

Analytic Hessian derivation for the quasi-one-dimensional Euler equations

Eyal Arian^a, Angelo Iollo^{b,*}

^a Computational Mathematics, The Boeing Company, MC 7L-21 P.O. Box 3707, Seattle, WA 98124-2207, USA

^b Institut de Mathématiques de Bordeaux, UMR 5251 CNRS, Université Bordeaux 1, INRIA Bordeaux Sud-Ouest, Equipe-Projet MC2, 33405 Talence cedex, France

ARTICLE INFO

Article history:

Received 6 March 2008

Received in revised form 8 September 2008

Accepted 24 September 2008

Available online 10 October 2008

Keywords:

Hessian

Adjoint

Compressible Euler equations

ABSTRACT

The Hessian for the quasi-one-dimensional Euler equations is derived. A pressure minimization problem and a pressure matching inverse problem are considered. The flow sensitivity, adjoint sensitivity, gradient and Hessian are calculated analytically using a direct approach that is specific to the model problems. For the pressure minimization problem we find that the Hessian exists and it contains elements with significantly larger values around the shock location. For the pressure matching inverse problem we find at least one case for which the gradient as well as the Hessian do not exist. In addition, two formulations for calculating the Hessian are proposed and implemented for the given problems. Both methods can be implemented in industrial applications such as large scale aerodynamic optimization.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Aerodynamic optimization problems are computationally massive and highly ill-conditioned. In the past two decades they have matured and are now an integral part of design in the aerospace industry. Mathematically these are optimal design problems governed by partial differential equations. In this work we assume that the initial guess of the shape design problem is in the basin of attraction of the desired global minimum, and therefore local gradient methods can be used to pursue the optimal solution. Jameson [1] introduced the idea of using Computational Fluid Dynamics (CFD) to solve the direct and adjoint flow models. By the adjoint method, the solution of a single adjoint problem allows to determine the gradient of an objective function with respect to the design parameters, regardless of the dimension of the design space. Beux and Dervieux [2] first gave examples of compressible flow shape derivatives computed by the adjoint of the discrete operator on unstructured meshes. Frank and Shubin [3] considered several shape optimization strategies and took as a test case the quasi-one-dimensional compressible Euler equations. Iollo and Salas [4] studied the case of flows with shocks. They found that a boundary condition at the shock is needed for the adjoint equation. Giles and Pierce [5] determined this boundary condition and the analytical adjoint solution for the quasi-one-dimensional compressible flow in a nozzle. Additional references and a more complete review can be found in [6].

In most practical scenarios quasi-Newton methods are used in combination with the adjoint method (or sensitivity equation method) to compute the search direction. However, for large scale ill-conditioned problems quasi-Newton methods are slow at the start, and may take many iterations to build a good Hessian approximation. In practice it may not be affordable to allow enough iterations for the quasi-Newton to reach a good Hessian approximation, resulting in poor convergence. An approximation of the Hessian that is constructed from the governing equations and the objective function, can serve as a starting point for quasi-Newton methods. Such an initial approximation will be much more accurate than the identity matrix

* Corresponding author.

E-mail address: iollo@math.u-bordeaux1.fr (A. Iollo).

that is used by default in practice. Also, other preconditioning techniques (not based on quasi-Newton) can be devised based on the Hessian form.

A good Hessian approximation is important in particular for large scale industrial optimal design problems where the governing equations are solved using CFD. The Hessian for such problems is typically highly ill-conditioned as analyzed for example by Arian and Ta'asan in [7]. The gradient serves as a very poor search direction, while the gradient pre-multiplied by the approximated inverse of the Hessian serves as an excellent search direction close enough to the solution as demonstrated by Arian and Vatsa in [8]. However, the techniques used in [7] and [8] are not practical in the industrial setting in which the governing state equations are more difficult to analyze. In practice we would like to use the available resources, such as sensitivities and adjoint solutions, to generate an approximated Hessian numerically. Another approach is to use automatic differentiation tools to generate the Hessian in the discrete level automatically (more details can be found in [9,10]).

In this paper we compute the Hessian analytically for two minimization problems governed by the quasi-one-dimensional Euler equations. An analytical form may reveal many unknowns on such problems, e.g., does the Hessian exist for transonic flow with a shock? What is the effect of the shock on conditioning? We examine three different methods to build the Hessian. The first method, called “the direct approach” in this paper, is specific to the problem at hand. Using the direct approach the analytical Hessian is derived for a transonic flow with a shock. The two additional methods are based on the adjoint and sensitivity equations, and can be applied in practice for large scale aerodynamic optimization problems. We demonstrate the application of these methods for supersonic flow and show numerically that all three methods result in the same Hessian matrix (up to a small truncation error). The two adjoint based methods have different computational cost. A preliminary analysis shows that one method might be numerically unstable when shocks are present in the flow. The method that is likely to be more stable is also more computationally intensive.

Our point of departure is the paper by Giles and Pierce [5]. In that paper the adjoint solution is derived analytically using Green's function technique. We recover the adjoint solution in [5] and in addition compute the sensitivity solutions.

