REDUCTION OF THE ADJOINT GRADIENT FORMULA IN THE CONTINUOUS LIMIT

Antony Jameson
Stanford University
Sangho Kim
Stanford University

41st AIAA Aerospace Sciences Meeting and Exhibit
Reno, NV
January 6-9, 2003

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

One Flow Solution
$+$ **One Adjoint Solution**

Adjoint Method for the Euler Equations

For example the **cost function** for the **inverse problem** can be defined as

$$I = \frac{1}{2} \iint_{B_W} \left(p - p_d\right)^2 d\xi_1 d\xi_2.$$

Taking the **first variation** of the **cost function**

$$\delta I = \iint_{B_W} \left(p - p_d\right) \delta p \;\; d\xi_1 d\xi_2.$$

Taking the **first variation** of the **governing equations**

$$\frac{\partial}{\partial \xi_i} \left(\delta F_i\right) = 0,$$

where

$$\delta F_i = S_{ij} \left[\frac{\partial f_j}{\partial w}\right] \delta w + \delta \left(S_{ij}\right) f_j$$

and

$$S_{ij} = J \frac{\partial \xi_i}{\partial x_j}.$$

Multiplying by a **Lagrange Multiplier** ψ and integrating by parts gives

$$\int_D \left(\frac{\partial \psi^T}{\partial \xi_i} S_{ij} \left[\frac{\partial f_j}{\partial w}\right] \delta w\right) \; d\xi + \int_D \left(\frac{\partial \psi^T}{\partial \xi_i} \delta S_{ij} f_j\right) \; d\xi + \int_B \left(\bar{n}_i \psi^T S_{ij} \frac{\partial f_j}{\partial w} \delta w\right) d\xi_B = 0$$

On the boundary

$$\delta F_2 = \left\{ \begin{array}{c} 0 \\ S_{21}\delta p \\ S_{22}\delta p \\ S_{23}\delta p \\ 0 \end{array} \right\} + p \left\{ \begin{array}{c} 0 \\ \delta\left(S_{21}\right) \\ \delta\left(S_{22}\right) \\ \delta\left(S_{23}\right) \\ 0 \end{array} \right\} \quad \text{on} \quad B_W.$$

Choose ψ to satisfy the **adjoint equation**

$$\frac{\partial \psi}{\partial t} - (S_{ij}\frac{\partial f_j}{\partial w})^T \frac{\partial \psi}{\partial \xi_i} = 0 \quad \text{in } D,$$

with the **boundary condition**

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

The **variation** in the **cost function** then becomes

$$\begin{aligned} \delta I = & - \int_D \frac{\partial \psi^T}{\partial \xi_i} \delta\left(S_{ij}\right) f_j d\xi_k \\ & - \iint_{B_W} \left\{ \psi_2 \delta\left(S_{21}\right) + \psi_3 \delta\left(S_{22}\right) + \psi_4 \delta\left(S_{23}\right) \right\} p \, d\xi_1 d\xi_2 \end{aligned}$$

This formula includes the **variaton** of the mesh throughout the **domain**. However, the **true gradient** must **depand only on** the **surface shape change**.

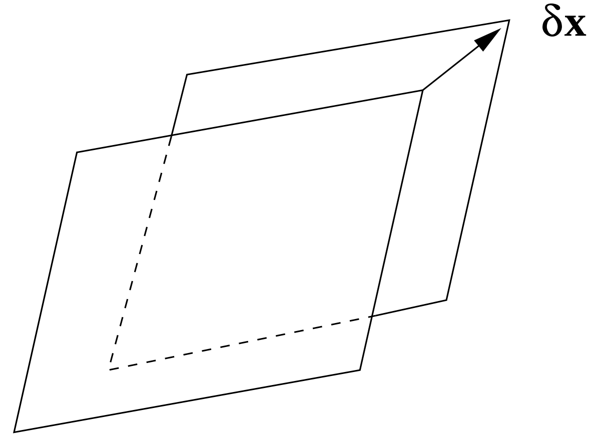
Reduced Gradient Formulation

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

$$\delta I = 0$$

but there is a **variation** in the **transformed flux**,

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



Here the **true solution** is **unchanged**, so the **variation** δw is due to the **mesh movement** δx at **fixed boundary configuration**. Therefore

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

and since

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

it follows that

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

Now

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

Here on the **wall boundary**

$$C_2 \delta w = \delta F_2 - \delta S_{2j} f_j.$$

Thus, by choosing ψ satisfying the **adjoint equation** and the **adjoint boundary condition**, we have finally the **reduced gradient formulation** that

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

It can be verified that **this relation holds in the general case with boundary movement.**

Reduced Gradient Formulation with Boundary Movement

Using the **permutation tensor** ϵ_{ijk} we can express the **elements of** S as

$$S_{ij} = \frac{1}{2}\epsilon_{jpq}\epsilon_{irs}\frac{\partial x_p}{\partial \xi_r}\frac{\partial x_q}{\partial \xi_s}. \tag{1}$$

Then

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

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

$$S_{1j} = \epsilon_{jpq}\frac{\partial x_p}{\partial \xi_2}\frac{\partial x_q}{\partial \xi_3}, \quad S_{2j} = \epsilon_{jpq}\frac{\partial x_p}{\partial \xi_3}\frac{\partial x_q}{\partial \xi_1}, \quad S_{3j} = \epsilon_{jpq}\frac{\partial x_p}{\partial \xi_1}\frac{\partial x_q}{\partial \xi_2}. \tag{3}$$

Now the **Euler equations** can be written as

$$J\frac{\partial w}{\partial t} + R\left(w\right) = 0 \tag{4}$$

where

$$R\left(w\right) = S_{ij}\frac{\partial f_j}{\partial \xi_i} = \frac{\partial}{\partial \xi_i}\left(S_{ij} f_j\right), \tag{5}$$

using (2).

