SRC TR 89-6

Optimality and Constrained Derivatives in Two-Level Design Optimation

by

S. Azarm

and

W. -C. Li

OPTIMALITY AND CONSTRAINED DERIVATIVES IN TWO-LEVEL DESIGN OPTIMIZATION

S. Azarm

and

W.-C. Li

Department of Mechanical Engineering

and

Systems Research Center
The University of Maryland
College Park, Maryland 20742

Abstract

The objective of this paper is twofold. First, an optimality test is presented to show that the optimality conditions for a two-level design optimization problem before and after its decomposition are the same. Second, based on identification of active constraints and exploitation of problem structure, a simple approach for calculating the gradient of a "second-level" problem is presented. This gradient is an important piece of information which is needed for solution of two-level design optimization problems. Three examples are given to demonstrate applications of the approach.

1. Introduction

Two-level design optimization problems can in general be formulated in the following integrated (or undecomposed) form (Azarm and Li, 1988a):

$$\begin{aligned}
&\underset{\underset{\sim}{y},\, \underset{\sim}{x}}{\text{Minimize}}\ f(\underset{\sim}{y}, \underset{\sim}{x}) = f_0(\underset{\sim}{y}) + \sum_{i=1}^{I} f_i(\underset{\sim}{y}, \underset{\sim}{x}_i) \\
&\text{subject to:} && (1) \\
&g_\ell(\underset{\sim}{y}) \le 0, && \ell = 1, \ldots, L \\
&g_{i,j}(\underset{\sim}{y}, \underset{\sim}{x}_i) \le 0\ , && i = 1, \ldots, I \\
& && j = 1, \ldots, J
\end{aligned}$$

where f is an integrated (overall) objective function, f_i is an objective function in subproblem i, g_ℓ is a constraint in a second-level problem, $g_{i,j}$ is a constraint in a first-level subproblem i, $\underset{\sim}{x}_i$ is a N-vector of design variables in a first-level subproblem i, $\underset{\sim}{y}$ is a T-vector of design variables in a second-level problem, i is an index corresponding to the number of first-level subproblems, j is an index corresponding to the number of constraints in a first-level subproblem i, and ℓ is an index corresponding to the number of constraints in a second-level problem. The formulation for a first-level subproblem i, i=1, ..., I, is:

$$\begin{aligned}
&\underset{\underset{\sim}{x}_i}{\text{Minimize}}\ f_i(\underset{\sim}{y}, \underset{\sim}{x}_i) \\
&\text{subject to:} && (2) \\
&g_{i,j}(\underset{\sim}{y}, \underset{\sim}{x}_i) \le 0, \quad j = 1, \ldots, J
\end{aligned}$$

where $\underset{\sim}{y}$ is fixed (found from the second-level problem) and $\underset{\sim}{x}_i$ is varied. Also, the formulation for the second-level problem is:

$$\underset{\underset{\sim}{y}}{\text{Minimize}}\ f(\underset{\sim}{y}, \underset{\sim}{x}) = f_0(\underset{\sim}{y}) + \sum_{i=1}^{I} f_i(\underset{\sim}{y}, \underset{\sim}{x}_i)$$

$$\text{subject to:} \tag{3}$$

$$g_\ell(\underset{\sim}{y}) \le 0, \quad \ell = 1,\ldots, L$$

where f is the second-level objective function, $\underset{\sim}{y}$ is varied and $\underset{\sim}{x}_i$ is fixed (found from the first-level subproblem i, i=1,..., I). The two-level problems, eqs. (2) and (3), can in general be solved by the following solution procedure (Kirsch, 1981):

1. choose an initial value for $\underset{\sim}{y}$,
2. for a given $\underset{\sim}{y}$, solve the first-level subproblem i for $\underset{\sim}{x}_i$, i=1,..., I,
3. modify $\underset{\sim}{y}$ so that f is reduced,
4. repeat steps (2) and (3) until f is minimized.

The aforementioned two-level formulation and solution procedure is useful when the integrated problem, eq. (1), must by its very nature be decomposed (Lasdon, 1970) into eqs. (2) and (3) or because of its size requires distributed or parallel processing capabilities (Lootsma and Ragsdell, 1988).

There exists a variety of solution procedures for handling eqs. (2) and (3) (Azarm and Li, 1988b; Haftka, 1984; Sobieski, 1982; Vanderplaats and Kim, 1988). It is important that a given procedure preserves the integrity of optimality conditions for the undecomposed problem, eq. (1). This means that an optimality test should be performed to see whether the optimality conditions for the undecomposed problem is the same as the ones for the decomposed subproblems, eqs. (2) and (3). Furthermore, it is crucial that the second-level problem is optimized as efficiently as possible. One important piece of information which is needed for optimization of the second-level problem, eq. (3), is the gradient of the second-level objective function with respect to its design variables, i.e., $\partial f/\partial \underset{\sim}{y}$. It will be shown here that by exploiting the problem structure, eqs. (2) and (3), and identification of active constraints in the first-level subproblems, an effective approach for calculating this gradient can be devised. This approach is based on a first-order gradient information, and does not require difficult to obtain optimum sensitivities (a second-order information) of first-level subproblems. Furthermore, as it will be demonstrated in the examples, the approach can be used to solve two-level design optimization problems analytically (i.e., to obtain a global optimum).

2. Optimality Test

It is important in two-level optimization that, after decomposition, the integrity of the overall optimality conditions for the undecomposed problem is preserved. In other words, we should be able to show that the optimality conditions for the two-level decomposed subproblems, eqs. (2) and (3), will constitute the ones corresponding to the undecomposed problem, eq. (1). Here, we assume that the objective and constraint functions are continuously differentiable. Furthermore, the regularity assumption of the points under consideration is that the gradient vectors of active constraints are linearly independent at the regular points.