We present results for two different objective functions. The first is a standard pressure minimization objective function as considered in [5]. The second corresponds to an inverse problem in which the difference in L^2 norm between the pressure and a given, target, pressure distribution is to be minimized.

The paper is organized as follows:

In Section 2 the mathematical formulation is presented.

In Section 3 the direct approach is discussed and key observations are made regarding the existence of the Hessian in the presence of a shock.

In Section 4 computational results of the direct method are discussed.

In Section 5 the two adjoint based methods are derived.

In Section 6 numerical results of the adjoint based methods are presented for supersonic flow conditions.

In Section 7 the paper is concluded with a discussion.

2. Formulation

The state variable, U , is defined to be the triplet composed of the density, mass flux, and total energy density:

$$U = (\rho, \rho q, \rho E)^T.$$

We consider the compressible Euler equations in a quasi-one-dimensional approximation. The flow takes place inside a nozzle of height $h(x)$, where x is the spatial coordinate (see Fig. 1). The optimal design problem consists of minimizing an objective function, $J(U)$, subject to state equations:

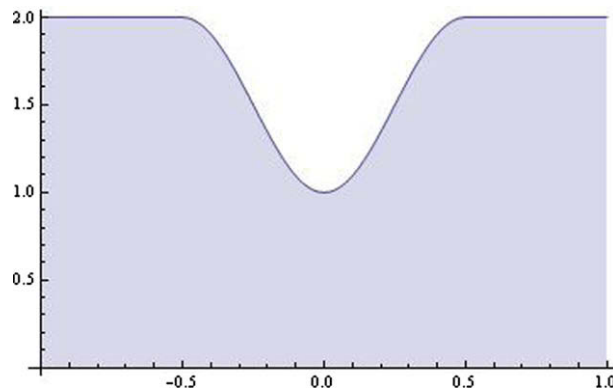


Fig. 1. The unperturbed nozzle.

$$\begin{aligned} \min_{h(x)} J(U) \\ R(U, h) = 0 \end{aligned} \quad (1)$$

with

$$R(U, h) = \frac{d}{dx}(hF) - \frac{dh}{dx}P, \quad (2)$$

and

$$F = (\rho q, \rho q^2 + p, \rho Hq)^T, \quad P = (0, p, 0),$$

where p is the static pressure, and ρH is the total enthalpy density.

In this paper the objective function is defined to be either the integral of the pressure distribution across the domain $\Omega \subset \mathbb{R}$

$$J(U) = \int_{\Omega} p dx. \quad (3)$$

or the L^2 norm of a pressure difference with respect to a reference pressure distribution p^* :

$$J(U) = \int_{\Omega} \frac{1}{2} (p - p^*)^2 dx. \quad (4)$$

These two objective functions are paradigms of typical aerodynamic optimization problems such as drag minimization and inverse design.

2.1. The perturbed problem

The gradient and Hessian are calculated by first and second order perturbations, respectively, to the non-linear problem at a given design point. The perturbation of the original minimization problem results in the following linear sub-problem:

$$\begin{aligned} \min_{\tilde{h}} J(U + u) \\ Lu - f = 0 \end{aligned} \quad (5)$$

where U is given, u is a perturbation to U , $f = f(\tilde{h})$, and \tilde{h} is perturbation of h .

The linear operator, L , has the following explicit form:

$$Lu = \frac{d}{dx}(hAu) - \frac{dh}{dx}Bu,$$

where $A = \partial F / \partial U$ and $B = \partial P / \partial U$. The right hand side of the linearized problem has the following explicit form:

$$f = \frac{d\tilde{h}}{dx}P - \frac{d}{dx}(\tilde{h}F). \quad (6)$$

The boundary conditions for the above perturbation equations, $Lu - f = 0$, depend on the boundary conditions of the unperturbed problem.

2.1.1. Boundary conditions

At the inlet, if the unperturbed problem is subsonic, total pressure and total enthalpy are fixed, and therefore their perturbation boundary condition is homogeneous. If the inlet is supersonic all the unperturbed flow variables are fixed, and therefore their perturbation is homogeneous. For a subsonic outlet, static pressure is imposed and hence its perturbation is homogeneous. For a supersonic outlet there are no boundary conditions on the perturbed variables.

2.1.2. The shock case

When a shock is present in the flow field, the above derivation is formal since U , A , and B are discontinuous at the shock. In this case, however, Bardos and Pironneau [11,12] showed that the linearized state equation in (5) is well defined in a generalized function setting. Moreover, their results coincide with those obtained by Giles and Pierce [5] splitting the domain in two regions where the solution is regular, i.e., $\Omega = \Omega^- \cup \Omega^+$ and $x_{sh} = \Omega^- \cap \Omega^+$. We will also limit ourselves to this framework, and hence the following derivation will apply before and after the shock, i.e., we will not address the issue of rigorously proving that the second order sensitivity equations are well defined in the distribution sense in case the domain is not split.