Using the formula(1), and the property (2)

$$\begin{aligned}
\frac{\partial}{\partial \xi_i} & (\delta S_{ij} f_j) \\
&= \frac{1}{2}\frac{\partial}{\partial \xi_i}\left\{\epsilon_{jpq}\epsilon_{irs}\left(\frac{\partial \delta x_p}{\partial \xi_r}\frac{\partial x_q}{\partial \xi_s} + \frac{\partial x_p}{\partial \xi_r}\frac{\partial \delta x_q}{\partial \xi_s}\right) f_j\right\} \\
&= \frac{1}{2}\epsilon_{jpq}\epsilon_{irs}\left(\frac{\partial \delta x_p}{\partial \xi_r}\frac{\partial x_q}{\partial \xi_s} + \frac{\partial x_p}{\partial \xi_r}\frac{\partial \delta x_q}{\partial \xi_s}\right)\frac{\partial f_j}{\partial \xi_i} \\
&= \frac{1}{2}\epsilon_{jpq}\epsilon_{irs}\left\{\frac{\partial}{\partial \xi_r}\left(\delta x_p \frac{\partial x_q}{\partial \xi_s}\frac{\partial f_j}{\partial \xi_i}\right)\right\} \\
&+ \frac{1}{2}\epsilon_{jpq}\epsilon_{irs}\left\{\frac{\partial}{\partial \xi_s}\left(\delta x_q \frac{\partial x_p}{\partial \xi_r}\frac{\partial f_j}{\partial \xi_i}\right)\right\} \\
&= \frac{\partial}{\partial \xi_r}\left(\delta x_p \epsilon_{pqj}\epsilon_{rsi}\frac{\partial x_q}{\partial \xi_s}\frac{\partial f_j}{\partial \xi_i}\right). \qquad (6)
\end{aligned}$$

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

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

Then we obtain

$$\frac{\partial}{\partial \xi_i}(\delta S_{ij} f_j) = \frac{\partial}{\partial \xi_r}\left(\epsilon_{pqj}\epsilon_{rsi}\frac{\partial x_p}{\partial \xi_k}\frac{\partial x_q}{\partial \xi_s}\frac{\partial f_j}{\partial \xi_i}\delta \xi_k\right). \qquad (7)$$

The term in $\frac{\partial}{\partial \xi_1}$ is

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

Here the term multiplying $\delta\xi_1$ is

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

According to the formulas(3) this may be recognized as

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

or, using the quasi-linear form(5) of the equation for steady flow, as

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

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

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

and

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

Thus the term in $\frac{\partial}{\partial \xi_1}$ is reduced to

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

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

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

Optimization Procedure : Descent Method with Smoothed Gradient

Let $\mathcal{F}$ represent the **design variable**, and $\mathcal{G}$ the **gradient**. An **improvement** could then be made with a **shape change**

$$\delta\mathcal{F} = -\lambda\mathcal{G}.$$

In order to preserve the **smoothness** we redefine the **gradient** to correspond to a **weighted Sobolev inner product** of the form

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

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

$$\delta I =< \bar{\mathcal{G}}, \delta\mathcal{F} > .$$

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

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

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

$$\bar{\mathcal{G}} - \frac{\partial}{\partial \xi_1} \epsilon \frac{\partial}{\partial \xi_1} \bar{\mathcal{G}} = \mathcal{G}. \tag{8}$$

In the multi-dimensional case, the smoothing is applied in product form. Finally we set

$$\delta \mathcal{F} = -\lambda \bar{\mathcal{G}}$$

with the result that

$$\delta I = -\lambda < \bar{\mathcal{G}}, \bar{\mathcal{G}} > \quad < 0,$$

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

When **second-order central differencing** is applied to (8), the equation at a given node, i, can be expressed as

$$\bar{\mathcal{G}}_i - \epsilon \left(\bar{\mathcal{G}}_{i+1} - 2\bar{\mathcal{G}}_i + \bar{\mathcal{G}}_{i-1}\right) = \mathcal{G}_i, \qquad 1 \le i \le n,$$

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

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

where A is the $n \times n$ tri-diagonal matrix such that

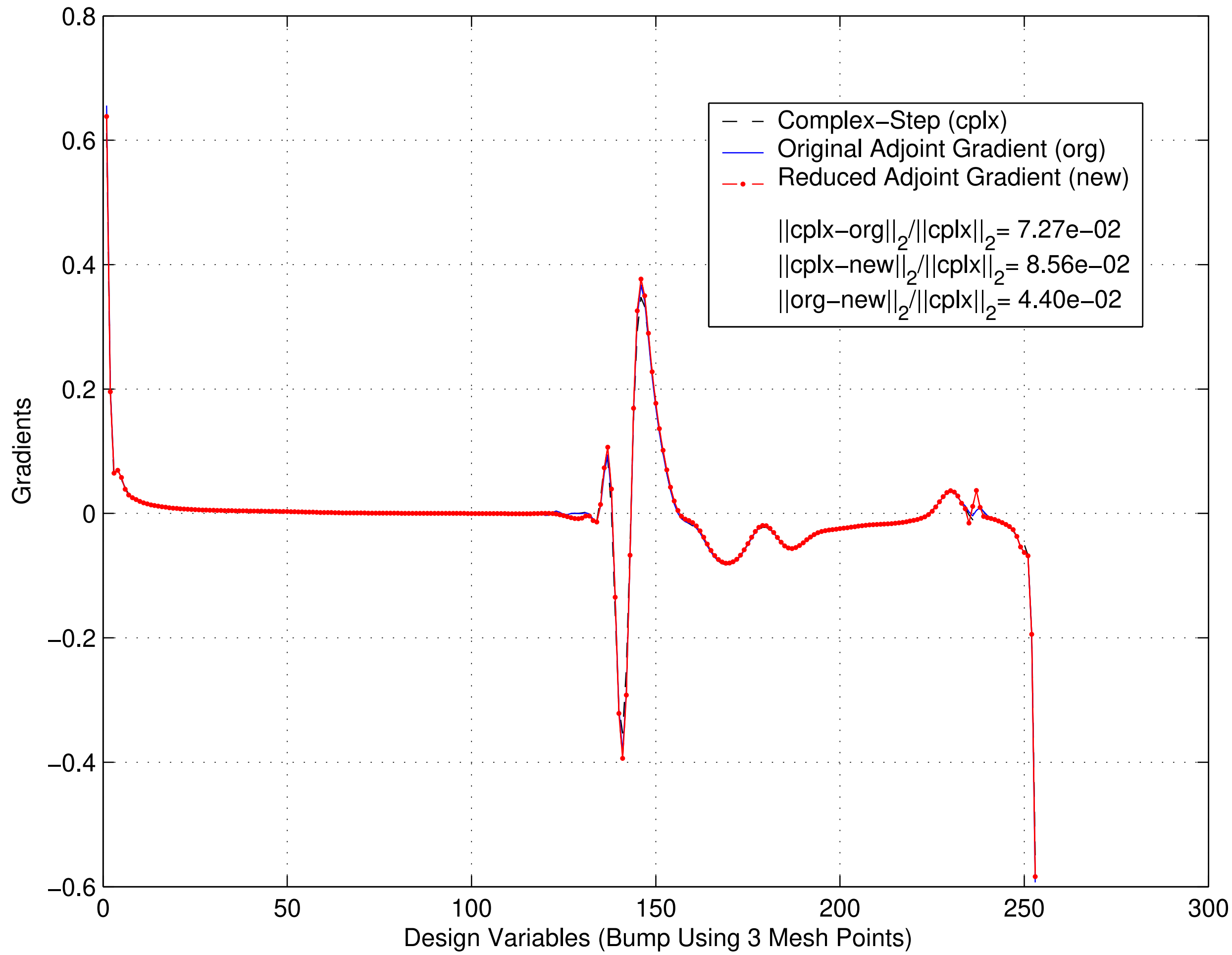
$$A^{-1} = \begin{bmatrix} 1+2\epsilon & -\epsilon & 0 & . & 0 \\ \epsilon & . & . & & \\ 0 & . & . & . & \\ . & & . & . & -\epsilon \\ 0 & & & \epsilon & 1+2\epsilon \end{bmatrix}.$$

