$l_2 = c_2(y_1, Q) + (\lambda_1 - \lambda_2 - 2K\hat{x}_2 - 2K\hat{z}_1)y_1 - (\lambda_2 + 2K\hat{z}_1)Q + 2Ky_1^2 + KQ^2 + 2KQy_1$$

The Hessian matrix of l_2 is

$$H_2 = \begin{Bmatrix} \frac{\partial^2 c_2}{\partial Q^2} + 2K & \frac{\partial^2 c_2}{\partial Q \partial y_1} + 2K \\ \frac{\partial^2 c_2}{\partial y_1 \partial Q} + 2K & \frac{\partial^2 c_2}{\partial y_1^2} + 4K \end{Bmatrix} = \begin{Bmatrix} d_1^* & e^* \\ e^* & d_2^* \end{Bmatrix}$$

Comparing this result to Avery and Foss, Equation (8), we see that

$$d_1^* = d_1 + 2K, \quad d_2^* = d_2 + 4K, \quad e^* = e + 2K$$

Then

$$\Delta^* \equiv 4K^2 + (4d_1 + 2d_2 - 4e)K + (d_1d_2 - e^2) = 4K^2 + \Gamma_2 K + \Gamma_3$$

Δ^* has roots

$$K = \frac{-\Gamma_2 \pm \sqrt{\Gamma_2^2 - 16\Gamma_3}}{8}$$

Avery and Foss, Equation (14) show that

$$\Gamma_3 = d_1d_2 - e^2 = -q^2p^2s^2(r-1)^2 \leq 0$$

thus

$$|\sqrt{\Gamma_2^2 - 16\Gamma_3}| \geq |\Gamma_2|$$

and the roots are real and of opposite sign. Let $K_1 \geq 0$ and $K_2 \leq 0$, then for $K > K_1$, $\Delta^* > 0$. Also for

$$K > -d_1/2 \equiv K_3 \quad d_1^* > 0$$

and for

$$K > -d_2/2 \equiv K_4 \quad d_2^* > 0$$

Therefore, if $K = \max\{K_1, K_3, K_4\}$ we find $H_2 > 0$ and the stationary point for unit 2 is a minimum as required. An algorithmic procedure, which implements numerically the previous approach to nonconvex problems, is given in Stephanopoulos and Westerberg (1973).

Thus we see that curvature of the sub-Lagrangian at the stationary point can be determined not only by the subunit cost function, c_2 in this case, but by the numerical value of the penalty factor K. It has been proved that under certain conditions which are commonly applicable in chemical engineering process design problems, K takes finite values to ensure convergence to the solution of the original problem. Therefore, by properly choosing a K we can secure that the stationary point for each subunit is a minimum and thus the stationary point of the Lagrangian L^* is a saddle point as required by the multilevel optimization method.

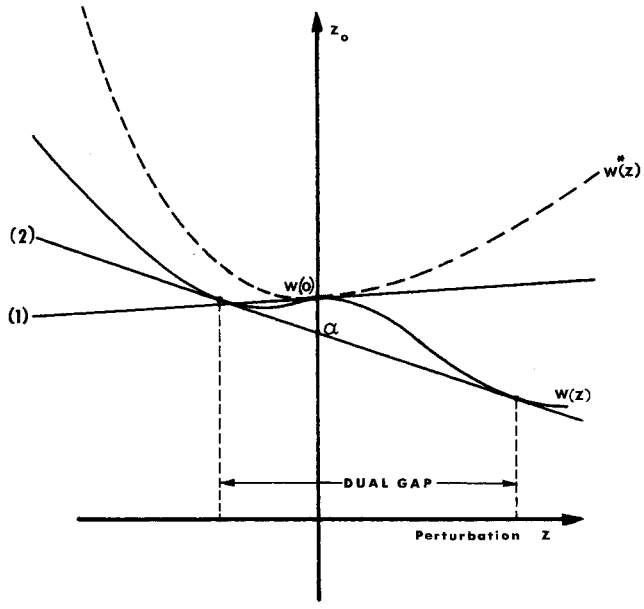

Fig. 1. Problem with a dual gap and its resolution. Line (1) is the supporting hyperplane for the set R^* (associated with w^* for the penalized problem) at the solution point. Point 'α' ($< w(o) = w^*(o)$) is the geometric location of $\max h(\lambda)$ for the unpenalized problem, and line (2) is the corresponding best supporting hyperplane for that problem.

