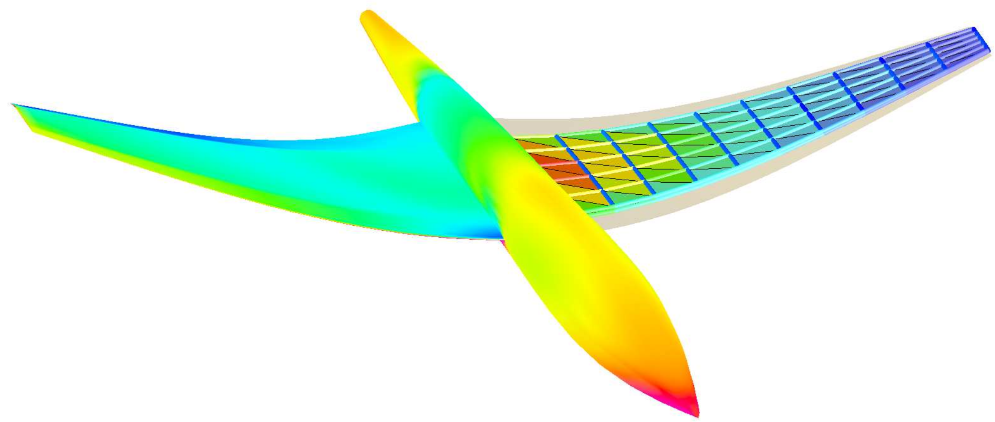
A Crash-Course on the Adjoint Method for Aerodynamic Shape Optimization



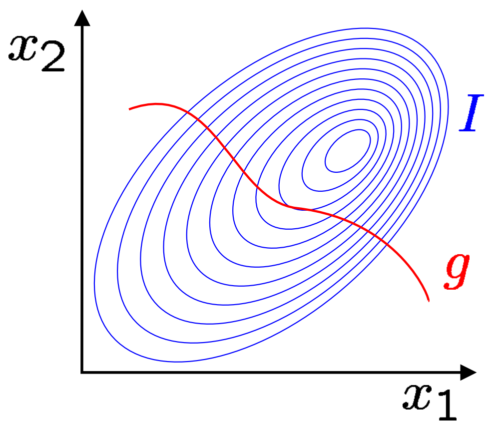
Juan J. Alonso
Department of Aeronautics & Astronautics
Stanford University
jjalonso@stanford.edu

Lecture 19

Outline

- Introduction to Optimization
  - Survey of available optimization methods
  - Approaches to sensitivity analysis
  - Performance of direct vs. adjoint method
- Theory of the adjoint method
- The adjoint system for the Euler equations
- Reduced gradient formulation
- Some results and examples

Introduction to Optimization

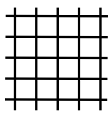


$$\begin{aligned} &\text{minimize} && I(x) \\ & x \in \mathbb{R}^n && \\ &\text{subject to} && g_m(x) \geq 0, \quad m = 1, 2, \ldots, N_g \end{aligned}$$

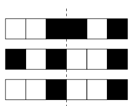
- I: objective function, output (e.g. structural weight).
- x_n: vector of design variables, inputs (e.g. aerodynamic shape); bounds can be set on these variables.
- g_m: vector of constraints (e.g. element von Mises stresses); in general these are nonlinear functions of the design variables.

Optimization Methods

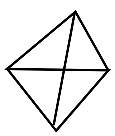
- **Intuition:** decreases with increasing dimensionality.



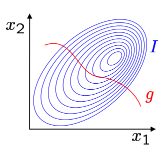
- **Grid or random search:** the cost of searching the design space increases rapidly with the number of design variables.



- **Evolutionary/Genetic algorithms:** good for discrete design variables and very robust; are they feasible when using a large number of design variables?

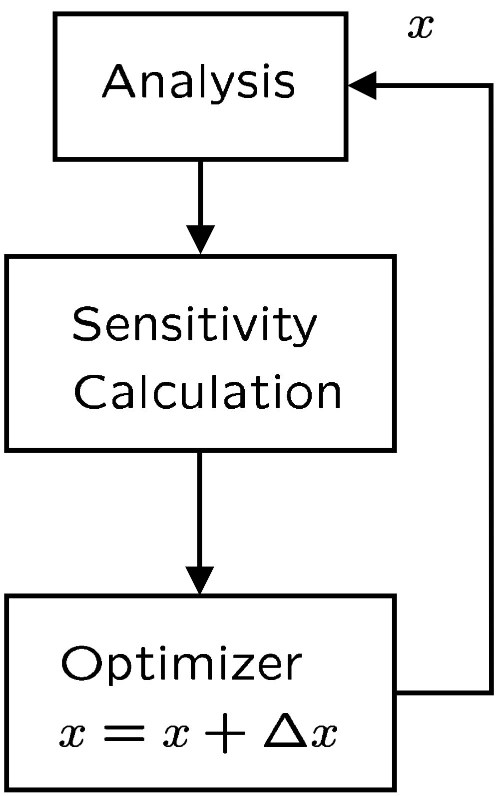


- **Nonlinear simplex:** simple and robust but inefficient for more than a few design variables.



