SENSITIVITY OF LUMPED CONSTRAINTS USING THE ADJOINT METHOD

Mehmet A. Akgün
Aeronautical Engineering Department
Middle East Technical University
Ankara 06531 Turkey

Raphael T. Haftka
Department of Aerospace Engineering, Mechanics and Engineering Science
University of Florida
Gainesville, FL 32611-6250

K. Chauncey Wu
and
Joanne L. Walsh
NASA Langley Research Center
Hampton, VA 23681

AIAA Paper No. 99–1314

Presented at
40th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference
St. Louis, Missouri
April 12-15, 1999
SENSITIVITY OF LUMPED CONSTRAINTS USING THE ADJOINT METHOD

Mehmet A. Akgün* (akgun@ae.metu.edu.tr)
Aeronautical Engineering Department, Middle East Technical University, Ankara 06531 Turkey
Raphael T. Haftka† (haftka@ufl.edu)
Department of Aerospace Engineering, Mechanics and Engineering Science
University of Florida, Gainesville, FL 32611-6250
K. Chauncey Wu‡ (k.c.wu@larc.nasa.gov), and Joanne L. Walsh§ (j.j.walsh@larc.nasa.gov)
NASA Langley Research Center, Hampton, VA 23681

Abstract
Adjoint sensitivity calculation of stress, buckling and
displacement constraints may be much less expensive
than direct sensitivity calculation when the number of
load cases is large. Adjoint stress and displacement
sensitivities are available in the literature. Expressions
for local buckling sensitivity of isotropic plate elements
are derived in this study. Computational efficiency of
the adjoint method is sensitive to the number of
constraints and, therefore, the method benefits from
constraint lumping. A continuum version of the
Kreisselmeier-Steinhauser (KS) function is chosen to
lump constraints. The adjoint and direct methods are
compared for three examples: a truss structure, a simple
HSCT wing model, and a large HSCT model. These
sensitivity derivatives are then used in optimization.

Introduction
Sensitivity of structural response quantities, such as
functions of stress or displacement components, to
design parameters is useful in structural applications of
optimization, system identification, and probabilistic
analysis. Finite difference derivatives are commonly
used, but they are expensive computationally and prone
to errors. Consequently, there is wide interest in
analytical sensitivity methods, commonly classified as
either direct or adjoint methods (e.g., Adelman and
Haftka, 1989).

The widely used direct method is obtained by direct
differentiation of the equations of equilibrium, which
are used to compute the structural responses or static
solutions. The number of static solutions required in
obtaining the derivatives is equal to the product of the
number of design variables and the number of load
cases. The adjoint method is obtained by differentiation
of constraint functionals. The number of static solutions
required to obtain the derivatives is equal to the number
of constraints of interest. For problems with multiple
load cases or multiple structural configurations, such as
damaged versions of a structure, the adjoint method can
be more efficient than the direct method. This
computational advantage is enhanced for derivatives of
stress functions because these derivatives can be
calculated directly rather than obtained from derivatives
of the displacements. Reduction of the number of
constraints through constraint deletion and/or lumping
enhances the efficiency of the adjoint method.

The computational efficiency of the adjoint method is
offset by greater implementation effort compared to the
direct method, especially when based on differentiation
of discretized structural models. Akgün et al. (1998)
showed that this implementation penalty is alleviated
for stress constraints when the adjoint method is
implemented based on the continuum equations. In this
case, the adjoint load is implemented as initial strains
imposed in the elements of interest.

Initial strains for the sensitivity of von Mises
constraints with and without constraint lumping were
derived in a previous work (Akgün et al., 1998).
Implementation of the adjoint method was
demonstrated for truss and plane-stress elements. A
continuum version of the Kreisselmeier-Steinhauser
(KS) functional that allowed lumping of the constraints
with the adjoint method was investigated via a twenty-
five-bar truss problem. The present study is a
continuation of that work.

