

Reduction of the Adjoint Gradient Formula for Aerodynamic Shape Optimization Problems

Antony Jameson* and Sangho Kim†
Stanford University, Stanford, California 94305

We present a new continuous adjoint method for aerodynamic shape optimization using the Euler equations, which reduces the computational cost of the gradients by reducing the volume integral part of the adjoint gradient formula to a surface integral. The savings are particularly significant for three-dimensional aerodynamic shape-optimization problems on general unstructured and overset meshes. To validate the concept, the new gradient equations have been tested for various aerodynamic shape-optimization problems, including an inverse problem for three-dimensional wing configurations, and drag-minimization problems of a single-element airfoil and a three-dimensional wing-fuselage configuration. To assess their accuracy, the results are compared with finite difference gradients, complex-step gradients, and gradients calculated by the previous adjoint method, which includes a volume integral.

Nomenclature

A_i	= Cartesian flow Jacobian matrix in the i th direction
CD, CL	= coefficients of drag and lift in three dimensions
C_d, C_l	= coefficients of drag and lift in two dimensions
C_i	= contravariant flow Jacobian matrix in the i th direction
C_p	= pressure coefficient
c	= chord length
E	= total energy
\mathcal{F}	= boundary shape
F_i	= contravariant inviscid fluxes
f_i	= Cartesian inviscid fluxes
\mathcal{G}	= gradient vector
H	= total enthalpy
I	= cost function
J	= $\det(K)$
K_{ij}	= mesh transformation Jacobian matrix components
M_∞	= freestream Mach number
p	= static pressure
p_d	= desired target static pressure
R	= governing equations or residual
Re	= Reynolds number
S_{ij}	= velocity transformation matrix
t	= time
U_i	= contravariant velocity components
u_i	= Cartesian velocity components
W	= conservative contravariant flow variables
w	= conservative flow variables in Cartesian coordinates
x_i	= Cartesian coordinates
y^+	= dimensionless wall distance
α	= angle of attack
α_{clmax}	= stall angle of attack
γ	= ratio of specific heats
δ	= first variation
δ_{ij}	= Kronecker delta function

ϵ_{ijk}	= permutation tensor
λ	= step size in steepest descent method
ξ_i	= computational coordinates
ρ	= density
ψ	= Lagrange multiplier, costate or adjoint variable

Introduction

RECENTLY the focus of computational-fluid-dynamics (CFD) applications has shifted to aerodynamic design.^{1–6} This shift has been facilitated by the availability of high-performance computing platforms and by the development of new and efficient analysis and design algorithms. In particular, automatic design procedures that use CFD combined with gradient-based optimization techniques have made it possible to remove difficulties in the decision-making process faced by the aerodynamicist.

Typically, in gradient-based optimization techniques a control function to be optimized (an airfoil shape, for example) is parameterized with a set of design variables, and a suitable cost function to be minimized or maximized is defined (drag coefficient, lift/drag ratio, difference from a specified pressure distribution, etc). Then, a constraint, the governing equations of the fluid flow, can be introduced in order to express the dependence between the cost function and the control function. The sensitivity derivatives of the cost function with respect to the design variables are calculated in order to get a direction of improvement. Finally, a step is taken in this direction, and the procedure is repeated until convergence to a minimum or maximum is achieved. Finding a fast and accurate way of calculating the necessary gradient information is essential to developing an effective design method because this can be the most time-consuming portion of the design algorithm. This is particularly true in problems that involve a very large number of design variables, as is the case in a typical three-dimensional wing shape design.

The control theory approach has dramatic computational cost advantages over the finite difference method of calculating gradients, for which the computational cost is proportional to the number of design variables. The control theory approach is often called the adjoint method because the necessary gradients are obtained via the solution of the adjoint equations of the governing equations of interest. The adjoint method is extremely efficient because the computational expense incurred in the calculation of the complete gradient is effectively independent of the number of design variables. The only cost involved is the calculation of one flow solution and one adjoint solution whose complexity is similar to that of the flow solution. Control theory was applied in this way to shape design for elliptic equations by Pironneau,⁷ and it was first used in transonic flow by Jameson.^{1,2,8} Since then, this method has become a popular choice for design problems involving fluid flow.^{4,9–11}

Presented as Paper 2003-0040 at the AIAA 41st Aerospace Sciences Meeting, Reno, NV, 6–9 January 2003; received 12 March 2003; accepted for publication 2 July 2003. Copyright © 2003 by Antony Jameson and Sangho Kim. Published by the American Institute of Aeronautics and Astronautics, Inc., with permission. Copies of this paper may be made for personal or internal use, on condition that the copier pay the \$10.00 per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923; include the code 0001-1452/03 \$10.00 in correspondence with the CCC.

*Thomas V. Jones Professor of Engineering, Department of Aeronautics and Astronautics, Fellow AIAA.

†Postdoctoral Fellow, Department of Aeronautics and Astronautics, Member AIAA.

Existing approaches to the adjoint method can be classified into two categories: the continuous adjoint and discrete adjoint methods. If the adjoint equations are directly derived from the governing equations and then discretized, they are termed continuous, and if instead they are directly derived from the discretized form of the governing equations then they are referred to as discrete. In theory, the discrete adjoint method should give gradients that are closer in value to exact finite difference gradients. On the other hand, the continuous adjoint method has the advantage that the adjoint system has a unique form independent of the scheme used to solve the flowfield system. Recently, Nadarajah and Jameson performed a detailed gradient comparison study of the continuous and discrete adjoint approaches using the Euler equations¹² and the Navier-Stokes equations¹³ and found that in typical shape-optimization problems in transonic flow the differences are small enough that they have no significant effect on the final result.

The work of the first author and his associates has mainly been based on the continuous adjoint method. In fact this method has been successfully used for the aerodynamic design of complete aircraft configurations.^{3,5,6} At an early stage Jameson realized that the true gradient should not depend on the way the mesh is modified, and thus there must be an adjoint-based formula for the gradient that is independent of the mesh modification. Motivated by the special case of a mesh variation with a fixed geometry, he derived a surface integral that replaces the volume integral part of the adjoint gradient (Jameson, A., "Gradient Calculations for Three-Dimensional Shape Modifications," unpublished note, Aug. 1995). During a visit to Sweden in 1999, he learned that reduced formulas for the gradient had also been derived by Weinerfelt and Enoksson.¹⁴ It was subsequently determined that these independently derived equations are essentially identical, although their expressions are different (Weinerfelt, P., and Enoksson, O., private communication, Nov. 1999). Because our existing software had been very carefully validated and the computational cost of evaluating the field integrals is negligible on structured meshes, we did not immediately pursue the alternative formulation. However, when considering extensions of the method to general unstructured and overset meshes the evaluation of formulas based on mesh movement can incur significant computational costs. Moreover, the equivalence of the alternate gradient formulas is true only in the continuous limit and is not exact for the discretized formulas. Thus we feel it is now useful to make a careful evaluation of the accuracy of the surface gradient formulas.

In the present paper the new adjoint gradient equations are implemented and tested for various two-dimensional and three-dimensional design problems, and the accuracy of the resulting derivative information is investigated by comparison with gradients from the finite difference, complex-step, and original adjoint methods. The finite difference method is extremely easy to implement. However, the cost of computing gradient information is proportional to the number of design variables because in the finite difference method small steps are taken in each of the design variables independently, and then a complete flow solution is computed for each of these steps in order to find the sensitivity of the cost function with respect to that design variable. Moreover, the accuracy of the finite difference gradient information depends strongly on the choice of step size (which is not known a priori) because a small step size is desired to reduce the truncation error, but too small a step size would also increase subtractive cancellation errors. The complex-step method, introduced by Lyness and Moler,¹⁵ overcomes this difficulty. Because the complex-step formula does not require any subtraction to yield the approximate derivative, its accuracy is independent of the choice of step size though the computational cost is unfortunately still proportional to the number of design variables. Detailed comparisons of finite difference vs adjoint gradients are available in the authors' previous work,¹⁶ using the earlier adjoint formulation. Also comparisons of finite difference vs complex-step gradients can be found in the work of Martins et al.¹⁷ The complex-step gradient formula has been implemented in the present work in order to provide a better standard for the present gradient accuracy study.

General Formulation of the Adjoint Approach to Optimal Design

For flow about an airfoil or wing, the aerodynamic properties that define the cost function are functions of the flowfield variables w and the physical location of the boundary, which can be represented by the function \mathcal{F} , say. Then

$$I = I(w, \mathcal{F})$$

and a change in \mathcal{F} results in a change

$$\delta I = \frac{\partial I^T}{\partial w} \delta w + \frac{\partial I^T}{\partial \mathcal{F}} \delta \mathcal{F} \quad (1)$$

in the cost function. Using control theory, the governing equations of the flowfield are introduced as a constraint in such a way that the final expression for the gradient does not require reevaluation of the flowfield. To achieve this, δw must be eliminated from Eq. (1). Suppose that the governing equation R , which expresses the dependence of w and \mathcal{F} within the flowfield domain D , can be written as

$$R(w, \mathcal{F}) = 0 \quad (2)$$

Then δw is determined from the equation

$$\delta R = \left[\frac{\partial R}{\partial w} \right] \delta w + \left[\frac{\partial R}{\partial \mathcal{F}} \right] \delta \mathcal{F} = 0 \quad (3)$$

Next, introducing a Lagrange multiplier ψ , we have

$$\begin{aligned} \delta I &= \frac{\partial I^T}{\partial w} \delta w + \frac{\partial I^T}{\partial \mathcal{F}} \delta \mathcal{F} - \psi^T \left(\left[\frac{\partial R}{\partial w} \right] \delta w + \left[\frac{\partial R}{\partial \mathcal{F}} \right] \delta \mathcal{F} \right) \\ &= \left\{ \frac{\partial I^T}{\partial w} - \psi^T \left[\frac{\partial R}{\partial w} \right] \right\} \delta w + \left\{ \frac{\partial I^T}{\partial \mathcal{F}} - \psi^T \left[\frac{\partial R}{\partial \mathcal{F}} \right] \right\} \delta \mathcal{F} \end{aligned}$$

Choosing ψ to satisfy the adjoint equation

$$\left[\frac{\partial R}{\partial w} \right]^T \psi = \frac{\partial I}{\partial w} \quad (4)$$

the first term is eliminated, and we find that

$$\delta I = \mathcal{G} \delta \mathcal{F} \quad (5)$$

where

$$\mathcal{G} = \frac{\partial I^T}{\partial \mathcal{F}} - \psi^T \left[\frac{\partial R}{\partial \mathcal{F}} \right]$$

The advantage is that Eq. (5) is independent of δw , with the result that the gradient of I with respect to an arbitrary number of design variables can be determined without the need for additional flowfield evaluations. In the case that Eq. (2) is a partial differential equation, the adjoint equation (4) is also a partial differential equation, and appropriate boundary conditions must be determined.

After making a step in the negative gradient direction, the gradient can be recalculated, and the process repeated to follow a path of steepest descent until a minimum is reached. To avoid violating constraints, such as a minimum acceptable wing thickness, the gradient can be projected into the allowable subspace within which the constraints are satisfied. In this way one can devise procedures that must necessarily converge at least to a local minimum and that can be accelerated by the use of more sophisticated descent methods such as conjugate gradient or quasi-Newton algorithms. There is the possibility of more than one local minimum, but in any case the method will lead to an improvement over the original design. Furthermore, unlike the traditional inverse algorithms any measure of performance can be used as the cost function.