- **Gradient-based:** the most efficient for a large number of design variables; assumes the objective function is "well-behaved". Convergence only guaranteed to a local minimum.

Gradient-Based Optimization: Design Cycle



- Analysis computes objective function and constraints (e.g. aero-structural solver)
- Optimizer uses the sensitivity information to search for the optimum solution (e.g. sequential quadratic programming)
- Sensitivity calculation is usually the bottleneck in the design cycle, particularly for large dimensional design spaces.
- Accuracy of the sensitivities is important, specially near the optimum.

Sensitivity Analysis Methods

- **Finite Differences:** very popular; easy, but lacks robustness and accuracy; run solver N_x times.

$$\boxed{\frac{\mathrm{d}f}{\mathrm{d}x_n} \approx \frac{f(x_n + h) - f(x)}{h} + \mathcal{O}(h)}$$

- **Complex-Step Method:** relatively new; accurate and robust; easy to implement and maintain; run solver N_x times.

$$\boxed{\frac{\mathrm{d}f}{\mathrm{d}x_n} \approx \frac{\mathrm{Im}\left[f(x_n + ih)\right]}{h} + \mathcal{O}(h^2)}$$

- **Algorithmic/Automatic/Computational Differentiation:** accurate; ease of implementation and cost varies.
- **(Semi)-Analytic Methods:** efficient and accurate; long development time; cost can be independent of N_x.

Finite-Difference Derivative Approximations

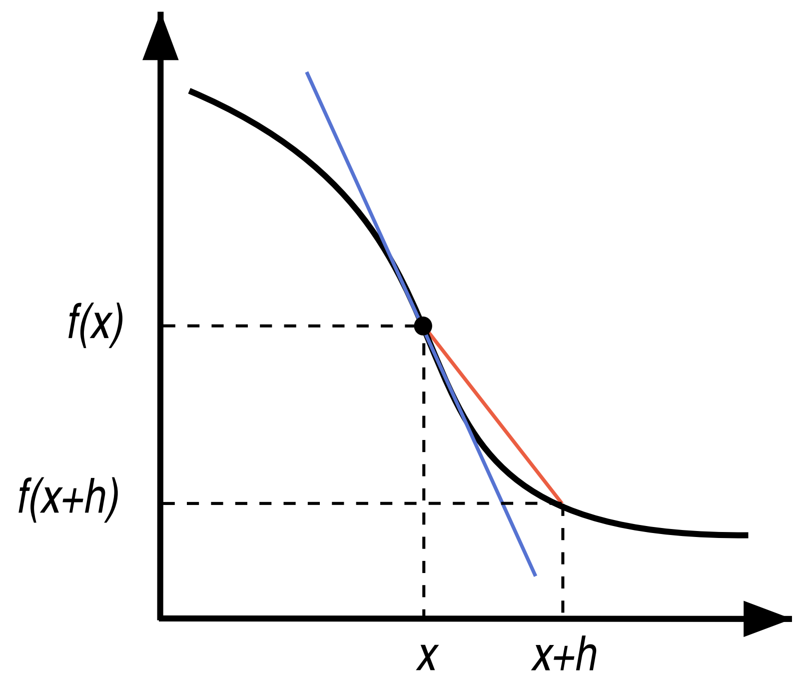
From Taylor series expansion,

$$f(x+h) = f(x) + hf'(x) + h^2\frac{f''(x)}{2!} + h^3\frac{f'''(x)}{3!} + \ldots.$$

Forward-difference approximation:

$$\Rightarrow \boxed{\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{f(x+h) - f(x)}{h} + \mathcal{O}(h).}$$

$f(x)$	$1.23456789012348$4
$f(x+h)$	1.234567890123456
Δf	0.000000000000028



Complex-Step Derivative Approximation

Can also be derived from a Taylor series expansion about x with a complex step ih:

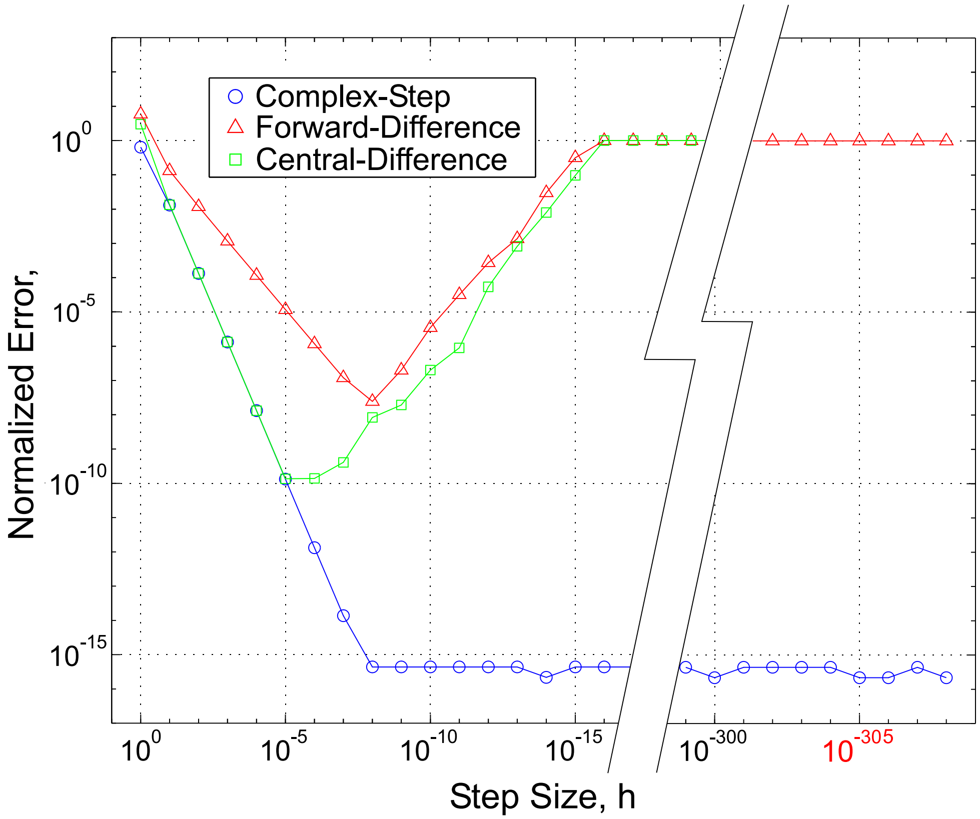
$$f(x+ih) = f(x) + ihf'(x) - h^2\frac{f''(x)}{2!} - ih^3\frac{f'''(x)}{3!} + \ldots$$

$$\Rightarrow f'(x) = \frac{\mathrm{Im}\left[f(x+ih)\right]}{h} + h^2\frac{f'''(x)}{3!} + \ldots$$

$$\Rightarrow \boxed{f'(x) \approx \frac{\mathrm{Im}\left[f(x+ih)\right]}{h}}$$

No subtraction! Second order approximation.

Simple Numerical Example



Estimate derivative at $x = 1.5$ of the function,

$$f(x) = \frac{e^x}{\sqrt{sin^3x + cos^3x}}$$

Relative error defined as:

$$\varepsilon = \frac{\left| f' - f'_{ref} \right|}{\left| f'_{ref} \right|}$$

Challenges in Large-Scale Sensitivity Analysis

- There are efficient methods to obtain sensitivities of many functions with respect to a few design variables - **Direct Method**.
- There are efficient methods to obtain sensitivities of a few functions with respect to many design variables - **Adjoint method**.
- Unfortunately, there are no known methods to obtain sensitivities of many functions with respect to many design variables.
- This is the *curse of dimensionality*.