Let $(\bar{\underset{\sim}{y}}, \bar{\underset{\sim}{x}})^t$ be a solution (a local minimum) of the undecomposed problem, eq. (1), where:

$$\bar{\underset{\sim}{y}} = (\bar{y}_1, \ldots, \bar{y}_t, \ldots, \bar{y}_T)^t \tag{4}$$

$$\bar{\underset{\sim}{x}} = (\bar{\underset{\sim}{x}}_1, \ldots, \bar{\underset{\sim}{x}}_i, \ldots, \bar{\underset{\sim}{x}}_I)^t \tag{5}$$

and $\bar{\underset{\sim}{x}}_i$ is a N-vector of variables in subproblem i:

$$\bar{\underset{\sim}{x}}_i = (\bar{x}_{i,1}, \ldots, \bar{x}_{i,N})^t \tag{6}$$

The Karush-Kuhn-Tucker (KKT) optimality conditions for the first-level subproblem i, eq. (2), can be written as follows (all evaluations are performed at $(\bar{\underset{\sim}{y}}, \bar{\underset{\sim}{x}})^t$):

$$\partial f_i/\partial \underset{\sim}{x}_i + (\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i)\, \underset{\sim}{u}_i = \underset{\sim}{0} \tag{7}$$

$$u_{i,j}\, g_{i,j} = 0, \qquad j = 1, \ldots, J \tag{8}$$

$$\underset{\sim}{u}_i \geq \underset{\sim}{0} \tag{9}$$

where, $\partial f_i/\partial \underset{\sim}{x}_i$ is a NX1 vector and $\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i$ is a NXJ matrix. Also, $\underset{\sim}{u}_i$ or $\underset{\sim}{g}_i$ is a JX1 vector:

$$\partial f_i/\partial \underset{\sim}{x}_i = (\partial f_i/\partial x_{i,1}, \ldots, \partial f_i/\partial x_{i,N})^t \tag{10}$$

$$\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i = (\partial \underset{\sim}{g}_i/\partial x_{i,1}, \ldots, \partial \underset{\sim}{g}_i/\partial x_{i,N})^t \tag{11}$$

$$\underset{\sim}{u}_i = (u_{i,1}, \ldots, u_{i,j}, \ldots, u_{i,J})^t \tag{12}$$

$$\underset{\sim}{g}_i = (g_{i,1}, \ldots, g_{i,j}, \ldots, g_{i,J})^t \tag{13}$$

Combining eqs. (7)-(9) for all of the first-level subproblems result in the KKT conditions for all the first-level subproblems:

$$(\partial f_i/\partial \underset{\sim}{x}_i) + (\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i)\, \underset{\sim}{u}_i = \underset{\sim}{0}, \qquad i = 1, \ldots, I \tag{14}$$

$$u_{i,j}\, g_{i,j} = 0, \qquad i=1,\dots, I, \qquad j = 1,\dots, J \tag{15}$$

$$\underset{\sim}{u}_i \geq \underset{\sim}{0} \qquad i = 1,\dots, I \tag{16}$$

Likewise, the KKT conditions for the second-level problem, eq. (3), can be written as follows (all evaluations are performed at $(\bar{\underset{\sim}{y}}, \bar{\underset{\sim}{x}})^t$):

$$\partial f/\partial \underset{\sim}{y} + \sum_{\ell=1}^{L} u_\ell(\partial g_\ell/\partial \underset{\sim}{y}) = \underset{\sim}{0} \tag{17}$$

$$u_\ell\, g_\ell = 0, \qquad \ell = 1,\dots, L \tag{18}$$

$$u_\ell \geq 0 \tag{19}$$

where:

$$\partial f/\partial \underset{\sim}{y} = \partial f_0/\partial \underset{\sim}{y} + \sum_{i=1}^{I} (\partial f_i/\partial \underset{\sim}{y}) + \sum_{i=1}^{I} (\partial \underset{\sim}{x}_i/\partial \underset{\sim}{y})(\partial f_i/\partial \underset{\sim}{x}_i) \tag{20}$$

in which $\partial f/\partial \underset{\sim}{y}$, $\partial f_0/\partial \underset{\sim}{y}$, and $\partial f_i/\partial \underset{\sim}{y}$ are TX1 vectors, and $\partial \underset{\sim}{x}_i/\partial \underset{\sim}{y}$ is a TXN matrix:

$$(\partial \underset{\sim}{x}_i/\partial \underset{\sim}{y}) = (\partial \underset{\sim}{x}_i/\partial y_1,\dots,\partial \underset{\sim}{x}_i/\partial y_T)^t \tag{21}$$

Suppose in subproblem i, i=1,..., I, there exists S active constraints ($S \leq N$) at the optimum. Without loss of generality, suppose that these active constraints are the first S constraints in subproblem i so that in a vector form they can be shown as:

$$\underset{\sim}{g}_i = \underset{\sim}{0} \tag{22}$$

Here, $\underset{\sim}{g}_i$ represents the S active constraints in subproblem i:

$$\underset{\sim}{g}_i = (g_{i,1},\dots, g_{i,S})^t \tag{23}$$

Then, for every feasible neighborhood for which the active constraints are unchanged, we should have from eq. (22):

$$\partial \underset{\sim}{g}_i = \underset{\sim}{0} \tag{24}$$

From which, to a first-order approximation, we should have:

$$(\partial \underset{\sim}{g}_i/\partial \underset{\sim}{y}) + (\partial \underset{\sim}{x}_i/\partial \underset{\sim}{y})(\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i) = \underset{\sim}{0} \tag{25}$$

