ACCURACY OF GRADIENT COMPUTATIONS FOR AERODYNAMIC SHAPE OPTIMIZATION PROBLEMS

Mattias Chevalier and Martin Berggren
FFA, The Aeronautical Research Institute of Sweden

Keywords: shape optimization, adjoint equations

Abstract

Applying nonlinear optimization techniques such as quasi-Newton methods to aerodynamic shape optimization problems requires the calculation of gradients of a given objective function. An effective way of calculating such gradients is through the use of the so-called adjoint equations. To achieve fast convergence in the optimization algorithm, accurately computed gradients are needed. In the computation of such gradients the discretization of the problem and the choice of boundary conditions are two important aspects. These issues are studied in the context of shape optimization of a quasi-1D nozzle using physically relevant boundary conditions. Entropy is enforced at the inlet boundary, and the static pressure is specified at the outlet boundary for subsonic flows. A cell-centered finite-volume discretization with a standard implementation of the boundary conditions is applied, and the corresponding numerical scheme and numerical boundary conditions for the adjoint equations are derived in a fully discrete sense.

Numerical experiments at subsonic and transonic speeds, show that the gradient evaluations are accurate enough to obtain satisfactory convergence of the quasi-Newton algorithm.

1 Introduction

As more and more powerful computers develop, the range of problems possible to solve numerically increases. One class of such problem is aerodynamic shape optimization using the Euler or Navier–Stokes equations as the flow model. Discretization issues in connection with aerodynamic shape optimization are discussed in this article. A simple model for nozzle flow is the quasi-1D Euler equations, in which the nozzle geometry is represented as a scalar function occurring in the coefficients of the equation. This is a standard model problem for transonic flow sharing many features with more complicated models, but having a known solution in terms of an implicit formula for the Mach number and the area function. An objective (or cost) function is introduced to measure, in a least-square sense, how far from the optimal design we are. To improve the geometry, we use optimization methods that utilize the objective function gradient, computed with the aid of the adjoint equations.

We will state the adjoint equations, derived from the state equation and objective function using physically relevant boundary conditions. A quasi-discrete form of the adjoint equations will also be stated. That is, the precise form of the discretization of the Euler equations as well as the precise way in which the boundary conditions are implemented are taken into account when the adjoint equations are derived. However, the precise form of the coefficients in the artificial dissipation terms is not reflected in the adjoint equations.

Copyright © 2000 by Mattias Chevalier and Martin Berggren. Published by the International Council of the Aeronautical Sciences, with permission.
This approach is just one among several ways of finding the gradient to the objective function. Using finite differences on each of the design variables is another alternative. A third alternative is to compute so-called flow sensitivities by repeatedly solve linearized versions of the equations. These alternatives are easier to implement but are computationally costly. The cost of computing the gradient from the adjoint equations has the advantage of being independent of the number of design variables.

Several authors, such as Iollo et al. [9], Narducci et al. [10], Ibrahim and Baysal [8], and Cliff et al. [4], have published works on shape optimization for the quasi-one-dimensional nozzle flow. In a very recent article, Giles and Pierce [6] also derive analytical expressions for solutions to the adjoint equations. Most of these articles concentrate on the particular difficulties that are associated with embedded shocks in the flow. In contrast to this, we limit ourselves to the case of smooth flow when deriving expressions for the gradient of the objective function. This is done to highlight the distinct features of the current investigation: the choice of boundary conditions together with the use of the quasi-discrete form of the adjoint equations.

The article is organized as follows. Section 2 introduces the governing equations and the shape optimization problem. Section 3 describes the numerical treatment of the equations involved. Section 4 presents computational results and is followed by the final discussion of section 5.