2.2. Finite dimension design space

We choose to solve the problem in a finite design space. The channel height, $h(x)$, is composed of a seed height function, $h_0(x)$, perturbed by a sum of N fixed shape basis functions, $h_i(x)$, with the coefficients, α_i , serving as the design variables:

$$h(x) = h_0(x) + \sum_{i=1}^N \alpha_i h_i(x).$$

The channel's seed shape, $h_0(x)$ is depicted in Fig. 1. We choose cubic B-Splines for the basis, $h_i(x)$:

$$h_i(x) = B_i^3(x).$$

Cubic B-Splines are widely used in applications and specifically in aerospace engineering. Also, for this study it will be useful to localize the effect of the shock wave on the Hessian and cubic B-Splines are local functions. It will be shown in the following section that the perturbation of the shape must be regular enough at the shock location for the gradient and the Hessian to exist (see Eqs. (17) and (18)). B-splines satisfy these regularity requirements.

3. Gradient and Hessian via the direct approach

Let us consider flow without shocks. For the pressure functional (3) we have

$$\frac{dJ}{d\alpha_i} = \int_{\Omega} \frac{\partial p}{\partial U} \frac{dU}{d\alpha_i} dx, \tag{7}$$

and

$$\frac{d^2J}{d\alpha_i d\alpha_j} = \int_{\Omega} \frac{\partial^2 p}{\partial U^2} \frac{dU}{d\alpha_i} \frac{dU}{d\alpha_j} dx + \int_{\Omega} \frac{\partial p}{\partial U} \frac{d^2U}{d\alpha_i d\alpha_j} dx. \tag{8}$$

Once the quantities $\frac{dU}{d\alpha_i}$ and $\frac{d^2U}{d\alpha_i d\alpha_j}$ are known, then first and second derivatives of the functional can be computed.

From the governing equations the sensitivity equation is easily derived ($L \equiv \frac{\partial R}{\partial U}$),

$$\frac{\partial R}{\partial U} \frac{dU}{d\alpha_i} + \frac{\partial R}{\partial h} h_i = 0, \tag{9}$$

which can be solved with homogeneous boundary conditions to get the sensitivities $dU/d\alpha_i$. Differentiating equation (9) with respect to α_j , the second order sensitivity equation is obtained:

$$\frac{\partial^2 R}{\partial U^2} \frac{dU}{d\alpha_i} \frac{dU}{d\alpha_j} + \frac{\partial R}{\partial U} \frac{d^2U}{d\alpha_i d\alpha_j} + \frac{\partial^2 R}{\partial h^2} h_i h_j = 0. \tag{10}$$

Once Eq. (9) is solved, we can solve equation (10) for $\frac{d^2U}{d\alpha_i d\alpha_j}$ using homogeneous boundary conditions.

Given N design variables, N sensitivity equations (9) need to be solved to obtain the gradient, and $\frac{1}{2}N(N + 1)$ Eq. (10) to obtain the Hessian of the objective function. With the methods presented in Section 5, $O(N)$ equations need to be solved to get the Hessian.

Note that Eqs. (9) and (10) are linear partial differential equations which differ only by the right hand side.

In the specific case of the quasi-one-dimensional Euler equation, $R(U, h) = 0$, with R defined in Eq. (2), it is not necessary to solve the sensitivity equations to get first and second derivatives of the solution with respect to the design variables (Section 5 treats the general case). For that specific case the solution is only a function of h , if the mass flow and the boundary conditions are fixed. Therefore for the model problem at hand the sensitivities, as well as the gradient and Hessian can be obtained directly (or explicitly as a function of h) as explained in detail in the following.

3.1. Sensitivities via the direct approach

Let us consider for example a subsonic flow through the nozzle. In this case the boundary conditions usually prescribed are total pressure p^0 , as well as total enthalpy H at the inlet and static pressure at the outlet. The flow solution is obtained by expressing all the variables as a function of the Mach number M and observing that the mass flow \dot{m} is conserved through the nozzle.

Specifically, the solution $U = (\rho, \rho q, \rho E)^T$ depends explicitly on the Mach number M as follows:

$$\begin{aligned} p(M) &= p^0 / (1 + (\gamma - 1)/2M^2)^{\gamma/(\gamma-1)} \\ \rho(M) &= (\gamma p(M)M^2 + 2\gamma/(\gamma - 1)p(M)) / (2H) \\ q(M) &= M \sqrt{\gamma p(M) / \rho(M)} \\ E(M) &= H - p(M) / \rho(M) \end{aligned} \tag{11}$$

Given the static pressure at the exit we can determine the Mach number at the exit and therefore if h is given \dot{m} is fixed, as can be shown from the following thermodynamic relation:

$$\dot{m} = \frac{\gamma p^0}{\sqrt{(\gamma - 1)H}} f(M) h, \tag{12}$$

where γ denotes the specific heat ratio and

$$f(M) = M \left(1 + \frac{\gamma - 1}{2} M^2 \right)^{-\frac{\gamma+1}{2(\gamma-1)}}.$$

Using Eq. (12) the solution in all regions can be determined.