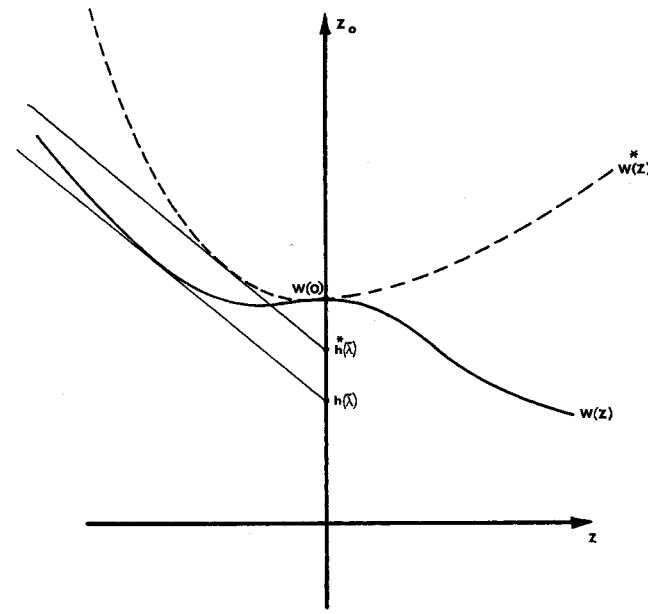

Fig. 2. Improvement of the dual bound $h^*(\tilde{\lambda})$ for the penalized problem over the dual bound $h(\tilde{\lambda})$ for the unpenalized problem for the same multipliers $\tilde{\lambda}$.

The introduction of the quadratic terms in the objective function offers another advantage. It desensitizes the dual function with respect to the multipliers λ. For given multipliers $\tilde{\lambda}$ the $h^*(\tilde{\lambda}) = \underset{x_i,u_i,i=1,\ldots,N}{\text{Min}} L^*$ is closer to the primal solution than the $h(\tilde{\lambda}) = \underset{x_i,u_i,i=1,\ldots,N}{\text{Min}} L$ (see Fig. 2). This is a characteristic of great importance for a dual bounding procedure such as the one used in structural sensitivity analysis (McGalliard and Westerberg, 1972).

The nonconvexity of the process design problems in chemical engineering is a very usual phenomenon, since the cost function typically includes a term $x^{0.6}$ where x is the throughput in a unit. The proposed approach helps to resolve this difficulty and makes the two-level optimization method applicable to such problems.

ACKNOWLEDGMENT

This work was supported by NSF Grant GK18633.

LITERATURE CITED

Avery, C. J., and A. S. Foss, "A Shortcoming of the Multilevel Optimization Technique," *AIChE J.*, 17, 998 (1971).

Brosilow, C., and L. Lasdon, "A Two-Level Optimization Technique for Recycle Processes," AIChE-IChE Joint Meeting, London, Section 4, 75 (1965).

Brosilow, C., and E. Nunez, "Multilevel Optimization Applied to a Catalytic Cracking Plant," *Can. J. Chem. Eng.*, 64, 205 (1968).

Everett, H., III, "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources," *Operation Res.*, 11, 399 (1963).

McGalliard, R. L., and A. W. Westerberg, "Structural Sensitivity Analysis in Design Synthesis," *Chem. Eng. J.*, 4, 127 (1972).

Stephanopoulos, G., and A. W. Westerberg, "The Use of Hestenes' Method of Multipliers to Resolve Dual Gaps in Engineering System Optimization," *J. Optimization Theory Applic.*, in press.

Manuscript received June 25, 1973; revision received August 13 and accepted August 14, 1973.