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

$$\delta\mathcal{F} = -\lambda A\mathcal{G}, \tag{9}$$

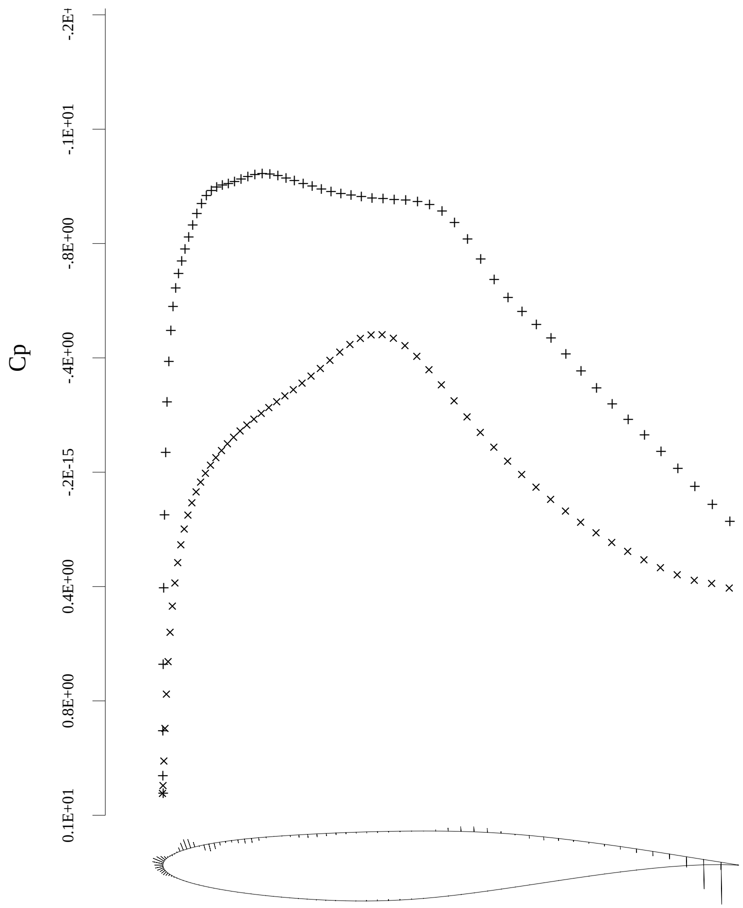
where the **smoothing operator A** may be regarded as a **Preconditioner.**