From here on we assume that the mass flow is a constraint of the optimization problem and therefore it is constant with respect to perturbations of the design variables,

$$\dot{m} = \text{const.} \tag{13}$$

This amounts to fixing $h(x)$ at the exit when the flow is subsonic. When the flow is supersonic three boundary conditions are given at the inlet, usually p^0 , H and M . Therefore requiring that the mass flow is constant amounts to requiring that $h(x)$ is fixed at the inlet. In the transonic case (with or without shocks) it amounts to fixing the throat section.

For a fixed mass flow and fixed exit pressure the solution does not depend explicitly on the design variables, i.e., $U = U(M(h))$ (see Eq. (11)), and hence

$$\frac{\partial U}{\partial \alpha_i} = 0. \tag{14}$$

Therefore,

$$\frac{dU}{d\alpha_i} = \frac{dU}{dM} M'(h) h_i. \tag{15}$$

Notice that also in the case of transonic flow with shocks, Eq. (14) implies that the solution can be computed as an explicit function of $h(x)$. Hence we can take advantage of that result to obtain the sensitivities directly. From Eq. (12) the derivative of M with respect to h is given by

$$M'(h) = -\frac{f(M)}{h f'(M)}. \tag{16}$$

When a shock is present in the flow field the total pressure is not constant across the shock. However, if the mass flow is constant and the exit pressure is fixed, the jump in total pressure across the shock is constant. Therefore Eq. (12) holds on both sides of the shock provided that one uses the correct value of the total pressure, as determined by the Rankine-Hugoniot conditions. As a consequence, also in the case of a shocked flow the solution only depends on h .

Let us denote h_{sh} the height of the channel at the shock location. This quantity does not depend on α_i from Eq. (14). Let x_{sh} be defined such that $h_{sh} = h(x_{sh})$, we obtain

$$\frac{dx_{sh}}{d\alpha_i} = -\left[\frac{h_i(x)}{h'(x)} \right]_{sh}, \tag{17}$$

and

$$\frac{d^2 x_{sh}}{d\alpha_i d\alpha_j} = \left[\frac{(h_i(x) h_j(x))'}{h'(x)^2} - \frac{h''(x)}{h(x)^3} h_i(x) h_j(x) \right]_{sh}. \tag{18}$$

Key observation

From the above equations we see that both the original geometry as well as the basis functions must be differentiable at the shock for the derivatives of the shock location with respect to the design variables to exist.

3.2. Gradient and Hessian for the shockless case

We are now in the position to explicitly compute the derivative of J with respect to the design variables. In the shockless cases we find

$$\frac{dJ}{d\alpha_i} = \int_{\Omega} \frac{dp}{dh} (h(x)) h_i(x) dx, \tag{19}$$

and

$$\frac{d^2 J}{d\alpha_i d\alpha_j} = \int_{\Omega} \frac{d^2 p}{dh^2} (h(x)) h_i(x) h_j(x) dx. \tag{20}$$

Fig. 2 depicts $dU/dh = dU/dM M'(h)$ for a transonic flow through the nozzle shown in Fig. 1. The density (solid) and total energy (dash-dotted) sensitivities are singular at the throat, while the mass flux (dashed) sensitivities are continuous.

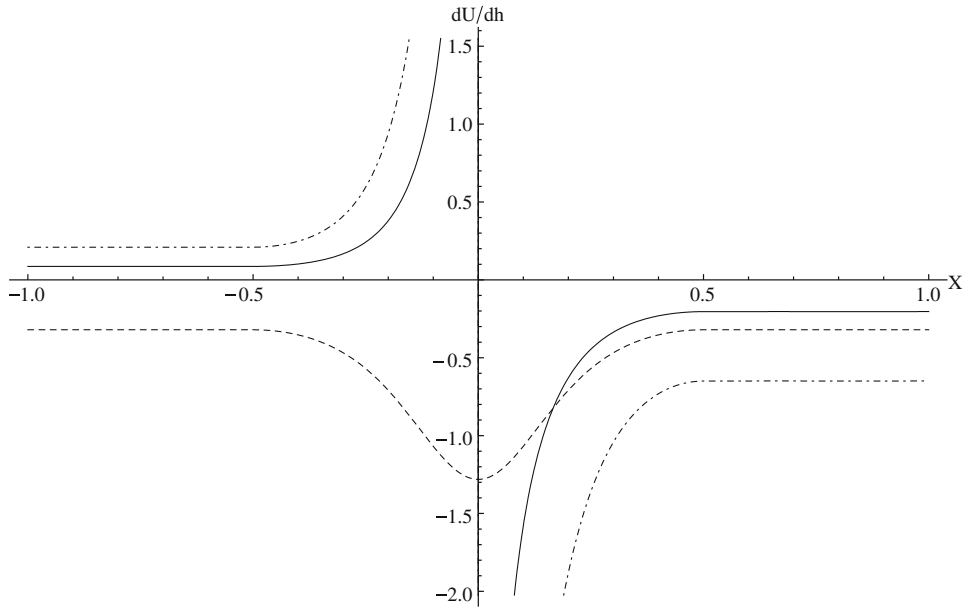


Fig. 2. Sensitivity solution for transonic flow. The solid line represent the density sensitivity, the dashed is the mass flux sensitivity, and the dash-dotted is the total energy density sensitivity.