Design Using the Euler Equations

The application of control theory to aerodynamic design problems is illustrated in this section for the case of three-dimensional wing design using the compressible Euler equations as the mathematical model. It proves convenient to denote the Cartesian coordinates and velocity components by x_1, x_2, x_3 and u_1, u_2, u_3 , and to use the convention that summation over $i = 1-3$ is implied by a repeated index i . Then, the three-dimensional Euler equations can be written as

$$\frac{\partial w}{\partial t} + \frac{\partial f_i}{\partial x_i} = 0 \quad \text{in} \quad D \quad (6)$$

where

$$w = \begin{Bmatrix} \rho \\ \rho u_1 \\ \rho u_2 \\ \rho u_3 \\ \rho E \end{Bmatrix}, \quad f_i = \begin{Bmatrix} \rho u_i \\ \rho u_i u_1 + p \delta_{i1} \\ \rho u_i u_2 + p \delta_{i2} \\ \rho u_i u_3 + p \delta_{i3} \\ \rho u_i H \end{Bmatrix} \quad (7)$$

Also,

$$p = (\gamma - 1)\rho \left\{ E - \frac{1}{2}(u_i^2) \right\} \quad (8)$$

$$\rho H = \rho E + p \quad (9)$$

Consider a transformation to coordinates ξ_1, ξ_2, ξ_3 , where

$$K_{ij} = \left[\frac{\partial x_i}{\partial \xi_j} \right], \quad J = \det(K), \quad K_{ij}^{-1} = \left[\frac{\partial \xi_i}{\partial x_j} \right], \quad S = J K^{-1}$$

The elements of S are the cofactors of K , and in a finite volume discretization they are just the face areas of the computational cells projected in the x_1, x_2 , and x_3 directions. Using the permutation tensor ϵ_{ijk} , we can express the elements of S as

$$S_{ij} = \frac{1}{2} \epsilon_{jpq} \epsilon_{irs} \frac{\partial x_p}{\partial \xi_r} \frac{\partial x_q}{\partial \xi_s} \quad (10)$$

Then

$$\begin{aligned} \frac{\partial}{\partial \xi_i} S_{ij} &= \frac{1}{2} \epsilon_{jpq} \epsilon_{irs} \left(\frac{\partial^2 x_p}{\partial \xi_r \partial \xi_i} \frac{\partial x_q}{\partial \xi_s} + \frac{\partial x_p}{\partial \xi_r} \frac{\partial^2 x_q}{\partial \xi_s \partial \xi_i} \right) \\ &= 0 \end{aligned} \quad (11)$$

Also in the subsequent analysis of the effect of a shape variation, it is useful to note that

$$\begin{aligned} S_{1j} &= \epsilon_{jpq} \frac{\partial x_p}{\partial \xi_2} \frac{\partial x_q}{\partial \xi_3}, & S_{2j} &= \epsilon_{jpq} \frac{\partial x_p}{\partial \xi_3} \frac{\partial x_q}{\partial \xi_1} \\ S_{3j} &= \epsilon_{jpq} \frac{\partial x_p}{\partial \xi_1} \frac{\partial x_q}{\partial \xi_2} \end{aligned} \quad (12)$$

Now, multiplying Eq. (6) by J and applying the chain rule,

$$J \frac{\partial w}{\partial t} + R(w) = 0 \quad (13)$$

where

$$R(w) = S_{ij} \frac{\partial f_j}{\partial \xi_i} = \frac{\partial}{\partial \xi_i} (S_{ij} f_j) \quad (14)$$

using Eq. (11). We can write the transformed fluxes in terms of the scaled contravariant velocity components

$$U_i = S_{ij} u_j$$

as

$$F_i = S_{ij} f_j = \begin{Bmatrix} \rho U_i \\ \rho U_i u_1 + S_{i1} p \\ \rho U_i u_2 + S_{i2} p \\ \rho U_i u_3 + S_{i3} p \\ \rho U_i H \end{Bmatrix}$$

Assume now that the new computational coordinate system conforms to the wing in such a way that the wing surface B_W is represented by $\xi_2 = 0$. Then the flow is determined as the steady-state solution of Eq. (13) subject to the flow tangency condition

$$U_2 = 0 \quad \text{on} \quad B_W \quad (15)$$

At the far-field boundary B_F conditions are specified for incoming waves, as in the two-dimensional case, whereas outgoing waves are determined by the solution.

The weak form of the Euler equations for steady flow can be written as

$$\int_D \frac{\partial \phi^T}{\partial \xi_i} F_i dD = \int_B n_i \phi^T F_i dB \quad (16)$$

where the test vector ϕ is an arbitrary differentiable function and n_i is the outward normal at the boundary. If a differentiable solution w is obtained to this equation, it can be integrated by parts to give

$$\int_D \phi^T \frac{\partial F_i}{\partial \xi_i} dD = 0$$

and because this is true for any ϕ the differential form can be recovered. If the solution is discontinuous, Eq. (16) can be integrated by parts separately on either side of the discontinuity to recover the shock jump conditions.

Suppose now that it is desired to control the surface pressure by varying the wing shape. For this purpose it is convenient to retain a fixed computational domain. Then variations in the shape result in corresponding variations in the mapping derivatives defined by K . As an example, consider the case of an inverse problem, where we introduce the cost function

$$I = \frac{1}{2} \iint_{B_W} (p - p_d)^2 d\xi_1 d\xi_3$$

The design problem is now treated as a control problem where the control function is the wing shape, which is to be chosen to minimize I subject to the constraints defined by the flow equations (13). A variation in the shape will cause a variation δp in the pressure and consequently a variation in the cost function:

$$\delta I = \iint_{B_W} (p - p_d) \delta p d\xi_1 d\xi_3 \quad (17)$$

Because p depends on w through the equation of state (8) and (9), the variation δp can be determined from the variation δw . Define the Jacobian matrices

$$A_i = \frac{\partial f_i}{\partial w}, \quad C_i = S_{ij} A_j \quad (18)$$

The weak form of the equation for δw in the steady state becomes

$$\int_D \frac{\partial \phi^T}{\partial \xi_i} \delta F_i dD = \int_B (n_i \phi^T \delta F_i) dB$$

where

$$\delta F_i = C_i \delta w + \delta S_{ij} f_j$$

which should hold for any differential test function ϕ . This equation can be added to the variation in the cost function, which can now be written as

$$\delta I = \iint_{B_W} (p - p_d) \delta p \, d\xi_1 \, d\xi_3 - \int_{\mathcal{D}} \left(\frac{\partial \phi^T}{\partial \xi_i} \delta F_i \right) d\mathcal{D} + \int_B (n_i \phi^T \delta F_i) dB \quad (19)$$

On the wing surface B_W , $n_1 = n_3 = 0$. Thus, it follows from Eq. (15) that

$$\delta F_2 = \begin{bmatrix} 0 \\ S_{21} \delta p \\ S_{22} \delta p \\ S_{23} \delta p \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \delta S_{21} p \\ \delta S_{22} p \\ \delta S_{23} p \\ 0 \end{bmatrix} \quad (20)$$

Because the weak equation for δw should hold for an arbitrary choice of the test vector ϕ , we are free to choose ϕ to simplify the resulting expressions. Therefore we set $\phi = \psi$, where the costate vector ψ is the solution of the adjoint equation

$$\frac{\partial \psi}{\partial t} - C_i^T \frac{\partial \psi}{\partial \xi_i} = 0 \quad \text{in } D \quad (21)$$

At the outer boundary incoming characteristics for ψ correspond to outgoing characteristics for δw . Consequently, one can choose boundary conditions for ψ such that

$$n_i \psi^T C_i \delta w = 0$$

Then, if the coordinate transformation is such that δS is negligible in the far field the only remaining boundary term is

$$- \iint_{B_W} \psi^T \delta F_2 \, d\xi_1 \, d\xi_3$$

Thus, by letting ψ satisfy the boundary condition

$$S_{21} \psi_2 + S_{22} \psi_3 + S_{23} \psi_4 = (p - p_d) \quad \text{on } B_W \quad (22)$$

we find finally that

$$\delta I = - \int_{\mathcal{D}} \frac{\partial \psi^T}{\partial \xi_i} \delta S_{ij} f_j \, d\mathcal{D} - \iint_{B_W} (\delta S_{21} \psi_2 + \delta S_{22} \psi_3 + \delta S_{23} \psi_4) p \, d\xi_1 \, d\xi_3 \quad (23)$$

Here the expression for the cost variation depends on the mesh variations throughout the domain, which appear in the field integral. However, the true gradient for a shape variation should not depend on the way in which the mesh is deformed, but only on the true flow solution. In the next section we show how the field integral can be eliminated to produce a reduced gradient formula that depends only on the boundary movement.

Reduced Gradient Formulation

Consider the case of a mesh variation with a fixed boundary. Then,

$$\delta I = 0$$

but there is a variation in the transformed flux

$$\delta F_i = C_i \delta w + \delta S_{ij} f_j$$

Here the true solution is unchanged. Thus, the variation δw is as a result of the mesh movement δx at fixed boundary configuration. Therefore,

$$\delta w = \nabla w \cdot \delta x = \frac{\partial w}{\partial x_j} \delta x_j (= \delta w^*)$$

and because

$$\frac{\partial}{\partial \xi_i} \delta F_i = 0$$

it follows that

$$\frac{\partial}{\partial \xi_i} (\delta S_{ij} f_j) = - \frac{\partial}{\partial \xi_i} (C_i \delta w^*) \quad (24)$$

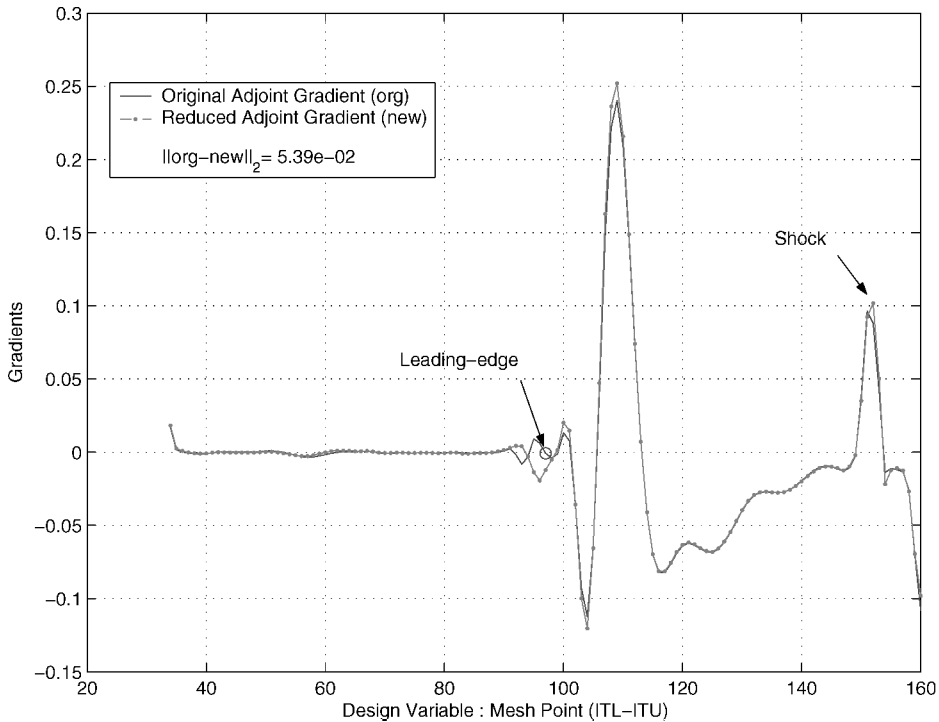
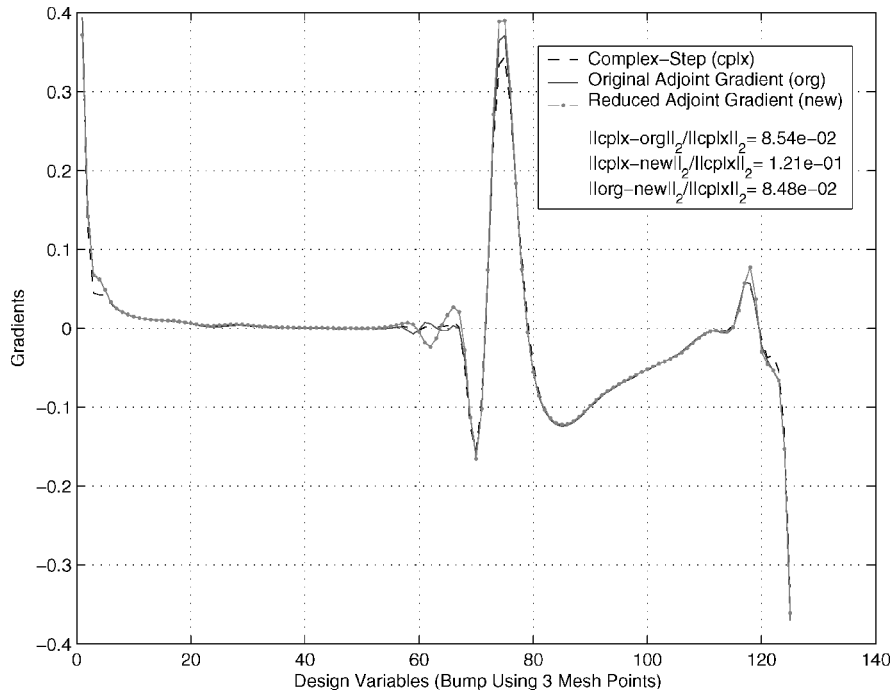
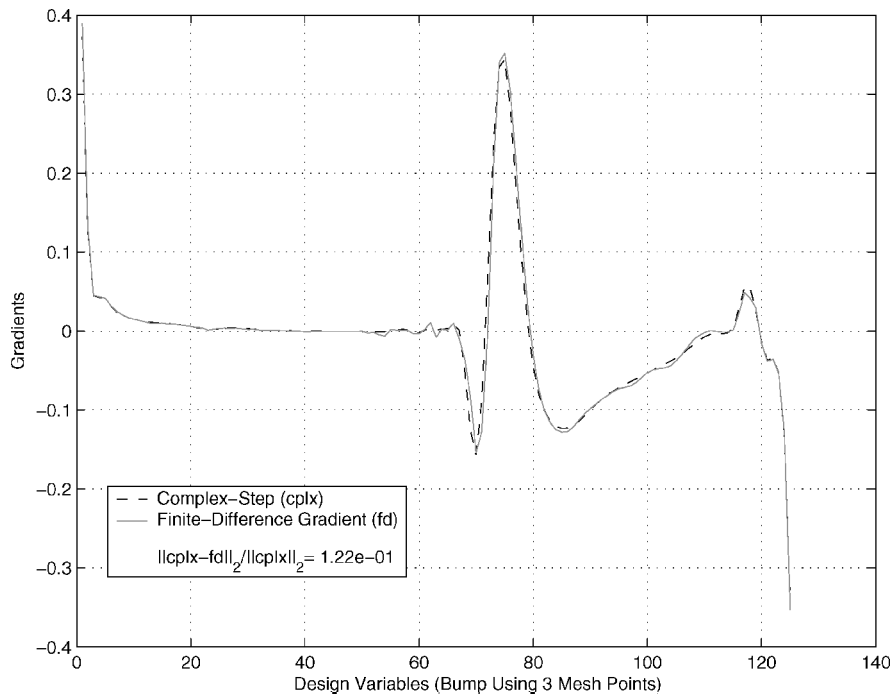