The objectives of this study are to derive the adjoint
initial strains for local buckling sensitivity computation
and to demonstrate the efficiency of the adjoint method
under multiple load cases with constraint lumping. It is
shown that buckling constraints satisfy a homogeneity
property that simplifies buckling sensitivity
computation by obviating the need to compute partial
derivatives of the buckling constraint. The example
problems used are the structural optimization of a truss
structure, a high-speed civil transport (HSCT) wing,
and a much larger HSCT model.

**Formulation of Buckling Sensitivity**

The discretized equations of equilibrium may be written
in terms of the stiffness matrix $K$, the force vector $f$, and the displacement vector $u$ as

$$ Ku = f $$

(1)

Differentiation of this equation with respect to a design
variable $p$ gives

$$ Ku_p = f^p = -K_p u $$

(2)

where subscript $p$ denotes derivative with respect to the
design variable, and $f^p$ is called the pseudo load. The
direct method solves for the displacement derivative
from Eq. (2) and, if necessary, computes stress
constraints from the displacement and its derivative. If
the number of design variables is $n_p$ and the number of
load cases (i.e., the number of vectors $f$) is $n_f$, Eq. (2)
has to be solved $n_p n_f$ times.

The adjoint method used here is based on a continuum
equation and requires the imposition of an adjoint load
in the form of an initial strain distribution. When the
method is applied to a discrete model, an equation of
the form

$$ Ku^a = f^a $$

(3)

has to be solved $n_c$ times where $n_c$ is the number of
constraints of interest and, therefore, the number of
adjoint load vectors $f^a$, $u^a$ is the adjoint displacement,
and $f^a$ is the force equivalent to the adjoint load. For
stress constraints, $f^a$ is derived from an applied initial
strain. A finite element code that has initial strain
capability can easily generate the adjoint force, $f^a$. The
use of the adjoint displacement $u^a$ to calculate
sensitivities is described subsequently.

Buckling sensitivity formulation will be carried out
next. A perfect isotropic rectangular flat plate of
thickness $t$ and sides $a$ and $b$ may buckle under biaxial
loading if the maximum value of

$$ \frac{1}{\lambda_n} = -\frac{1}{\Delta} \left[ N_x \left( \frac{m}{a} \right)^2 + N_y \left( \frac{n}{b} \right)^2 \right] $$

(4)

exceeds unity, where an applied stress resultant $N_i$ is
positive when tensile, and

$$ \Delta = \pi^2 t^3 H \left( \frac{m^2}{a^2} + \frac{n^2}{b^2} \right)^2 $$

(5)

with

$$ H = E/12(1 - \nu^2) $$

(6)

An isotropic infinite strip of width $b$ having two simply
supported sides, on the other hand, may buckle under
pure in-plane shear if
\[
\frac{1}{\lambda_i} = \frac{|N_{x|t}|b^2}{52.7t^2H}
\]  

(7)

exceeds unity (Whitney, 1987, p. 121). When Eq. (7) is
applied to a rectangular flat plate, \( b \) is taken to be the
shorter edge. Buckling under combined loading may
occur if the value of the interaction equation
\[
\frac{1}{\lambda_{\text{comb}}} = \frac{1}{\lambda_n} + \frac{1}{\lambda_t^2}
\]

(8)

reaches unity (Lekhnitskii, 1968). A buckling
constraint for an isotropic plate under combined loading
may be therefore written as
\[
g(N, t) = g_1(N, t) + g_2(N, t) - 1 \leq 0
\]

(9)

where \( N \) is the vector of stress resultants and
\[
g_1(N, t) \equiv \frac{1}{\lambda_n}, \quad g_2(N, t) \equiv \frac{1}{\lambda_t^2}
\]

(10)

If \( \lambda_n \) is negative, \( g_1 \) is taken to be zero.

The sensitivity of a buckling constraint is derived more
conveniently if the constraint is expressed as
\[
G = \frac{1}{A} \int g(N, t) dA
\]

(11)

where \( A \) is the plate surface area. That is, \( G \) is the value
of the buckling constraint averaged over an area \( A \).