RAE 2822 Drag Minimization : 2D Gradient Comparison

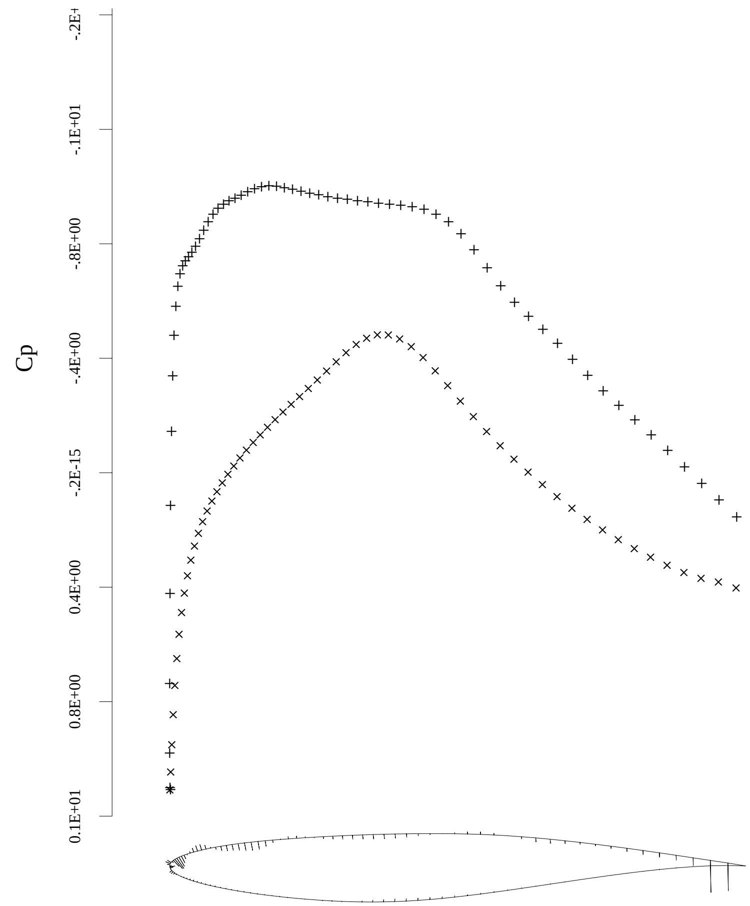


Reduced adjoint vs. original adjoint and complex-step gradient

RAE 2822 Drag Minimization : $M_\infty = 0.75$, $C_l = 0.6$

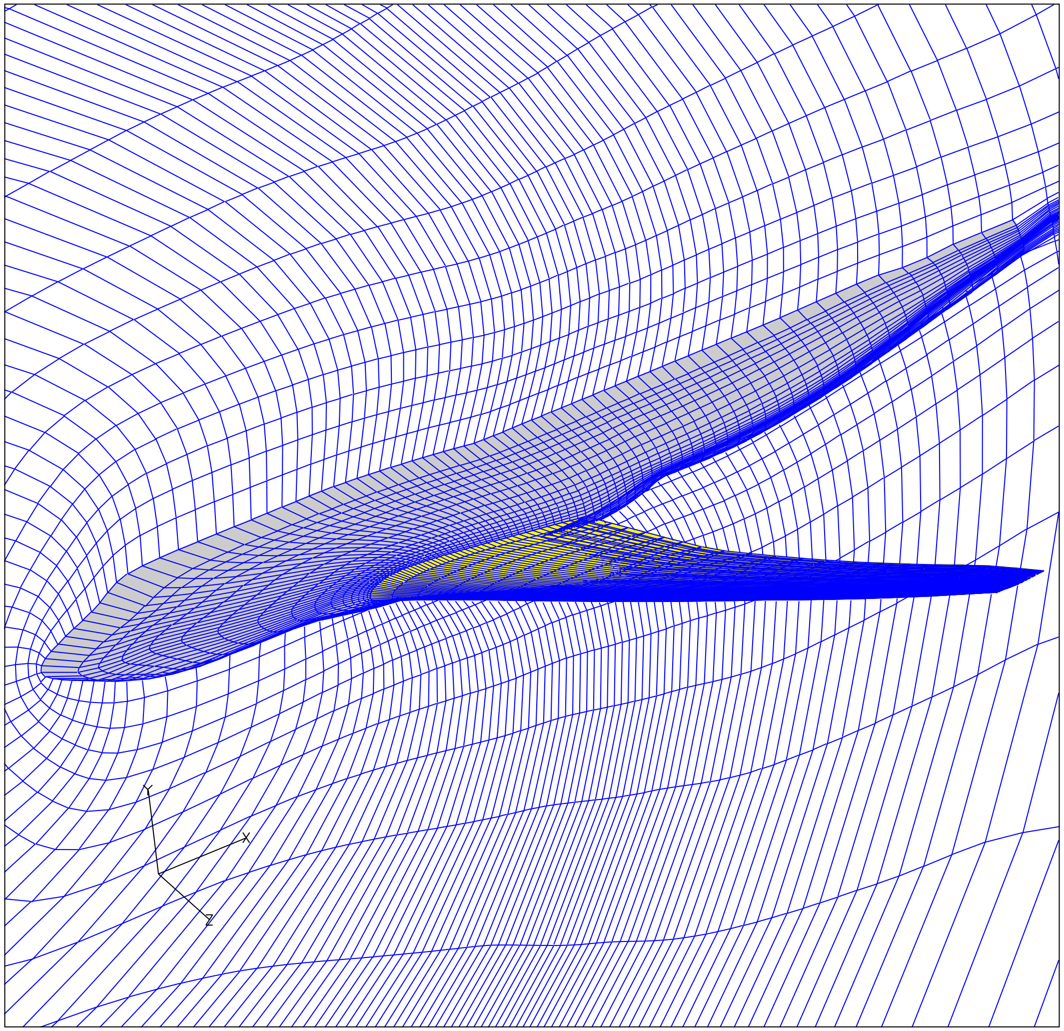


Original adjoint gradient: After 12 design iterations, $C_d = 0.0081 \rightarrow C_d = 0.0019$

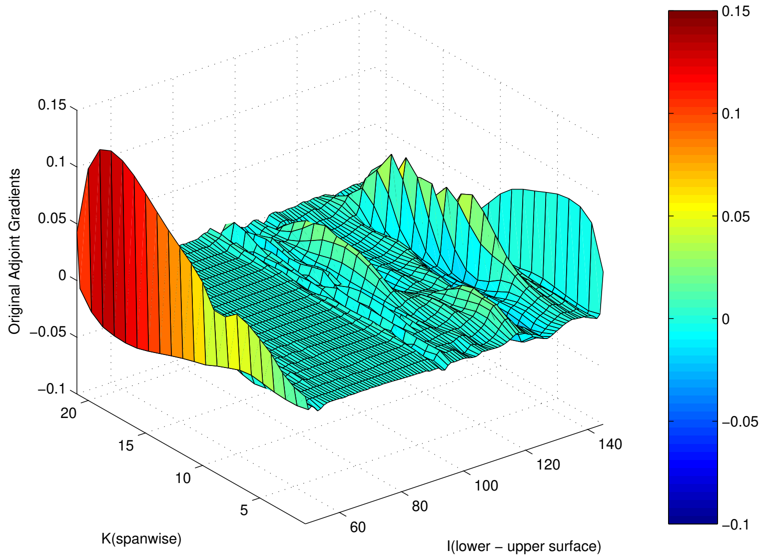


Reduced adjoint gradient: After 12 design iterations, $C_d = 0.0081 \rightarrow C_d = 0.0019$

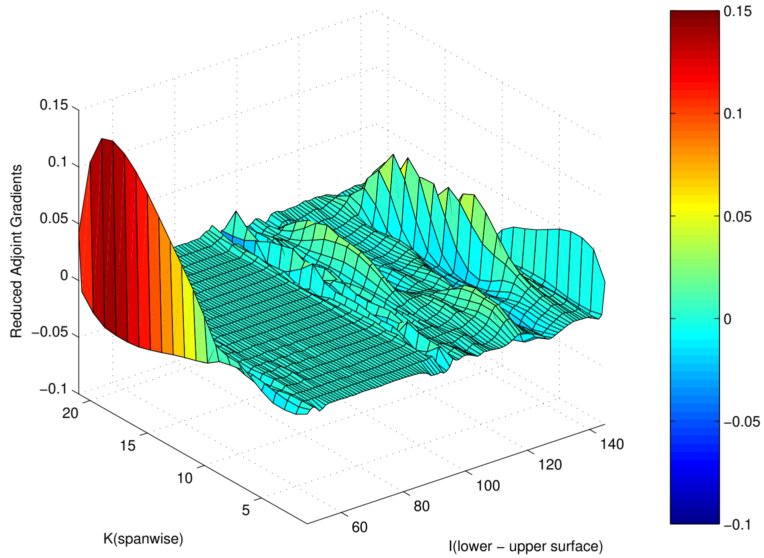
Wing Redesign of Boeing 747 : $192 \times 32 \times 32$ C-H Grid