Fig. 1 Euler drag minimization for RAE 2822: comparison of original adjoint and reduced adjoint gradients using mesh points as design variable. Grid: 192 × 32.



a) Reduced adjoint vs original adjoint and complex step



b) Finite difference vs complex step

Fig. 2 Euler drag minimization for RAE 2822: comparison of original adjoint, reduced adjoint, complex-step, and finite difference gradients using three mesh-point bumps as design variables. Grid: 192 × 32.

It is verified in the following paragraph that this relation holds in the general case with boundary movement. Now

$$\begin{aligned} \int_{\mathcal{D}} \phi^T \delta R \, d\mathcal{D} &= \int_{\mathcal{D}} \phi^T \frac{\partial}{\partial \xi_i} C_i (\delta w - \delta w^*) \, d\mathcal{D} \\ &= \int_{\mathcal{B}} \phi^T C_i (\delta w - \delta w^*) \, d\mathcal{B} - \int_{\mathcal{D}} \frac{\partial \phi^T}{\partial \xi_i} C_i (\delta w - \delta w^*) \, d\mathcal{D} \quad (25) \end{aligned}$$

Here on the wall boundary

$$C_2 \delta w = \delta F_2 - \delta S_{2j} f_j \quad (26)$$

Thus, by choosing ϕ to satisfy the adjoint equation (21) and the adjoint boundary condition (22) we have finally the reduced gradient formulation that

$$\begin{aligned} \delta I &= \int_{\mathcal{B}_w} \psi^T (\delta S_{2j} f_j + C_2 \delta w^*) \, d\xi_1 \, d\xi_3 \\ &\quad - \iint_{\mathcal{B}_w} (\delta S_{21} \psi_2 + \delta S_{22} \psi_3 + \delta S_{23} \psi_4) p \, d\xi_1 \, d\xi_3 \quad (27) \end{aligned}$$

For completeness the general derivation of Eq. (24) is presented here. Using formula (10) and property (11),

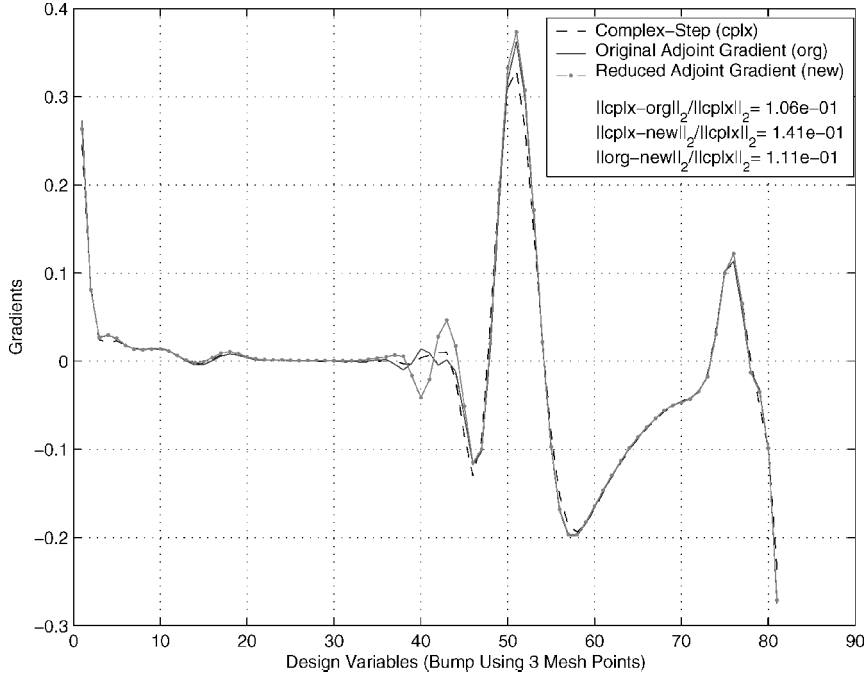


Fig. 3 Euler drag minimization for RAE 2822: comparison of original adjoint and reduced adjoint gradients using three mesh-point bumps as design variables. Coarse grid: 128×32 .

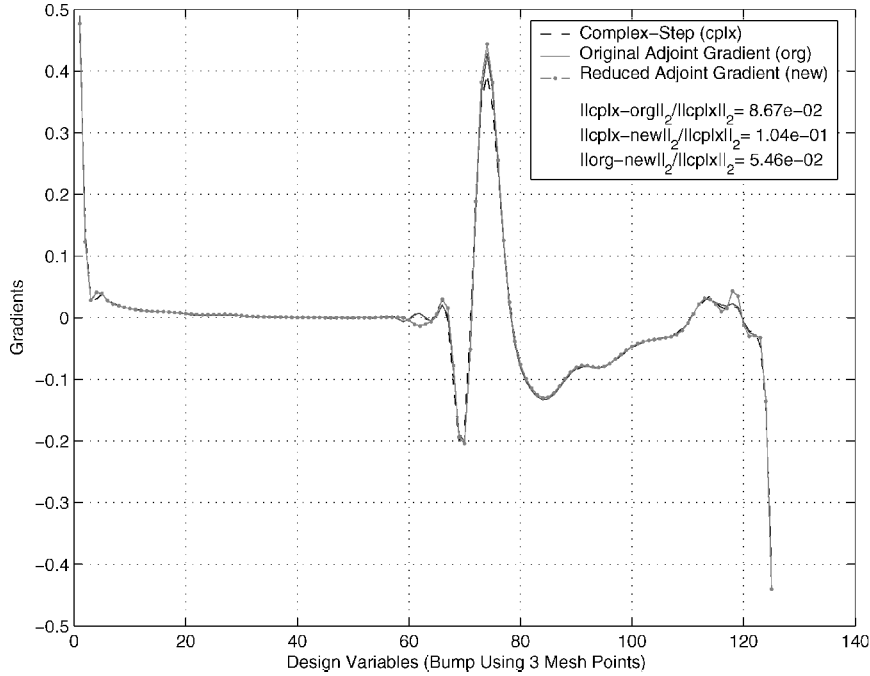


Fig. 4 Euler drag minimization for RAE 2822: comparison of original adjoint, reduced adjoint, and finite difference gradients using three mesh-point bumps as design variables. Medium grid: 192×64 .

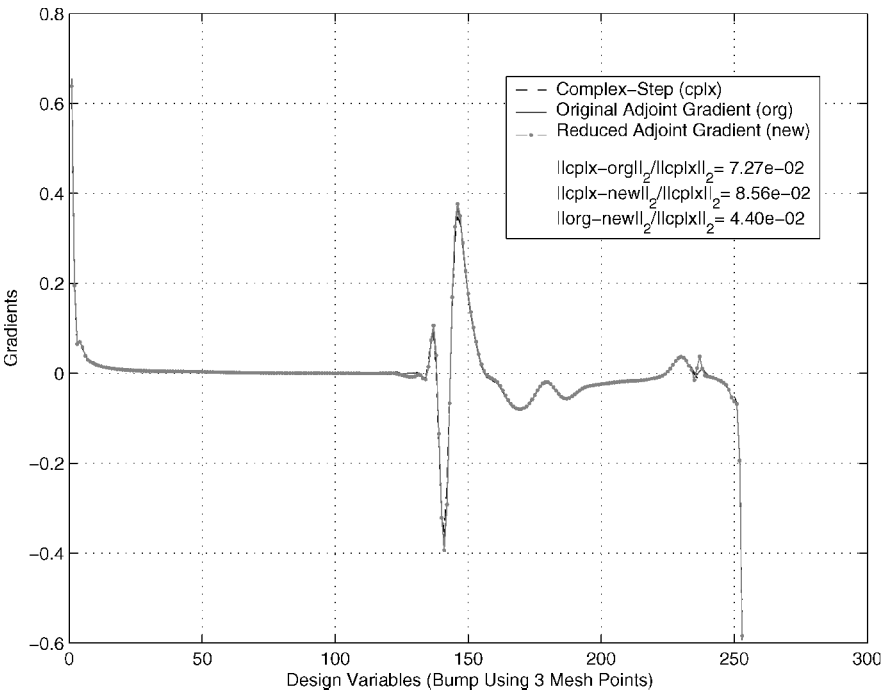
$$\frac{\partial}{\partial \xi_i} (\delta S_{ij} f_j) = \frac{1}{2} \frac{\partial}{\partial \xi_i} \left[\epsilon_{jpq} \epsilon_{irs} \left(\frac{\partial \delta x_p}{\partial \xi_r} \frac{\partial x_q}{\partial \xi_s} + \frac{\partial x_p}{\partial \xi_r} \frac{\partial \delta x_q}{\partial \xi_s} \right) f_j \right] = \frac{\partial}{\partial \xi_r} \left(\delta x_p \epsilon_{pqj} \epsilon_{rsi} \frac{\partial x_q}{\partial \xi_s} \frac{\partial f_j}{\partial \xi_i} \right) \quad (28)$$

Now express δx_p in terms of a shift in the original computational coordinates:

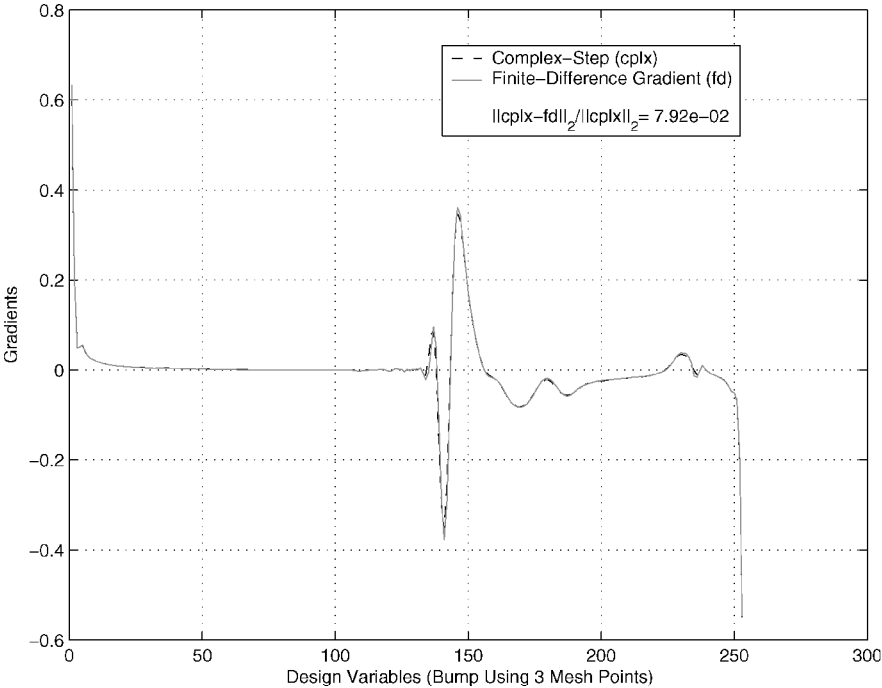
$$\delta x_p = \frac{\partial x_p}{\partial \xi_k} \delta \xi_k$$

Then we obtain

$$\frac{\partial}{\partial \xi_i} (\delta S_{ij} f_j) = \frac{\partial}{\partial \xi_r} \left(\epsilon_{pqj} \epsilon_{rsi} \frac{\partial x_p}{\partial \xi_k} \frac{\partial x_q}{\partial \xi_s} \frac{\partial f_j}{\partial \xi_i} \delta \xi_k \right) \quad (29)$$



a) Reduced adjoint vs original adjoint and complex step



b) Finite difference vs complex step

Fig. 5 Euler drag minimization for RAE 2822: comparison of original adjoint, reduced adjoint, complex-step, and finite difference gradients using three mesh-point bump as design variable. Fine grid: 384 × 96.

The term in $\partial/\partial\xi_1$ is

$$\epsilon_{123}\epsilon_{pqj}\frac{\partial x_p}{\partial\xi_k}\left(\frac{\partial x_q}{\partial\xi_2}\frac{\partial f_j}{\partial\xi_3}-\frac{\partial x_q}{\partial\xi_3}\frac{\partial f_j}{\partial\xi_2}\right)\delta\xi_k$$

Here the term multiplying $\delta\xi_1$ is

$$\epsilon_{jpq}\left(\frac{\partial x_p}{\partial\xi_1}\frac{\partial x_q}{\partial\xi_2}\frac{\partial f_j}{\partial\xi_3}-\frac{\partial x_p}{\partial\xi_1}\frac{\partial x_q}{\partial\xi_3}\frac{\partial f_j}{\partial\xi_2}\right)$$

According to formulas (12), this can be recognized as

$$S_{2j}\frac{\partial f_1}{\partial\xi_2}+S_{3j}\frac{\partial f_1}{\partial\xi_3}$$

or, using the quasi-linear form (14) of the equation for steady flow, as

$$-S_{1j}\frac{\partial f_1}{\partial\xi_1}$$

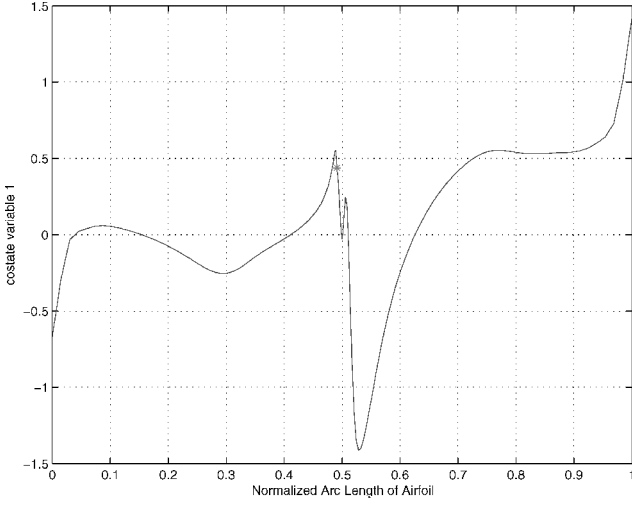
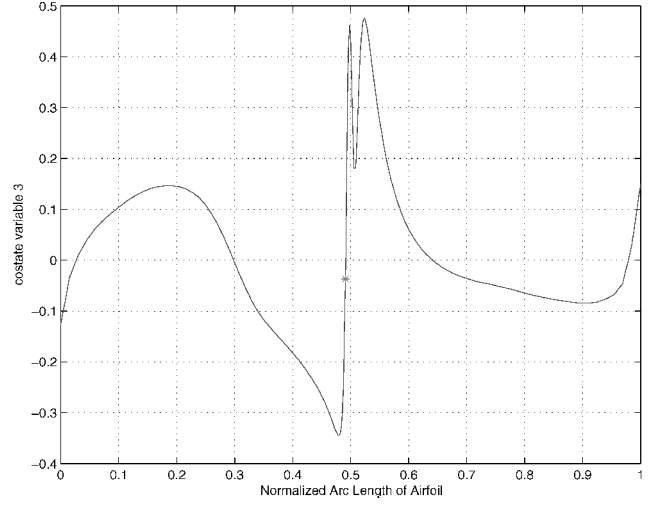
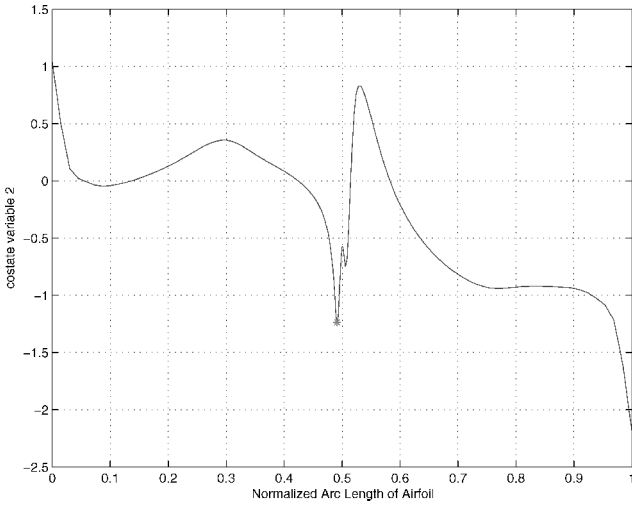
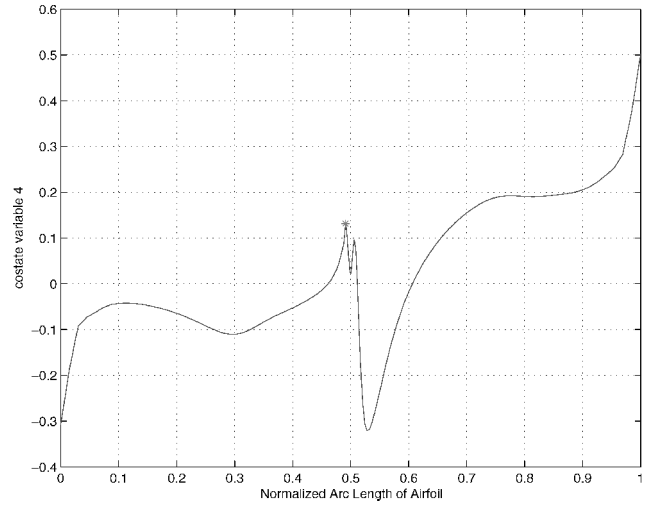
a) ψ_1 c) ψ_3 b) ψ_2 d) ψ_4

Fig. 6 Rapid variation of adjoint variables across the stagnation point.

The terms multiplying $\delta\xi_2$ and $\delta\xi_3$ are

$$\epsilon_{j pq} \left(\frac{\partial x_p}{\partial \xi_2} \frac{\partial x_q}{\partial \xi_2} \frac{\partial f_j}{\partial \xi_3} - \frac{\partial x_p}{\partial \xi_2} \frac{\partial x_q}{\partial \xi_3} \frac{\partial f_j}{\partial \xi_2} \right) = -S_{1j} \frac{\partial f_1}{\partial \xi_2}$$

$$\epsilon_{j pq} \left(\frac{\partial x_p}{\partial \xi_3} \frac{\partial x_q}{\partial \xi_2} \frac{\partial f_j}{\partial \xi_3} - \frac{\partial x_p}{\partial \xi_3} \frac{\partial x_q}{\partial \xi_3} \frac{\partial f_j}{\partial \xi_2} \right) = -S_{1j} \frac{\partial f_1}{\partial \xi_3}$$

Thus the term in $\partial/\partial\xi_1$ is reduced to

$$-\frac{\partial}{\partial \xi_1} \left(S_{1j} \frac{\partial f_1}{\partial \xi_k} \delta \xi_k \right)$$

Finally, with similar reductions of the terms in $\partial/\partial\xi_2$ and $\partial/\partial\xi_3$ we obtain

$$\frac{\partial}{\partial \xi_i} (\delta S_{ij} f_j) = -\frac{\partial}{\partial \xi_i} \left(S_{ij} \frac{\partial f_j}{\partial \xi_k} \delta \xi_k \right) = -\frac{\partial}{\partial \xi_i} (C_i \delta w^*)$$

as was to be proved.

Optimization Procedure

The search procedure used in this work is a descent method using a smoothed gradient. Let \mathcal{F} represent the design variable and \mathcal{G} the gradient. An improvement could then be made with a shape change:

$$\delta \mathcal{F} = -\lambda \mathcal{G} \quad (30)$$

In fact, however, the gradient \mathcal{G} is generally of a lower smoothness class than the shape \mathcal{F} , with the result that this process might fail to converge or even become unstable. To preserve the smoothness, we redefine the gradient to correspond to a weighted Sobolev inner product of the form

$$\langle u, v \rangle = \int \left(uv + \epsilon \frac{\partial u}{\partial \xi} \frac{\partial v}{\partial \xi} \right) d\xi$$

Thus, we define a modified gradient $\bar{\mathcal{G}}$ such that

$$\delta I = \langle \bar{\mathcal{G}}, \delta \mathcal{F} \rangle$$

In the one-dimensional case, taking $\bar{\mathcal{G}} = 0$ at the endpoints, integration by parts yields

$$\delta I = \int \left(\bar{\mathcal{G}} - \frac{\partial}{\partial \xi_1} \epsilon \frac{\partial \bar{\mathcal{G}}}{\partial \xi_1} \right) \delta \mathcal{F} d\xi$$

Then $\bar{\mathcal{G}}$ is obtained by solving the smoothing equation

$$\bar{\mathcal{G}} - \frac{\partial}{\partial \xi_1} \epsilon \frac{\partial \bar{\mathcal{G}}}{\partial \xi_1} = \mathcal{G} \quad (31)$$

