Optimization and Engineering, 6, 33–62, 2005

A Coupled-Adjoint Sensitivity Analysis Method for High-Fidelity Aero-Structural Design

JOAQUIM R.R.A. MARTINS
University of Toronto Institute for Aerospace Studies, Toronto, ON M3H 5T6, Canada
email: martins@utias.utoronto.ca

JUAN J. ALONSO
Stanford University, Stanford, CA 94305, USA
email: jjalonso@stanford.edu

JAMES J. REUTHER
NASA Ames Research Center, Moffett Field, CA 95035, USA
email: jreuther@mail.arc.nasa.gov

Received August 19, 2002; Revised September 24, 2003

Abstract. This paper presents an adjoint method for sensitivity analysis that is used in an aero-structural aircraft design framework. The aero-structural analysis uses high-fidelity models of both the aerodynamics and the structures. Aero-structural sensitivities are computed using a coupled-adjoint approach that is based on previously developed single discipline sensitivity analysis. Alternative strategies for coupled sensitivity analysis are also discussed. The aircraft geometry and a structure of fixed topology are parameterized using a large number of design variables. The aero-structural sensitivities of aerodynamic and structural functions with respect to these design variables are computed and compared with results given by the complex-step derivative approximation. The coupled-adjoint procedure is shown to yield very accurate sensitivities and to be computationally efficient, making high-fidelity aero-structural design feasible for problems with thousands of design variables.

Keywords: sensitivity analysis, adjoint method, aero-structural design, complex-step method

1. Introduction

During the past decade the advancement of numerical methods for the analysis of complex engineering problems such as those found in fluid dynamics and structural mechanics has reached a mature stage: many difficult numerically intensive problems are now readily solved with modern computer facilities. In fact, the aircraft design community is increasingly using computational fluid dynamics (CFD) and computational structural mechanics (CSM) tools to replace traditional approaches based on simplified theories and wind tunnel testing. With the advancement of these numerical analysis methods well underway, the focus for engineers is shifting toward integrating these analysis tools into numerical design procedures.

These design procedures are usually based on computational analysis methods that evaluate the relative merit of a set of feasible designs. The performance of a design is quantified

by an *objective function* that is computed using one or several analysis tools, while the design parameters are controlled by an optimization algorithm.

There are a large number of optimization algorithms, but they all fall into one of two main categories. In the first category we have the *zeroth order methods* which include grid searches, genetic algorithms, neural networks, random searches and simulated annealing. None of these approaches rely on any information other than the value of the objective function. Their most significant limitation is that, as the number of design parameters increases, the number of function evaluations needed to reach the optimum rapidly increases beyond what is computationally feasible.

Optimization algorithms in the second category use not only the value of the objective function but also its gradient with respect to the design parameters. The main advantage of these *gradient methods* is that they converge to the optimum with a significantly smaller number of function evaluations. Unfortunately, these methods only work well when the objective function varies smoothly within the design space. Furthermore, these methods only guarantee convergence to a local optimum.

Both classes of optimization algorithms have a role in solving engineering problems. In a problem with a limited number of design variables that has multiple local minima or discontinuities, it is clear that a zeroth order method is more suitable. On the other hand, many single-discipline high-fidelity aircraft design problems are characterized by having a large number of design variables and smooth design spaces. These problems are amenable to the use of gradient-based optimization algorithms. In particular, gradient methods are used extensively for aerodynamic shape optimization (ASO) problems because these problems are often parameterized with hundreds of design variables and usually require computationally expensive high-fidelity analyses. With a few notable exceptions (Sasaki et al., 2001; Obayashi and Sasaki, 2002) these requirements make the use of zeroth order methods infeasible for high-fidelity ASO problems.

The field of sensitivity analysis emerged to address the need for computing gradients accurately and efficiently. A sensitivity analysis method that is very commonly used is finite differencing, where for each design variable, the design is perturbed and analyzed to determine the new value of the objective function. Although finite differences are not known for being particularly accurate or computationally efficient, they are extremely easy to implement. The complex-step derivative approximation (Martins et al., 2003) is a recently developed method for calculating sensitivities that maintains the simplicity of finite differences but is much more accurate. Yet another technique, algorithmic differentiation (Bischof et al., 1992; Griewank, 2000), automatically differentiates a given algorithm by adding source code that computes the required sensitivities. With the exception of the reverse mode of algorithmic differentiation (Griewank, 2000), the main disadvantage of the approaches mentioned thus far is that the cost of computing sensitivities is directly proportional to the number of design variables. Thus, even though a forward-based gradient calculation method is computationally more efficient than a zeroth order method, the cost of computing gradients can still be prohibitive when optimizing a design parameterized with a large number of variables.

Fortunately, there are other techniques for computing sensitivities whose cost is independent of the number of design variables. In particular, the adjoint method discussed in this

article is such an alternative. Adjoint methods allow the computation of sensitivities for an arbitrary number of design variables, at a cost that is similar to that of a single function evaluation. Applications of the adjoint method are now well known in CSM (Adelman and Haftka, 1986) and CFD (Jameson, 1989; Reuther et al., 1999a). However, only recently has the this method been applied to the computation of sensitivities of coupled high-fidelity systems (Martins et al., 2001, 2002; Maute et al., 2001).

Despite revolutionary accomplishments in single-discipline applications, progress towards the development of high-fidelity, multidisciplinary design optimization (MDO) methods has been slow. The level of coupling between disciplines is highly problem dependent and significantly affects the choice of algorithm. Multiple difficulties also arise from the heterogeneity among design problems: an approach that is applicable to one discipline may not be compatible with the others.

An important feature that characterizes the various solution strategies for MDO problems is the allowable level of disciplinary autonomy in the analysis and optimization components. Excellent discussions of these issues are presented by Sobieszczanski-Sobieski and Haftka (1996) and Alexandrov and Lewis (1999). The allowable level of disciplinary autonomy is usually inversely proportional to the bandwidth of the interdisciplinary coupling. Thus, for highly coupled problems it may be necessary to resort to fully integrated MDO, while for more weakly coupled problems, modular strategies may hold an advantage in terms of ease of implementation. With these constraints in mind, a number of ideas for solving complex MDO problems have been developed. These ideas include multilevel optimization strategies (Alexandrov and Dennis, 1994; Kodiyalam and Sobieszczanski-Sobieski, 2002), collaborative optimization (Braun and Kroo, 1996; Kroo, 1996; DeMiguel and Murray, 2000), individual discipline-feasible methods (Cramer et al., 1994), as well as tightly coupled optimization procedures. The main difference between the different MDO strategies is the degree of coupling that is required between the disciplines in both the analysis and the optimization procedures.

In the particular case of high-fidelity aero-structural optimization, the coupling between disciplines has a very high bandwidth. Furthermore, the values of the objective functions and constraints depend on highly coupled multidisciplinary analyses (MDA). Therefore, we believe that a tightly coupled MDO environment is more appropriate for aero-structural optimization.

The difficulty in formulating this type of MDO problem is that there are significant technical challenges when implementing tightly coupled analysis and design procedures. Not only must MDA be performed at each design iteration but, in the case of gradient-based optimization, the coupled sensitivities must also be computed at each iteration.

This work presents a tightly coupled approach to high-fidelity aero-structural MDO that uses CFD and CSM. Section 2 presents the analysis framework and explains in detail how the CFD and CSM software are integrated to obtain accurate and efficient aero-structural solutions. Section 3 describes the optimization problem that we propose to solve, placing into context the sensitivity calculations that are the primary focus of this article. Sensitivity analysis theory is presented in Section 4. In addition to the coupled-adjoint sensitivity analysis that we implement, a brief discussion of alternative coupled sensitivity analysis methods and their associated costs is presented. Finally, in Section 5, we show a sensitivity

validation study that demonstrates the accuracy and efficiency of this newly developed method. Timings are included for cases with thousands of design variables.

2. Aero-structural analysis framework

The foundation for both the coupled analysis procedure and the coupled-adjoint solver is an aero-structural design framework previously developed by the authors (Reuther et al., 1999c; Martins, 2002). The framework consists of an aerodynamic analysis and design module (which includes a geometry engine and a mesh perturbation algorithm), a linear finite-element structural analysis and design module, an aero-structural coupling procedure (for both the analyses and the sensitivities), and various pre-processing tools that are used to setup aero-structural design problems. The multidisciplinary nature of this framework is illustrated in Figure 1 where the aircraft geometry, the CFD mesh and flow solution, and the primary structure inside the wing are shown.

The aerodynamic analysis and design module, SYN107-MB (Reuther et al., 1999a), is a multiblock parallel flow solver that is applicable to both the Euler and the Reynolds averaged Navier–Stokes equations. This solver represents the state-of-the art, being accurate and efficient for the computation of the flow around full aircraft configurations (Reuther et al., 1997; Yao et al., 2001). SYN107-MB also includes an adjoint solver for aerodynamic sensitivity analysis that relies on the same mesh and solution strategy as the flow solver (Reuther et al., 1996). Completing the aerodynamics module are a parametric geometry engine and a mesh perturbation package (Reuther et al., 1999a, b).

The structural analysis package is FESMEH, a linear finite-element solver developed by Holden (1999). The package includes two element types that are suitable for computing

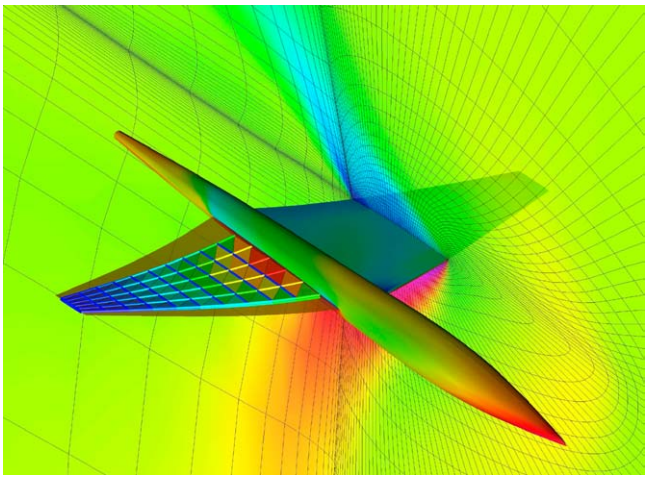

Figure 1. Aero-structural model and solution of a supersonic business jet configuration, showing a slice of the grid and the internal structure of the wing.

structural displacements and stresses of wing-like structures. Although this solver is not as general as some commercially-available packages, it is still representative of the challenges that are encountered when using large models with tens of thousands of degrees of freedom.