or

$$(\partial \underset{\sim}{g}_i/\partial \underset{\sim}{y}) = -(\partial \underset{\sim}{x}_i/\partial \underset{\sim}{y})(\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i) \tag{26}$$

where

$$(\partial \underset{\sim}{g}_i/\partial \underset{\sim}{y}) = (\partial g_{i,1}/\partial \underset{\sim}{y}, \ldots, \partial g_{i,S}/\partial \underset{\sim}{y}) \tag{27}$$

and

$$(\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i) = (\partial g_{i,1}/\partial \underset{\sim}{x}_i, \ldots, \partial g_{i,S}/\partial \underset{\sim}{x}_i) \tag{28}$$

$$(\partial \underset{\sim}{x}_i/\partial \underset{\sim}{y}) = (\partial x_{i,1}/\partial \underset{\sim}{y}, \ldots, \partial x_{i,N}/\partial \underset{\sim}{y}) \tag{29}$$

Also from eq. (7), for the S active constraints in subproblem i, we have:

$$\partial f_i/\partial \underset{\sim}{x}_i = -(\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i)\, \underset{\sim}{u}_i \tag{30}$$

Considering eqs. (26) and (30), eq. (20) can be written as:

$$\partial f/\partial \underset{\sim}{y} = \partial f_0/\partial \underset{\sim}{y} + \sum_{i=1}^{I} (\partial f_i/\partial \underset{\sim}{y}) + \sum_{i=1}^{I} (\partial \underset{\sim}{g}_i/\partial \underset{\sim}{y})\, \underset{\sim}{u}_i \tag{31}$$

Equations (14)-(19) together with eq. (31) are, in fact, the KKT conditions of the undecomposed problem, eq. (1). Note that in eq. (31), $\underset{\sim}{g}_i$ may also include the inactive constraints ($g_{i,j} < 0$), in which case the corresponding Lagrange multipliers should be zero ($u_{i,j} = 0$).

3. Constrained Derivatives

The method of constrained derivatives within single-level design optimization was introduced by Wilde and Beightler (1967). Abadie and Carpentier (1969) referred to a similar method as generalized reduced gradient. This method essentially depends upon calculation of the gradient of objective function in the subspace of active constraints. An active constraint refers to a constraint which has a direct effect on the location of the optimum. For an inequality constraint this means that it should be satisfied as an equality at the optimum.

In two-level design optimization, the method of constrained derivatives can be used to calculate an important piece of information, namely, the gradient of the second-level objective function ($\partial f/\partial \underset{\sim}{y}$) which is also called the second-level constrained-derivatives. It will be shown here that this gradient can be obtained without explicit calculations of Lagrange multipliers ($\underset{\sim}{u}_i$) or optimum sensitivities ($\partial \underset{\sim}{x}_i/\partial \underset{\sim}{y}$) of first-level subproblems.

The argument is based on the assumption that in the neighborhood of a minimum solution the active constraints of subproblem i are unchanged, the so-called Jacobian uniqueness conditions (Lootsma and Ragsdell, 1988). Suppose that the number of active constraints in subproblem i are S ($S \leq N$). We then partition the N variables in subproblem i into two groups, namely, the dependent or solution variables ($\underset{\sim}{x}_i^S$), and the independent or decision variables ($\underset{\sim}{x}_i^D$). In subproblem i, without loss of generality, the first group which has S components is selected as the solution variables and the second group which has D components is selected as the decision variables (i.e., N = S+D):

$$\underset{\sim}{x}_i = \begin{cases} \underset{\sim}{x}_i^S & \text{S solution variables} \\ \underset{\sim}{x}_i^D & \text{D decision variables} \end{cases} \tag{32}$$

Following this partitioning, eq. (7) can be written in the following form for the S active constraints:

$$\partial f_i/\partial \underset{\sim}{x}_i^S + (\partial \underset{\approx}{g}_i/\partial \underset{\sim}{x}_i^S)\ \underset{\sim}{u}_i = \underset{\sim}{0} \tag{33}$$

and

$$\partial f_i/\partial \underset{\sim}{x}_i^D + (\partial \underset{\approx}{g}_i/\partial \underset{\sim}{x}_i^D)\ \underset{\sim}{u}_i = \underset{\sim}{0} \tag{34}$$

where $\partial f_i/\partial \underset{\sim}{x}_i^S$ is a SX1 vector, $\partial f_i/\partial \underset{\sim}{x}_i^D$ is a DX1 vector, $\partial \underset{\approx}{g}_i/\partial \underset{\sim}{x}_i^S$ is a SXS matrix, $\partial \underset{\approx}{g}_i/\partial \underset{\sim}{x}_i^D$ is a DXS matrix, and $\underset{\sim}{u}_i$ is a SX1 vector. Note that the selection of solution and decision variables are arbitrary as long as linear independence in the matrix $\partial \underset{\approx}{g}_i/\partial \underset{\sim}{x}_i^S$ is preserved (Wilde and Beightler, 1967). From eq. (33), we have:

$$\underset{\sim}{u}_i = -(\partial \underset{\approx}{g}_i/\partial \underset{\sim}{x}_i^S)^{-1}\ (\partial f_i/\partial \underset{\sim}{x}_i^S) \tag{35}$$

which can be substituted into eq. (31) to obtain:

$$\partial f/\partial \underset{\sim}{y} = \partial f_o/\partial \underset{\sim}{y} + \sum_{i=1}^{I} (\partial f_i/\partial \underset{\sim}{y})$$

$$- \sum_{i=1}^{I} [(\partial \underset{\approx}{g}_i/\partial \underset{\sim}{y})(\partial \underset{\approx}{g}_i/\partial \underset{\sim}{x}_i^S)^{-1}(\partial f_i/\partial \underset{\sim}{x}_i^S)] \tag{36}$$