In the multidimensional case the smoothing is applied in product form. Finally, we set

$$\delta \mathcal{F} = -\lambda \bar{\mathcal{G}} \quad (32)$$

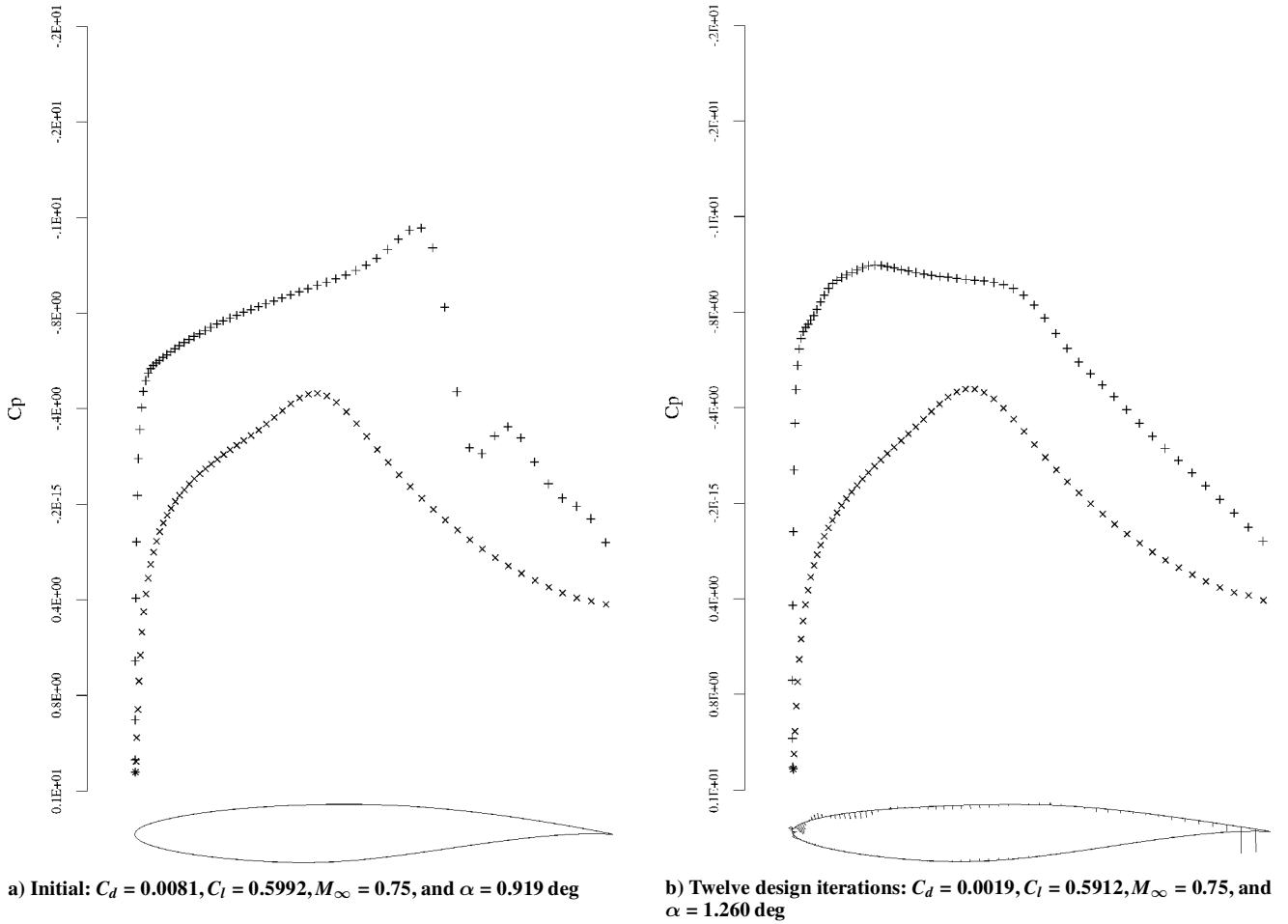


Fig. 7 Euler drag minimization at fixed $C_l = 0.6$ using reduced gradient formula, RAE 2822 airfoil.

with the result that

$$\delta I = -\lambda \langle \bar{\mathcal{G}}, \bar{\mathcal{G}} \rangle < 0$$

unless $\bar{\mathcal{G}} = 0$, and correspondingly $\mathcal{G} = 0$.

When second-order central differencing is applied to Eq. (31), the equation at a given node i can be expressed as

$$\bar{\mathcal{G}}_i - \epsilon(\bar{\mathcal{G}}_{i+1} - 2\bar{\mathcal{G}}_i + \bar{\mathcal{G}}_{i-1}) = \mathcal{G}_i, \quad 1 \leq i \leq n$$

where \mathcal{G}_i and $\bar{\mathcal{G}}_i$ are the point gradients at node i before and after the smoothing respectively and n is the number of design variables equal to the number of mesh points in this case. Then,

$$\bar{\mathcal{G}} = A\mathcal{G}$$

where A is the $n \times n$ tridiagonal matrix such that

$$A^{-1} = \begin{bmatrix} 1+2\epsilon & -\epsilon & 0 & \cdot & 0 \\ \epsilon & \cdot & \cdot & \cdot & \\ 0 & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & -\epsilon \\ 0 & \cdot & \cdot & \epsilon & 1+2\epsilon \end{bmatrix}$$

Now using the steepest descent method in each design iteration, a step $\delta\mathcal{F}$ is taken such that

$$\delta\mathcal{F} = -\lambda A\mathcal{G} \quad (33)$$

As can be seen from the form of this expression, implicit smoothing can be regarded as a preconditioner that allows the use of much larger steps for the search procedure and leads to a large reduction in the number of design iterations needed for convergence. Our software also includes an option for Krylov acceleration.¹⁸ We have found this to be particularly useful for inverse problems.

Results

This section presents the results of numerical tests of the new gradient formula. Its accuracy is assessed by comparison with the original adjoint-based, complex-step, and finite difference gradients. Examples are presented for a two-dimensional airfoil design, for a three-dimensional wing inverse problem, and for the wing redesign of a wing-fuselage combination. In all cases the flow is modeled by the Euler equations.

Gradient Comparison for Two-Dimensional Design

Figure 1 shows the comparison of the gradients obtained using the reduced adjoint method and the original adjoint method. Computations were performed for an RAE 2822 airfoil at a fixed coefficient of lift $C_l = 0.6$ and $M_\infty = 0.75$, using a C mesh of size 192×32 for both the flow and adjoint solutions. Every mesh point on the airfoil was used as a design variable, and the pressure drag C_d was used for the cost function. In this case a comparison with finite difference gradients was not made because accurate finite difference gradients cannot be calculated with the shape discontinuity caused by movement of a single mesh point. As shown in Fig. 1, the gradients obtained using both the original and new adjoint formulas agree well. Some discrepancies exist around the leading-edge region where the curvature is large. We believe that the new formulas might be sensitive to details of their discretization in regions of high curvature. Also, because the reduced adjoint gradient equations are integrals over a surface only the new formulas are more sensitive to details of their integration over regions of steep pressure gradient. In fact, small differences in gradients are observed both in the regions of high pressure gradient near the leading edge and in the regions where the shock appears.

The following figures show a comparison between the gradients obtained using the adjoint-based methods, the finite difference

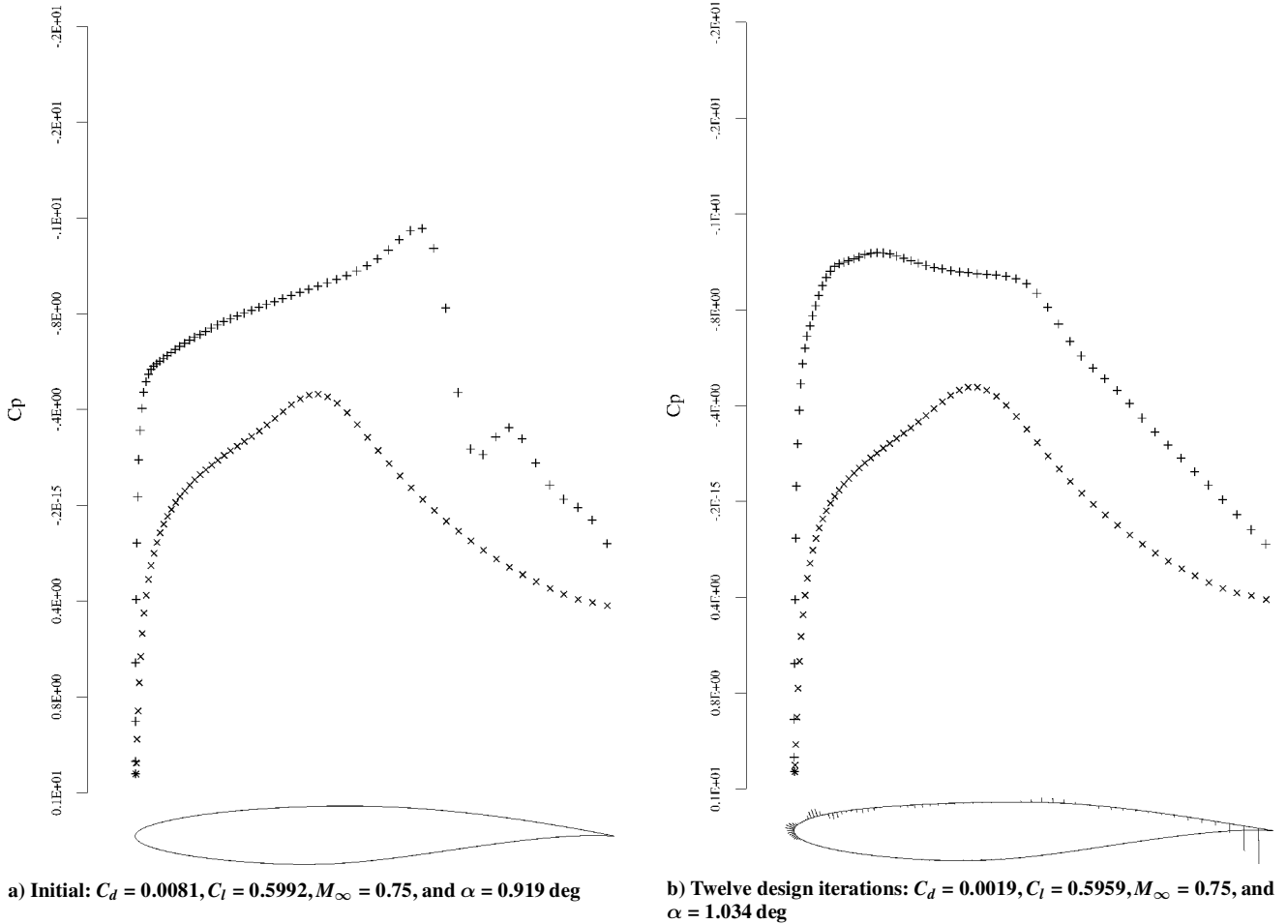


Fig. 8 Euler drag minimization at fixed $C_l = 0.6$ using original gradient formula, RAE 2822 airfoil.