3.3. Gradient and Hessian for the shock case

When shocks are present in the domain we take $\Omega = \Omega^- \cup \Omega^+$ such that $x_{sh} = \Omega^- \cap \Omega^+$. In this case, the objective function (3) can be written as follows:

$$J = \int_{\Omega^-} p^-(h(x))dx + \int_{\Omega^+} p^+(h(x))dx,$$

where $p^+(h(x))$ and $p^-(h(x))$ indicate the subsonic and supersonic pressure function to the right and to the left of the shock respectively.

The function $p^-(x > x_{sh})$ can be interpreted as the extension of the supersonic pressure after x_{sh} , while the function $p^+(x < x_{sh})$ as the extension of the subsonic pressure before x_{sh} . These two functions are continuous at the shock. Let us define

$$J^-(\omega) = \int_a^\omega p^-(h(x))dx \quad \text{and} \quad J^+(\omega) = \int_\omega^b p^+(h(x))dx,$$

where $\Omega = [a, b]$, $a < x_{sh} < b$, and $a < \omega < b$. The integrals $J^-(\omega)$ and $J^+(\omega)$ are smooth functions of ω for the pressure minimization problem. For the case of inverse design with a target pressure distribution that contains a shock, this argument does not always hold, as will be discussed in the next subsection.

The gradient is given by

$$\frac{dJ}{d\alpha_i} = \int_{\Omega^-} \frac{dp^-}{dh} h_i(x)dx + \int_{\Omega^+} \frac{dp^+}{dh} h_i(x)dx - (p^+(h(x_{sh})) - p^-(h(x_{sh}))) \frac{dx_{sh}}{d\alpha_i}. \tag{21}$$

The Hessian is obtained by taking the partial derivative with respect to α_j of Eq. (21). When taking the derivative of the jump term we make use of the fact that the shock intensity does not change with the design variables, $p^+(h(x_{sh})) - p^-(h(x_{sh})) = const$, although the shock location does move, therefore,

$$\begin{aligned} \frac{d^2J}{d\alpha_i d\alpha_j} &= \int_{\Omega^-} \frac{d^2p^-}{dh^2} h_i(x)h_j(x)dx + \int_{\Omega^+} \frac{d^2p^+}{dh^2} h_i(x)h_j(x)dx \\ &\quad - \left(\frac{dp^+}{dh}(h(x_{sh})) - \frac{dp^-}{dh}(h(x_{sh})) \right) h_i(x_{sh}) \frac{dx_{sh}}{d\alpha_j} - (p^+(h(x_{sh})) - p^-(h(x_{sh}))) \frac{d^2x_{sh}}{d\alpha_i d\alpha_j}. \end{aligned} \tag{22}$$

3.4. The inverse design problem

Let us now consider the objective function (4). The above formulas for the gradient and Hessian hold with minor modifications:

$$\frac{dJ}{d\alpha_i} = \int_{\Omega^-} (p^-(h(x)) - p^*(x)) \frac{dp^-}{dh} h_i(x) dx + \int_{\Omega^+} (p^+(h(x)) - p^*(x)) \frac{dp^+}{dh} h_i(x) dx - \frac{1}{2} \left((p^+(h(x_{sh})) - p^*(x_{sh}))^2 - (p^-(h(x_{sh})) - p^*(x_{sh}))^2 \right) \frac{dx_{sh}}{d\alpha_i} \tag{23}$$

$$\begin{aligned} \frac{d^2J}{d\alpha_i d\alpha_j} &= \int_{\Omega^-} \left(\left(\frac{dp^-}{dh} \right)^2 + (p^-(h(x)) - p^*(x)) \frac{d^2p^-}{dh^2} \right) h_i(x) h_j(x) dx \\ &+ \int_{\Omega^+} \left(\left(\frac{dp^+}{dh} \right)^2 + (p^+(h(x)) - p^*(x)) \frac{d^2p^+}{dh^2} \right) h_i(x) h_j(x) dx \\ &- \left((p^+(h(x_{sh})) - p^*(x_{sh})) \frac{dp^+}{dh}(h(x_{sh})) - (p^-(h(x_{sh})) - p^*(x_{sh})) \frac{dp^-}{dh}(h(x_{sh})) \right) h_i(x_{sh}) \frac{dx_{sh}}{d\alpha_j} \\ &- \frac{1}{2} \left((p^+(h(x_{sh})) - p^*(x_{sh}))^2 - (p^-(h(x_{sh})) - p^*(x_{sh}))^2 \right) \frac{d^2x_{sh}}{d\alpha_i d\alpha_j} \end{aligned} \tag{24}$$

The target pressure distribution p^* is in principle arbitrary. In the case that it is discontinuous at x_{sh} , then the functional integrands $(p^-(h(x)) - p^*(x))^2$ and $(p^+(h(x)) - p^*(x))^2$ contain a jump at the shock. Therefore the integrals

$$J^-(\omega) = \int_a^\omega \frac{1}{2} (p^-(h(x)) - p^*(x))^2 dx \quad \text{and} \quad J^+(\omega) = \int_\omega^b \frac{1}{2} (p^+(h(x)) - p^*(x))^2 dx,$$

have a discontinuous derivative at $\omega = x_{sh}$. As a result both the gradient and the Hessian do not exist at that point. An example of such a situation is given in Section 4.2.