In aero-structural analysis, there is a clear interdependence in the equilibrium state of the two systems: the flow solution depends on deflections calculated by the structures, and the structural solution depends on the loads calculated by the flow solver. In this section we consider three aspects of the high-fidelity coupling between the aerodynamic and structural analyses: the geometry engine and the transfer of loads and displacements. The procedures for the transfer of displacements and loads are based on work by Brown (1997). Alternative procedures have been developed by other researchers (Maman and Farhat, 1995; Cebral and Löhner, 1997a, b).

2.1. Geometry engine and database

The aircraft is surrounded by fluid which is separated from the structure by the fluid-structure interface. Therefore, there is a well-defined surface in three-dimensional space which constitutes the outer-mold line (OML). Because of the importance of the OML in aero-structural analysis and design problems, a separate utility—*Aerosurf*—is used to generate and manage the OML. *Aerosurf* was specifically created for the analysis and design of aircraft configurations (Reuther et al., 1999a, b).

The baseline geometry of an aircraft configuration is given to *Aerosurf* in the form of separate components, each one being described by a series of cross-sections in three-dimensional space. These components can be fuselages, pylons, nacelles, and wing-like surfaces. *Aerosurf* intersects these components and divides the resulting surface into a series of patches. At this stage, *Aerosurf* creates a parametric description of each patch and then distributes points on their surface, forming a fine structured watertight mesh. Thus, the set of points formed by the grids of all patches represents a discretization of the OML within *Aerosurf*.

In addition to providing a high-fidelity description of the aircraft geometry, *Aerosurf* also manages a centralized database for the analysis and design of the aircraft. During analysis, any information that needs to be exchanged through the fluid-structure interface—such as aerodynamic pressures and structural displacements—is interpolated onto the OML points. Changes in the OML shape can be due to either structural displacements during aero-structural analysis or changes in shape design variables between design cycles. While the OML changes due to structural displacements are transferred directly to the OML points, changes due to shape design variables are applied to the un-intersected components first and then these components are re-intersected, creating a new discretized representation of the OML.

2.2. Displacement transfer

The displacements calculated by the CSM solver are first transferred onto the OML grid, and then onto the CFD surface mesh. Each OML point is associated with a point on the surface of the CSM model in a pre-processing step, as shown in Figure 2. The association

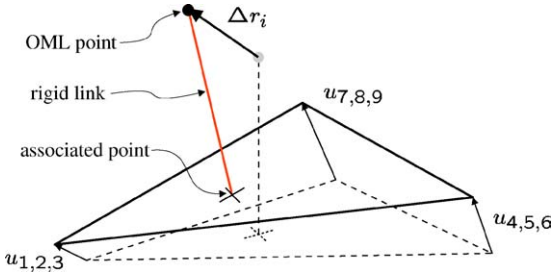

Figure 2. Displacement extrapolation procedure.

is performed by locating the point on the CSM model surface that is closest to each OML point.

During aero-structural analyses, the displacement of each associated point is computed by interpolating the CSM nodal displacements of the element containing that point. For consistency, the structural finite-element shape functions are used to perform this interpolation. (If the element basis functions are not known, as in the case of commercial CSM software, an isoparametric approximation is sufficiently accurate.) The displacement is then transferred onto the OML point using extrapolation functions that emulate a rigid link between the OML point and the associated point on the surface of the structural model. For small angular deflections we use the linear relationship

$$\Delta r_i = \eta_{ij} u_j, \tag{1}$$

where we use index notation to write the product of the matrix η_{ij} with the vector of CSM node displacements u_j.

Unlike the nodes of the CSM model, the CFD surface mesh points are assumed to exist on the OML. Figure 3 shows a representation of both the OML and CFD meshes. The parametric coordinates of the CFD surface mesh points on the corresponding OML patches are calculated in a pre-processing step via closest point projection. Therefore the patch number and the parametric coordinates of the associated point uniquely define the

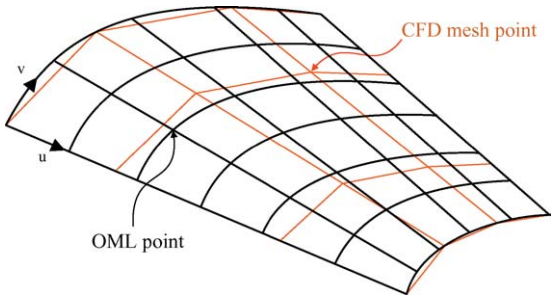

Figure 3. OML and CFD surface meshes on an *Aerosurf* patch.

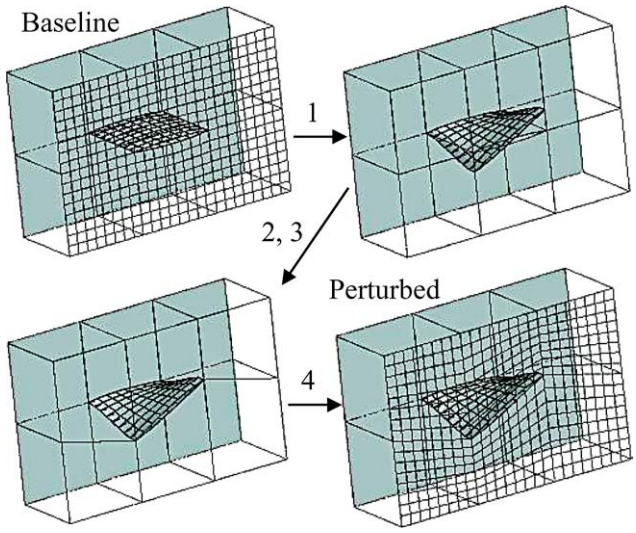

Figure 4. Mesh perturbation procedure used by WARPMB.

transfer operator. The CFD points are assumed to be "tied" to these parametric locations and any displacement of the OML, due to either design variable perturbations or structural displacements, is transferred to the CFD surface mesh points by evaluating their parametric locations on the corresponding *Aerosurf* patches.

Once a perturbation is applied to the surface of the CFD mesh, it must be propagated throughout the whole multiblock mesh. This volume mesh perturbation is achieved very efficiently by using the WARPMB algorithm. WARPMB perturbs the volume mesh in four stages and is illustrated in Figure 4. The procedure is as follows:

1. Block faces directly affected by the surface mesh movement—the active faces—are explicitly perturbed.
2. Edges with end points in contact with active faces—in the same or in adjacent blocks—are implicitly perturbed using an arc-length attenuation method.
3. The interiors of faces that are bordered by implicitly perturbed edges or share common edges with adjacent active faces are implicitly perturbed with WARP3QD, a three-dimensional in-plane mesh perturbation algorithm.
4. A final routine, WARP3D, is used to perturb the interiors of blocks that have at least one active or perturbed face.

2.3. Load transfer

The load transfer procedure consists in converting the pressures calculated by the CFD algorithm to the structural nodes through the OML points.

In order to transfer pressures from the CFD surface mesh to the OML points, we identify, in a pre-processing step, the appropriate "donor cell" and the parametric location of each

OML point within this cell. The pressures at the OML points are then calculated using bilinear interpolation. The underlying assumption that ensures the accuracy of this simple transfer is that the OML mesh is of comparable or better fidelity than that of the CFD surface mesh, and that the two surface representations are consistent and contiguous.

In translating interpolated pressures from the OML surface into CSM nodal forces, it is crucial that both consistency and conservation be maintained. The property of consistency specifies that the resultant forces and moments due to the pressure field, must be equal to the sum of the nodal forces and moments. There are an infinite number of CSM load vectors that satisfy this requirement. However, we also require that the load transfer be conservative. Conservation stipulates that the virtual work performed by the load vector, f_j, undergoing a virtual displacement of the structural model, δu_j, must be equal to the work performed by the pressure field, p, undergoing the equivalent displacement of the OML mesh, δr_i. The virtual work in the CSM model is given by the dot product

$$\delta W_{\text{CSM}} = f_j \delta u_j, \tag{2}$$

while the virtual work performed by the fluid acting on the surface of the OML mesh is given by the surface integration

$$\delta W_{\text{CFD}} = \int_S p \; n_i \delta r_i \; \mathrm{d}S, \tag{3}$$

where the integral is taken over the entire OML and n_i represents the the unit vector normal to the OML. For a conservative scheme, $\delta W_{\text{CFD}} = \delta W_{\text{CSM}}$, and a consistent and conservative load vector can be shown to be given by

$$f_j = \int_S p \; n_i \eta_{ij} \, \mathrm{d}S, \tag{4}$$

where we used the linear relationship (1) for the virtual displacements δr_i. In Figure 5 we can see how the pressure field (which has been interpolated from the CFD mesh to the points on the OML) is integrated over an OML patch to produce a force vector that is translated into the nodal forces of a CSM element using equation (4).

As mentioned in Section 2.2, the transfer matrix η_{ij} is calculated in a pre-processing step. Note that this matrix plays a dual role: it provides the appropriate weighting factors for both the transfer of OML pressures to CSM load vectors (4) and the transfer of the CSM displacements to OML point displacement (1).

2.4. Aero-structural iteration

The aerodynamic and structural solvers are coupled by exchanging information at regular intervals during the convergence process. This coupling is greatly simplified by the fact that we only consider static aeroelastic solutions, and hence time accuracy is not an issue.

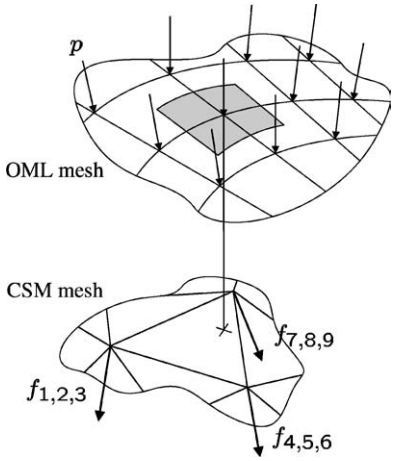

Figure 5. Transfer of the pressure on the OML points to the nodal forces on a given finite element.