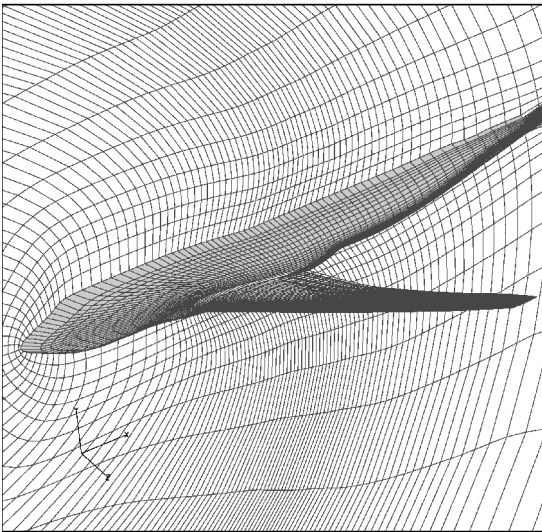


Fig. 9 C-H grid ($192 \times 32 \times 32$) for wing redesign of Boeing 747 wing-body configuration.

method, and the complex-step technique. To avoid geometric discontinuities, bumps were generated such that, while the same movement of each mesh point in turn was made as before, one-fourth of the movement was distributed to its two neighboring points, corresponding to local B splines. Although more continuous bumps could be used, this approach was selected in order to generate gradient information similar to the preceding comparison. Figure 2 illustrates the values of the gradients obtained from the reduced adjoint, original

Table 1 Difference between adjoint and complex-step gradient

Grid size	Original	New	Original/new
128×32	$1.06e-01$	$1.41e-01$	$1.11e-01$
192×32	$8.54e-02$	$1.21e-01$	$8.48e-02$
384×32	$1.62e-01$	$2.02e-01$	$8.07e-02$
192×64	$8.67e-02$	$1.04e-01$	$5.46e-02$
192×96	$9.56e-02$	$1.09e-01$	$4.93e-02$
384×96	$7.27e-02$	$8.56e-02$	$4.40e-02$

adjoint, complex-step, and finite difference methods. To eliminate any possible effect of the flow solution convergence on the finite difference and complex-step gradient accuracy, flow solutions were converged to more than seven orders of magnitude. A step size h of 10^{-4} of the distance of first node from the surface was chosen for the finite difference gradient. Because the complex-step method has no roundoff error in subtraction, a step size of 10^{-6} of the step size used for the finite difference gradient was chosen in order to get an accurate complex-step gradient, with a negligible truncation error $\mathcal{O}(h^2)$. Figure 2a shows that the three gradients from the reduced adjoint, original adjoint, complex-step methods agree quite well, although there is a discrepancy near the leading edge and at the location of the shock. The difference between the original adjoint and the complex-step gradient is slightly smaller than that between the reduced and the complex-step gradient. The difference between the adjoint-based gradients is smaller than that between the complex-step and adjoint-based gradients. Figure 2b shows the gradient comparison between the finite difference and complex-step methods. The difference is the greatest in this case, presumably because of the sensitivity of the finite difference gradient to the step size.

Table 1 shows values of the normalized difference (for example, $\epsilon = \|\text{adjoint} - \text{complex}\|_2 / \|\text{complex}\|_2$) between the adjoint

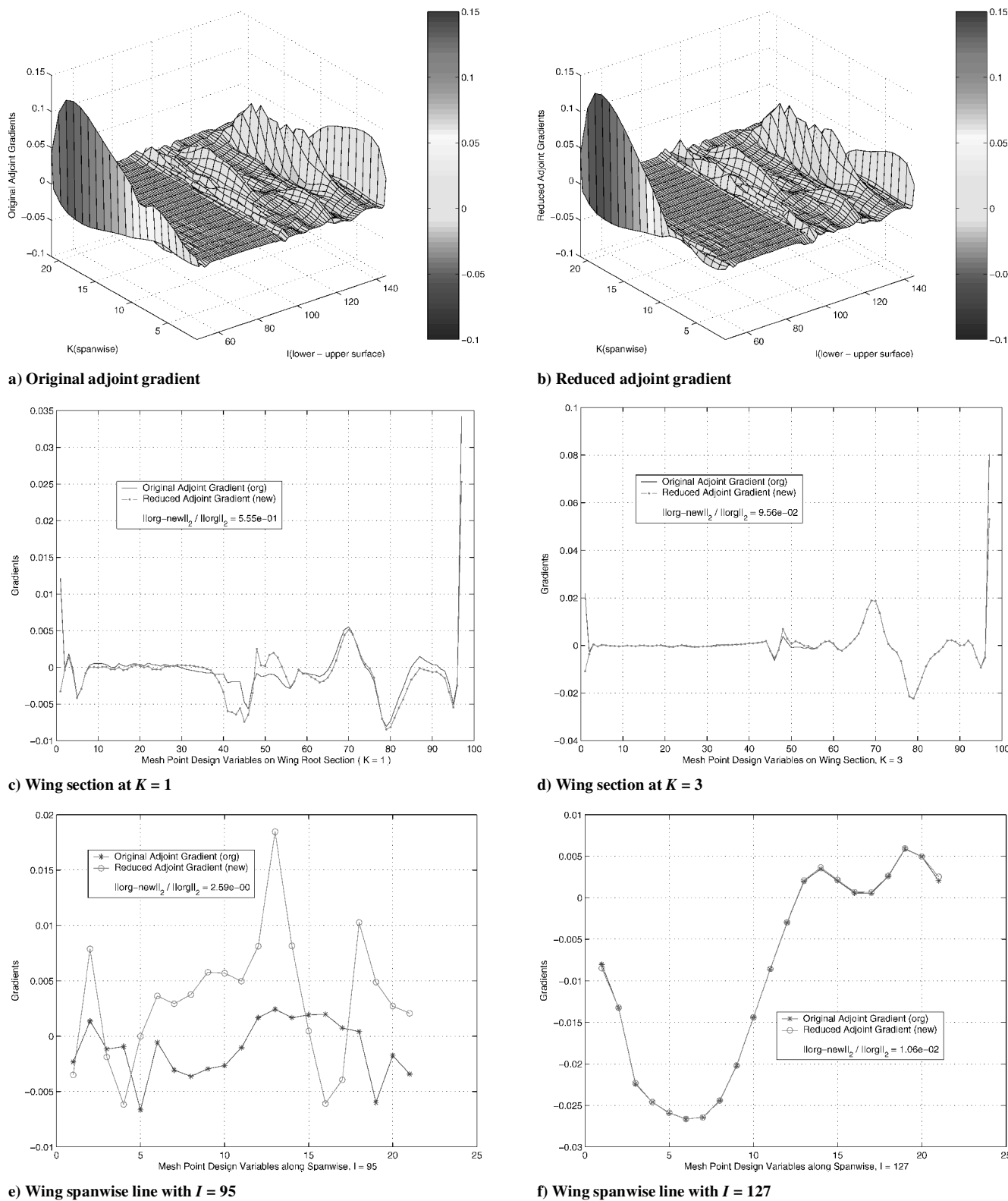


Fig. 10 Comparison between original adjoint and reduced adjoint gradients for a wing. Test case: Boeing 747-200, wing-body configuration, $M_\infty = 0.87$, and fixed $\alpha = 2.3$ deg.

and complex-step gradients using different mesh sizes. The second column depicts the difference between the original adjoint and complex-step gradients, the third column depicts the difference between the reduced adjoint and complex-step gradients, and the last column depicts the difference between the original adjoint and reduced adjoint gradients. The normalized difference between the complex-step gradients and two adjoint-based gradients does not decrease much as the mesh interval is decreased, possibly because

the bumps are too abrupt to produce reliable flow variations. However we can see first that the difference between the original adjoint and complex-step gradients is always somewhat smaller than the difference between the reduced adjoint and complex-step gradients, and second, that the difference between the original adjoint and reduced adjoint gradients decreases as the mesh interval is decreased. Figures 3–5 show three of the gradient comparisons summarized in Table 1. It can be seen that the difference near the leading edge is

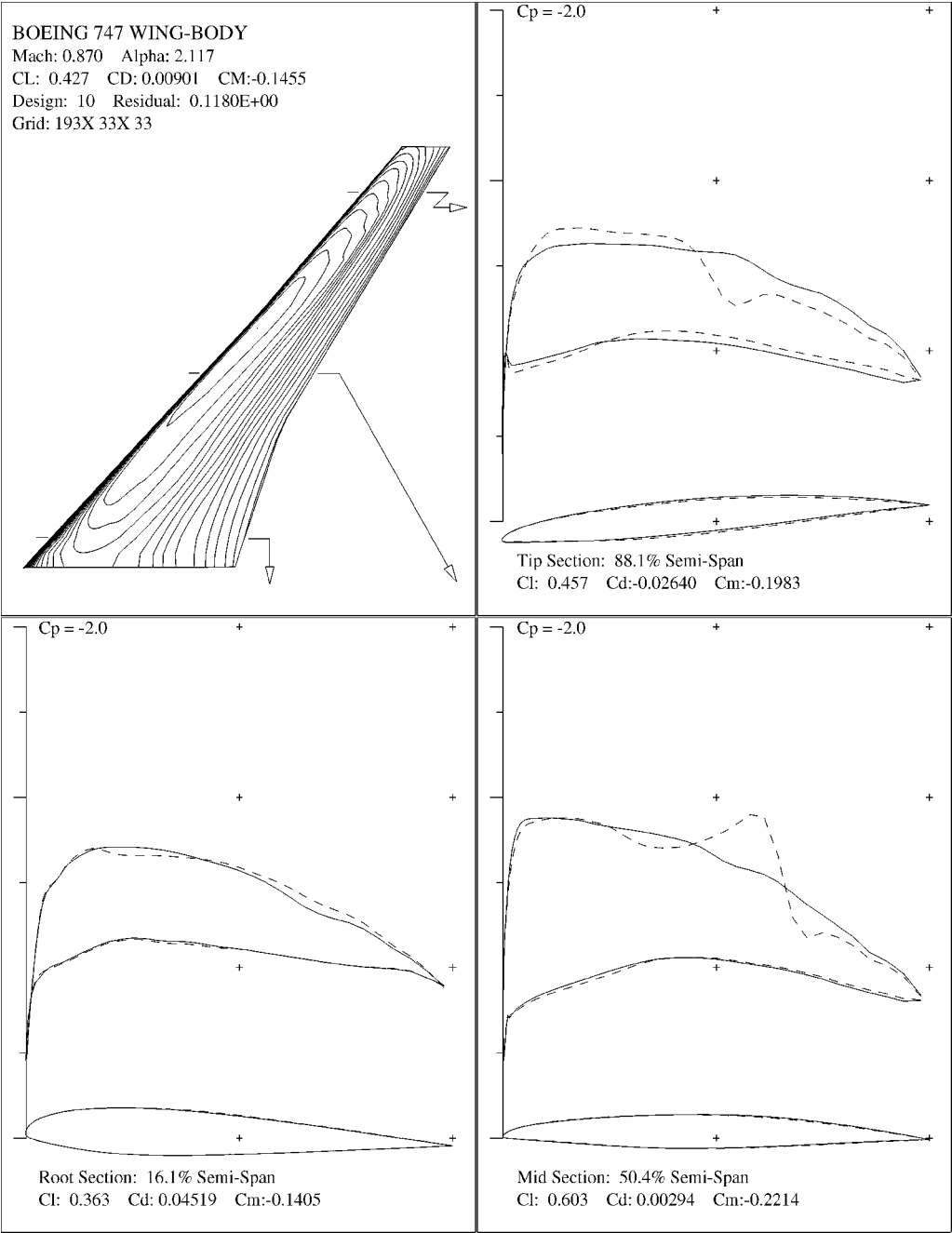


Fig. 11 Solution after 10 design iterations; Euler drag minimization at fixed $CL = 0.425$ using reduced gradient formula; Boeing 747 wing body: ---, initial C_p and —, redesigned C_p .