4. Computational results using the direct method

The flow conditions for all the cases below are $p^0 = 2, H = 4$ at the inlet. For the cases with a shock $p = 1.1$ at the outlet. The geometry is defined by $\Omega = [-1, 1]$ and $h_0(x) = 2$ for $-1 \leq x < -1/2$, $h_0(x) = 1 + \sin^2(\pi x)$ for $-1/2 \leq x \leq 1/2$, and $h_0(x) = 2$ for $1/2 < x \leq 1$. With these conditions a shock is present in the flow field at $x_{sh} = 0.3941$, corresponding to a channel height of 1.893.

The B-splines basis functions are uniformly distributed starting after the throat and in the diverging part of the nozzle. We denote by N the number of collocation points in the interval $[0.1, 0.5]$. The support of the basis functions is compact, $h_i(x)$ is non-zero only between the collocation point $i - 2$ and $i + 2$. This allows the study of the effect induced by localized geometrical perturbations on the flow.

4.1. Pressure minimization

We compute the condition number of the Hessian for the pressure minimization objective function (3) as a function of the number of design variables. We consider three cases: a subsonic flow, a transonic flow with supersonic exit, and a transonic shocked flow, see Fig. 3. In the subsonic case $p = 1.98$ at the outlet, whereas in the transonic case the nozzle is adapted. In the supersonic case $M = 3$ at the inlet.

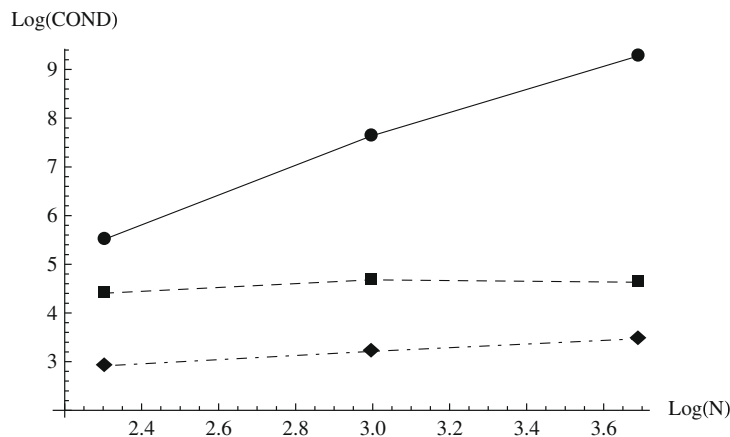


Fig. 3. Hessian Condition number (Cond) as a function of design space size (N). Subsonic case (dot-dashed), transonic (dashed), shocked (solid).

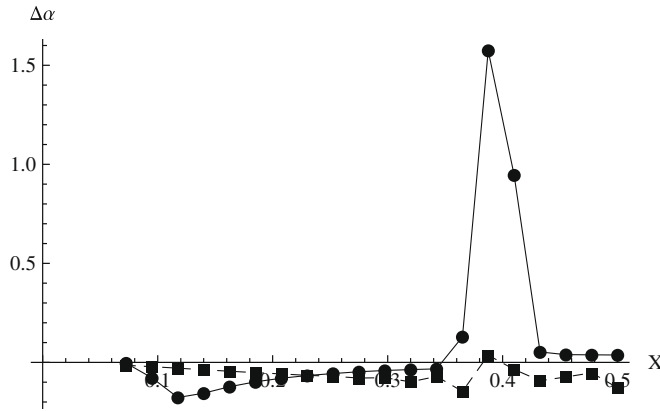


Fig. 4. Pressure minimization problem. The solid line (circles) represents the gradient, the dashed (squares) represents the gradient pre-multiplied by the inverse of the Hessian.

For the shocked case the condition number increases significantly as the number of design variables increases, whereas for the subsonic or the transonic cases it stays almost unchanged. In the presence of a shock, larger and larger numbers are obtained close to the Hessian diagonal as the number of design variables increases (see Fig. 5) resulting in a larger condition number. This behavior is due to the fact that the splines become more and more localized, steeper (non-differentiable at the limit), and therefore the large values around the diagonal at the shock location is a direct consequence of Eq. (18). For the shocked case, in Fig. 4, we show the gradient as well as the gradient pre-multiplied by the inverse of the Hessian as a function of the corresponding B-spline collocation point. The gradient curve is proportional to the shape update ($\Delta\alpha$) in a steepest descent optimization method. On the other hand, the plot of the gradient pre-multiplied by the inverse of the Hessian inverse represents the correction to a given shape in a Newton's method. The two curves are strikingly different.