2 Theory

2.1 The shape optimization problem

The quasi-1D Euler equations for steady flow are (references Hirsch [2] and Anderson [5])

\[ f_x + \xi g = 0, \]  

where the following vector notation is introduced

\[ f = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ (\rho e + p)u \end{pmatrix}, \quad g = \begin{pmatrix} \rho u \\ \rho u^2 \\ (\rho e + p)u \end{pmatrix}, \]  

The equation of state,

\[ p = (\gamma - 1) \left( \rho e - \frac{u^2}{2} \right), \]  

closes the system. Here, \( x \) is the streamwise coordinate, \( \rho \) is the density, \( u \) is the fluid velocity, \( p \) is the pressure, \( e \) is the total energy per unit mass, and \( A \) is the area function. We use \( \gamma = 1.4 \) (air and standard conditions) for all simulations.

Perhaps the physically most natural choice of inlet boundary conditions for nozzle flow is to specify constant stagnation conditions. When the flow is subsonic at the outlet, we also need to supply a boundary condition there. That is done through the back pressure, a given constant static pressure at the outlet. With this choice of boundary conditions, the state equation reads

\[ f_x + \xi g = 0 \quad \text{in } (0, 1), \]

\[ p \left( 1 + \frac{\gamma - 1}{2} M^2 \right)^{\gamma/\gamma - 1} = p_s \quad \text{at } x = 0, \]

\[ T \left( 1 + \frac{\gamma - 1}{2} M^2 \right) = T_s \quad \text{at } x = 0, \]

\[ \text{If } M < 1 \quad p = p_{\text{out}} \quad \text{at } x = 1, \]

where \( \xi = A/A_x \) and where \( M = u/c \) is the Mach number; the speed of sound is given by the relation \( c^2 = \gamma p/\rho \). The constants \( p_s \) and \( T_s \) are the given values of the stagnation pressure and stagnation temperature respectively, and \( p_{\text{out}} \) is the given static pressure at the outlet.

Two independent set of variables, conservative, \( w \), and primitive, \( v \), will be used:

\[ w = \begin{pmatrix} \rho \\ \rho u \\ \rho e \end{pmatrix}, \quad v = \begin{pmatrix} \rho \\ u \\ p \end{pmatrix}. \]  

To exert control on the nozzle flow, the shape of the nozzle will be adjusted. The shape enters the Euler equations through \( \xi \) and the most obvious choice of design parameter is \( \xi \). Note however that we could also have used the area, \( A \), and computed \( \xi \) from \( A \).

The aim of the optimization is to force the nozzle to mimic a target distribution of some flow
quantity, in our case, the pressure. To quantify this constraint we introduce an objective function

$$I(\xi) = \frac{1}{2} \int_0^1 (p - p')^2 \, dx,$$  \hspace{1cm} (6)

where $p'$ is the target pressure distribution and $p$ is the pressure distribution computed from $\xi$ by solving the state equation.

The shape optimization problem corresponds to finding the design $\xi$, which minimizes the objective function $I$ in (6). Given a target pressure distribution $p' = p'(x)$, the problem is

Find $\xi^* \in \mathcal{U}_{ad}$ such that

$$I(\xi^*) \leq I(\xi) \quad \forall \xi \in \mathcal{U}_{ad}.$$ \hspace{1cm} (7)

The set of admissible designs is denoted $\mathcal{U}_{ad}$ and is a subset of bounded functions on $[0, 1]$. That choice for $\mathcal{U}_{ad}$ leads to a well-posed problem for the Euler equations. Cliff et al. [4, § 3] give an example of a closed and convex set of this kind.

### 2.2 The gradient

Most minimization algorithms, such as steepest descent, conjugate-gradient, and (quasi-)Newton methods, utilize gradient information. The gradient $\nabla I$ of the objective function (6) is defined through taking the directional derivative in the $\delta \xi$ direction:

$$\delta I = \langle \nabla I, \delta \xi \rangle = \lim_{s \to 0} \left| \frac{I(\xi + s \delta \xi) - I(\xi)}{s} \right|,$$ \hspace{1cm} (8)

where $\delta \xi$ is an arbitrary variation of the shape. To supply gradient information to a quasi-Newton algorithm, we use the adjoint equation approach.