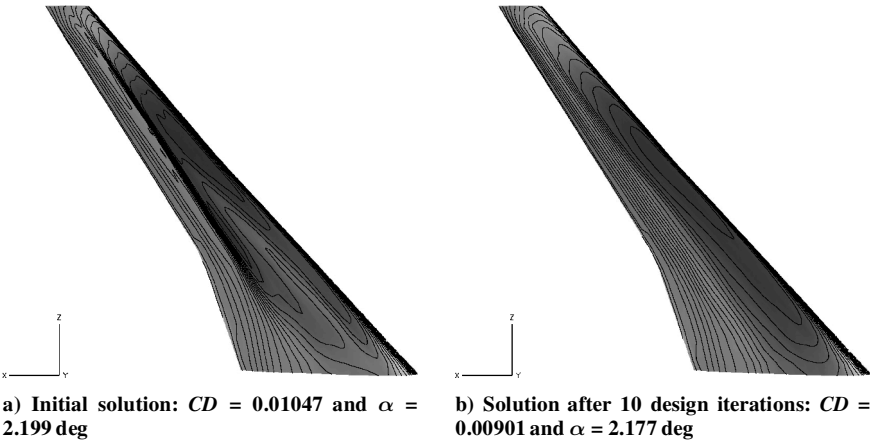
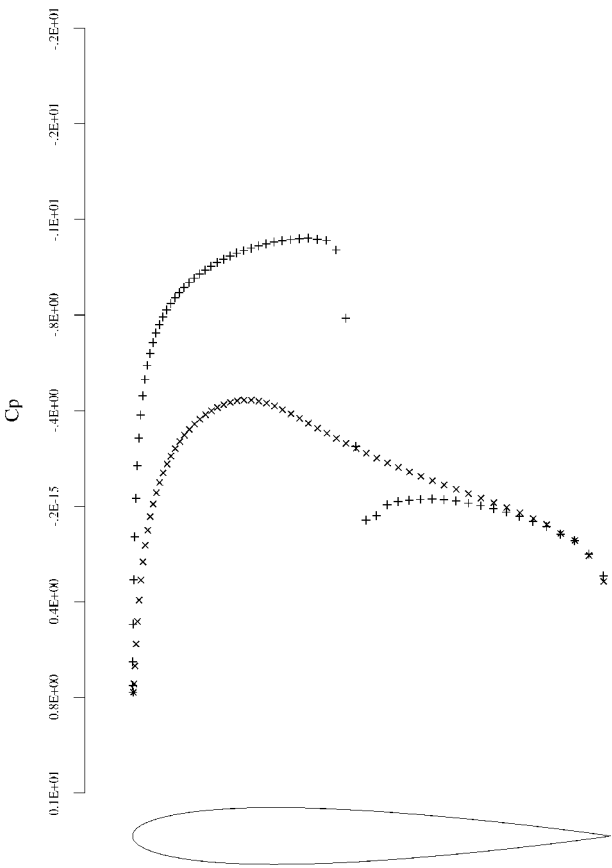
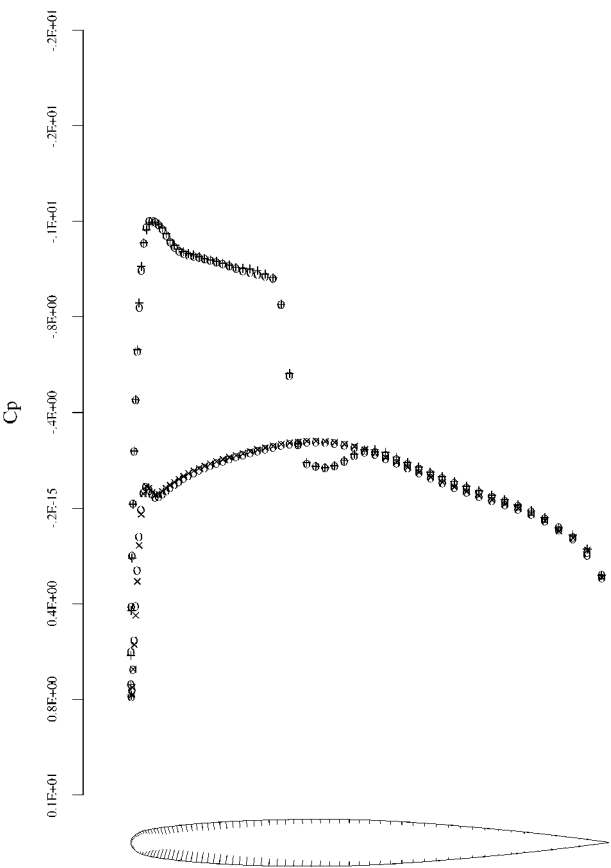


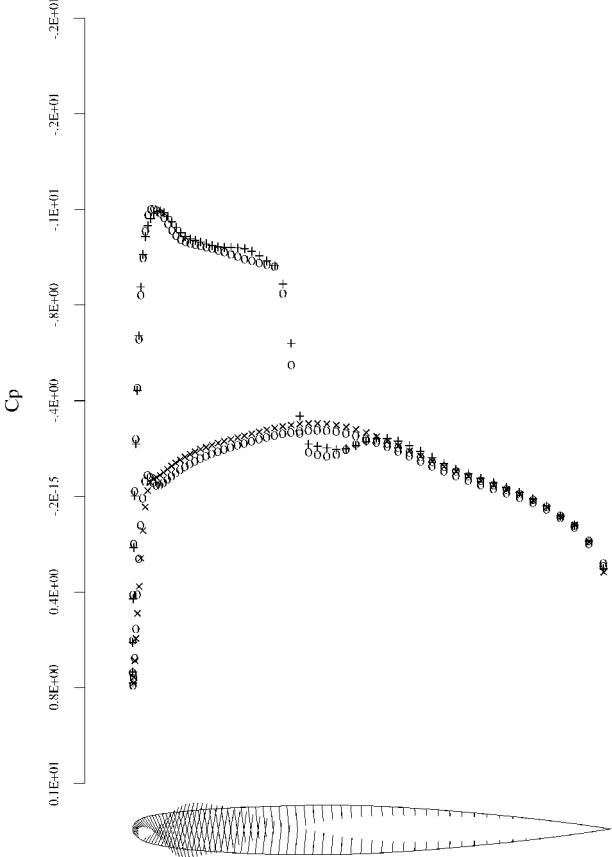
Fig. 12 Pressure contours over the upper surface of the wing. Test case: Boeing 747-200, wing-body configuration, $M_\infty = 0.87$, and fixed $CL = 0.425$.



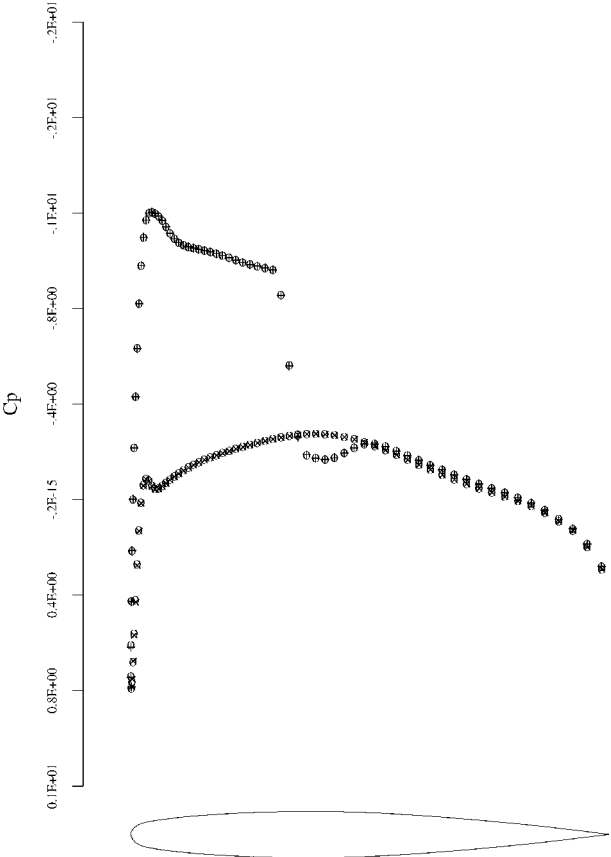
a) Initial solution of NACA 0012 wing: $CL = 0.305$, $CD = 0.020$, and $\alpha = 2.954$ deg



c) After 20 design iterations: $CL = 0.305$, $CD = 0.014$, and $\alpha = 3.151$ deg



b) After 10 design iterations, $CL = 0.304$, $CD = 0.015$, and $\alpha = 3.149$ deg



d) After 100 design iterations, $CL = 0.304$, $CD = 0.014$, and $\alpha = 3.137$ deg

Fig. 13 Euler inverse design of NACA 0012 wing to Onera M6 wing: $M_\infty = 0.84$, fixed $CL = 0.305$, and 85.9% semispan.

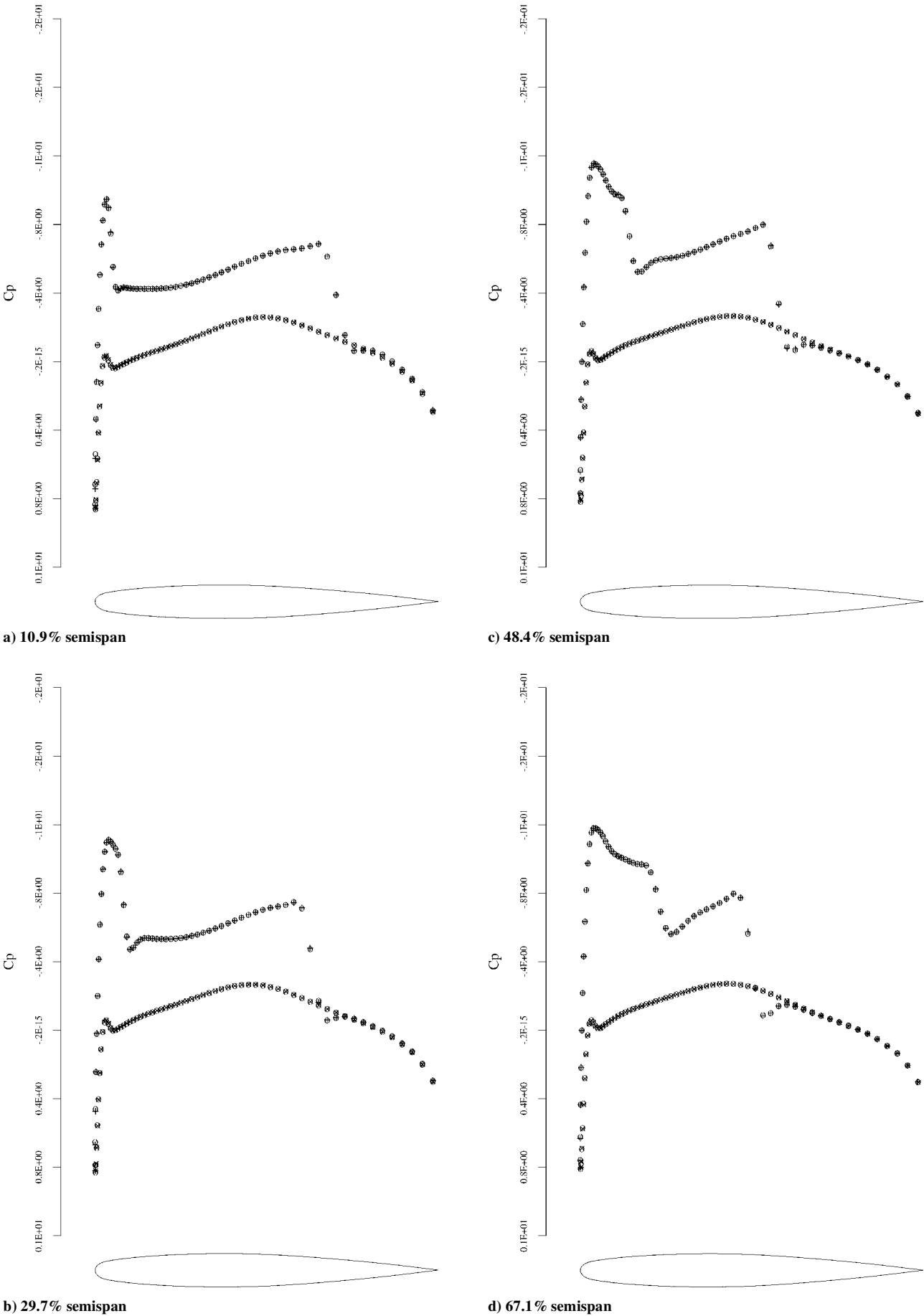


Fig. 14 Pressure distributions of redesigned Onera M6 wing: \circ , target and +, computed. $M_\infty = 0.84$, $CL = 0.304$, $CD = 0.014$, and $\alpha = 3.137$ deg.