The constitutive equation for the plate is
\[
N = \begin{bmatrix} N_x \\ N_y \\ N_{xy} \end{bmatrix} = D \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{bmatrix}
\]

(12)

where \( D \) is the constitutive matrix. The sensitivity of \( G \)
to the design variable \( p \), where \( p \) may, for example, be
the plate's own thickness or the property of another
structural element, is obtained by differentiating Eq.
(11) with respect to \( p \) to yield
\[
G_p = \frac{1}{A} \left[ \left( g_p + g_N \cdot (N_p + N_e \varepsilon_p) \right) \right] dA
\]

(13)

where a dot indicates inner product, a subscript
following a comma shows a partial derivative, and a
subscript without a comma denotes a total derivative.

\[N_p = D\varepsilon; \quad N_e = tD.\]

(14)

In this formulation, it can be shown (Haftka and
Gürdal, 1992, pp. 312-316) that the adjoint load \( f^a \)
needed to compute sensitivities is obtained from an
initial strain given by \( g_N/A \), regardless of what \( p \) is, so
that the loading depends on the buckling equation but
not on the details of the finite element. The initial
strain for the buckling constraint is thus given by (assuming
\( \lambda_n > 0 \))
\[
\varepsilon_{in} = \frac{1}{A} \left[ -\frac{m^2}{a^2\Delta} , -\frac{n^2}{b^2\Delta} , \frac{2N_{x}b^4}{(52.7H)^2 t^6} \right]
\]

(15)

This loading can be easily applied, then, in any program
that has initial strain loading capabilities. Under the
application of the adjoint load, the resulting
displacement field is called the adjoint displacement \( u^a \).

If we assume that \( g \) is constant over \( A \) it can be shown
that
\[
G_t = g_t = g_{,N} \cdot N_t + f^p \cdot u^a
\]

(16)

where \( f^p \) is the pseudo load which is also used in the
direct method (Eq. 2).

To simplify Eq. (16), a property of homogeneous
functions is introduced next with a brief digression. If \( h \)
is a homogeneous function of degree \( n \) in the arguments
\( N \) and \( t \), that is, if
\[
h(cN, ct) = c^n h(N, t)
\]

(17)

and since
\[
\frac{N}{t} = N_t
\]  
(18)

it can be shown that

\[
h_t + h_N \cdot N_t = \frac{n}{t} 
\]  
(19)

(Akgün et al., 1998). In the present case, the constraint \( g \) in Eq. (9) is not homogeneous as a whole but its components are. For \( g_1, \ n = -2 \) and for \( g_2, \ n = -4 \). Equation (16) hence becomes

\[
G_t = g_t = -\frac{1}{t} \left(2g_1 + 4g_2 \right) + f^p \cdot u^a 
\]  
(20)

The homogeneity property (Eq. 19) hence does away with the need to implement and compute partial derivatives of the buckling constraint, and allows us to use instead the constraint value, which is already available. When \( p \) is a property of a structural element other than the plate whose sensitivity is being computed, the buckling sensitivity is simply given by

\[
G_p = g_p = f^p \cdot u^a 
\]  
(21)

It should be noted that the homogeneity is not limited to the constraints for isotropic plates; it is valid for composite laminates as well.

**Constraint Lumping**

The adjoint method requires computation of the adjoint displacements for every potentially active constraint. Reducing the number of constraints by lumping them into groups can increase the savings brought about by the method. The continuum version of the KS (Akgün et al., 1998) function is used here to lump a set of buckling constraints for a number of flat plates into a single constraint and is defined here as

\[
KS(g(N,t)) = \frac{1}{\rho} \ln G(g(N,t)) 
\]  
(22)

with \( \rho \) being a positive parameter and where

\[
G(g(N,t)) = \int \frac{1}{A} e^{\rho (g(N,t))} dA 
\]  
(23)

Here the integral is taken over all the plate elements being lumped together, and \( A \) is the surface area of the corresponding plate. In the implementation for this study the integrand in Eq. (23) is constant within each plate. The derivative of the KS function is given as

\[
KS_p = G_p / \rho G 
\]  
(24)