Outline

- Introduction to Optimization
  - Survey of available optimization methods
  - Approaches to sensitivity analysis
  - Performance of direct vs. adjoint method
- Theory of the adjoint method
- The adjoint system for the Euler equations
- Reduced gradient formulation
- Some results and examples

Symbolic Development of the Adjoint Method

Let I be the cost (or objective) function

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

where

$$\begin{aligned} w &= \text{flow field variables} \\ \mathcal{F} &= \text{grid variables} \end{aligned}$$

The first variation of the cost function is

$$\delta I = \frac{\partial I}{\partial w}^T \delta w + \frac{\partial I}{\partial \mathcal{F}}^T \delta \mathcal{F}$$

The flow field equation and its first variation are

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

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

Introducing a Lagrange Multiplier, ψ, and using the flow field equation as a constraint

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

By choosing ψ such that it satisfies the adjoint equation

$$\left[\frac{\partial R}{\partial w}\right]^T \psi = \frac{\partial I}{\partial w},$$

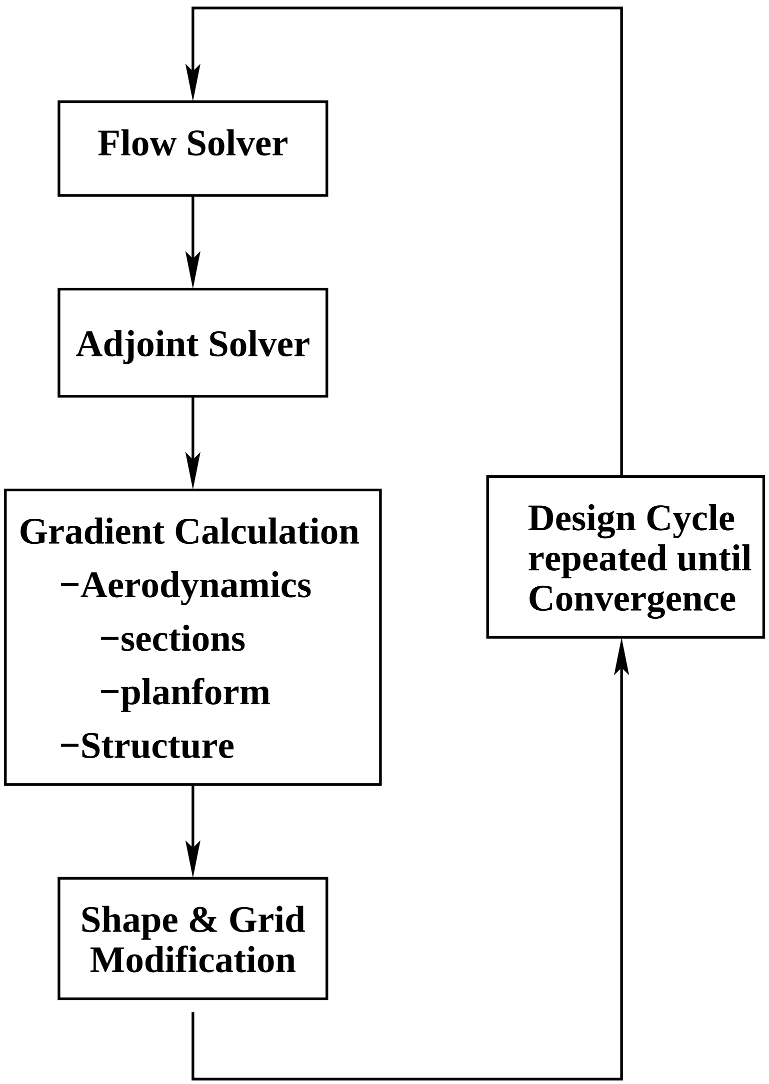
we have

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

This reduces the gradient calculation for an arbitrarily large number of design variables at a single design point to

$$\begin{aligned} &\textbf{One Flow Solution} \\ + \quad &\textbf{One Adjoint Solution} \end{aligned}$$

Design Cycle



Outline

- Introduction to Optimization
  - Survey of available optimization methods
  - Approaches to sensitivity analysis
  - Performance of direct vs. adjoint method
- Theory of the adjoint method
- The adjoint system for the Euler equations
- Reduced gradient formulation
- Some results and examples

Design Using the Euler Equations

In a body-fitted coordinate system, the Euler equations can be written as

$$\frac{\partial W}{\partial t} + \frac{\partial F_i}{\partial \xi_i} = 0 \quad \text{in } D, \tag{1}$$

where

$$W = Jw,$$

and

$$F_i = S_{ij} f_j.$$

Assuming that the surface being designed, B_W, conforms to the computational plane $\xi_2 = 0$, the flow tangency condition can be written as

$$U_2 = 0 \quad \text{on } B_W. \tag{2}$$

Formulation of the Design Problem

Introduce the cost function

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

A variation in the shape will cause a variation δp in the pressure and consequently a variation in the cost function

$$\delta I = \iint_{B_W} (p - p_d) \, \delta p \;\; d\xi_1 d\xi_3. \tag{3}$$

Since p depends on w through the equation of state the variation δp can be determined from the variation δw. Define the Jacobian matrices