Wing Redesign of Boeing 747 : 3D Gradient Comparison

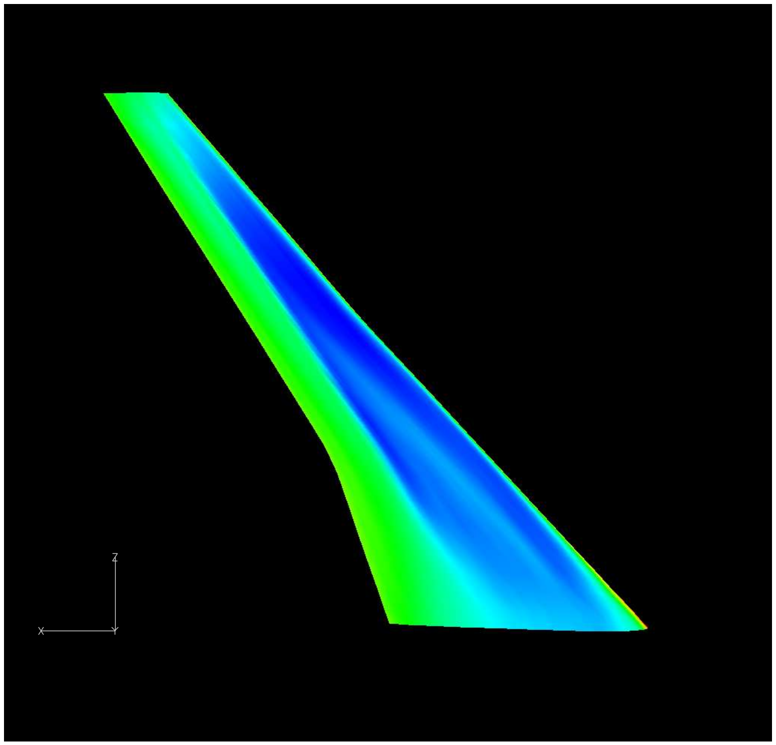


Original adjoint gradient

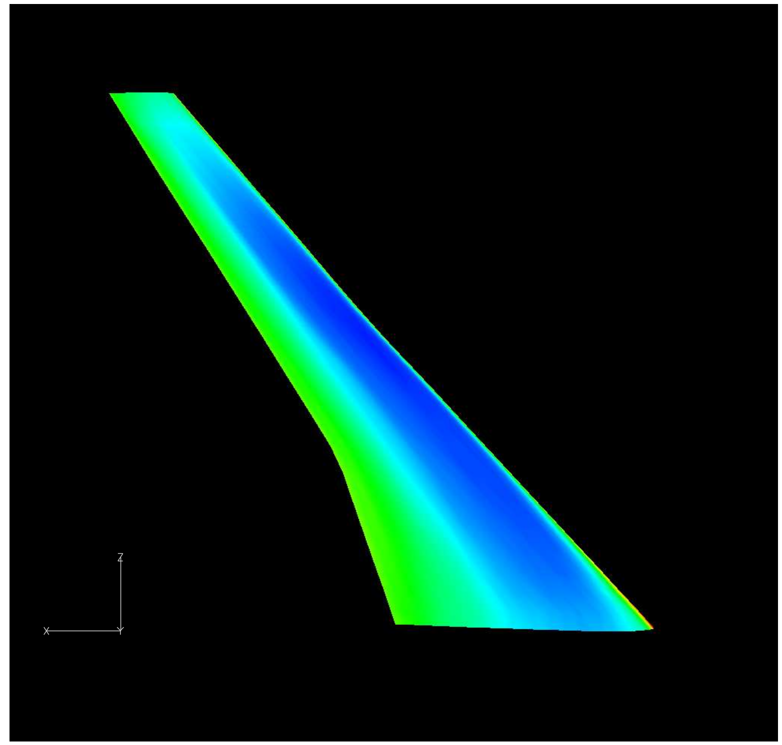


Reduced adjoint gradient

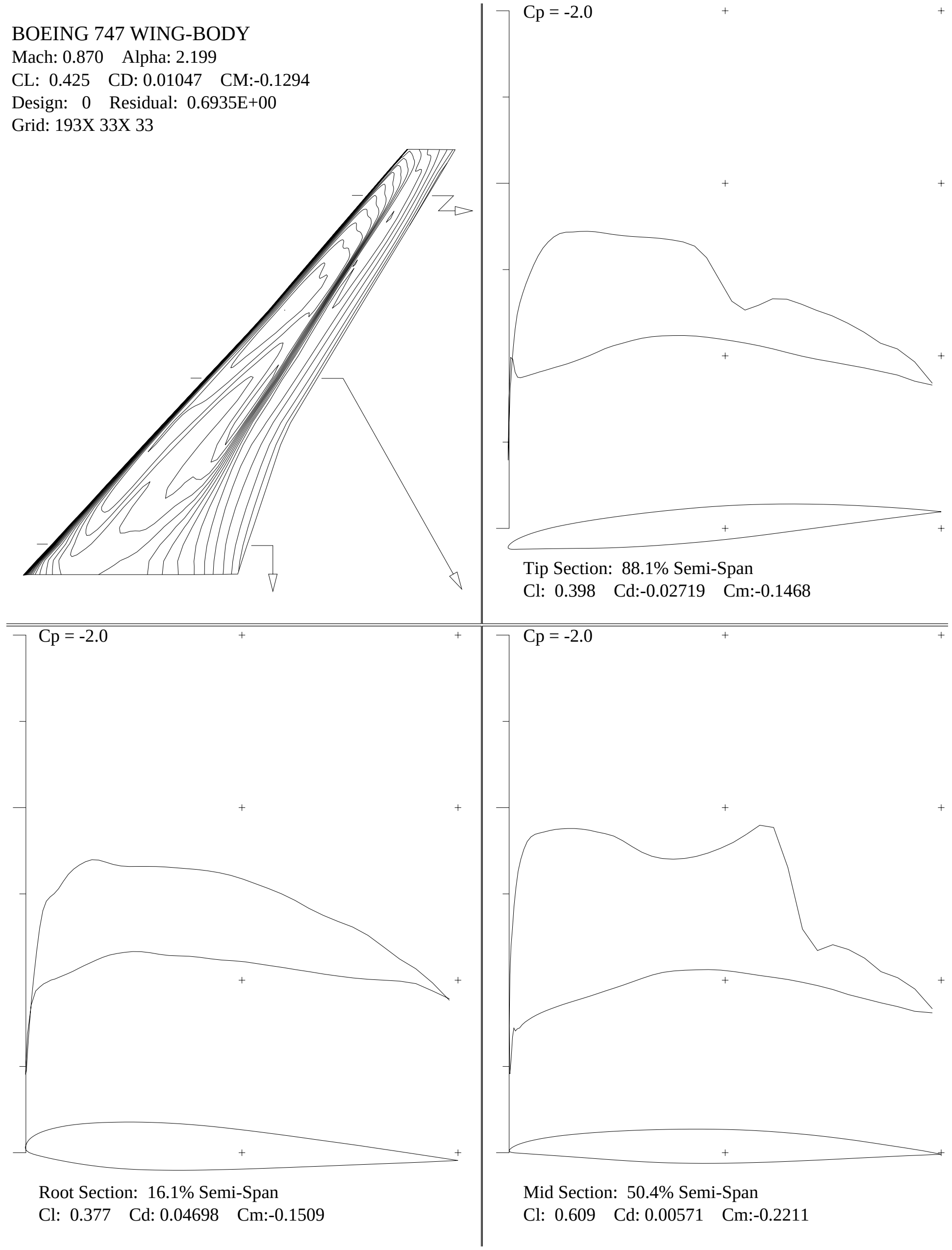
Wing Redesign of Boeing 747 : $M_\infty = 0.87, CL = 0.425$