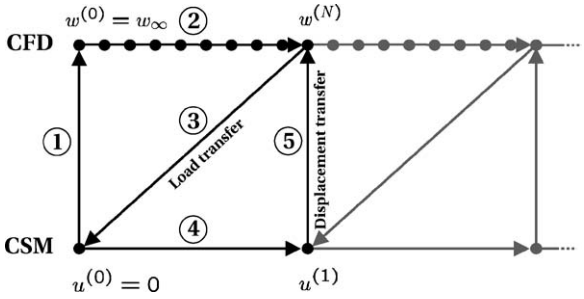

Figure 6. Schematic representation of the aero-structural iteration procedure.

A diagram representing the aero-structural iteration is shown in Figure 6. The first time the flow solver is called, the displacement field of the structure is initialized to zero. After N iterations of the flow solver, the surface pressures are translated into nodal forces and the structural solver is called. The new displacement field is then translated to a movement of the CFD mesh and N more flow solver iterations are performed. The process continues until the state of the flow and the structure have converged as determined by the norm of the flow solver and structural displacement residuals. In our case, N typically corresponds to 10 iterations.

For the configuration shown in Figure 1, running the aero-structural solver in Euler mode requires 86 multigrid cycles to reduce the average density residual by five orders of magnitude. This represents only a 15% increase when compared with the number of multigrid cycles that are required for a rigid calculation.

Another factor that must be considered when comparing the cost of an aero-structural solution to an aerodynamics-only solution is the computational cost incurred by the structural solver. For the linear finite-element models we use, most of this cost is due

to the factorization of the stiffness matrix. However, since for linear structures the stiffness matrix does not change unless the structure is modified, only one factorization is necessary for each aero-structural solution. During the aero-structural iteration the load vector changes periodically, and the displacement field can be quickly updated in a back-solve operation.

In cases with hundreds of thousands of degrees of freedom, for which it is impractical to factorize the stiffness matrix explicitly, the cost of the structural solution becomes significant, and we would have to resort to efficient solution methods for multiple right-hand sides.

3. Aero-structural design methodology

In this section we describe the optimization problem we propose to solve, in order to put into context the sensitivities whose computation is the focus of this article.

3.1. Design parameterization

The aircraft design is parameterized using two types of design variables. The first type of variable controls the OML of the configuration while the second type of variable dictates the sizing of underlying structure. The OML design variables can be applied to any of the un-intersected components used to define the aircraft geometry. For each wing-like component (main wing, canard, horizontal tail, etc.), the shape is modified at a number of specified airfoil sections. Each of these sections is independently modified and the spanwise resolution is controlled by the number and the position of the sections. The airfoil shape modifications are linearly lofted between stations. Various types of design variables may be applied to the airfoils: twist, leading and trailing edge droop, and Hicks–Henne bump functions, among others. The Hicks–Henne functions are of the form

$$b(\zeta) = x_n \left[\sin\left(\pi \zeta^{\frac{\log(1/2)}{\log(t_1)}}\right)\right]^{t_2}, \tag{5}$$

where t_1 is the location of the maximum of the bump in the range $0 \leq \zeta \leq 1$ at $\zeta = t_1$, since the maximum occurs when $\zeta^{\alpha} = 1/2$, where $\alpha = \log(1/2)/\log t_1$. The parameter t_2 controls the width of the bump. The advantage of these functions is that when they are applied to a smooth airfoil, that airfoil remains smooth.

The structural design variables are the thicknesses of the structural finite elements. The topology of the structure remains unchanged, i.e., the number of spars and ribs and their position are fixed throughout the optimization. Using any of these discrete parameters as design variables results in a discontinuous design space that is not compatible with the gradient-based design approach that we use. Note that since the OML determines the location of the nodes of the structural model, variations of the OML have an effect on the depth of the spars and ribs of the wing box.

3.2. *Objective function*

From the detailed mission analysis of a particular aircraft it is possible to find the correct trade-off between aerodynamic drag and structural weight. This means that we can optimize a design by minimizing the objective function,

$$I = \alpha C_D + \beta W, \tag{6}$$

where C_D is the drag coefficient, W is the structural weight and α and β are scalar parameters.

To perform gradient-based optimization, we need the sensitivities of the objective function (6) with respect to all the design variables. Since this objective function is a linear combination of the drag coefficient and the structural weight, its sensitivity can be written as

$$\frac{\mathrm{d}I}{\mathrm{d}x_n} = \alpha \frac{\mathrm{d}C_D}{\mathrm{d}x_n} + \beta \frac{\mathrm{d}W}{\mathrm{d}x_n}. \tag{7}$$

The computation of the structural weight sensitivity, $\mathrm{d}W/\mathrm{d}x_n$ is easy, since the weight calculation is independent of the aero-structural solution. This gradient is calculated analytically for the structural thickness variables and by finite differences for the OML variables. The drag coefficient sensitivity, $\mathrm{d}C_D/\mathrm{d}x_n$, is not this simple since it does depend on the aero-structural solution.

3.3. *Constraints*

In addition to minimizing the objective function (6), the aero-structural optimization problem is subject to a number of constraints, the most significant of which are the constraints on the structural stress of each element in the CSM model.

In our methodology, and for reasons that we discuss in Section 4, the structural constraints are lumped into a single Kreisselmeier–Steinhauser (KS) function. Suppose that we have the following constraint for each structural finite element,

$$g_m = 1 - \frac{\sigma_m}{\sigma_y} \geq 0, \tag{8}$$

where σ_m is the von Mises stress in the element m, and σ_y is the yield stress of the material. The corresponding KS function for the complete structure is defined as

$$\mathrm{KS} = -\frac{1}{\rho} \ln \left[\sum_m \mathrm{e}^{-\rho g_m} \right]. \tag{9}$$

This function represents a lower bound envelope of all the constraint inequalities and ρ is a positive parameter that expresses how close this bound is to the actual minimum of

the constraints. The use of KS functions constitutes a viable alternative to treating constraints separately. The effectiveness of the method has been demonstrated for structural optimization problems with thousands of constraints (Akgün et al., 1999).

In the section that follows we focus our attention on the main goal of this article: the efficient computation of the drag coefficient gradients, $\mathrm{d}C_D/\mathrm{d}x_n$, and the KS function gradients, $\mathrm{dKS}/\mathrm{d}x_n$, with respect to both aerodynamic and structural design variables.

4. Coupled sensitivity analysis

4.1. General formulation

The main objective is to calculate the sensitivity of a multidisciplinary function of interest with respect to a number of design variables. The function of interest can be either the objective function or any of the constraints specified in the optimization problem. In general, such functions depend not only on the design variables, but also on the physical state of the multidisciplinary system. Thus we can write the function as

$$I = I(x_n, y_i), \tag{10}$$

where x_n represents the vector of design variables and y_i is the state variable vector.

For a given vector x_n, the solution of the governing equations of the multidisciplinary system yields a vector y_i, thus establishing the dependence of the state of the system on the design variables. We denote these governing equations by

$$\mathcal{R}_k\,(x_n, y_i\,(x_n)) = 0. \tag{11}$$

The first instance of x_n in the above equation indicates the fact that the residual of the governing equations may depend *explicitly* on x_n. In the case of a structural solver, for example, changing the size of an element has a direct effect on the stiffness matrix. By solving the governing equations we determine the state, y_i, which depends *implicitly* on the design variables through the solution of the system. These equations may be non-linear, in which case the usual procedure is to drive residuals, $\mathcal{R}_k$, to zero using an iterative method.

Since the number of equations must equal the number of state variables, the ranges of the indices i and k are the same, that is, $i, k = 1, \ldots, N_{\mathcal{R}}$. In the case of a structural solver, for example, $N_{\mathcal{R}}$ is the number of degrees of freedom, while for a CFD solver, $N_{\mathcal{R}}$ is the number of mesh points multiplied by the number of state variables at each point. In the more general case of a multidisciplinary system, $\mathcal{R}_k$ represents *all* the governing equations of the different disciplines, including their coupling.

A graphical representation of the system of governing equations is shown in Figure 7, with the design variables x_n as the inputs and I as the output. The two arrows leading to I illustrate the fact that the objective function typically depends on the state variables and may also be an explicit function of the design variables.

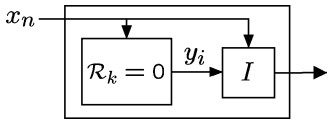

Figure 7. Schematic representation of the governing equations ($\mathcal{R}_k = 0$), design variables (x_n), state variables (y_i), and objective function (I), for an arbitrary system.

As a first step toward obtaining the derivatives that we ultimately want to compute, we use the chain rule to write the total sensitivity of I as

$$\frac{\mathrm{d}I}{\mathrm{d}x_n} = \frac{\partial I}{\partial x_n} + \frac{\partial I}{\partial y_i}\frac{\mathrm{d}y_i}{\mathrm{d}x_n}, \tag{12}$$

for $i = 1, \ldots, N_{\mathcal{R}}$, $n = 1, \ldots, N_x$. Index notation is used to denote the vector dot products. It is important to distinguish the total and partial derivatives in this equation. The partial derivatives can be directly evaluated by varying the denominator and re-evaluating the function in the numerator. The total derivatives, however, require the solution of the multidisciplinary problem. Thus, all the terms in the total sensitivity equation (12) are easily computed except for $\mathrm{d}y_i/\mathrm{d}x_n$.

Since the governing equations must always be satisfied, the total derivative of the residuals (11) with respect to any design variable must also be zero. Expanding the total derivative of the governing equations with respect to the design variables we can write,

$$\frac{\mathrm{d}\mathcal{R}_k}{\mathrm{d}x_n} = \frac{\partial \mathcal{R}_k}{\partial x_n} + \frac{\partial \mathcal{R}_k}{\partial y_i}\frac{\mathrm{d}y_i}{\mathrm{d}x_n} = 0, \tag{13}$$

for all $i, k = 1, \ldots, N_{\mathcal{R}}$ and $n = 1, \ldots, N_x$. This expression provides the means for computing the total sensitivity of the state variables with respect to the design variables. By rewriting equation (13) as

$$\frac{\partial \mathcal{R}_k}{\partial y_i}\frac{\mathrm{d}y_i}{\mathrm{d}x_n} = -\frac{\partial \mathcal{R}_k}{\partial x_n}, \tag{14}$$

we can solve for $\mathrm{d}y_i/\mathrm{d}x_n$ and substitute this result into the total derivative equation (12), to obtain

$$\frac{\mathrm{d}I}{\mathrm{d}x_n} = \frac{\partial I}{\partial x_n} - \underbrace{\frac{\partial I}{\partial y_i}\overbrace{\left[\frac{\partial \mathcal{R}_k}{\partial y_i}\right]^{-1}}^{-\mathrm{d}y_i/\mathrm{d}x_n}}_{-\Psi_k}\frac{\partial \mathcal{R}_k}{\partial x_n}. \tag{15}$$

The inverse of the Jacobian $\partial \mathcal{R}_k/\partial y_i$ is not necessarily explicitly calculated. In the case of large iterative problems neither this matrix nor its factorization are usually stored due to their prohibitive size.