$$A_i = \frac{\partial f_i}{\partial w}, \quad C_i = S_{ij} A_j. \tag{4}$$

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

$$\int_{\mathcal{D}} \frac{\partial \psi^T}{\partial \xi_i} \delta F_i d\mathcal{D} = \int_{\mathcal{B}} (n_i \psi^T \delta F_i) d\mathcal{B},$$

where

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

Adding to the variation of the cost function

$$\begin{aligned} \delta I = & \iint_{B_W} (p - p_d)\, \delta p \;\; d\xi_1 d\xi_3 \\ & - \int_{\mathcal{D}} \left(\frac{\partial \psi^T}{\partial \xi_i} \delta F_i \right) d\mathcal{D} \\ & + \int_{\mathcal{B}} \left(n_i \psi^T \delta F_i \right) d\mathcal{B}, \end{aligned} \tag{5}$$

which should hold for an arbitrary choice of ψ. In particular, the choice that satisfies the adjoint equation

$$\frac{\partial \psi}{\partial t} - C_i^T \frac{\partial \psi}{\partial \xi_i} = 0 \quad \text{in } D, \tag{6}$$

subject to far field boundary conditions

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

and solid wall conditions

$$S_{21}\psi_2 + S_{22}\psi_3 + S_{23}\psi_4 = (p - p_d) \quad \text{on } B_W, \tag{7}$$

yields and expression for the gradient that is *independent* of the variation in the flow solution δw:

$$\begin{aligned} \delta I &= -\int_{\mathcal{D}} \frac{\partial \psi^T}{\partial \xi_i} \delta S_{ij} f_j d\mathcal{D} \\ &- \iint_{B_W} \left(\delta S_{21}\psi_2 + \delta S_{22}\psi_3 + S_{23}\psi_4 \right) p \, d\xi_1 d\xi_3. \end{aligned} \tag{8}$$

Outline

- Introduction to Optimization
  - Survey of available optimization methods
  - Approaches to sensitivity analysis
  - Performance of direct vs. adjoint method
- Theory of the adjoint method
- The adjoint system for the Euler equations
- Reduced gradient formulation
- Some results and examples

The Reduced Gradient Formulation

Consider the case of a field 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, so the variation δw is due to the field mesh movement δx. Therefore

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

and since

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

it follows that

$$\int_{\mathcal{D}} \psi^T \frac{\partial}{\partial \xi_i} \left(\delta S_{ij} f_j \right) d\mathcal{D} = - \int_{\mathcal{D}} \psi^T \frac{\partial}{\partial \xi_i} \left(C_i \delta w^* \right) d\mathcal{D}. \qquad (9)$$

A similar relationship has been derived in the general case with boundary movement and the complete derivation will be presented in un upcoming conference paper. Now

$$\begin{aligned} \int_{\mathcal{D}} \psi^T \delta R d\mathcal{D} &= \int_{\mathcal{D}} \psi^T \frac{\partial}{\partial \xi_i} C_i \left(\delta w - \delta w^* \right) d\mathcal{D} \\ &= \int_{\mathcal{B}} \psi^T C_i \left(\delta w - \delta w^* \right) d\mathcal{B} \\ &- \int_{\mathcal{D}} \frac{\partial \psi^T}{\partial \xi_i} C_i \left(\delta w - \delta w^* \right) d\mathcal{D}. \qquad (10) \end{aligned}$$

By choosing ψ to satisfy the adjoint equation (6) and the adjoint boundary condition (7), we have finally the reduced gradient formulation

$$\begin{aligned} \delta I = & + \int_{B_W} \psi^T \left(\delta S_{2j} f_j + C_2 \delta w^*\right) d\xi_1 d\xi_3 \\ & - \iint_{B_W} \left(\delta S_{21}\psi_2 + \delta S_{22}\psi_3 + S_{23}\psi_4\right) p \, d\xi_1 d\xi_3, \end{aligned} \tag{11}$$

which only involves surface integrals.

We have tested this formulation in two- and three-dimensional flows and the results are encouraging for both direct gradient comparisons and actual optimization.