Once the active constraints in subproblem i, i=1,..., I, have been identified, then eq. (36) can be used to obtain $\partial f/\partial \underset{\sim}{y}$. Likewise, eq. (36) can be used when the first-level subproblems are solved by an optimization method which does not yield the $\underset{\sim}{u}_i$ values. It appears that this approach is computationally less expensive than the one in which $\underset{\sim}{u}_i$ is calculated by (Sobieski et al. 1982):

$$\underset{\sim}{u}_i = -[(\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i)^t(\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i)]^{-1} (\partial g_i/\partial \underset{\sim}{x}_i)^t (\partial f_i/\partial \underset{\sim}{x}_i) \tag{37}$$

where $\partial \underset{\sim}{g}_i/\partial \underset{\sim}{x}_i$ is a NXS matrix.

4. Examples

Three examples with increasing degree of difficulty will be presented to demonstrate the applications of eq. (36).

4.1. Example 1 - A Two-Bar Truss Problem

This simple example is selected from the literature (Kirsch, 1981). It is a two-bar truss problem (Fig. 1) subject to a single vertical load of 100 kN at joint C. The variables are the cross-sectional areas of the bars, x_1, x_2, and the vertical coordinate of the joint, y. The constraints are: an upper-limit of 100,000 kN/m^2 for the stress, the interval 1.0-3.0 m for y, and a positive value for the cross-sectional area. For a minimum-volume criterion, the problem is formulated as follows:

$$\begin{aligned}
&\text{Minimize } f(y, \underset{\sim}{x}) = x_1 (16 + y^2)^{1/2} + x_2 (1 + y^2)^{1/2} \\
&\text{subject to:} \\
&20(16 + y^2)^{1/2} - 100{,}000\ yx_1 \le 0 \\
&80(1 + y^2)^{1/2} - 100{,}000\ yx_2 \le 0 \\
&1 \le y \le 3 \\
&(x_1, x_2)^t > \underset{\sim}{0}
\end{aligned} \tag{38}$$

The problem is decomposed into two levels:

First-Level Subproblem 1: Find x_1 (y is fixed)

$$\underset{x_1}{\text{Minimize}}\ f(y, x_1) = x_1\,(16 + y^2)^{1/2}$$

subject to: (39)

$g_{1,1}$: $20(16 + y^2)^{1/2} - 100{,}000\, yx_1 \le 0$

$x_1 > 0$

First-Level Subproblem 2: Find x_2 (y is fixed)

$$\underset{x_2}{\text{Minimize}}\ f_2(y, x_2) = x_2(1 + y^2)^{1/2}$$

subject to: (40)

$g_{2,1}$: $80(1 + y^2)^{1/2} - 100{,}000\, yx_2 \le 0$

$x_2 > 0$

Second-Level Problem: Find y (x_1 and x_2 are fixed)

$$\underset{y}{\text{Minimize}}\ f(y, \underset{\sim}{x}) = f_1(y, x_1) + f_2(y, x_2)$$

subject to: (41)

$1 \le y \le 3$

In subproblems 1 and 2, it can be easily verified that constraints $g_{1,1}$ and $g_{2,1}$ should be active (for example, by the monotonicity analysis (Papalambros and Wilde, 1988)) at the optimum, so that:

$$x_1 = 20(16 + y^2)^{1/2}/(100{,}000\ y) \tag{42}$$

$$x_2 = 80(1 + y^2)^{1/2}/(100{,}000\ y) \tag{43}$$

Assuming that $x^S_{1,1} = x_1$ and $x^S_{2,1} = x_2$ to be the solution variables in subproblems 1 and 2 (note that there is no decision variable in the subproblems). Then from eq. (36) we have:

$$\partial f/\partial y = \sum_{i=1}^{2} (\partial f_i/\partial y) - \sum_{i=1}^{2} [(\partial g_{i,1}/\partial y)(\partial g_{i,1}/\partial x^S_{i,1})^{-1}(\partial f_i/\partial x^S_{i,1})] \tag{44}$$

where:

$$\partial f_1/\partial y = yx_1/(16 + y^2)^{1/2} \tag{45}$$

$$\partial f_2/\partial y = yx_2/(1 + y^2)^{1/2} \tag{46}$$

$$\partial g_{1,1}/\partial y = 20y/(16 + y^2)^{1/2} - 100{,}000x_1 \tag{47}$$

$$\partial g_{2,1}/\partial y = 80y/(1 + y^2)^{1/2} - 100{,}000x_2 \tag{48}$$

$$\partial g_{1,1}/\partial x^S_{1,1} = -100{,}000y \tag{49}$$

$$\partial g_{2,1}/\partial x^S_{2,1} = -100{,}000y \tag{50}$$

$$\partial f_1/\partial x^S_{1,1} = (16 + y^2)^{1/2} \tag{51}$$

$$\partial f_2/\partial x^S_{2,1} = (1 + y^2)^{1/2} \tag{52}$$

Considering eqs. (42) and (43), we substitute the right-hand sides of eqs. (45)-(52) into eq. (44) to obtain:

$$\partial f/\partial y = \frac{1}{1000}(y^2 - 4) \tag{53}$$

If we set $\partial f/\partial y = 0$, then $y^*=2$ will be the global optimum solution which is within $1 \le y \le 3$. From eqs. (42) and (43), we obtain $(x_1^*, x_2^*) = (4.48, 8.96)10^{-4}$. This solution is identical to the one reported by Kirsch (1981).