The approach where we first calculate $\mathrm{d}y_i/\mathrm{d}x_n$ using equation (14) and then substitute the result in the expression for the total sensitivity (15) is called the *direct* method. Note that solving for $\mathrm{d}y_i/\mathrm{d}x_n$ requires the solution of the matrix equation (14) *for each design variable* x_n. A change in the design variable affects only the right-hand side of the equation, so for problems where the matrix $\partial\mathcal{R}_k/\partial y_i$ can be explicitly factorized and stored, solving for multiple right-hand-side vectors by back substitution would be relatively inexpensive. However, for large iterative problems—such as the ones encountered in CFD—the matrix $\partial\mathcal{R}_k/\partial y_i$ is never factorized explicitly and the system of equations requires an iterative solution which is usually as costly as solving the governing equations. When we multiply this cost by the number of design variables, the total cost for calculating the sensitivity vector may become unacceptable.

Returning to the total sensitivity equation (15), we observe that there is an alternative option for computing the total sensitivity $\mathrm{d}I/\mathrm{d}x_n$. The auxiliary vector Ψ_k can be obtained by solving the *adjoint equations*

$$\frac{\partial\mathcal{R}_k}{\partial y_i}\Psi_k = -\frac{\partial I}{\partial y_i}. \tag{16}$$

The vector Ψ_k is usually called the *adjoint vector* and is substituted into equation (15) to find the total sensitivity. In contrast with the direct method, the adjoint vector does not depend on the design variables, x_n, but instead depends on the function of interest, I.

We can now see that the choice of the solution procedure (direct vs. adjoint) to obtain the total sensitivity (15) has a substantial impact on the cost of sensitivity analysis. Although all the partial derivative terms are the same for both the direct and adjoint methods, the order of the operations is not. Notice that once $\mathrm{d}y_i/\mathrm{d}x_n$ is computed, it is valid for any function I, but must be recomputed for each design variable (direct method). On the other hand, Ψ_k is valid for all design variables, but must be recomputed for each function (adjoint method).

The cost involved in calculating sensitivities using the adjoint method is therefore practically independent of the number of design variables. After having solved the governing equations, the adjoint equations are solved only once for each I. Moreover, the cost of solution of the adjoint equations is similar to that of the solution of the governing equations since they are of similar complexity and the partial derivative terms are easily computed.

Therefore, if the number of design variables is greater than the number of functions for which we seek sensitivity information, the adjoint method is computationally more efficient. Otherwise, if the number of functions to be differentiated is greater than the number of design variables, the direct method would be a better choice.

The adjoint method has been widely used for single discipline sensitivity analysis and examples of its application include structural sensitivity analysis (Adelman and Haftka, 1986) and aerodynamic shape optimization (Jameson, 1989; Jameson et al., 1998).

4.2. Aero-structural sensitivity equations

Although the theory we have just presented is applicable to multidisciplinary systems (provided that the governing equations for all disciplines are included in $\mathcal{R}_k$) we now explicitly

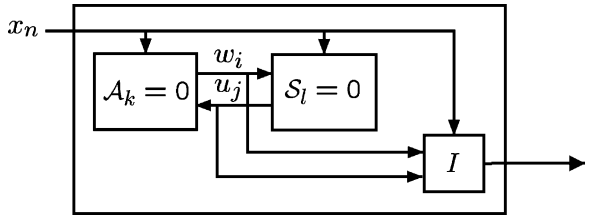

Figure 8. Schematic representation of the aero-structural system.

discuss the sensitivity analysis of multidisciplinary systems, using aero-structural optimization as an example. This example illustrates the fundamental computational cost issues that motivate our choice of strategy for sensitivity analysis. The following equations and discussion can easily be generalized for cases with additional disciplines.

In the aero-structural case we have coupled aerodynamic ($\mathcal{A}_k$) and structural ($\mathcal{S}_l$) governing equations, and two sets of state variables: the flow state vector, w_i, and the vector of structural displacements, u_j. In the following expressions, we split the vectors of residuals, states and adjoints into two smaller vectors corresponding to the aerodynamic and structural systems

$$\mathcal{R}_{k'} = \begin{bmatrix} \mathcal{A}_k \\ \mathcal{S}_l \end{bmatrix}, \quad y_{i'} = \begin{bmatrix} w_i \\ u_j \end{bmatrix}, \quad \Psi_{k'} = \begin{bmatrix} \psi_k \\ \phi_l \end{bmatrix}. \tag{17}$$

Figure 8 shows a diagram representing the coupling in this system.

Using this new notation, the direct sensitivity equation (14) for an aero-structural system can be written as

$$\begin{bmatrix} \frac{\partial \mathcal{A}_k}{\partial w_i} & \frac{\partial \mathcal{A}_k}{\partial u_j} \\ \frac{\partial \mathcal{S}_l}{\partial w_i} & \frac{\partial \mathcal{S}_l}{\partial u_j} \end{bmatrix} \begin{bmatrix} \frac{\mathrm{d}w_i}{\mathrm{d}x_n} \\ \frac{\mathrm{d}u_j}{\mathrm{d}x_n} \end{bmatrix} = - \begin{bmatrix} \frac{\partial \mathcal{A}_k}{\partial x_n} \\ \frac{\partial \mathcal{S}_l}{\partial x_n} \end{bmatrix}. \tag{18}$$

This equation was first written for a multidisciplinary system by Sobieski (1990). In his paper, Sobieski also presents an alternative approach to the problem, which he shows is equivalent to (18), i.e.,

$$\begin{bmatrix} \mathcal{I} & -\frac{\partial w_i}{\partial u_j} \\ -\frac{\partial u_j}{\partial w_i} & \mathcal{I} \end{bmatrix} \begin{bmatrix} \frac{\mathrm{d}w_i}{\mathrm{d}x_n} \\ \frac{\mathrm{d}u_j}{\mathrm{d}x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial w_i}{\partial x_n} \\ \frac{\partial u_j}{\partial x_n} \end{bmatrix}, \tag{19}$$

where $\mathcal{I}$ denotes the identity matrix. Solving either of these equations (18) and (19) yields the total sensitivity of the state variables with respect to the design variables. This result can then be substituted into the aero-structural equivalent of the total sensitivity equation (12),

$$\frac{\mathrm{d}I}{\mathrm{d}x_n} = \frac{\partial I}{\partial x_n} + \frac{\partial I}{\partial u_j}\frac{\mathrm{d}u_j}{\mathrm{d}x_n} + \frac{\partial I}{\partial w_i}\frac{\mathrm{d}w_i}{\mathrm{d}x_n}. \tag{20}$$

The biggest disadvantage of these direct approaches, as we discussed earlier, is that the sensitivity equation must be solved for each design variable x_n. For large iterative coupled systems, the cost of computing the total sensitivities with respect to many design variables becomes prohibitive, and this approach is impractical.

In the alternate direct approach (19), the partial derivatives of the state variables of a given system with respect to the variables of the other system ($\partial w_i/\partial u_j$, $\partial u_j/\partial w_i$) and the partial derivatives of the state variables with respect to the design variables ($\partial w_i/\partial x_n$, $\partial u_j/\partial x_n$) have a different meaning from the partial derivatives we have seen so far. In this formulation, the partial derivatives of the state variables of a given system take into account the solution of that system. Although the solution of the coupled system is not required, this differs significantly from the partial derivatives of the residuals in the formulation (18), which do not require the solution of even the single discipline.

The adjoint approach to sensitivity analysis is also applicable to multidisciplinary systems. In the case of the aero-structural system, the adjoint equation (16) can be written as

$$\begin{bmatrix} \frac{\partial \mathcal{A}_k}{\partial w_i} & \frac{\partial \mathcal{A}_k}{\partial u_j} \\ \frac{\partial \mathcal{S}_l}{\partial w_i} & \frac{\partial \mathcal{S}_l}{\partial u_j} \end{bmatrix}^T \begin{bmatrix} \psi_k \\ \phi_l \end{bmatrix} = - \begin{bmatrix} \frac{\partial I}{\partial w_i} \\ \frac{\partial I}{\partial u_j} \end{bmatrix}. \tag{21}$$

Note that the matrix in the coupled adjoint equation is the same as in the coupled direct method (18). In addition to containing the diagonal terms that appear when we solve the single discipline adjoint equations, this matrix includes off-diagonal terms that express the sensitivity of one discipline to the state variables of the other. The details of the partial derivative terms of this matrix are described in Section 4.3.

Finally, for completeness, we note that there is an alternative formulation for the coupled-adjoint method which is parallel to the alternate direct equations (19),

$$\begin{bmatrix} \mathcal{I} & -\frac{\partial w_i}{\partial u_j} \\ -\frac{\partial u_j}{\partial w_i} & \mathcal{I} \end{bmatrix}^T \begin{bmatrix} \bar{\psi}_i \\ \bar{\phi}_j \end{bmatrix} = \begin{bmatrix} \frac{\partial I}{\partial w_i} \\ \frac{\partial I}{\partial u_j} \end{bmatrix}, \tag{22}$$

where the partial derivatives have the same meaning as in the alternate direct sensitivity equations (19) and therefore, as previously discussed, the uncoupled solution of each discipline is required. This is a disadvantage relative to the standard coupled-adjoint approach (21), where none of the partial derivatives require the solution of governing equations. The alternate direct and adjoint formulations are, however, not without their advantages. They are the only suitable approach when disciplinary solvers are available just as black boxes, that is, if one has access only to the input and output. Furthermore, the size of the matrix of state variable sensitivities can be reduced since one only needs to consider the coupling variables.

The alternate adjoint vector, $\bar{\psi}_k$ differs from the standard adjoint and therefore requires a different total sensitivity equation,

$$\frac{\mathrm{d}I}{\mathrm{d}x_n} = \frac{\partial I}{\partial x_n} + \bar{\psi}_i \frac{\partial w_i}{\partial x_n} + \bar{\phi}_j \frac{\partial u_j}{\partial x_n}, \tag{23}$$

where the partial derivatives of the state variables ($\partial w_i/\partial x_n$, $\partial u_j/\partial x_n$) also require the solution of the corresponding governing equations.