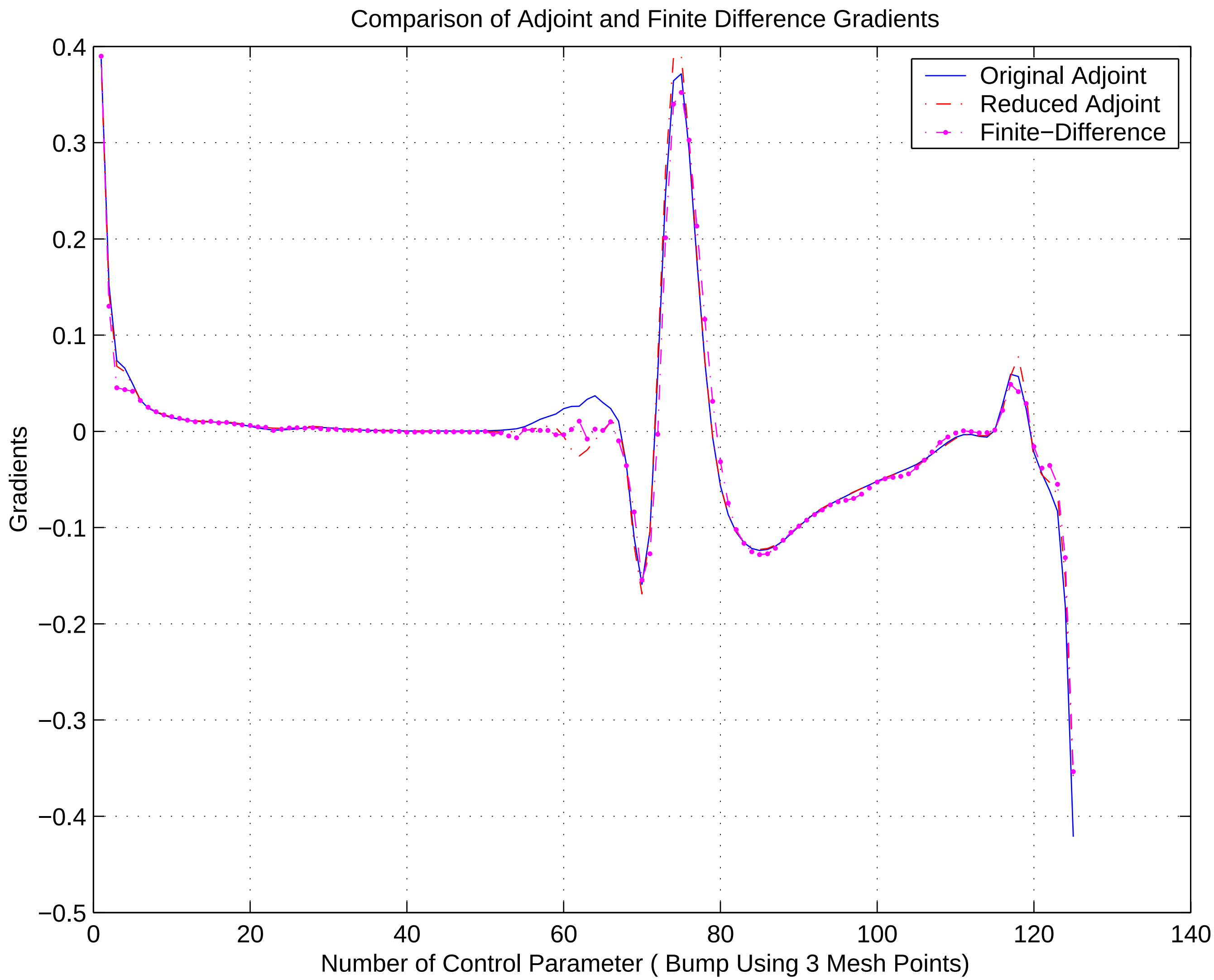
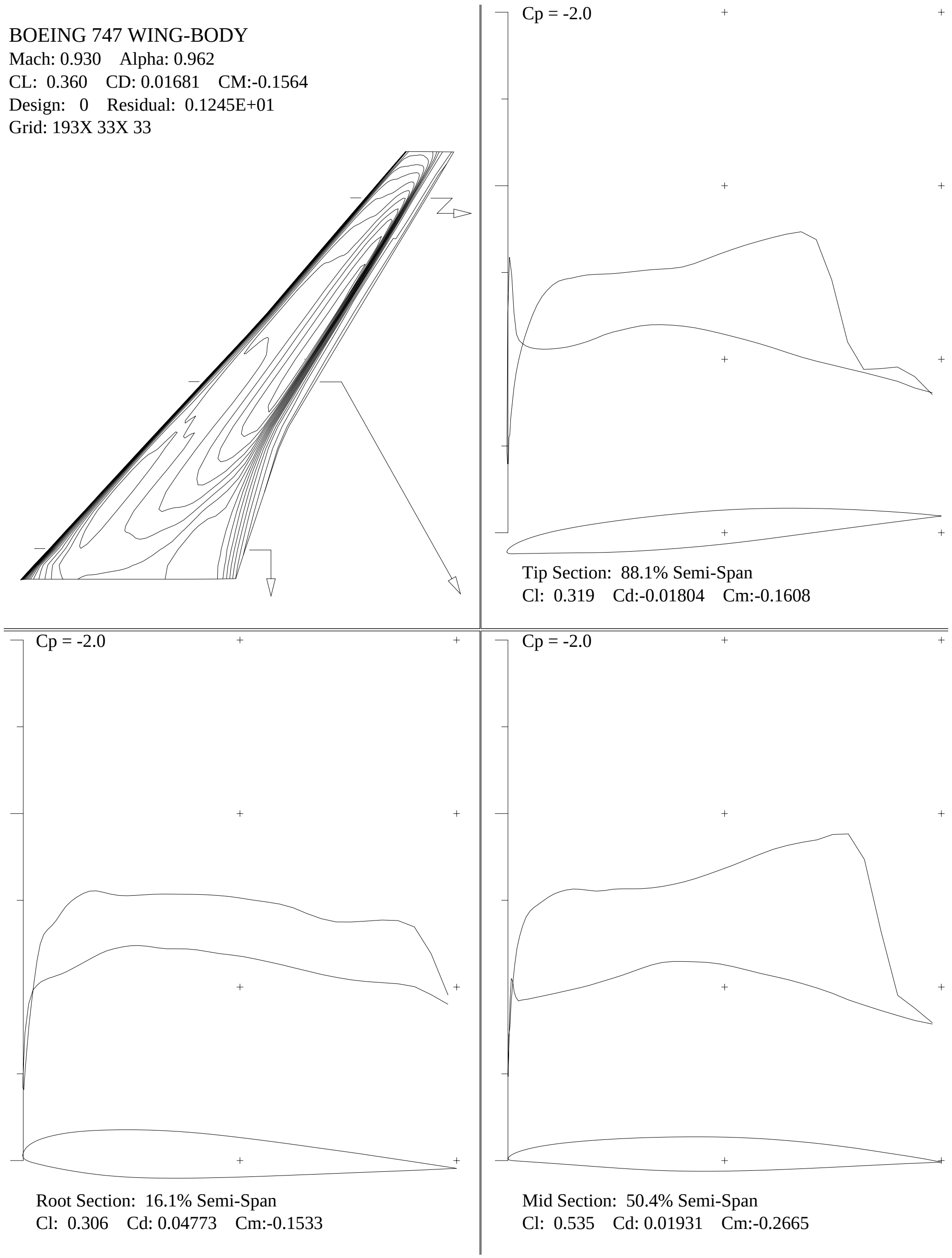