Steps similar to those leading to Eqs. (16) and (20) from Eq. (10) yield

\[
G_p = -\rho \sum_{j=1}^{m} e^{\rho g_{j} \delta} (2g_{1j} + 4g_{2j}) \delta_{ij} + f^p \cdot u^a 
\]  
(25)

where \( m \) is the number of buckling constraints lumped together and \( \delta \) is a Kronecker delta. \( \delta \) is unity if design variable \( p \) is the thickness of the \( i \)th plate and zero if design variable \( p \) is not the thickness of the \( i \)th plate. The adjoint displacement \( u^a \) is now due to an initial strain state, imposed simultaneously in each of the plate elements being lumped. The strain state to be imposed in the \( i \)th plate is given by

\[
\varepsilon^T_{in} = \rho e^{\rho g_{in}} \frac{g_N}{A} 
\]  
(26)

where \( g_{in}/A \) is given by the right hand side of Eq. (15), which is the strain state imposed in a single element in the absence of lumping.

When multiple load cases are present, one may want to lump all constraints for the group of \( m \) members under all the load cases together into the same KS function. This is the approach we used with the direct method in the comparison studies described below. For such an approach, however, we could not find an initial strain state for sensitivity calculation with the adjoint method. Instead of lumping, approach adopted here for the adjoint method is to select the most critical load case for buckling and von Mises constraints separately for each group of elements and to lump each type of
constraints under the respective load case together. The constraints under the other load cases are ignored. The initial strain state given by Eq. (26) is then computed at the most critical load case.

**Applications**

Three examples of varying complexity are used to compare the efficiency of the adjoint and direct methods. The three examples are a 108-bar truss, a simple HSCT wing model, and a large HSCT model. Sensitivity derivatives are obtained using both methods and then used in optimization. The 108-bar truss example is used to compare the effect of lumping on the two methods. The efficiency of both methods is compared by varying the number of load cases for the two HSCT examples.

Both the adjoint and direct methods are implemented in a commercial finite element program, EAL (Engineering Analysis Language, EISI-EAL, 1983), for plane-stress elements with von Mises stress constraints and buckling constraints and for truss elements. The optimization code MINOS (Murtagh and Saunders, 1993) has been incorporated into EAL. The linear programming option of MINOS (based on the simplex method) is used in a sequential linear programming (SLP) optimization approach for this study.

Examples 1 and 2 were run on a Pentium 166 MHz PC with a single processor with a Windows NT operating system at the University of Florida. Example 3 was run on a dual processor Pentium 400 MHz PC with a Windows NT operating system at the NASA Langley Research Center. For all three examples, a finite difference approach was used to compute the derivative of the stiffness matrix, design variables were constrained by move limits which were initially 50 percent of the starting design variable values, and when lumping was used the value of the KS function parameter \( r \) was 50.

The lumping strategy used with the adjoint method was to lump each type of constraints (i.e., von Mises and buckling types where applicable) in a design region under the most critical load case into a single KS function whereas the strategy with the direct method was to lump each type of constraints under all the load cases into a single KS function. The two lumping strategies were thus different even though the resulting optima were very close. No constraint deletion strategy was used with either method with or without lumping.

**Example 1: 108-bar truss**

A 45-node, 108-bar, 80-dof plane truss, shown in Fig. 1, was divided into 32 design regions, with bars in each region having a common cross-section, thus giving 32 design variables. Optimization runs were made with seven load cases using the direct and the adjoint methods, with and without lumping. When lumping was used, bar stress constraints in each region were lumped into a continuum KS function. Hence there were 32 KS functions in either method. When no lumping was used, there were 756 constraints. The direct method with and without lumping required the solution of Eq. (2) 224 times, with the stiffness matrix already factored, whereas the adjoint method required the solution of Eq. (3), of the same size, 108 and 32 times without and with lumping, respectively. The reason for 108 solutions instead of 756 is that the adjoint load vector \( f^a \) for a given bar for different load cases differs only by a scalar factor.

The optimum weight and the CPU time using the adjoint method and the direct method are given in Table 1. The CPU times are for the entire optimization run (i.e., 8 SLP cycles). The optimum weights in the table are the weights at the end of eight optimization cycles representing convergence to 10 significant digits.