Given that the aerodynamic analysis in our case (CFD) is rather costly, and that the number of coupling variables (all surface pressures and all nodal forces) is $\mathcal{O}(10^4)$, the standard coupled-adjoint approach (21) is adopted.

Solving the coupled-adjoint equations (21) by factorizing the full matrix would be extremely costly, so our approach is to decouple the two disciplines, much like we did for the aero-structural solution. To solve the coupled-adjoint equations iteratively, the adjoint vectors are *lagged* and the two different sets of equations are solved separately. For the calculation of the adjoint vector of one discipline, we use the adjoint vector of the other discipline from the previous iteration, i.e., we solve

$$\underbrace{\frac{\partial \mathcal{A}_k}{\partial w_i}\psi_k = -\frac{\partial I}{\partial w_i}}_{\text{Aerodynamic adjoint}} - \frac{\partial \mathcal{S}_l}{\partial w_i}\tilde{\phi}_l, \tag{24}$$

$$\underbrace{\frac{\partial \mathcal{S}_l}{\partial u_j}\phi_l = -\frac{\partial I}{\partial u_j}}_{\text{Structural adjoint}} - \frac{\partial \mathcal{A}_k}{\partial u_j}\tilde{\psi}_k, \tag{25}$$

where $\tilde{\psi}_k$ and $\tilde{\phi}_l$ are the lagged aerodynamic and structural adjoint vectors respectively. Upon convergence, the final result given by this system is the same as that given by the original coupled-adjoint equations (21). We call this the *lagged-coupled adjoint* (LCA) method for computing sensitivities of coupled systems. Note that these equations look like the single discipline adjoint equations for the aerodynamic and structural solvers, with the addition of forcing terms in the right-hand side that contain the off-diagonal terms of the residual sensitivity matrix. This allows us to use existing single-discipline adjoint sensitivity analysis methods. Note also that, even for more than two disciplines, this iterative solution procedure is nothing more than the well-known block-Jacobi method.

Once both adjoint vectors have converged, we can compute the final sensitivities of the objective function by using the following expression

$$\frac{\mathrm{d}I}{\mathrm{d}x_n} = \frac{\partial I}{\partial x_n} + \psi_k \frac{\partial \mathcal{A}_k}{\partial x_n} + \phi_l \frac{\partial \mathcal{S}_l}{\partial x_n}, \tag{26}$$

which is the coupled version of the total sensitivity equation (15).

The approach for solving the coupled system of sensitivity equations by lagging can also be used to solve the direct equations (18) and (19) but the disadvantages of these methods for problems that require iterative methods and are parameterized with a large number of design variables remain the same.

For the aero-structural optimization problem at hand the aerodynamic portion is usually characterized by a single objective function, a few aerodynamic constraints, and a large number of design variables. On the other hand, the structural portion of the optimization problem involves a large number of constraints dictating that the stress in each element of

the finite-element model is not to exceed the material yield stress for a set of load conditions. Constrained gradient optimization methods generally require that the user provide the gradient of both the cost function and each nonlinear constraint with respect to all of the design variables in the problem. Using the adjoint approach, the evaluation of the gradient of each constraint would require an independent coupled solution of a large adjoint system. Since the number of structural constraints is similar to the number of design variables in the problem ($\mathcal{O}(10^3)$ or larger), the usefulness of the adjoint approach is questionable in this case.

The remaining alternatives—the direct and finite-difference methods—are not advantageous either since they both require a number of solutions that is comparable to the number of design variables. In the absence of other choices that can efficiently evaluate the gradient of a large number of constraints with respect to a large number of design variables, it is necessary to reduce the size of the problem either through a reduction in the number of design variables or through a reduction in the number of nonlinear constraints.

The reason for using KS functions to lump the structural constraints now becomes clear. By employing KS functions, the number of structural constraints for the problem can be reduced from $\mathcal{O}(10^3)$ to just a few. In some problems, a single KS function may suffice. If this constraint lumping methodology is effective, an adjoint method would be very efficient for computing MDO sensitivities.

4.3. Partial derivative term details

In this section, we describe the calculation of the partial derivative terms in the aero-structural adjoint equations (24, 25) and the total sensitivity equations (26). This description is divided into four sections. The first two sections discuss the terms involving the partial derivatives of the aerodynamic and structural equations respectively. The last two sections cover the partial derivatives of the drag coefficient (C_D) and the KS function. These four terms are differentiated with respect to the state vectors (w_i, u_j) and the vector of design variables (x_n).

4.3.1. Partial derivatives of the aerodynamic governing equations. A number of publications describe in detail the terms involved in the aerodynamic adjoint equation and its associated boundary conditions (Jameson, 1989; Jameson et al., 1998; Reuther, 1996; Reuther et al., 1999a). Although we do not describe these terms in comparable detail, we do explain the meaning of all the terms, specially those that arise from the inclusion of structural deformations.

The aerodynamic adjoint equation can be written as

$$\frac{\partial \mathcal{A}_k}{\partial w_i}\psi_k = -\frac{\partial I}{\partial w_i}. \tag{27}$$

The components of the right-hand side vector are usually non-zero only for those points on the CFD surface mesh. The aerodynamic adjoint used in our work is based on a *continuous* formulation that is derived from the partial differential equations that govern the flow. The

adjoint partial differential equations are then discretized using the same scheme and mesh as the flow equations. This is in contrast with the *discrete* approach, where the governing equations of the flow are first discretized and an adjoint version of these discrete equations is then constructed by taking the transpose of $\partial \mathcal{A}_k / \partial w_i$.

The Jacobian $\partial \mathcal{A}_k / \partial w_i$ in the adjoint equation (27) represents the variation of the residuals for each cell of the CFD mesh due to changes in the flow solution for every cell in the mesh. When a flow variable at a given cell center is perturbed the residuals of that cell and other cells in its vicinity are modified. The extent of the influence of these flow variable perturbations depends on the stencil used in the flow solver: in our case, a single-level halo of cells is affected. Therefore, even though $\partial \mathcal{A}_k / \partial w_i$ is a very large square matrix, it is also extremely sparse and its non-zero terms can be easily calculated using finite differences. In our coupled-adjoint solver this matrix is never stored explicitly and the adjoint equation (24) is solved iteratively, much like the flow equations.

The off-diagonal term $\partial \mathcal{A}_k / \partial u_j$ in the LCA equation (25) represents the effect that the structural displacements have on the residuals of the CFD equations through the perturbation of the CFD mesh. When a given structural node moves, both the surface and the interior of the CFD grid must be perturbed, thus affecting a large number of CFD mesh points. Even though the flow variables are constant in the calculation of this partial derivative, the change in the mesh geometry affects the sum of the fluxes, whose variation is easily obtained by recalculating the residuals for the perturbed cells. Because the actual term we want to compute in equation (25) is the product of this matrix, $\partial \mathcal{A}_k / \partial u_j$, with the lagged aerodynamic adjoint vector, $\tilde{\psi}_k$, it is possible to multiply each column j by the adjoint vector as it is calculated. This approach eliminates the need to store the complete matrix, since we only need to store a vector with the same dimension as that of the adjoint vector.

The term $\partial \mathcal{A}_k / \partial x_n$ in the total sensitivity equation (26) represents the direct effect of the design variables on the CFD residuals of all cells in the mesh. For finite-element thickness design variables, this term is identically zero, since these design variables do not affect the CFD residuals explicitly. For shape design variables, this Jacobian is similar to $\partial \mathcal{A}_k / \partial u_j$, since a change in an OML design variable also perturbs the CFD grid. In the present work, this term is calculated by finite differencing, since the mesh perturbation algorithm is very efficient. Again, the term we ultimately want is the vector that results from the product $\psi_k \, \partial \mathcal{A}_k / \partial x_n$, and the matrix multiplication can be performed as each row of the matrix is calculated.

The number of CFD mesh perturbations required for the calculation of the aerodynamic equation sensitivities is equal to $N_x + N_S M$, where N_x is the number of design variables, N_S is the number of surface degrees of freedom of the structural model and M is the number of times the adjoint vectors are exchanged in the iterative solution of the LCA equations (24) and (25). The cost of computing $\psi_k \, \partial \mathcal{A}_k / \partial x_n$ is proportional to N_x, while the computation of $\partial \mathcal{A}_k / \partial u_j \, \tilde{\psi}_k$ is proportional to N_S and must be performed M times. Since $N_x + N_S M$ can be very large, it is extremely important that the mesh perturbation procedure be efficient. This is achieved in the present work because the mesh perturbation algorithm is completely algebraic. The fact that the CFD mesh is structured makes it possible to attenuate perturbations applied to the surface throughout the volume mesh, as described

in Section 2.2. In Section 5.2 we compare the cost of the complete sensitivity analysis to the cost of the mesh perturbations.

4.3.2. Partial derivatives of the structural governing equations. When using linear finite-element models for structural analysis, the discretized governing equations are given by

$$\mathcal{S}_l = K_{lj} u_j - f_l = 0, \tag{28}$$

where, K_{lj} is the global stiffness matrix of the structure, u_j is the vector of nodal displacements, and f_l is the vector of applied nodal forces. In our case, the structural model has a relatively small number of degrees of freedom—$\mathcal{O}(10^3)$—and a Cholesky factorization is appropriate to solve for the unknown displacements. The factorization is explicitly stored and is used to solve the structural equations multiple times with different load vectors. For large finite-element models, where the number of degrees of freedom exceeds $\mathcal{O}(10^5)$, alternative approaches to solving the structural equations (28) are used.

To calculate structural sensitivities using the adjoint method we need the partial derivative of the structural governing equations (28) with respect to the displacements, which is nothing more than the global stiffness matrix, i.e.,

$$\frac{\partial \mathcal{S}_l}{\partial u_j} = K_{lj}. \tag{29}$$

Hence, the adjoint equations for the structural system are

$$K_{jl}\phi_l = -\frac{\partial I}{\partial u_j}. \tag{30}$$

Since the stiffness matrix is symmetric ($K_{lj} = K_{jl}$), the structural adjoint equations (30) have the same matrix as the structural governing equations (28), i.e., the system is *self adjoint*. The only difference between these two sets of equations is the vector in right-hand side: instead of the load vector in the governing equations (28), the adjoint equations (30) have a vector (often referred to as *pseudo load*) related to the function of interest, I. As previously mentioned, the stiffness matrix is factorized once when solving for the displacements and therefore, it is possible to use this factorization to solve for ϕ_l with only a small additional cost.