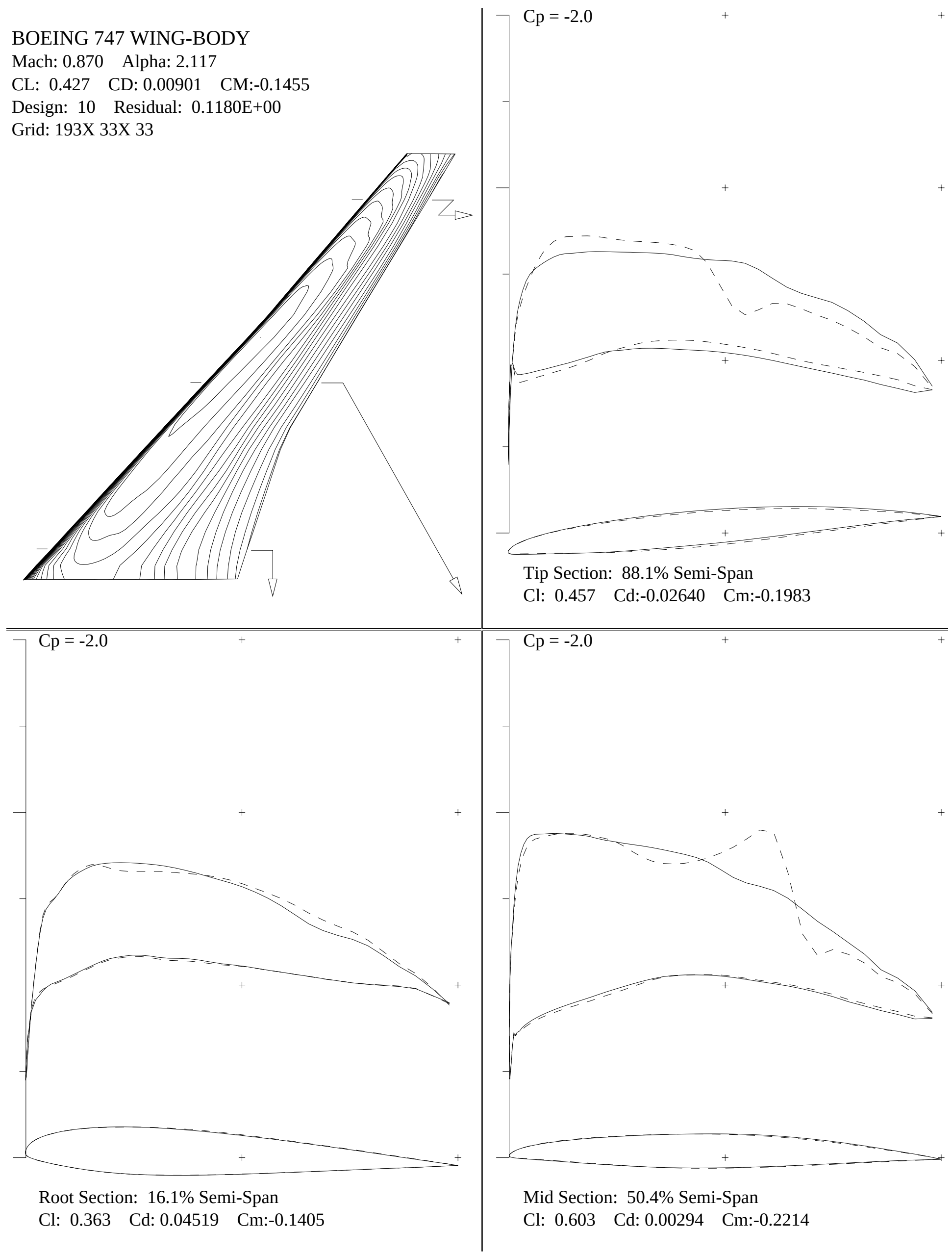
Initial solution :
$CD = 0.01047, \alpha = 2.199°$