4.2. The inverse problem

We calculate the Hessian for the inverse problem consisting in minimizing the objective function defined in Eq. (4). The minimum is given by the values of the design variables that were used to compute the target pressure p^* . In this case the target pressure is that relative to the unperturbed geometry $h_0(x)$. Hence, in addition to what was shown for the previous case, we are able to show the plot of the error, i.e., the difference between the present value of the design variable and the value corresponding to the minimum.

The number of design variables considered is $N = 20$. The position of the shock can be perturbed by shape function number 4, 5, and 6 (due to their collocation). Indeed, in Fig. 5 we can observe the presence of large values on the 4th, 5th and 6th columns. In Fig. 6 one can see that the inverse of the Hessian gives an appropriate correction to the gradient by shifting the curve's peak to the left, the location of the only non-zero component of the error. In fact, we have only perturbed shape function 6.

Finally, we study the case where the target pressure distribution p^* has a shock located at an identical location as the shock of the pressure, p , for the current geometry. As discussed earlier the objective function integrand does not have continuous prolongations from the left and from the right at the shock location, and hence the objective function is not differentiable with respect to those shape functions that can perturb the shock. We took 10 B-spline evenly distributed between $x = 0.05$ and $x = 1/2$. The pressure at the exit is $p_{ex} = 1.1$ and the reference pressure $p^*(x)$ corresponds to the unperturbed nozzle shape. Under these conditions the shock is again located at $x_{sh} = 0.3941$. In Fig. 7 we show the functional value when α_2 is varied between -0.01 and 0.01 . This design variable does not perturb the shock. In contrast, Fig. 8 depicts the value of

0.762381	0.64422	0.053092	0.000265035	0.	0.	0.	0.	0.	0.
0.64422	1.06888	0.33664	0.0211841	0.000114226	0.	0.	0.	0.	0.
0.053092	0.33664	0.441555	0.142793	0.00963464	0.0000553631	0.	0.	0.	0.
0.000265035	0.0211841	0.142793	0.876425	5.03465	1.74494	-0.0422927	0.	0.	0.
0.	0.000114226	0.00963464	5.03465	20.9205	-5.08484	-0.981995	0.0000387477	0.	0.
0.	0.	0.0000553631	1.74494	-5.08484	-15.9449	-0.838875	0.00365294	0.0000242098	0.
0.	0.	0.	-0.0422927	-0.981995	-0.838875	0.0479238	0.0294292	0.00245602	0.0000176755
0.	0.	0.	0.	0.0000387477	0.00365294	0.0294292	0.0499779	0.0213544	0.00193724
0.	0.	0.	0.	0.	0.0000242098	0.00245602	0.0213544	0.0392941	0.01821
0.	0.	0.	0.	0.	0.	0.0000176755	0.00193724	0.01821	0.0360422

Fig. 5. The Hessian for the inverse problem minimizing the integral of $(p - p^*)^2/2$. Since the shape functions are local B-Splines, it is possible to relate the location of the shock wave with the 4th, 5th and 6th column of the Hessian having significantly larger values.

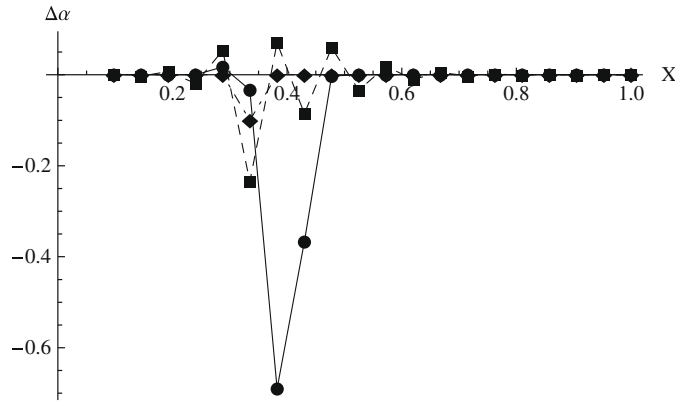


Fig. 6. The solid line (circles) represents the gradient, dashed (squares) represents the gradient pre-multiplied by the inverse of the Hessian, and the dot-dashed (diamonds) represents the error. (Inverse problem with a design space consisting of $N = 20$ design variables.)

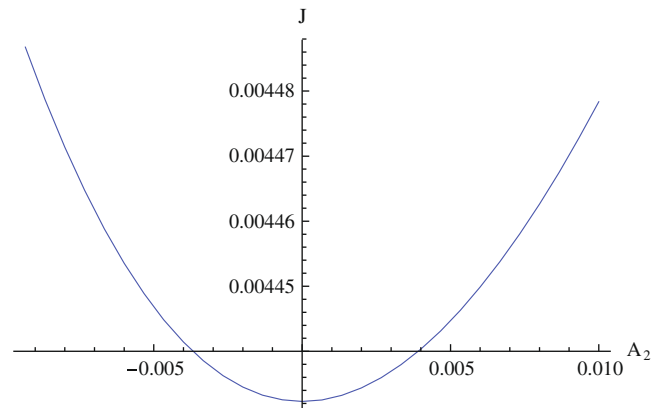


Fig. 7. The objective function as a function of the single design variable α_2 . The objective function in this case is the inverse problem minimizing the integral of $(p - p^*)^2/2$. The two shocks (in both p and p^*) are aligned and the design variable does not perturb the shock location. The resulting curve is smooth indicating the existence of the Hessian in this case.

the functional when $-0.01 \leq \alpha_8 \leq 0.01$. The corresponding shape function does perturb the shock and when the two shocks are aligned, for $\alpha_8 = 0$, the functional is not differentiable. However, we speculate that in many practical computations the shock wave is smeared on a few grid points due to artificial or physical dissipation, and the issue of non-differentiability may not be observed.