The derivative of the structural governing equations with respect to the flow variables, $\partial \mathcal{S}_l / \partial w_i$, is the other off-diagonal term in the aero-structural adjoint equations (21). In the LCA equation (25) we need this term to compute the lagged term $\partial \mathcal{S}_l / \partial w_i \; \tilde{\phi}_l$. The only term in the governing equations (28) that the flow variables affect directly is the applied force, and thus

$$\frac{\partial \mathcal{S}_l}{\partial w_i} = -\frac{\partial f_l}{\partial w_i} = -\frac{\partial f_l}{\partial p_{i'}} \frac{\partial p_{i'}}{\partial w_i}, \tag{31}$$

where we note that the flow variables affect the structural forces via the surface pressures, $p_{i'}$. Although the matrix $\partial p_{i'} / \partial w_i$ is rather large, it is very sparse since the surface pressures

depend only on a small subset of the flow variables. The matrix $\partial f_l/\partial p_{i'}$ is calculated analytically by examining the procedure that integrates the pressures in the CFD mesh and transfers them to the structural nodes to obtain the applied forces. The resulting matrix, $\partial \mathcal{S}_l/\partial w_i$, is rather large, of $\mathcal{O}(10^3 \times 10^6)$, but is never stored explicitly. Since the term we want is actually the vector $\partial \mathcal{S}_l/\partial w_i\, \tilde{\phi}_l$, we calculate one row at a time and perform the dot product with the structural adjoint vector.

Finally, we also need the partial derivative with respect to the design variables, $\partial \mathcal{S}_l/\partial x_n$. The shape design variables have a direct effect on both the stiffness matrix and the load vector. Although this partial derivative assumes a constant surface pressure field, a variation in the OML affects the transfer of these pressures to structural loads. Hence,

$$\frac{\partial \mathcal{S}_l}{\partial x_n} = \frac{\partial K_{lj}}{\partial x_n} u_j - \frac{\partial f_l}{\partial x_n}. \tag{32}$$

The element thickness design variables also affect the stiffness matrix, but not the force, and therefore $\partial f/\partial x_n = 0$ in this case. As in the case of $\partial \mathcal{S}_l/\partial w_i$, this matrix is computed by finite differences and multiplied by the structural adjoint vector, one row at a time, eliminating unnecessary storage overhead.

4.3.3. Partial derivatives of the drag coefficient. When solving the adjoint equations for $I = C_D$, we need the partial derivative $\partial C_D/\partial w_i$ to calculate the right-hand side of the first aero-structural adjoint equation (24). The value of C_D only depends on the flow variables corresponding to those cells that lie on the surface of the aircraft, so this vector is very sparse. The non-zero sensitivities in this vector are obtained analytically by differentiating the numerical integration procedure of the surface pressures that produces C_D.

The second aero-structural adjoint equation (25) contains another partial derivative of C_D, but this one is taken with respect to the structural displacements. The vector $\partial C_D/\partial u_j$ represents the change in the drag coefficient due to the displacement of the wing while keeping the pressure field constant. The structural displacements affect the drag directly, since they change the wing surface over which the pressure is integrated. This vector of sensitivities is efficiently computed by finite differencing.

Finally, in order to calculate the total sensitivity of the drag coefficient using equation (26), we need the term $\partial C_D/\partial x_n$. This represents the change in the drag coefficient due to design variable perturbations, while keeping the pressure and displacement fields constant. In the case of shape perturbations, $\partial C_D/\partial x_n$ is analogous to $\partial C_D/\partial u_j$ because these design variables change the surface of integration. This vector is also inexpensively calculated using finite differences. For structural design variables this term is zero because they do not affect the OML directly.

4.3.4. Partial derivatives of the KS function. As discussed in Section 3.3, the other set of sensitivities we are ultimately interested in is that of the KS function (9), i.e., when $I = \mathrm{KS}$. Since this function depends directly on the stresses we use the chain rule to write,

$$\frac{\partial \mathrm{KS}}{\partial u_j} = \frac{\partial \mathrm{KS}}{\partial g_m} \frac{\partial g_m}{\partial \sigma_m} \frac{\partial \sigma_m}{\partial u_j}. \tag{33}$$

Differentiating the KS function (9) we can write the first term as

$$\frac{\partial \mathrm{KS}}{\partial g_m} = \left[\sum_{m'} \mathrm{e}^{-\rho g_{m'}} \right]^{-1} \mathrm{e}^{-\rho g_m}. \tag{34}$$

The second term is easily derived from the definition of the stress constraints (8),

$$\frac{\partial g_m}{\partial \sigma_m} = -\frac{1}{\sigma_y}. \tag{35}$$

To obtain the third term of equation (33) we consider the expression that relates the stresses to the displacement field,

$$\sigma_m = S_{mj} u_j. \tag{36}$$

Given the linear nature of this relationship, the partial derivative we need is simply

$$\frac{\partial \sigma_m}{\partial u_j} = S_{mj}. \tag{37}$$

Using these results we can rewrite the partial derivative (33) as

$$\frac{\partial \mathrm{KS}}{\partial u_j} = -\left[\sigma_y \sum_{m'} \mathrm{e}^{-\rho g_{m'}} \right]^{-1} \mathrm{e}^{-\rho g_m} S_{mj}. \tag{38}$$

We use this term in the right-hand side of the structural adjoint equation (30)—or equation (25) in the aero-structural case—to solve for the adjoint vector that corresponds to the sensitivities of the KS function.

For the case where $I = \mathrm{KS}$, the right-hand-side of the aerodynamic adjoint equation (24) includes $\partial \mathrm{KS}/\partial w_i$. This term is zero, since the stresses do not depend explicitly on the loads. They only depend on the loads implicitly, through the displacements.

Finally, the last partial derivative of the KS function, $\partial \mathrm{KS}/\partial x_n$, appears in the total sensitivity equation (26). This term represents the variation of the lumped stresses for fixed loads and displacements. As in the case of the partial derivative with respect to the displacements (33), we can use the chain rule to write

$$\frac{\partial \mathrm{KS}}{\partial x_n} = \frac{\partial \mathrm{KS}}{\partial g_m} \frac{\partial g_m}{\partial \sigma_m} \frac{\partial \sigma_m}{\partial x_n}. \tag{39}$$

Since we have derived the two first partial derivative terms (34, 35) we are left with only one new term, the partial derivative of the stresses with respect to the design variables. Taking the derivative of the stress-displacement relationship (36) yields

$$\frac{\partial \sigma_m}{\partial x_n} = \frac{\partial S_{mj}}{\partial x_n} u_j, \tag{40}$$

where $\partial S_{mj}/\partial x_n$ is calculated using finite differences. For element thickness design variables, $\partial \sigma_m/\partial x_n = 0$, since S_{mj} does not depend on these variables for the type of finite elements we use. However, when the OML is perturbed, S_{mj}—and hence the stresses—can vary in a given element if its shape is distorted.

5. Results

In this section we compare the aero-structural sensitivities of a supersonic business jet configuration given by the LCA with both finite differences and the complex-step method. This business jet configuration is being developed by the ASSET Research Corporation and it is designed to achieve a large percentage of laminar flow on a low-sweep wing, resulting in decreased friction drag (Kroo et al., 2002). The aircraft is to fly at Mach 1.5 and have a range of 5, 300 nautical miles.

To compute the flow for this configuration, we use the CFD mesh shown in Figure 1. This is a multiblock Euler mesh with 36 blocks and a total of 220,000 mesh points.

The structural model of the wing is also shown in Figure 1 and consists of a wing box with six spars evenly distributed from 15% to 80% of the chord. Ribs are distributed along the span at every tenth of the semispan. A total of 640 finite elements are used in the construction of this model.

5.1. Aero-structural sensitivity validation

To gain confidence in the effectiveness of the aero-structural coupled-adjoint sensitivities for use in design optimization, we must ensure that the values of the gradients are accurate. For this purpose, we chose to validate the four sets of sensitivities discussed below. For comparison purposes, we compute the exact discrete value of these sensitivities using the complex-step derivative approximation (Martins et al., 2003).

In this sensitivity study two different functions are considered: the aircraft drag coefficient, C_D, and the KS function (9). The sensitivities of these two quantities with respect to both OML shape design variables and structural design variables are validated. The design variables are as described in Section 3.1.

5.1.1. Drag coefficient sensitivities. The aero-structural sensitivities of the drag coefficient with respect to shape perturbations are shown in Figure 9. The ten shape perturbations were chosen to be Hicks–Henne bumps distributed chordwise on the upper surface of two adjacent airfoils around the quarter span. Note that the lines connecting the points in this graphs do not have a physical meaning and are drawn solely for the sake of readability.

The plot shows very good agreement between the coupled-adjoint and the complex-step results, with an average relative error between the two of only 3.5%. This error is partially due to the fact that we use a discretization of the continuous adjoint equations that is only consistent with the values given by the complex-step method (or finite differences) in the limit of very fine meshes. This fact has been demonstrated in comparisons between different approaches to solving the adjoint equations (Nadarajah and Jameson, 2000). The other source of error is the fact that the partial derivative terms described in Section 4.3 are

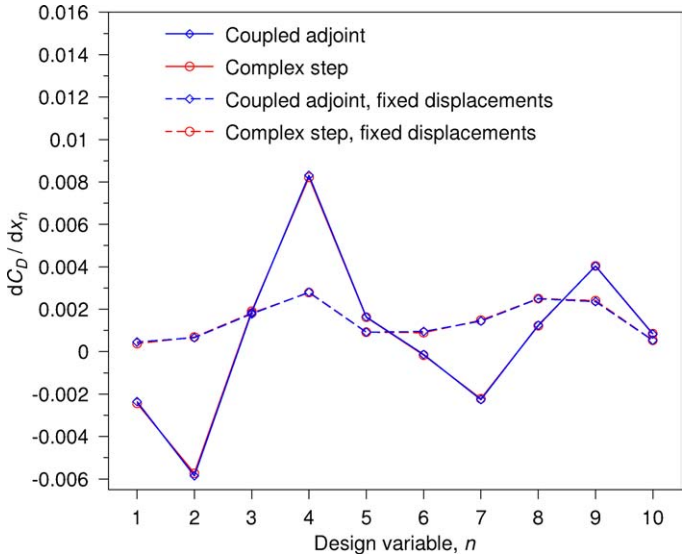

Figure 9. Sensitivities of the drag coefficient with respect to shape perturbations.

estimated by finite differences and are therefore subject to both Taylor series truncation and subtractive cancellation errors.