After 10 design iterations :
$CD = 0.00901, \alpha = 2.117°$

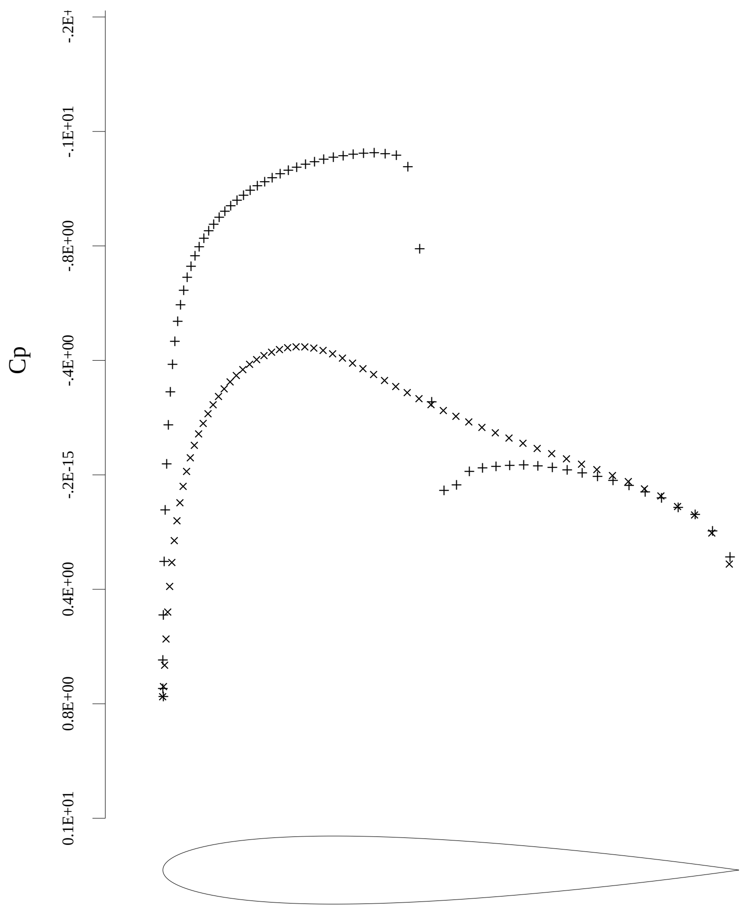


Initial C_p distribution on the upper surface of the wing

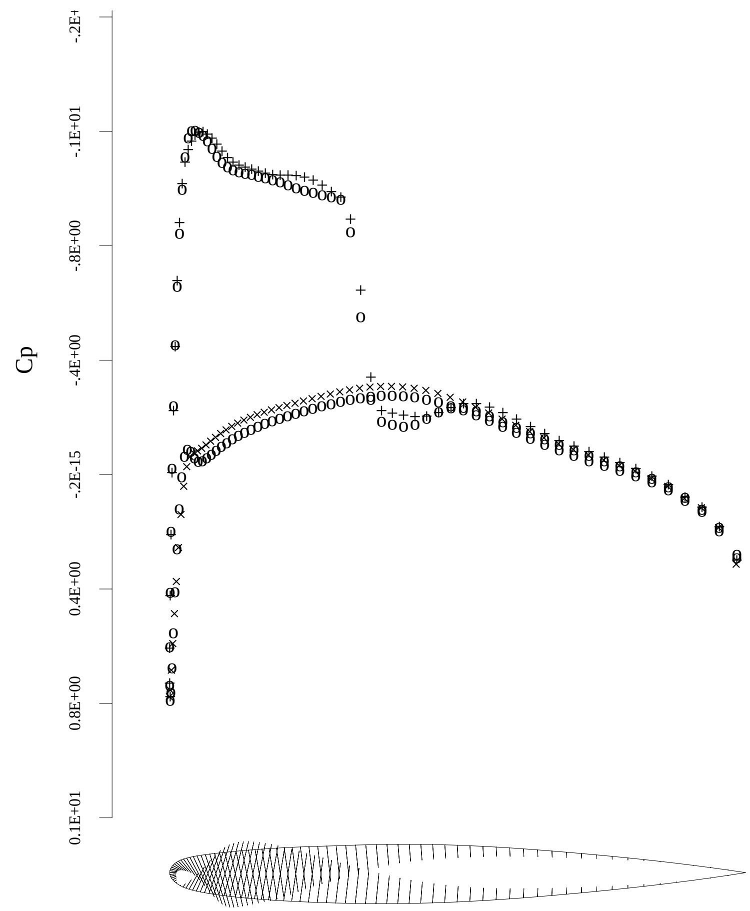


Initial C_p : $---$, Redesigned C_p : —–

Inverse Design : NACA 0012 Wing → ONERA M6 Wing

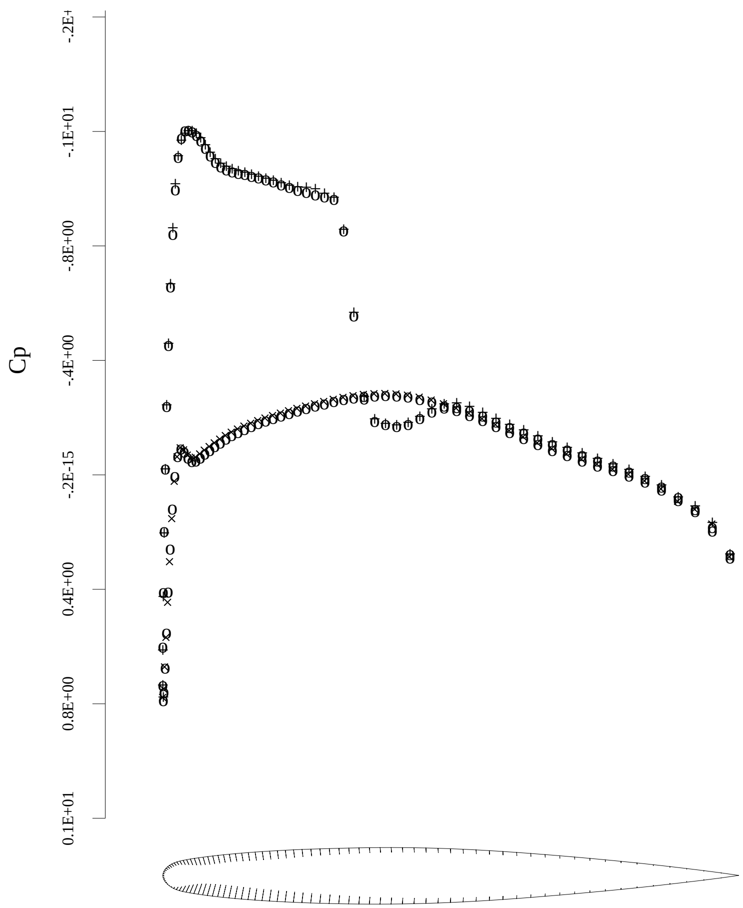


Initial solution