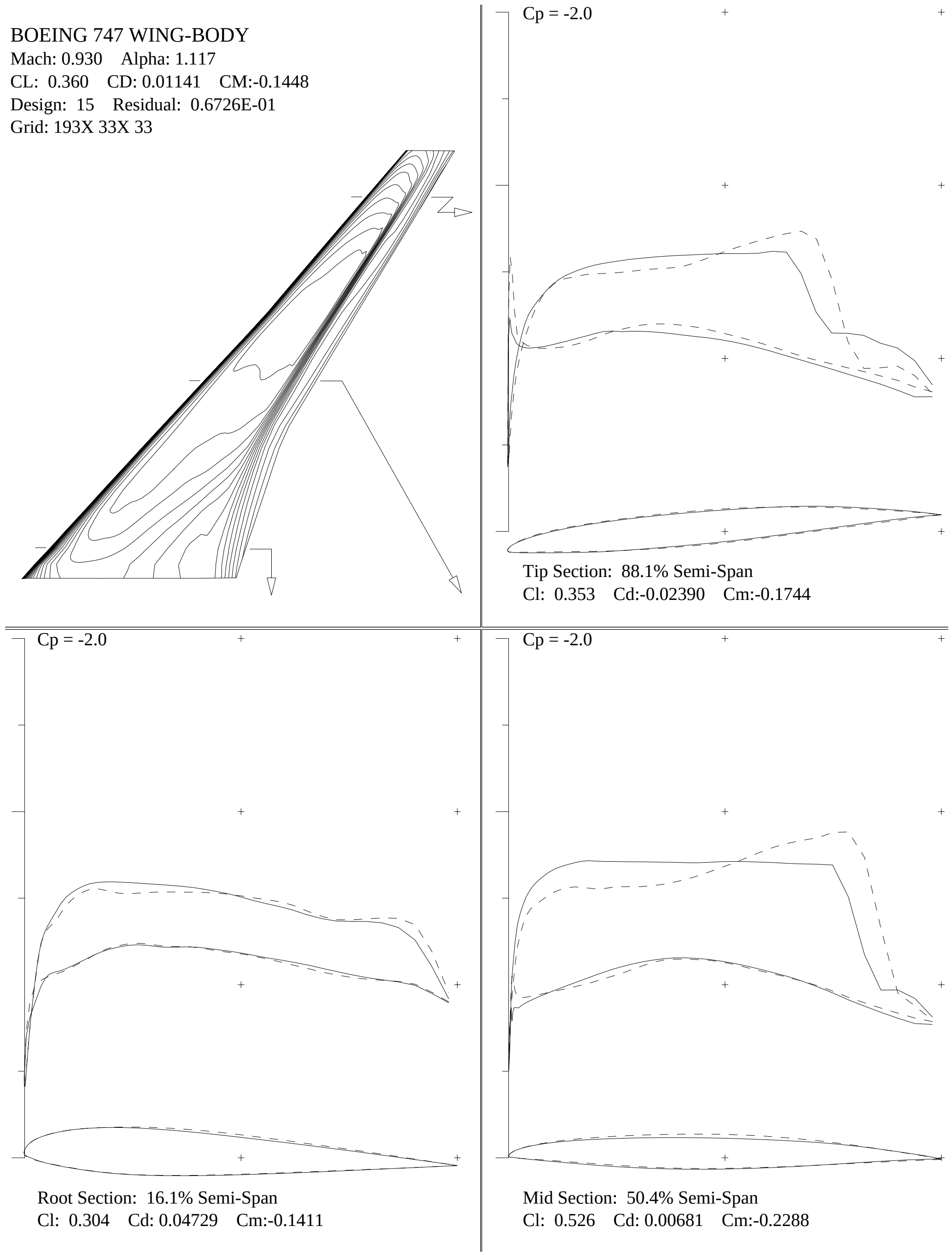