Note that all these sensitivities are *total* sensitivities in the sense that they account for the coupling between aerodynamics and structures. To verify the need for taking the coupling into account, the same set of sensitivities was calculated for fixed structural displacements, where the displacement field is frozen after the aero-structural solution. This is equivalent to assuming that the wing, after the initial aeroelastic deformation, is infinitely rigid as far as the computation of sensitivities is concerned. When calculating these sensitivities using the complex step, the reference solution is aero-structural, but only the flow solver is called for each shape perturbation. When using the adjoint method, this is equivalent to solving only the aerodynamic adjoint in (24) and omitting the partial derivatives of $\mathcal{S}_l$ in the gradient calculation (26). Figure 9 shows that the single-system sensitivities exhibit significantly lower magnitudes and even opposite signs for many of the design variables when compared with the coupled sensitivities. The use of single-discipline sensitivities would therefore lead to erroneous design decisions.

Figure 10 also shows the sensitivity of the drag coefficient, this time with respect to the thicknesses of five skin groups and five spar groups distributed along the span. The agreement in this case is even better: the average relative error is only 1.6%. Even though these are sensitivities with respect to internal structural variables that do not modify the jig OML, coupled sensitivity analysis is still required.

5.1.2. KS function sensitivities. The sensitivities of the KS function with respect to the two sets of design variables described above are shown in Figures 11 and 12. The results show that the coupled-adjoint sensitivities are extremely accurate, with average relative errors of

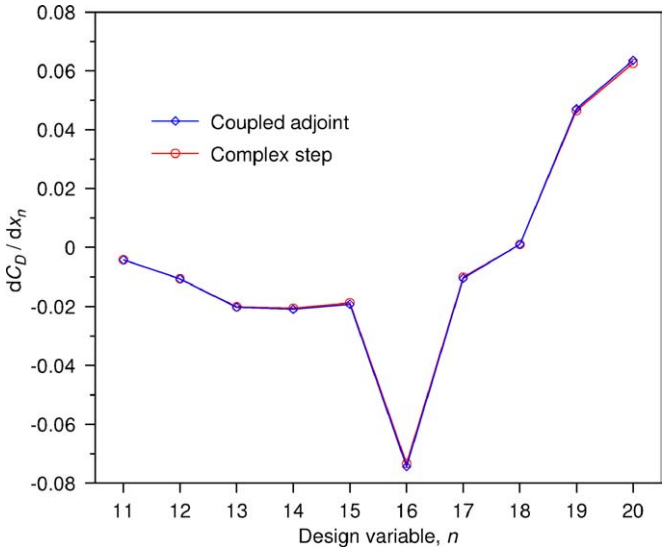

Figure 10. Sensitivities of the drag coefficient with respect to structural thicknesses.

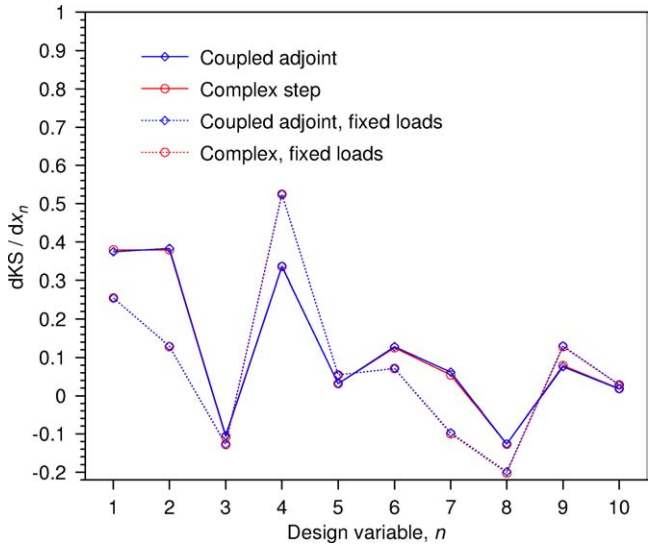

Figure 11. Sensitivities of the KS function with respect to shape perturbations.

2.9% and 1.6%, respectively. In Figure 12 we observe that the sensitivity of the KS function with respect to the first structural thickness is much higher than the remaining sensitivities. This markedly different magnitude is due to the fact that this particular structural design variable corresponds to the thickness of the top and bottom skins of the wing bay closest to the root, where the stress is the highest at this particular load condition.

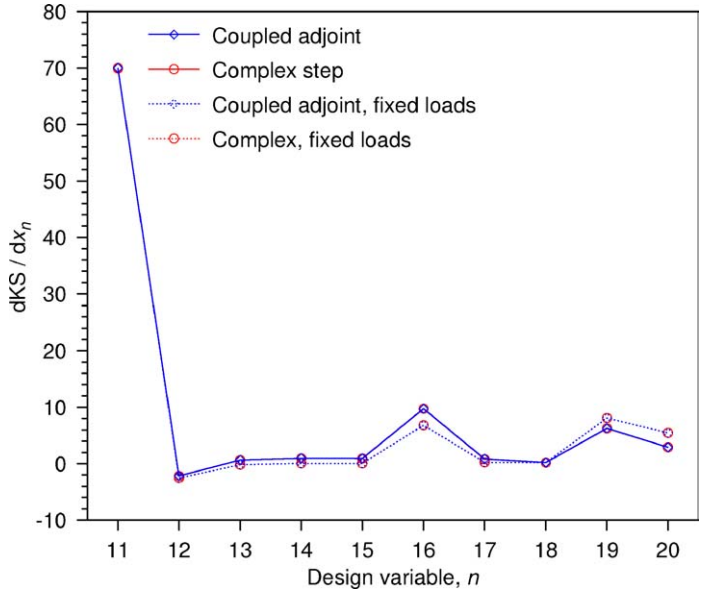

Figure 12. Sensitivities of the KS function with respect to structural thicknesses.

The sensitivities of the KS function for fixed loads are also shown in Figures 11 and 12. Using the complex-step method, these sensitivities were calculated by calling only the structural solver after the initial aero-structural solution has converged, which is equivalent to using just equations (25, 26) without the partial derivatives of $\mathcal{A}_k$ after solving the aero-structural system. The difference in these sensitivities when compared to the coupled ones is not as dramatic as in the fixed displacements case shown in Figure 9, but it is still significant.

5.2. *Computational efficiency*

5.2.1. Comparison of the coupled adjoint with finite differencing. The cost of calculating a gradient vector using either the finite-difference or the complex-step methods is expected to be linearly dependent on the number of design variables. This expectation is confirmed in Figure 13 where the gradient calculation times are shown for increasing numbers of design variables. The time axis is normalized with respect to the time required for a single aero-structural solution (98 seconds on 9 processors of an SGI Origin 2000).

The cost of a finite-difference gradient evaluation can be linearly approximated by the equation $1.0+0.38\times N_x$, where N_x is the number of design variables. Notice that one might expect this method to incur a computational cost equivalent to one aero-structural solution per additional design variable. The cost per additional design variable is lower than this because each additional aero-structural calculation does not start from a uniform flow-field initial condition, but from the previously converged solution, which is closer to the final solution.

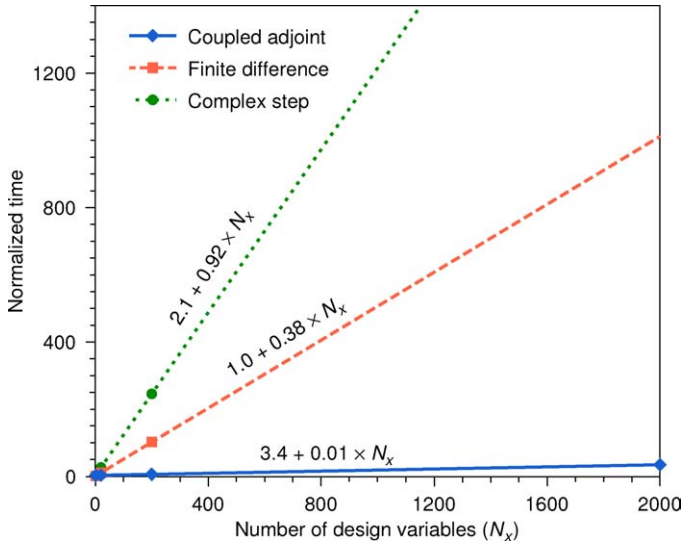

Figure 13. Computational time vs. number of design variables for finite differencing, complex step and coupled adjoint. The time is normalized with respect to the time required for one aero-structural solution.

The same applies to the cost of the complex-step method. Because the function evaluations require complex arithmetic, the cost of the complex step method is, on average, 2.4 times higher than that of finite differencing. However, this cost penalty is worthwhile since there is no need to find an acceptable step size *a priori*, as is the case for finite-difference approximations (Martins et al., 2003).

The cost of computing sensitivities using the coupled-adjoint procedure is in theory independent of the number of variables. Using our implementation, however, some of the partial derivatives in the total sensitivity equation (26) are calculated using finite differences and therefore, there is a small dependence on the number of design variables. The line representing the cost of the coupled adjoint in Figure 13 has a slope of 0.01 which is between one and two orders of magnitude less than the slope for the other two lines.

In short, the cost of computing sensitivities with respect to hundreds or even thousands of variables is acceptable when using the coupled-adjoint approach, while it is impractical to use finite-differences or the complex-step method for such a large number of design variables, even with current state-of-the-art parallel computing systems.

5.2.2. Coupled-adjoint solution. The constant terms in the equations for the straight lines of Figure 13 represent the cost of each procedure when no sensitivities are required. For the finite-difference case, this is equivalent to one aero-structural solution, and hence the constant is 1.0. When performing the aero-structural solution using complex arithmetic, the cost rises to 2.1 times the real arithmetic solution.

The cost of computing the coupled-adjoint vectors (without computing the gradients) is 3.4. This cost includes the aero-structural solution—which is required before solving the adjoint equations—and hence the aero-structural adjoint computation alone incurs a cost of

Table 1. Computational times for solving the LCA equations (24) and (25).

Aero-structural solution	1.000
Aero-structural adjoint	
Aerodynamic adjoint equation (24)	0.597
RHS of equation (24)	0.642
Structural adjoint equation (25)	<0.001
RHS of equation (25)	1.203

2.4. To gain a better understanding of how this cost it divided, we timed the computation for four different components of the aero-structural adjoint equations (24) and (25) as shown in Table 1.