As can be seen from Table 1, the final weights show some dependence on lumping and the choice of method. This is attributed to optimization history and to the effect of the KS function, which applies a slightly more conservative constraint. It can also be seen that the lumping was more beneficial to the direct method than the adjoint method, even though one would expect the opposite. The reason for this anomaly is that the savings associated with the adjoint method is in the solution of the equations of equilibrium with fewer right hand sides. This solution increases superlinearly with problem size, and is very small for small
problems. The additional cost due to implementation discussed earlier, is proportional to the number of elements in the problem, and is therefore more significant than the solution time for small problems.

Table 1. Optimum weights of the 108-bar truss and total (8 cycles) CPU times with the direct and adjoint methods with and without KS lumping under seven load cases.

<table>
<thead>
<tr>
<th>Method</th>
<th>Weight, lb (Time, sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct No Lump</td>
<td>36.523367 (312)</td>
</tr>
<tr>
<td>Direct Lump</td>
<td>37.359306 (169)</td>
</tr>
<tr>
<td>Adjoint No Lump</td>
<td>37.022500 (292)</td>
</tr>
<tr>
<td>Adjoint Lump</td>
<td>37.353403 (256)</td>
</tr>
</tbody>
</table>

Example 2: Simple HSCT wing
A simple high-speed civil transport (HSCT) wing model developed by Balabanov et al. (1996), whose planform view is shown in Fig. 2, is investigated here. The 3-D model has 193 nodes, 449 2-node bar elements, 383 triangular isotropic membrane elements and 129 isotropic shear panel elements with a total of 533 degrees of freedom. Figure 2 shows the design regions selected on the top skin of the HSCT model, the bottom skin being the mirror image of those. Skin regions 1, 2, 10, 11, 12, and 13, which cover the trailing edge and strake of the wing, are governed by the minimum gauge-thickness constraint set at 0.0035 ft. Skin thickness within each design region is uniform. The minimum gauge area used for the bar elements is 0.0040 ft². Not all of the finite elements are designed. There are 40 design variables consisting of 26 membrane thicknesses and 14 bar cross-sectional areas.

A series of optimization runs with different numbers of load cases was made. There were five basic load cases which were due to pull-up maneuvers, climb, cruise and taxiing. Additional fifteen load cases were derived as a combination of the basic five cases. Stress constraints were based on von Mises equivalent stresses in the membrane elements and axial stresses in the bars.

The implementation of the adjoint method for the sensitivity of lumped buckling constraints was first verified. The model was then optimized under one, five and twenty load cases in different runs subject to buckling and von Mises stress constraints for membrane elements and yield limits for bar elements. There were 66 KS functions with both methods regardless of the number of load cases. Equation (2) had to be solved 40\(n\) times with the direct method where \(n\) is the number of load cases and Eq. (3) 66 times with the adjoint method regardless of the number of load cases. The optimum weights and the CPU times are given in Table 2. The CPU times are for the entire optimization run (i.e., 8 SLP cycles). The optimum weights in the table are the weights at the end of eight optimization cycles representing convergence to 13 significant digits.

Table 2. Optimum weights of the simple HSCT model and total (8 cycles) CPU times with the two methods using KS lumping.

<table>
<thead>
<tr>
<th>No of Load Cases</th>
<th>Direct W, lb (t, sec)</th>
<th>Adjoint W, lb (t, sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>119746.87495532 (750)</td>
<td>119746.87495534 (1410)</td>
</tr>
<tr>
<td>5</td>
<td>149327.13548552 (1410)</td>
<td>149327.13548557 (1560)</td>
</tr>
<tr>
<td>20</td>
<td>149327.13548552 (2962)</td>
<td>149327.13548557 (1627)</td>
</tr>
</tbody>
</table>

As can be seen from Table 2, the final weights agree to 13 digits accuracy. Also, for this larger problem, the improved efficiency with number of load cases of this implementation of the adjoint method compared to the direct method is apparent. For five load cases the two
methods are comparable, and for 20 load cases the adjoint method is about 45% cheaper.