Figure 1: Euler Drag Minimization for RAE2822: Comparison of Original Adjoint, Reduced Adjoint and Finite-Difference Gradients Using 3 Mesh-Point Bump as Design Variable.


BOEING 747 WING-BODY
Mach: 0.930   Alpha: 0.962
CL: 0.360   CD: 0.01681   CM:-0.1564
Design: 0   Residual: 0.1245E+01
Grid: 193X 33X 33
Cp = -2.0
Tip Section: 88.1% Semi-Span
Cl: 0.319   Cd:-0.01804   Cm:-0.1608
Cp = -2.0
Root Section: 16.1% Semi-Span
Cl: 0.306   Cd: 0.04773   Cm:-0.1533
Cp = -2.0
Mid Section: 50.4% Semi-Span
Cl: 0.535   Cd: 0.01931   Cm:-0.2665


BOEING 747 WING-BODY
Mach: 0.930 Alpha: 1.117
CL: 0.360 CD: 0.01141 CM:-0.1448
Design: 15 Residual: 0.6726E-01
Grid: 193X 33X 33
Cp = -2.0
Tip Section: 88.1% Semi-Span
Cl: 0.353 Cd:-0.02390 Cm:-0.1744
Cp = -2.0
Root Section: 16.1% Semi-Span
Cl: 0.304 Cd: 0.04729 Cm:-0.1411
Cp = -2.0
Mid Section: 50.4% Semi-Span
Cl: 0.526 Cd: 0.00681 Cm:-0.2288

Adjoint Design Software

The adjoint for the Euler and Navier-Stokes equations has been implemented into the following codes:

- SYN87: Wing-alone Euler C-H mesh.
- SYN88: Wing-body Euler C-H mesh.
- SYN107: Wing-body N-S C-H mesh.
- SYN87-MB: Arbitrary configuration, Euler, multiblock mesh.
- SYN107-MB: Arbitrary configuration, N-S, multiblock mesh.
- SYNPLANE: Arbitrary configuration, Euler, unstructured mesh.

Adjoint Design Projects

Adjoint methods (Euler and/or N-S) have been used with the following configurations:

- Boeing 747-200
- McDonnell Douglas MDXX
- Raytheon Premier I
- NASA High Speed Civil Transport
- Reno Air Racer
- IPTN N2130
- Other projects at BAE, DLR, NLR, etc.
- Research configurations (subsonic, transonic, supersonic)