The cost of solving the lagged aerodynamic adjoint equation (24)—not including the computation of the right-hand-side vector—is about 0.6 times the aero-structural solution. This is expected, since the cost of solving the adjoint equations of a given system is usually similar to the cost of solving the corresponding governing equations. The cost of computing right-hand-side of the same equation, is also about 0.6. This cost is almost exclusively due to the lagged term $\partial \mathcal{S}_l / \partial w_i \; \tilde{\phi}_l$, which is partially computed using finite differences, as explained in Section 4.3.2. Note that the cost of computing this term is proportional to the number of OML points, which is 4,200 in our calculations.

The computation time for solving the lagged structural adjoint equation (25) is negligible. Again, this does not account for the computation of the right-hand side of that equation. The cost of solving this equation is so low because the factorization of the matrix has already been computed and only one back-solve operation is required. The computation time for the right-hand side of this equation, however, is rather high: 1.2. Again, this is almost solely due to the lagged term, which is $\partial \mathcal{A}_k / \partial u_j \; \tilde{\psi}_k$ in this case. As explained in Section 4.3.1 this term is computed using finite differences and therefore its cost is proportional to the number of structural surface degrees of freedom, which in this case is equal to 396.

6. Conclusions

An adjoint method for coupled sensitivity analysis of high-fidelity aero-structural systems was presented. The aero-structural adjoint sensitivity equations were compared with two different direct formulations to show that, as in the case of single disciplines, the adjoint formulation is preferred when the number of design variables is significantly larger than the number of functions of interest. An alternate adjoint formulation, which had not been previously published, was also presented.

The sensitivities computed by the lagged-coupled adjoint method were compared to sensitivities given by the complex-step derivative approximation and shown to be extremely accurate, having an average relative error of 2%. The coupled aero-structural sensitivities were also compared to single-discipline sensitivities. Large deviations from the coupled results show that the true fully-coupled sensitivities are essential for aero-structural design optimization.

In realistic aero-structural design problems with hundreds or even thousands of design variables, there is a considerable reduction in computational cost when using the coupled-adjoint method as opposed to either finite differences or the complex step. This is due to the fact that the cost associated with the adjoint method is practically independent of the number of design variables.

We believe that this integrated approach to aero-structural design is a first step in the development of high-fidelity MDO environments, since aerodynamics and structures are two very tightly coupled core disciplines in aircraft design. In the future, this framework is expected to be complemented by additional disciplines that may or may not use the same integration strategies, depending on the bandwidth requirements of the problem.

Acknowledgments

The first author acknowledges the support of the *Fundação para a Ciência e a Tecnologia* from the Portuguese government and the Stanford University Charles Lee Powell Fellowship. The second author has benefited greatly from the support of the AFOSR under Grant No. AF-F49620-01-1-0291 and the Raytheon Aircraft Preliminary Design Group. Finally we would like to thank the ASSET Research Corporation for providing the geometry and specifications for the natural laminar flow supersonic business jet.

References

H. M. Adelman and R. T. Haftka, "Sensitivity analysis of discrete structural systems," *AIAA Journal* vol. 24, no. 5, pp. 823–832, 1986.

M. A. Akgün, R. T. Haftka, K. C. Wu, and J. L. Walsh, "Sensitivity of lumped constraints using the adjoint method," *AIAA Paper* 99-1314, 1999.

N. M. Alexandrov and J. E. Dennis, Jr., "Multilevel algorithms for nonlinear optimization," in *Optimal Design and Control*, J. Borggaard, J. Burkardt, M. Gunzburger, and J. Peterson (eds.), Birkhauser, 1994, pp. 1–22.

N. M. Alexandrov and R. M. Lewis, "Comparative properties of collaborative optimization and other approaches to MDO," in *Proceedings of the First ASMO UK/ISSMO Conference on Engineering Design Optimization*, 1999.

C. Bischof, A. Carle, G. Corliss, A. Grienwank, and P. Hoveland, 'ADIFOR: Generating derivative codes from fortran programs," *Scientific Programming* vol. 1, no. 1, pp. 11–29, 1992.

R. Braun and I. Kroo, "Development and application of the collaborative optimization architecture in a multidisciplinary design environment," in *Multidisciplinary Design Optimization: State of the Art*, N. M. Alexandrov and M. Y. Hussaini (eds.), SIAM, 1997, pp. 98–116.

S. A. Brown, "Displacement extrapolation for CFD+CSM aeroelastic analysis," *AIAA Paper* 97-1090, 1997.

J. Cebral and R. Löhner, "Conservative load projection and tracking for fluid-structure problems," *AIAA Journal* vol. 35, no. 4, pp. 687–692, 1997a. Also: AIAA-96-0797 (1996).

J. Cebral and R. Löhner, "Fluid-structure coupling: Extensions and improvements," *AIAA Paper* 97-0858, 1997b.

E. J. Cramer, J. E. Dennis, P. D. Frank, R. M. Lewis, and G. R. Shubin, "Problem formulation for multidisciplinary optimization," *SIAM Journal on Optimization* vol. 4, pp. 754–776, 1994.

A. DeMiguel and W. Murray, "An analysis of collaborative optimization methods," *AIAA Paper* 2000-4720, 2000.

A. Griewank, *Evaluating Derivatives*, SIAM: Philadelphia, 2000.

M. E. Holden, "Aeroelastic optimization using the collocation method," Ph.D. thesis, Stanford University, Stanford, CA, 1999.

A. Jameson, "Aerodynamic design via control theory," *Journal of Scientific Computing* vol. 3, no. 3, pp. 233–260, 1989.

A. Jameson, L. Martinelli, and N. A. Pierce, "Optimum aerodynamic design using the Navier–Stokes equations," *Theoretical and Computational Fluid Dynamics* vol. 10, pp. 213–237, 1998.

S. Kodiyalam and J. Sobieszczanski-Sobieski, "Bilevel integrated system synthesis with response surfaces," *AIAA Journal* vol. 38, no. 8, pp. 1479–1485, 2002.

I. M. Kroo, "Decomposition and collaborative optimization for large scale aerospace design," in *Multidisciplinary Design Optimization: State of the Art*, SIAM, 1996.

I. Kroo, R. Tracy, J. Chase, and P. Sturdza, "Natural laminar flow for quiet and efficient supersonic aircraft," *AIAA Paper* 2002-0146, 2002.

N. Maman and C. Farhat, "Matching fluid and structure meshes for aeroelastic computations: A parallel approach," *Comput. and Struc.* vol. 54, pp. 779–785, 1995.

J. R. R. A. Martins, "A coupled-adjoint method for high-fidelity aero-structural optimization," Ph.D. thesis, Stanford University, Stanford, California, 2002.

J. R. R. A. Martins, J. J. Alonso, and J. Reuther, "Aero-structural wing design optimization using high-fidelity sensitivity analysis," in H. Honlinger (ed.): *Proceedings—CEAS Conference on Multidisciplinary Aircraft Design Optimization, Cologne, Germany*. Bonn, 2001, pp. 211–226, Lilienthal-Oberth e.V.

J. R. R. A. Martins, J. J. Alonso, and J. J. Reuther, "Complete configuration aero-structural optimization using a coupled sensitivity analysis method," *AIAA Paper* 2002-5402, 2002.

J. R. R. A. Martins, P. Sturdza, and J. J. Alonso, "The complex-step derivative approximation," *ACM Transactions on Mathematical Software* vol. 29, no. 3, pp. 245–262, 2003.

K. Maute, M. Nikbay, and C. Farhat, "Coupled analytical sensitivity analysis and optimization of three-dimensional nonlinear aeroelastic systems," *AIAA Journal* vol. 39, no. 11, pp. 2051–2061, 2001.

S. Nadarajah and A. Jameson, "A comparison of the continuous and discrete adjoint approach to automatic aerodynamic optimization," *AIAA Paper* 2000-0667, 2000.

S. Obayashi and D. Sasaki, "Self-organizing map of pareto solutions obtained from multiobjective supersonic wing design," *AIAA Paper* 2002-0991, 2002.

J. J. Reuther, "Aerodynamic shape optimization using control theory," Ph.D. thesis, University of California Davis. also NASA-CR-201064, 1996.

J. Reuther, J. J. Alonso, A. Jameson, M. Rimlinger, and D. Saunders, "Constrained multipoint aerodynamic shape optimization using an adjoint formulation and parallel computers: Part I," *Journal of Aircraft* vol. 36, no. 1, pp. 51–60, 1999a.

J. Reuther, J. J. Alonso, A. Jameson, M. Rimlinger, and D. Saunders, "Constrained multipoint aerodynamic shape optimization using an adjoint formulation and parallel computers: Part II," *Journal of Aircraft* vol. 36, no. 1, pp. 61–74, 1999b.

J. Reuther, J. J. Alonso, J. R. R. A. Martins, and S. C. Smith, "A coupled aero-structural optimization method for complete aircraft configurations," *AIAA Paper* 99-0187, 1999c.

J. Reuther, J. J. Alonso, J. C. Vassberg, A. Jameson, and L. Martinelli, "An efficient multiblock method for aerodynamic analysis and design on distributed memory systems," *AIAA Paper* 97-1893, 1997.

J. Reuther, A. Jameson, J. Farmer, L. Martinelli, and D. Saunders, "Aerodynamic shape optimization of complex aircraft configurations via an adjoint formulation," *AIAA Paper* 96-0094, 34th Aerospace Sciences Meeting and Exhibit, Reno, Nevada, 1996.

D. Sasaki, S. Obayashi, and K. Nakahashi, "Navier–stokes optimization of supersonic wings with four design objectives using evolutionary algorithm," *AIAA Paper* 2001-2531, 2001.

J. Sobieszczanski-Sobieski, "Sensitivity of complex, internally coupled systems," *AIAA Journal* vol. 28, no. 1, pp. 153–160, 1990.

J. Sobieszczanski-Sobieski and R. T. Haftka, "Multidisciplinary aerospace design optimization: Survey of recent developments," *AIAA Paper* 96-0711, 1996.

J. Yao, J. J. Alonso, A. Jameson, and F. Liu, "Development and validation of a massively parallel flow solver for turbomachinery flow," *Journal of Propulsion and Power* vol. 17, no. 3, pp. 659–668, 2001.