In [3] we derive, by use of standard perturbation analysis applied on the objective function (6) and the state equation (4), an expression for the gradient in terms of the solution to an auxiliary problem, the adjoint equations. Here we only state the final expressions for the gradient and the adjoint equations for subsonic in- and outlet conditions:

$$\nabla I(\xi) = \psi^T g \quad \text{in } L^2(0, 1),$$  \hspace{1cm} (9)

where $\psi$ is the solution to the adjoint equations

$$-\mathbf{J}^T \psi_x + \mathbf{K}^T \psi \xi + \theta (p - p') = 0, \quad \text{in } (0, 1),$$

$$\mathbf{I}^T \mathbf{J}^T \psi = 0, \quad \text{at } x = 0,$$

$$\left( \begin{array}{c} J_1^T \\ J_2^T \end{array} \right) \psi = 0, \quad \text{at } x = 1.$$  \hspace{1cm} (10)

Here, $\mathbf{I}$ is a vector consisting of data computed from flow quantities on the boundary and total quantities assuming isentropic process. The vector $\theta$ is the pressure differentiated with respect to $w$,

$$\theta = \left( \frac{\partial p}{\partial w} \right)^T.$$

The matrices $\mathbf{J}$, $\mathbf{K}$, and $\mathbf{J}$ are defined as

$$\mathbf{J} = \frac{\partial \mathbf{f}}{\partial \mathbf{w}}, \quad \mathbf{K} = \frac{\partial \mathbf{g}}{\partial \mathbf{w}}, \quad \mathbf{J} = \frac{\partial \mathbf{f}}{\partial \mathbf{v}},$$

and $\mathbf{j}_1$, $\mathbf{j}_2$, and $\mathbf{j}_3$ are the column vectors of $\mathbf{J}$.

Note that in the case of subsonic in- and outlet conditions, the Euler equations have two downstream and one upstream characteristics. For the adjoint equations we have the opposite situation: one downstream and two upstream characteristics due to the negative sign on the Jacobian. This is consistent with the boundary conditions in (10); there are two conditions supplied at $x = 1$ and one at $x = 0$.

This all summarizes into the following procedure to compute the gradient of $I$:

1. Solve the state equation (4) given a design $\xi$.

2. Solve the adjoint equations (10) using the solution obtained above.

3. Compute the gradient from expression (9).

Note that to compute the gradient we basically have to solve only two equally expensive, measured in computer time, systems of equations regardless of how many design parameters we are using.
3 Discretization

A cell-centered finite-volume scheme is applied for the spatial discretization. The step size is constant and denoted $\Delta x = 1/N_x$ where $N_x$ is the number of cells in the domain. The stationary problem is solved by marching the corresponding non-stationary problem to steady state using a five step Runge–Kutta scheme.

The solution vectors, for the whole domain, for conservative and primitive variables are denoted $\{w_i\}_{i=1}^{N_x}$ and $\{v_i\}_{i=1}^{N_x}$, respectively. We also define

$$f_i = f(w_i),$$

$$f_{i\pm 1/2} = \frac{1}{2}(f_i + f_{i\pm1}),$$

$$g_i = g(w_i),$$

(11)

where $f$ and $g$ are the functions of (2). Integer index $i$ denotes cell-centered values and $i \pm 1/2$ denotes node-centered values.

The discrete state equation is

$$\frac{f_{i+1/2} - f_{i-1/2}}{\Delta x} + \xi_i g_i = d, \quad i = 1, \ldots, N_x,$$

$$\begin{pmatrix}
\rho_0 \\
u_0 \\
\rho_0
\end{pmatrix}
= \begin{pmatrix}
2\rho_1/2 - \rho_1 \\\n2u_1/2 - u_1 \\\n2\rho_1/2 - p_1
\end{pmatrix},$$

$$\begin{pmatrix}
\rho_{N_x+1} \\
u_{N_x+1} \\
p_{N_x+1}
\end{pmatrix}
= \begin{pmatrix}
\rho_{N_x} \\
u_{N_x} \\
2p_{\text{out}} - p_{N_x}
\end{pmatrix},$$

(12)

where the vector $\xi_h = \{\xi_i\}_{i=1}^{N_x}$ now is our design variable.