**Example 3: Large HSCT**

The third example is a half-symmetric model of an entire high-speed transport aircraft\(^1\) shown in Fig. 3. This model was originally presented by Scotti (1995). Results were performed at NASA Langley Research Center. The present EAL adjoint runstreams cannot handle four-noded membrane elements and, therefore, each four-noded element in the original model’s design regions was replaced with two three-noded elements, which increased the total number of elements in the model to 16434. In addition, all composite honeycomb plate sections were replaced with isotropic membrane sections, which reduced the number of design variables in the model to 61. Both stress and buckling constraints are computed for the triangular elements in the design regions.

There are seven basic load cases, six arising from supersonic maneuvers at +2.5g and -1.0g and the seventh represents a taxi condition. Additional fourteen load cases were generated from various combinations of the basic six maneuver load cases. For timing purposes, the model was run for four optimization cycles using both the adjoint and direct method. For each method, analyses were performed with one, seven, 14, and 21 load cases. The CPU time required to complete an optimization cycle is plotted against the number of load cases in Fig. 4. The lines shown in the figure are best-fit through the average of the run times for the four cycles. The actual run times for individual cycles are also indicated on the figure.

For a single load case, the direct method is twice as fast as the adjoint method. However, as the number of load cases increases, the adjoint method becomes more efficient. Because this example is larger than the prior two examples, the efficiency of the adjoint method as the number of load cases increases is greater, and the break-even point now appears to be around three or four load cases. For 21 load cases, the problem solution with the adjoint method requires only about one-third of the time required by the direct method.

In addition, the optimization was run to convergence using both methods with one load case and seven load cases. Figure 5 shows the convergence history for the normalized objective function. It is seen that the direct and adjoint methods converge to comparable weights. The values of the design variables were also similar.

**Discussion and Conclusions**

Sensitivity of normal and shear buckling constraints in membranes has been formulated and implemented using the adjoint method. Sensitivity of lumped constraints has also been implemented. The present formulation utilizes a characteristic of homogenous functions to avoid calculation of partial derivatives of the constraints and the shear resultants. Test cases including a 108-bar truss and two high speed civil transport (HSCT) models of different complexities were used to compare the direct and the adjoint methods of sensitivity computation. The two methods have been compared in terms of efficiency for various numbers of load cases in optimization runs performed with the EAL finite element software.

The direct method was found to be more efficient for a small number of load cases. The adjoint method outperforms the direct method in runs with a large number of load cases. The advantage of the adjoint method results from requiring static solutions equal in number to the number of active constraints, while the direct method requires a number equal to the number of design variables times the number of load cases. The advantage of the adjoint method increases with problem size because the solution cost per load case increases superlinearly with problem size. It must be noted, however, that the implement EAL (Engineering Analysis Language) optimization code was originally developed for the direct method and the adjoint method has been recently added to the code; there is room for

---

\(^1\) The structural model for this example has been supplied by the Boeing Company and the results are presented without absolute scales in this paper under the conditions of a NASA Langley Property Loan Agreement, Loan Control Number I922931.
improving the efficiency of the adjoint method as currently implemented.

The implementation of the direct method in EAL has been aided with specialized processors developed especially for this method. The implementation of the adjoint method used standard EAL runstream commands, which may accrue higher overhead during execution. Thus, for a single load case the adjoint method required substantially more time than the direct method.

No constraint deletion strategy was used with either the direct or the adjoint method. It may be worthwhile to investigate the effect of constraint deletion in conjunction with lumping or without lumping.

Acknowledgement
This work was supported in part by NASA Cooperative Agreement NCC-1-268. The first author acknowledges the fellowship from Turkish Scientific and Technical Research Council that supported part of his visit to the University of Florida and the sabbatical leave of absence granted by Middle East Technical University. Help from V. O. Balabanov from Vanderplaats R&D with the HSCT model is gratefully acknowledged.

References


Figure 1. 108-bar truss.

Figure 2. Simple HSCT wing model and the design regions on the top skin.
Figure 3. Half-symmetric finite-element model of the large high-speed civil transport.

Figure 4. CPU timing as a function of load cases for the large high-speed civil transport.
Figure 5. Optimization history for the large high-speed civil transport.