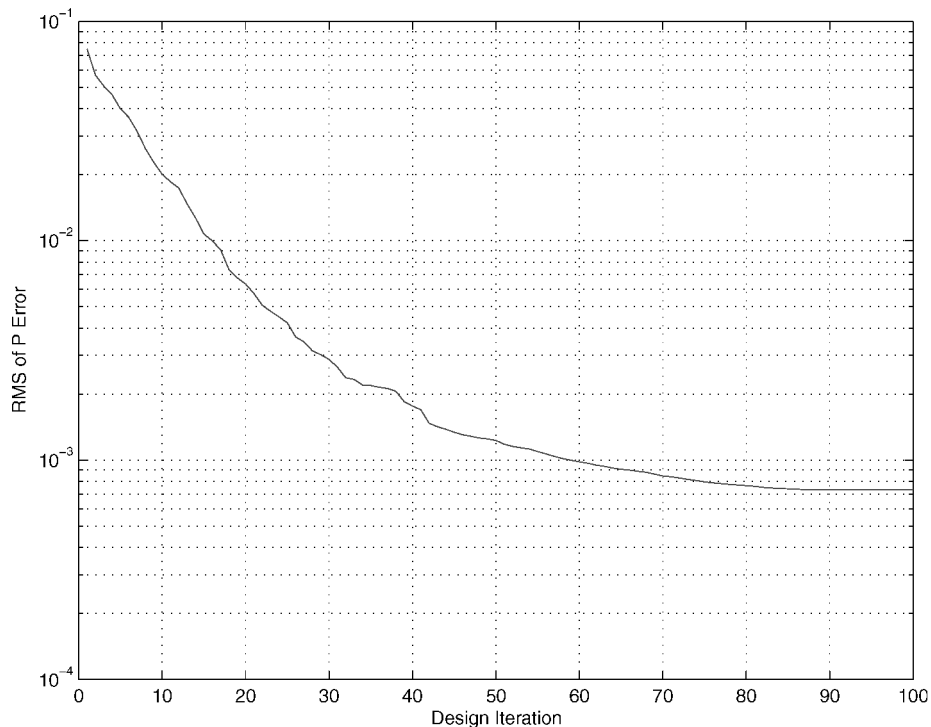


Fig. 15 Convergence history of pressure difference for the inverse wing design.

reduced and that the differences are more localized to the regions of high pressure gradient as the mesh is refined. This verifies the proper formulation of the new reduced adjoint gradient, while confirming that the new formulas are more sensitive to details of their discretization in regions of high curvature and high surface-pressure gradient. Moreover, as shown in Fig. 6, there are rapid variations in the costate variables across the stagnation point (indicated by * in the figure), which is very close to the leading-edge point on the airfoil. The authors believe that the new method might be more sensitive to large gradients in the costate solution, and this might explain the remaining difference at the leading-edge regions visible in Fig. 5. A detailed study of adjoint solution behavior can be found in Giles and Pierce's previous work.¹⁹

Drag Minimization for RAE 2822 Airfoil

To validate the usefulness of the reduced adjoint gradient formula in aerodynamic shape design, a drag minimization for the RAE 2822 airfoil has been performed, and the result has been compared with the result using the original adjoint gradients. The C mesh of size 192×32 was used, and every mesh point on the airfoil was chosen as a design variable. The gradients computed by both the reduced adjoint and the original adjoint formula were modified by the implicit smoothing procedure just described. Figure 7 illustrates the drag minimization of the RAE 2822 airfoil using the reduced adjoint formulation at $M_\infty = 0.75$ and fixed $C_l = 0.6$. After 12 design iterations the initial shock was completely removed, and the coefficient of drag has been reduced from $C_d = 0.0081$ to 0.0019. Figure 8 shows the same design case using the original adjoint gradients. The lines drawn normal to the airfoils of Figs. 7b and 8b represent the gradients calculated. It can be seen that the differences in the gradients from the two methods result in slightly different airfoils, showing small differences in the C_p distributions, particularly around the leading edge. However, the performance improvements from the two methods are almost identical, with a reduction of 62 drag counts being achieved in the same number of design iterations.

Gradient Comparison for Wing Redesign of Boeing 747 Wing-Body Configuration

A gradient comparison was performed for the Boeing 747 wing-body configuration in order to verify the proper implementation of the reduced gradient formulas for three-dimensional design

applications. Gradients with respect to mesh points on the wing of the Boeing 747 were computed using both the original adjoint and reduced adjoint methods on a C-H grid of size $192 \times 32 \times 32$. The surface mesh is shown in Fig. 9. The results from the original adjoint and reduced adjoint are shown respectively in Figs. 10a and 10b. Overall, the gradients agreed very well all over the wing surface. In the figures I is the index of mesh, which runs from lower trailing edge to leading edge and to upper trailing edge, and K is the index of mesh running spanwise from root wing section to tip wing section. Figure 10c shows the comparison of the gradients at the root wing section, where the maximum differences were observed, and Figure 10d shows gradient comparison at the wing section of $K = 3$, where the differences in gradients are the least. Similarly, Figs. 10e and 10f show the gradients of maximum ($I = 95$) and minimum ($I = 127$) differences at a fixed I , respectively. As in the two-dimensional case, the maximum differences were observed at the leading edge of the wing.

Drag Minimization for Wing Redesign of Boeing 747 Wing-Body Configuration

To validate the use of the reduced adjoint formulation in three-dimensional design, a drag minimization of the Boeing 747 wing-body configuration was performed by modifying the wing with the fuselage shape fixed. An Euler calculation was carried out for the wing-body configuration at a fixed coefficient of lift $CL = 0.425$ and $M_\infty = 0.87$ using the same C-H grid of size $192 \times 32 \times 32$. Figure 11 shows a comparison of the flow over the initial configuration with that over the improved configuration after 10 design cycles. The final pressure contours on the upper surface are displayed in the upper-left quadrant, while the other three quadrants superpose the C_p distributions at root, mid-, and tip sections. Color contours of the upper surface pressure before and after the redesign are displayed in Fig. 12. While keeping CL fixed close to 0.425, the total wing drag was reduced from $CD = 0.01047$ to 0.00901. The initial shock was completely removed over the entire wing, while rather small modifications of the each section shape were required.

Wing Inverse Problem

The last demonstration is an application of the new adjoint method for inverse design. The pressure distribution of the ONERA M6 wing at $\alpha = 3$ deg and $M_\infty = 0.84$ is taken as the target pressure. Starting

from an initial wing with an NACA 0012 section, the ONERA M6 wing is to be recovered by minimizing the pressure difference from the target pressure. This is a hard problem because it calls for the recovery of a smooth symmetric profile from an asymmetric pressure distribution with a double-shock pattern. The calculations were performed on a $192 \times 32 \times 48$ C-H grid. Figure 13a shows the solution for the wing section of NACA 0012 airfoil at $M_\infty = 0.84$ and $CL = 0.305$. In the following figures the \circ denotes the target pressure distribution, the $+$ denotes the current upper surface pressure, and the \times denotes the lower surface-pressure distribution. After only 10 design cycles the general shape of the target section at 86% semispan was recovered, as shown in Fig. 13b. Furthermore, 93% of the total reduction in the cost function was achieved in only 20 design cycles. After 100 design iterations the desired target section was completely obtained. Figure 13d shows a perfect C_p match even inside the shock. Similar point-to-point matches in the C_p distribution were observed at all span stations. The results for wing sections at 10.9, 29.7, 48.4, and 67.1% semispan are shown in Fig. 14. As can be seen in the figure, the double shock in each section (λ shock of the ONERA M6 wing) was completely captured by the redesign. Finally, Fig. 15 illustrates the convergence history: full convergence was obtained in 83 design cycles.

Conclusions

The methods described in this work are the culmination of ongoing studies over the past 14 years, since the adjoint method was first formulated for shape optimization in transonic flow.¹ The numerical tests establish that accurate gradient information can be obtained by the new reduced adjoint formulation. Although the original adjoint gradients agree slightly better than the new reduced adjoint gradients with complex-step gradients, the difference between the original adjoint and reduced adjoint gradients decreases as the mesh interval decreases. Although the formulas are not exactly equivalent for the discretized equations, the optimization process works just as well with the reduced gradient formula as it did with the original formula. By eliminating the dependence of the gradient formulas on the mesh perturbation, the new formulas provide the opportunity to make a drastic simplification of gradient calculations on overset and unstructured meshes.

Acknowledgment

This work has benefited greatly from the support of the U.S. Air Force Office of Scientific Research under Grant AF F49620-98-1-2002.

References

- ¹Jameson, A., "Aerodynamic Design via Control Theory," *Journal of Scientific Computing*, Vol. 3, 1988, pp. 233–260.
- ²Jameson, A., "Optimum Aerodynamic Design Using CFD and Control Theory," AIAA Paper 95-1729, June 1995.
- ³Reuther, J., Jameson, A., Farmer, J., Martinelli, L., and Saunders, D., "Aerodynamic Shape Optimization of Complex Aircraft Configurations via an Adjoint Formulation," AIAA Paper 96-0094, Jan. 1996.
- ⁴Reuther, J., Alonso, J. J., Vassberg, J. C., Jameson, A., and Martinelli, L., "An Efficient Multiblock Method for Aerodynamic Analysis and Design on Distributed Memory Systems," AIAA Paper 97-1893, June 1997.
- ⁵Reuther, J. J., Jameson, A., Alonso, J. J., Rimlinger, M., and Saunders, D., "Constrained Multipoint Aerodynamic Shape Optimization Using an Adjoint Formulation and Parallel Computers: Part I," *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 51–60.
- ⁶Reuther, J. J., Jameson, A., Alonso, J. J., Rimlinger, M., and Saunders, D., "Constrained Multipoint Aerodynamic Shape Optimization Using an Adjoint Formulation and Parallel Computers: Part II," *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 61–74.
- ⁷Pironneau, O., *Optimal Shape Design for Elliptic Systems*, Springer-Verlag, New York, 1984.
- ⁸Jameson, A., "Re-Engineering the Design Process Through Computation," *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 36–50.
- ⁹Reuther, J., Alonso, J. J., Rimlinger, M. M., and Jameson, A., "Aerodynamic Shape Optimization of Supersonic Aircraft Configurations via an Adjoint Formulation on Parallel Computers," *Computers and Fluids*, Vol. 28, No. 4, 1999, pp. 675–700.
- ¹⁰Baysal, O., and Eleshaky, M. E., "Aerodynamic Design Optimization Using Sensitivity Analysis and Computational Fluid Dynamics," AIAA Paper 91-0471, Jan. 1991.
- ¹¹Anderson, W. K., and Venkatakrishnan, V., "Aerodynamic Design Optimization on Unstructured Grids with a Continuous Adjoint Formulation," *Computers and Fluids*, Vol. 28, No. 4, 1999, pp. 443–480.
- ¹²Nadarajah, S., and Jameson, A., "A Comparison of the Continuous and Discrete Adjoint Approach to Automatic Aerodynamic Optimization," AIAA Paper 2000-0667, Jan. 2000.
- ¹³Nadarajah, S., and Jameson, A., "Studies of the Continuous and Discrete Adjoint Approaches to Viscous Automatic Aerodynamic Optimization," AIAA Paper 2001-2530, June 2001.
- ¹⁴Weinfelt, P., and Enoksson, O., "Numerical Methods for Aerodynamic Optimization," Technical Rept., 8th International Symposium on Computational Fluid Dynamics, Bremen, Germany, Sept. 1999.
- ¹⁵Lyness, J. N., and Moler, C. B., "Numerical Differentiation of Analytic Functions," *SIAM Journal on Numerical Analysis*, Vol. 4, 1967, pp. 202–210.
- ¹⁶Kim, S., Alonso, J. J., and Jameson, A., "A Gradient Accuracy Study for the Adjoint-Based Navier-Stokes Design Method," AIAA Paper 99-0299, Jan. 1999.
- ¹⁷Martins, J. R. R. A., Kroo, I. M., and Alonso, J. J., "An Automated Method for Sensitivity Analysis Using Complex Variables," AIAA Paper 2000-0689, Jan. 2000.
- ¹⁸Jameson, A., and Vassberg, J. C., "Studies of Alternative Numerical Optimization Methods Applied to the Brachistochrone Problem," *Computational Fluid Dynamics*, Vol. 9, 2000, pp. 281–296.
- ¹⁹Giles, M. B., and Pierce, N. A. M., "Adjoint Equations in CFD: Duality, Boundary Conditions and Solution Behaviour," AIAA Paper 97-1850, June 1997.

M. Ahmadian
Associate Editor