Data for the boundary conditions are supplied through the back pressure $p_{\text{out}}$ and through $\rho_{1/2}$ and $p_{1/2}$, which are computed from the isentropic assumptions:

$$\rho_{1/2} \left(1 + \frac{\gamma - 1}{2}M_1^2 \right)^{1/(\gamma - 1)} = \rho_s,$$

$$p_{1/2} \left(1 + \frac{\gamma - 1}{2}M_1^2 \right)^{\gamma/(\gamma - 1)} = p_s,$$

(13)

(14)

where $\rho_s$ and $p_s$ are the given stagnation density and pressure respectively. The symbol $d$ on the right-hand side of equation (12) represents artificial dissipation which is needed to stabilize central schemes of this kind. We use the Jameson-style combined second- and fourth-order dissipation.

Solving equation (12), we obtain grid functions like the pressure $p_i$, $i = 1, \ldots, N_x$. Therefore, it is reasonable to approximate the objective function (6) with

$$I_h(\xi) = \frac{1}{2} \Delta x \sum_{i=1}^{N_x} (p_i(\xi) - p_i')^2,$$

(15)

where $\{p_i\}_{i=1}^{N_x}$ is obtained from the finite-volume solution below and $p_i' \approx p_i'((i - 1/2)\Delta x)$ approximates the target pressure.

The discrete counterpart to optimization problem (7) is

Find $\xi^*_h \in \mathcal{U}_d$, $i = 1, \ldots, N_x$, such that

$$I(\{\xi^*_i\}_{i=1}^{N_x}) \leq I(\{\xi_i\}_{i=1}^{N_x}) \quad \forall \xi_i \in \mathcal{U}_d.$$

(16)

3.1 Gradient computations in the discrete case

One approach to compute the objective-function gradient is to directly discretize the adjoint equations (10) and the gradient expression (9). However, once discretizations of the Euler equations and the objective function are selected, this implicitly defines the discrete adjoint equations from which we obtain the expression for the exact gradient of the discretized objective function. This “discrete” adjoint equation may not coincide with a straight-forward discretization of equation (10), particularly not in the implementation of the boundary conditions. Using this discrete gradient minimizes numerical errors in the gradient directions. This may be important since highly accurate gradient directions are typically needed in the quasi-Newton algorithm. However, note that artificial dissipation needs to be added in the adjoint equation for stability and these effects are not taken into account in the derivation.

The discretized equations are derived in detail in [3]; here we merely state the results. The
discrete gradient is
\[ \nabla I_h = \Delta x \{ \psi^T \mathbf{g}_i \}_{i=1}^{N_x}, \]  
where \( \{ \psi^T \}_{i=1}^{N_x} \) is the solution to the discrete adjoint equations
\[ -\mathbf{J}_i^T \left( \psi_{i+1} - \psi_{i-1} \right) \frac{1}{2 \Delta x} + \xi_i \mathbf{K}_i^T \psi_i = - \theta_i (p_i - p'_i), \]
\[ \mathcal{I}_i \psi_0 = \mathcal{M}_i \psi_1, \]
\[ \mathcal{I}_r \psi_{N_x+1} = \mathcal{M}_r \psi_{N_x}, \]  
where \( i = 1, \ldots, N_x \). The matrices \( \mathcal{I} \) and \( \mathcal{M} \), defined and derived in [3], consist of surprisingly complicated algebraic combinations of flow data at and around the boundaries.

4 Numerical experiments

For the optimization we used the limited-memory quasi-Newton algorithm of Byrd et al. [1], publicly available at Netlib/toms/778. Two flow cases are considered, defined by the boundary data in table 1. These correspond to a fully subsonic and a shock-free transonic case, respectively, using the area function of figure 1. This particular area function is obtained from a cubic polynomial in \( \xi \) using the coefficients in the second column of table 2. A 200 grid-point mesh is used in all reported experiments.