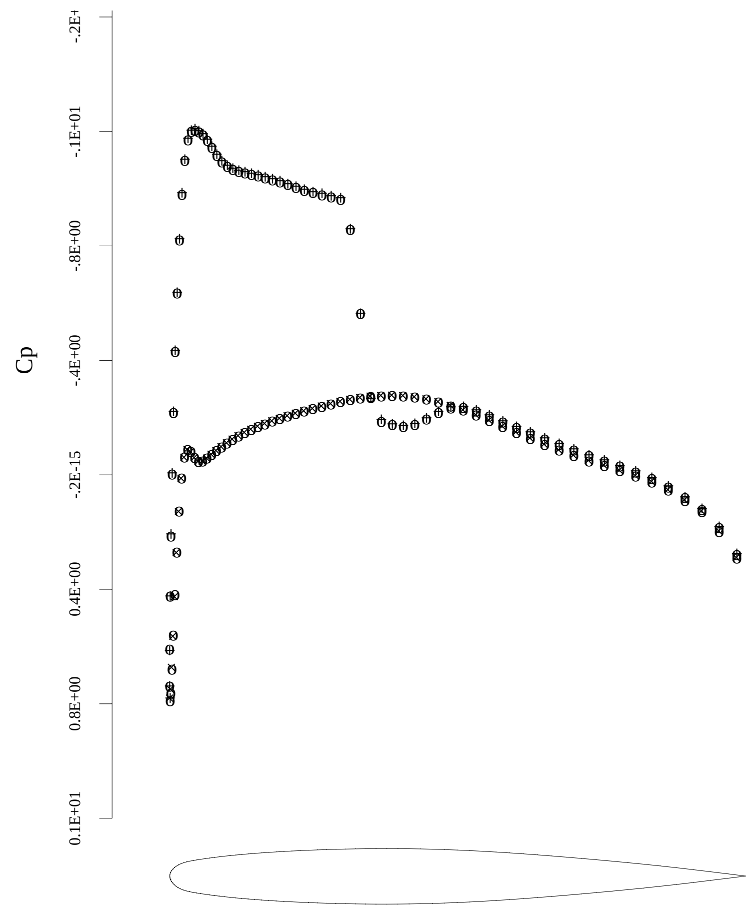


After 10 design iterations

Inverse Design : NACA 0012 Wing → ONERA M6 Wing

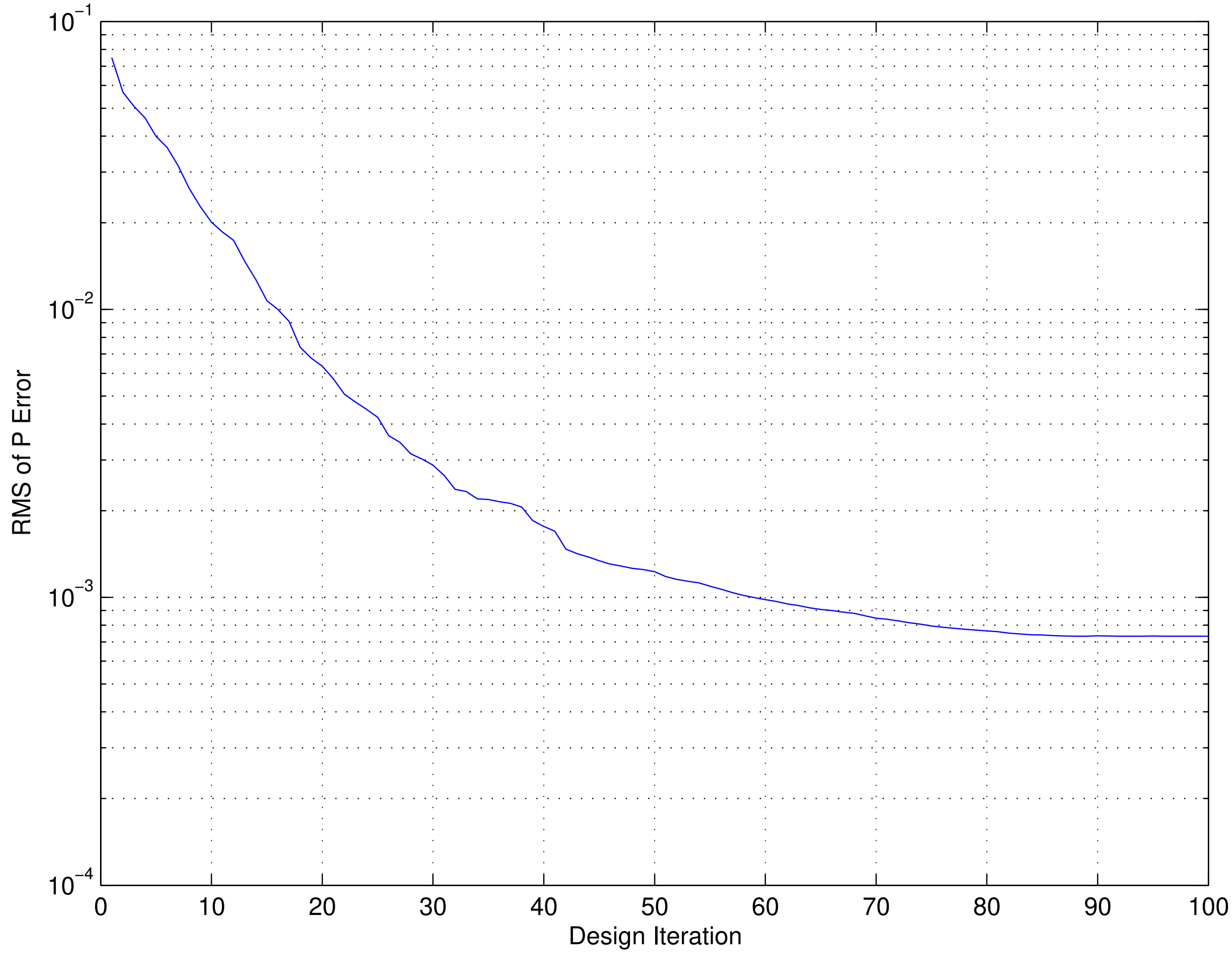


After 20 design iterations



After 100 design iterations

Inverse Design : NACA 0012 Wing → ONERA M6 Wing



Convergence history of pressure difference for the inverse wing design, P_{err} : $0.0749 \rightarrow 0.0007$