4.2. Example 2 - A Modification to Example 1

Here we have made a modification to the two-bar truss problem to demonstrate a case in which there is a decision variable (in addition to a solution variable) in each subproblem. Variables x_3 and x_4 are introduced into the problem so that this example is formulated as follows:

$$\text{Minimize } f(y, \underset{\sim}{x}) = x_1(x_3^2 + y^2)^{1/2} + x_2(x_4^2 + y^2)^{1/2} - 0.01\,(x_3 + x_4)$$

subject to: (54)

$$20(x_3^2 + y^2)^{1/2} - 100{,}000\ yx_1 \le 0$$

$$80(x_4^2 + y^2)^{1/2} - 100{,}000\ yx_2 \le 0$$

$$1 \le y \le 3$$

$$(x_1, x_2, x_3, x_4)^t > \underset{\sim}{0}$$

Again, the problem is decomposed into two-levels:

First-Level Subproblem 1: Find x_1 and x_3 (y is fixed):

$$\begin{aligned} &\underset{x_1, x_3}{\text{Minimize}}\ f_1(y, x_1, x_3) = x_1(x_3^2 + y^2)^{1/2} - 0.01\ x_3 \\ &\text{subject to:} \\ &g_{1,1}: \quad 20(x_3^2 + y^2)^{1/2} - 100{,}000\ yx_1 \le 0 \\ &\qquad\quad (x_1, x_3)^t > \underset{\sim}{0} \end{aligned} \tag{55}$$

First-Level Subproblem 2: Find x_2 and x_4 (y is fixed)

$$\begin{aligned} &\underset{x_2, x_4}{\text{Minimize}}\ f_2(y, x_2, x_4) = x_2(x_4^2 + y^2)^{1/2} - 0.01\ x_4 \\ &\text{subject to:} \\ &g_{2,1}: \quad 80(x_4^2 + y^2)^{1/2} - 100{,}000\ yx_2 \le 0 \\ &\qquad\quad (x_2, x_4)^t > \underset{\sim}{0} \end{aligned} \tag{56}$$

Second-Level Problem: Find y (x_1, x_2, x_3, and x_4 are fixed)

$$\begin{aligned} &\underset{y}{\text{Minimize}}\ f(y, \underset{\sim}{x}) = f_1(y, x_1, x_3) + f_2(y, x_2, x_4) \\ &\text{subject to:} \\ &1 \le y \le 3 \end{aligned} \tag{57}$$

Again, in subproblems 1 and 2 it can be easily verified that constraints $g_{1,1}$ and $g_{2,1}$ are active, respectively. Also, in subproblem 1, $x_{1,1}^S = x_1$ and $x_{1,1}^D = x_3$ are selected as the solution and decision variables, respectively.

Likewise, in subproblem 2, $x^S_{2,1} = x_2$ and $x^D_{2,1} = x_4$ are selected as the solution and decision variables, respectively. Furthermore, for a given y, we can solve subproblems 1 and 2 to obtain:

$$x^S_{1,1} = (626)^{1/2}/5000 \tag{58}$$

$$x^S_{2,1} = (646)^{1/2}/5000 \tag{59}$$

$$x^D_{1,1} = 25y \tag{60}$$

$$x^D_{2,1} = 25y/4 \tag{61}$$

Also, considering eqs. (58)-(61), we have:

$$\partial f_1/\partial y = (20) \times 10^{-5} \tag{62}$$

$$\partial f_2/\partial y = (80) \times 10^{-5} \tag{63}$$

$$(\partial g_{1,1}/\partial y)(\partial g_{1,1}/\partial x^S_{1,1})^{-1} (\partial f_1/\partial x^S_{1,1}) = (12500) \times 10^{-5} \tag{64}$$

$$(\partial g_{2,1}/\partial y)(\partial g_{2,1}/\partial x^S_{2,1})^{-1} (\partial f_2/\partial x^S_{2,1}) = (3125) \times 10^{-5} \tag{65}$$

substitute eq. (62)-(65) into eq. (44) to obtain:

$$\partial f/\partial y = -(15525) \times 10^{-5} < 0 \tag{66}$$

Therefore, by applying the monotonicity analysis (Papalambros and Wilde, 1988) on the second-level problem we obtain $y^* = 1$, and from eqs. (58)-(61): $(x_1^*, x_2^*, x_3^*, x_4^*) = ((626)^{1/2}/5000, (646)^{1/2}/5000, 25, 25/4)$.

4.3. Example 3 - A Gear-Reducer

In this section, we present a well-known gear reducer example, Figure 2, which was first formulated by Golinski (1970) and solved by several optimization methods (Azarm and Li, 1988b; Li and Papalambros, 1985). Here we present the final design optimization model in which the design objective is to minimize the overall volume (or weight). The design variables for the example are as follows:

x_1 = gear face width (cm)

x_2 = teeth module (cm)

x_3 = number of teeth of pinion

x_4 = distance between bearings 1 (cm)

x_5 = distance between bearings 2 (cm)

x_6 = diameter of shaft 1 (cm)

x_7 = diameter of shaft 2 (cm)

And, the constraints are as follows:

g_1	: Upper bound on the bending stress of the gear tooth.
g_2	: Upper bound on the contact stress of the gear tooth.
g_3-g_4	: Upper bounds on the transverse deflection of the shaft.
g_5-g_6	: Upper bounds on the stresses of the shaft.
g_7-g_{23}	: Dimensional restrictions based on space and/or experience.
g_{24}-g_{25}	: Design condition for the shaft based on experience.