Table 1 Boundary data for the simulated flow types. All quantities are given in SI-units.

<table>
<thead>
<tr>
<th>Flow</th>
<th>( p_s )</th>
<th>( T_s )</th>
<th>( \rho_s )</th>
<th>( p_{out} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subsonic</td>
<td>2000000</td>
<td>300</td>
<td>2.32</td>
<td>174488</td>
</tr>
<tr>
<td>Transonic</td>
<td>2000000</td>
<td>300</td>
<td>2.32</td>
<td>51159</td>
</tr>
</tbody>
</table>

As a first test, we define a target pressure from solving the Euler equation with the area function of figure 1 and the subsonic data of table 1. Then the coefficients in the polynomial defining \( \xi \) are perturbed (table 2), and we attempt to recover the target area function by solving the optimization problem (16). This problem is solved using two different parameterizations of the design variable \( \xi \): (i) the coefficients in a cubic polynomials, and (ii) the value of \( \xi \) at each grid point. The dimension of the design space is 4 in the first case and 200 in the second. Note that the computational effort needed for each quasi-Newton iteration is essentially the same in the two cases, since the adjoint-equation approach is used.

Figure 2 shows the value of the objective function and the norm of the gradient versus iteration number. The convergence behavior for the two different parameterizations of the design are similar for about the 25 first iterations, after which the 4-degrees-of-freedom case appears to enter a region of superlinear convergence for the quasi-Newton method. Figure 3 compares the pressure distribution using the different parameterizations at a few stages in the optimization.

The second test is the same as the first, ex-
Table 2 Initial and final values of the polynomial coefficients $\alpha$ defining $\xi$. Here, $k$ denotes the degree of corresponding monomial.

<table>
<thead>
<tr>
<th>$k$</th>
<th>Initial $\alpha$</th>
<th>Target $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.9474</td>
<td>-0.8574</td>
</tr>
<tr>
<td>1</td>
<td>1.1376</td>
<td>1.2376</td>
</tr>
<tr>
<td>2</td>
<td>1.7380</td>
<td>1.5980</td>
</tr>
<tr>
<td>3</td>
<td>-1.4525</td>
<td>-1.3525</td>
</tr>
</tbody>
</table>

Except that the transonic boundary data of table 1 are used instead. Figure 4 shows the convergence behavior for the different parameterizations and figure 5 compares the pressure distributions at different stages in the optimization. The convergence behavior is similar to the subsonic case, but the superlinear convergence appears later, after about 35 iterations.

The tests above may give the impression that parameterizing with low-order polynomials is better than using many degrees of freedom for the design variables. That this is not at all the case in general is demonstrated in the next test. Note that the cases above use target pressures that are reachable by a cubic $\xi$, that is, one particular cubic yields exactly the target pressure. (This means also that the objective function is zero at the optimum.) In a third test, we picked a target pressure distribution which is (most likely) not reachable by any $\xi$, cubic or not. Figure 7 shows the convergence behavior when using 4 and 200 design variables, respectively. Note that both the convergence behavior and the final value of the objective function is better when using a higher degree of the design space. Figure 8 depicts the target pressure and the pressure distribution at different stages in the optimization.

In a last test case, using data from the first test problem above with 200 design variables, we test the influence of the boundary conditions at $x = 0$ for the adjoint equation. We compare the use of the “exact” form (18), derived by exact transposition of corresponding boundary conditions of equation (12) with one implementation in which the boundary condition at $x = 0$ in equation (10) are supplemented with “numerical” boundary conditions: extrapolation of two of the variables in $\psi$. From figure 6, we see that the difference in the convergence between using these approaches is surprisingly small, considering the elements of arbitrariness in the second approach (which variables should be extrapolated, e.g.?). Inaccuracies in the gradient direction will certainly be introduced in the second approach. That this does not degrade the convergence rate more than indicated in figure 6 somewhat contradicts the experience of the authors from other studies ([7]), in which the convergence rate of a related optimization problem was sensitive to small inaccuracies introduced in the gradient directions.
There are at least three possible reasons for this. The cases in which we have formerly noted significant effects on the accuracy of the gradient by changes in the implementation have all concerned parabolic or elliptic state equations. The objective function in such cases is often quite insensitive to small changes in the control (or design), which means that precise gradient information is crucial since the objective function will hardly decrease otherwise. In the present case, the state equation is hyperbolic, and the objective function is quite sensitive to small changes in the design. Thus, precise gradient information may be less important in this case, since even a slightly off gradient direction may greatly reduce the objective function.