Finally, the formulation for this example is presented (Azarm and Li, 1988b):

$$\text{Minimize } f(x) = 0.7854x_1x_2^2(3.3333x_3^2 + 14.9334x_3 - 43.0934) - 1.508x_1(x_6^2 + x_7^2)$$

$$+ 7.477(x_6^3 + x_7^3) + 0.7854(x_4x_6^2 + x_5x_7^2)$$

subject to: (67)

g_1: $27x_1^{-1}x_2^{-2}x_3^{-1} \le 1$

g_2: $397.5x_1^{-1}x_2^{-2}x_3^{-2} \le 1$

g_3: $1.93x_2^{-1}x_3^{-1}x_4^3x_6^{-4} \le 1$

g_4: $1.93x_2^{-1}x_3^{-1}x_5^3x_7^{-4} \le 1$

g_5: $A_1/B_1 \le 1100$

$$A_1 = [(\frac{745x_4}{x_2x_3})^2 + (16.9)10^6]^{0.5}$$

$$B_1 = 0.1x_6^3$$

g_6: $A_2/B_2 \le 850$

$$A_2 = [(\frac{745x_5}{x_2x_3})^2 + (157.5)10^6]^{0.5}$$

$$B_2 = 0.1x_7^3$$

g_7: $x_2x_3 \le 40$

g_8: $5 \le x_1/x_2 \le 12$: g_9

g_{10}: $2.6 \le x_1 \le 3.6$: g_{11}

g_{12}: $0.7 \le x_2 \le 0.8$: g_{13}

g_{14}: $17 \le x_3 \le 28$: g_{15}

g_{16}: $7.3 \le x_4 \le 8.3$: g_{17}

g_{18}: $7.3 \le x_5 \le 8.3$: g_{19}

g_{20}: $2.9 \le x_6 \le 3.9$: g_{21}

g_{22}: $5.0 \le x_7 \le 5.5$: g_{23}

g_{24}: $(1.5x_6 + 1.9)x_4^{-1} \le 1$

g_{25}: $(1.1x_7 + 1.9)x_5^{-1} \le 1$.

4.3.1. Two-Level Decomposition

The gear reducer considered here, Figure 2, consists of two subsystems, namely, shaft and bearings 1 and shaft and bearings 2. These two subsytems are selected to correspond to subproblems 1 and 2, respectively.

First-Level Subproblem 1: Find x_4 and x_6 (x_1, x_2, and x_3 are fixed)

$$\underset{x_4,x_6}{\text{Minimize}}\ f_1(x_1,x_2,x_3,x_4,x_6) = -1.508x_1x_6^2 + 7.477x_6^3 + 0.7854x_4x_6^2$$

subject to: (68)

g_3: $1.93x_2^{-1}x_3^{-1}x_4^{3}x_6^{-4} \le 1$

g_5: $A_1/B_1 \le 1100$

g_{16}: $7.3 \le x_4 \le 8.3$:g_{17}

g_{20}: $2.9 \le x_6 \le 3.9$:g_{21}

g_{24}: $(1.5x_6 + 1.9)x_4^{-1} \le 1$

First-Level Subproblem 2: Find x_5 and x_7 (x_1, x_2, and x_3 are fixed)

$$\underset{x_5,x_7}{\text{Minimize}}\ f_2(x_1,x_2,x_3,x_5,x_7) = -\ 1.508x_1x_7^2 + 7.477x_7^3 + 0.7854x_5x_7^2$$

subject to: (69)

g_4: $1.93\ x_2^{-1}\ x_3^{-1}x_5^{3}x_7^{-4} \le 1$

g_6: $A_2/B_2 \le 850$

g_{18}: $7.3 \le x_5 \le 8.5$ $:g_{19}$

g_{22}: $5 \le x_7 \le 5.5$ $:g_{23}$

g_{25}: $(1.1\ x_7 + 1.9)\ x_5^{-1} \le 1$

Second-Level Problem: Find x_1, x_2, and x_3 (x_4, x_5, x_6 and x_7 are fixed)

$$\underset{x_1,x_2,x_3}{\text{Minimize}}\ f(x) = f_1(x_1, x_2, x_3, x_4, x_6) + f_2(x_1, x_2, x_3, x_5, x_7)$$

subject to: (70)

g_1: $27x_1^{-1}x_2^{-2}x_3^{-1} \le 1$

g_2: $397.5\ x_1^{-1}x_2^{-2}x_3^{-2} \le 1$

g_7: $x_2x_3 \le 40$

g_8: $5 \le x_1/x_2 \le 12$ $:g_9$

g_{10}: $2.6 \le x_1 \le 3.6$ $:g_{11}$

g_{12}: $0.7 \le x_2 \le 0.8$ $:g_{13}$

g_{14}: $17 \le x_3 \le 28$ $:g_{15}$

4.3.2. Active Constraints in the Subproblems

In subproblem 1, it can be easily verified that the objective function is increasing w.r.t. x_4 and x_6 within the feasible range of $x_4 \ge 7.3$, $x_6 \ge 2.9$, and $x_1 \le 3.6$. Hence, according to the first rule of monotonicity analysis,[12] w.r.t. x_6 constraints g_3, g_5, and g_{20} are the candidate active constraints. In order to find the dominant active constraint, we rearrange g_3, g_5, and g_{20} as follows:

$$g_3: \qquad x_6 \ge (1.93\, x_2^{-1} x_3^{-1} x_4^{3})^{1/4} \tag{71}$$

$$g_5: \qquad x_6 \ge (A_1/110)^{1/3} \tag{72}$$

$$g_{20}: \qquad x_6 \ge 2.9 \tag{73}$$

Then, we find the lower and upper bounds of the right-hand sides of eqs. (71) and (72) using the available bounds on the variables x_2, x_3, and x_4:

$$2.406 \le \quad (1.93 x_2^{-1} x_3^{-1} x_4^{3})^{1/4} \quad \le 3.103 \tag{74}$$

$$3.346 \le \quad (A_1/110)^{1/3} \quad \le 3.352 \tag{75}$$

From equations (71)-(75), we conclude that g_5 is the active constraint and g_3, g_{20} are the redundant constraints. Likewise, it can be verified that in subproblem 1, w.r.t. x_4 constraints g_{16} and g_{24} are the candidate active constraints. If we rearrange g_{24} as follows:

$$g_{24}: \qquad x_4 \ge 1.5x_6 + 1.9 \tag{76}$$

Since g_5 is active, by substituting the lower and upper bounds of x_6 from eq. (75), we can find the lower and upper bounds of the right-hand side of eq. (76):

$$6.918 \le 1.5x_6 + 1.9 \le 6.928 \tag{77}$$

which if compared with $x_4 \ge 7.3$, will result in g_{16} as the active constraint and g_{24} as the redundant constraint.

Therefore, in subproblem 1 constraints g_5 and g_{16} are found to be active:

$$g_5: \qquad x_6 = \{[(745\ x_4/(x_2x_3))^2 + 16.9 \times 10^6]^{0.5} \ /\ 110\}^{1/3} \qquad (78)$$

$$g_{16}: \qquad x_4 = 7.3 \qquad (79)$$

In subproblem 2, w.r.t. x_7 constraints g_4, g_6, g_{22} are the candidate active constraints. By a similar analysis, as in subproblem 1, we have:

$$g_4: \qquad x_7 \geq (1.93x_2^{-1}x_3^{-1}x_5^{3})^{1/4} \qquad (80)$$

$$g_6: \qquad x_7 \geq (A_2/85)^{1/3} \qquad (81)$$

$$g_{22}: \qquad x_7 \geq 5 \qquad (82)$$

where

$$2.406 \leq (1.93x_2^{-1}x_3^{-1}x_5^{3})^{1/4} \leq 3.103 \qquad (83)$$

$$5.28568 \leq (A_2/85)^{1/3} \leq 5.28686 \qquad (84)$$

From equations (80)-(84) we conclude that g_6 is active. Likewise, w.r.t. x_5 constraints g_{18} and g_{25} are the candidate active constraints. However, since g_6 is active, then we can use the lower bound of x_7 from equation (84) into the right-hand side of g_{25}:

$$g_{25}: \qquad x_5 \geq 1.1x_7 + 1.9 \geq 1.1\ (5.28568) + 1.9 \approx 7.714 \qquad (85)$$

which if compared with g_{18}, results in g_{25} as the active constraint.

Therefore, in subproblem 2, constraints g_6 and g_{25} are found to be active:

$$g_6: \qquad [(745x_5/(x_2x_3))^2 + 157.5\text{x}10^6]^{0.5} \ /0.1x_7 = 850 \qquad (86)$$

$$g_{25}: \quad (1.1x_7 + 1.9)x_5^{-1} = 1 \tag{87}$$

4.3.3. Solution

In the previous section, we found the active constraints in subproblem 1 (constraints g_5 and g_{16}) and subproblem 2 (constraints g_6 and g_{25}). We now assume that $x^S_{1,1} = x_4$, $x^S_{1,2} = x_6$ to be the solution variables in subproblem 1. Also, we assume that $x^S_{2,1} = x_5$, $x^S_{2,2} = x_7$ to be the solution variables in subproblem 2. Note that there is no decision variables in subproblems 1 and 2. Furthermore, in the second-level problem, we have: $(y_1, y_2, y_3) = (x_1, x_2, x_3)$. Also, in the second-level problem, we can show that g_1, g_2, g_7, g_9, g_{10}, and g_{13} are redundant (Li and Papalambros, 1985). We then calculate the second-level constrained-derivative, $\partial f/\partial \underset{\sim}{y}$ (eq. (36)), which can be shown to be positive w.r.t. y_1, y_2, and y_3 within the feasible domain. Therefore, in the second-level problem g_8, g_{12}, and g_{14} have to be active. The final global optimum solution is $\underset{\sim}{x}^* = (3.5, 0.7, 17, 7.3, 7.71, 3.35, 5.29)^t$ for which constraints g_5, g_6, g_8, g_{12}, g_{14}, g_{16}, and g_{25} are active. This solution is identical to the one reported by Li and Papalambros (Li and Papalambros, 1985).

5. Concluding Remarks

For a two-level design optimization problem, we have proved that the optimality conditions for the undecomposed problem is the same as the ones corresponding to the decomposed subproblems. Extension of this proof, which in essence examines the integrity of optimality conditions before and after decomposition, to a multi-level design optimization problem should be easy.

We also demonstrated that an effective approach for calculating the gradient of objective function in the second-level problem can be devised. This approach consists of exploitation of problem structure and identification of active constraints in the first-level subproblems. The approach is based on a first-order gradient information, does not require calculation of optimum sensitivities (a difficult to obtain second-order information) of first-level subproblems, and should be easy to implement.

6. Acknowledgement

The work reported in this paper was supported in part by a grant from Westinghouse Electric Corporation. Special thanks to Dr. William Fourney, Chairman of Mechanical Engineering Department, for his encouragement and support of this work.

7. References

Abadie, J., and Carpentier, J., 1969, "Generalization of the Wolfe Reduced Gradient Method to the Case of Nonlinear Constraints," Optimization (R. Fletcher, Eds.).