A second reason could be that other inaccuracies dominate. For instance, we do not consider the exact form of the artificial dissipation of the state equation when deriving the adjoint equations.

A third reason could be effects of the zeroth-order extrapolation used for the Mach number in defining the pressure and density at $x = 0$ (expression (14)). This affects the solution in the form of small, local “kinks” close to $x = 0$. We observed no spread of these disturbances downstream in the solution to the state equation. How-

---

**Fig. 3** Pressure distributions for polynomial (upper) and full (lower) descriptions of $\xi$ at 1 (blue/solid), 3 (green/dashed), and 10 (red/dash-dotted) iterations for subsonic flow.

**Fig. 4** The objective function (upper) and the gradient (lower) as functions of iteration number for both polynomial (green/dash-dotted) and full (red/solid) description of $\xi$ for transonic flow.
However, since the boundary conditions for the adjoint equation are derived from the actual boundary conditions used in the state equation, it may well be that the effects on the adjoint equations of the zeroth-order extrapolation in the state equation is significant. Evidence for this claim is that we noted oscillations in the solution of the adjoint equations originating at the boundary \( x = 0 \). This could cause an increase of the conditioning of the discrete optimization problem (16) of purely numerical origin.

Fig. 5  Pressure distributions for polynomial (upper) and full (lower) descriptions of \( \xi \) at 1 (blue/solid), 3 (green/dashed), and 10 (red/dash-dotted) iterations for transonic flow.

Fig. 6  The objective function as a function of iteration number for the case when using “exact” boundary conditions (green/dash-dotted) and when using “numerical” boundary conditions (red/solid) for the adjoint equations.

5 Conclusions and outlook

We have derived, implemented and tested a quasi-discrete form of the adjoint equations to the quasi-1D Euler equations for nozzle flow in order to compute gradients in a shape optimization procedure. Physically relevant boundary conditions are used. For the adjoint equations, we apply corresponding boundary conditions, derived by transposing the exact form in which the boundary conditions for the Euler equations are implemented.

The gradient computed in this way could be successfully used in an optimization procedure to recover a reachable pressure distribution as well as finding area functions that yield a pressure distribution that well approximates a nonreachable pressure distribution. We demonstrated this in the subsonic as well as the transonic, shock-free regime. Numerical experiments, not reported here, were also performed for cases with embedded shocks. These cases worked surprisingly well, considering that the important effects of the artificial dissipation in the vicinity of the shock were not considered at all in the gradient derivation. However, the convergence rate of the opti-
mization algorithm was not as good as in the reported cases. To implement a dissipation mechanism in the adjoint equations in a similar “discrete” way as the boundary conditions studied here is an obvious, but nontrivial candidate for a next stage in the development.

As discussed in section 4, we noted some oscillations in the solution to the adjoint equations originating at the boundary \( x = 0 \). We conjecture that the cause of this is the zeroth-order extrapolation used for the Mach number in the state equation at \( x = 0 \). Using instead a first-order extrapolation of the Mach number, the local “kinks” in the vicinity of the boundary can be avoided altogether. It would be interesting to see if this also improves the smoothness of the adjoint solution and the conditioning of the optimization problem. However, the derivation of the corresponding adjoint boundary condition is complicated and tedious.
Acknowledgments

Part of the research was done when the first author was a student for Professor Roland Glowinski which is greatfully acknowledged.

References