Azarm, S., Li, W.-C., 1988a, "A Two-Level Decomposition Method for Design Optimization," Engineering Optimization, Vol. 13, pp. 211-224.

Azarm, S., Li, W.C., 1988b, "A Multi-Level Optimization-Based Design Procedure Using Global Monotonicity Analysis," Advances in Design Automation (S.S. Rao, Eds.), ASME Pub. No. DE-Vol. 14, pp. 115-120.

Golinski, J., 1970, "Optimal Synthesis Problems Solved by Means of Nonlinear Programming and Random Methods," Journal of Mechanisms, Vol. 5, pp. 287-309.

Haftka, R.T., 1984, "An Improved Computational Approach for Multi-Level Optimum Design," Journal of Structural Mechanics, 12(2), pp. 245-261.

Kirsch, U., 1981, Optimal Structural Design, McGraw-Hill.

Lasdon, L.S., 1970, Optimization Theory for Large Systems, Macmillan, London.

Li, H.L., Papalambros, P., 1985, "A Production System for Use of Global Optimization Knowledge," Trans. ASME, Journal of Mechanisms, Transmissions, and Automation in Design, Vol. 107, No. 2, pp. 277-284.

Lootsma, F.A., Ragsdell, K.M., 1988, "State-of-the-Art in Parallel Nonlinear Optimization," Parallel Computing, 6, pp. 133-155.

Papalambros, P., Wilde, D.J., 1988, Principles of Optimal Design: Modelling and Computation, Cambrdige University Press.

Sobieski, J., 1982, "A Linear Decomposition Method for Large Optimization Problems," NASA-TM-83248, NASA Langley Research Center, Hampton, VA.

Sobieski, S., Barthelemy, J.-F., Riley, K.-M., 1982, "Sensitivity of Optimum Solutions of Problem Parameters," AIAA Journal, September, pp. 1291-1299.

Vanderplaats, G.N., Kim, D.S., 1988, "Recent Development in Multi-Level Optimization," Proceedings of Second NASA/Air-Force Symposium in Multidisciplinary Analysis and Optimization, Hampton, VA.

Wilde, D.J., Beightler, C., 1967, Foundations of Optimization, Prentice-Hall, N.J.

List of Figures:

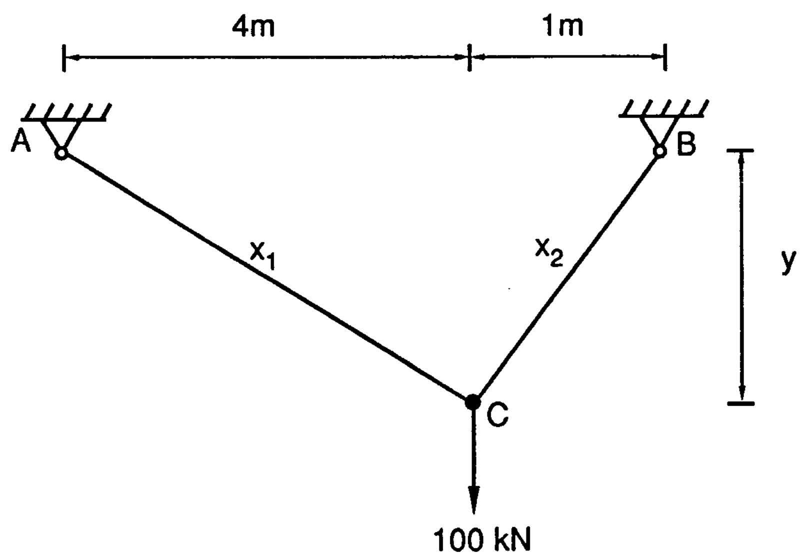

Figure 1 A Two-Bar Truss Problem

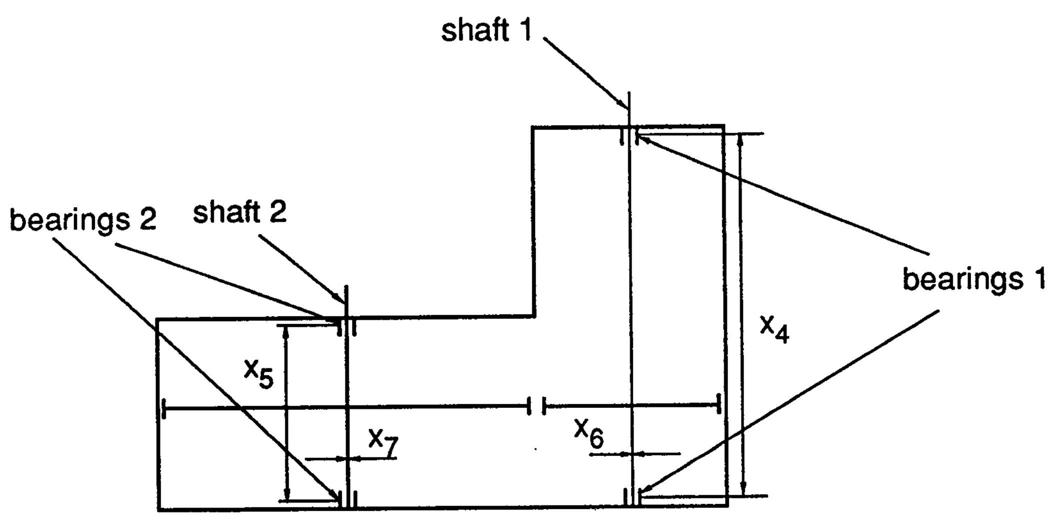

Figure 2 A Gear-Reducer