5. Gradient and Hessian derivatives via adjoint equation approach

The direct method presented in Sections 3 and 4 was very useful to analyze the model problem at hand, but it is not practical for industrial applications. In general we can not represent the solution as an explicit function of the geometry.

In this section we develop two methods that are viable for industrial applications, and that are based on the adjoint method.

5.1. The gradient via the adjoint method

The adjoint representation of the gradient is given by

$$\frac{dJ}{d\alpha_i} = \int_{\Omega} v^T f_i dx, \tag{25}$$

where (recall that $L = \frac{\partial R}{\partial U}$ and $-f_i = \frac{\partial R}{\partial \alpha_i}$)

$$f_i = \left(\frac{\partial h_i}{\partial x} p - \frac{d}{dx} (h_i F) \right). \tag{26}$$

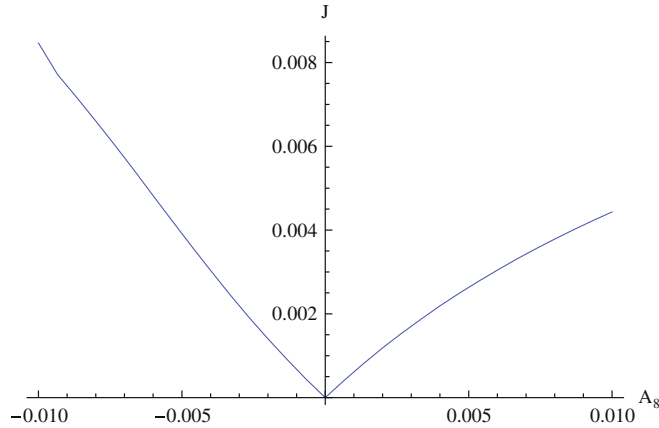


Fig. 8. The objective function as a function of a single design variable α_8 . The objective function in this case is the inverse problem minimizing the integral of $(p - p^*)^2/2$. The two shocks (in both p and p^*) are aligned and the design variable does perturb the shock location. The resulting curve is not smooth at the shock indicating that the Hessian (and gradient) do not exist in this case.

The adjoint v satisfies the adjoint equation

$$L^* v = \frac{\partial p}{\partial U}, \tag{27}$$

where the application of the adjoint linearized operator is given explicitly by,

$$L^* v \equiv -hA^T \frac{dv}{dx} - \frac{dh}{dx} B^T v. \tag{28}$$

5.2. Analytical derivation of the Hessian method I: $O(2N + 1)$ linear solutions

The formal expression for the element (i, j) of the Hessian is given by:

$$\frac{d^2 J}{d\alpha_i d\alpha_j} = \int_{\Omega} \left[\left(\frac{dv}{d\alpha_j} \right)^T f_i + v^T \frac{df_i}{d\alpha_j} \right] dx. \tag{29}$$

We observe that applying the above formula to a case where shocks are present in the flow, the term f_i is discontinuous at the shock (Heaviside function) and the adjoint v has a log singularity at the throat but is continuous at the shock [5]. The term $\frac{df_i}{d\alpha_j}$ contains a delta function at the shock since the pressure, p , contains a Heaviside function.

The term f_i depends on the design variables implicitly via the sensitivities. Its derivative with respect to α_j is given by,

$$\frac{df_i}{d\alpha_j} = - \frac{\partial L}{\partial \alpha_i} \frac{dU}{d\alpha_j}. \tag{30}$$

We take the derivative of the adjoint equation to obtain an equation for the adjoint sensitivity, $dv/d\alpha_j$:

$$L^* \frac{dv}{d\alpha_j} = - \frac{dL^*}{d\alpha_j} v + \frac{dU}{d\alpha_j} \frac{\partial g}{\partial U}, \tag{31}$$

where g denotes the right hand side (RHS) of the adjoint equation (27) given by

$$g = \frac{\partial p}{\partial U}. \tag{32}$$

Eq. (31) is similar to the adjoint equation but with a different right hand side. Denoting the RHS of Eq. (31) by G_j , we have

$$L^* \frac{dv}{d\alpha_j} = G_j. \tag{33}$$

In terms of G_j the Hessian can be written as follows:

$$\frac{d^2 J}{d\alpha_i d\alpha_j} = \int_{\Omega} [f_i L^* G_j dx + v^T \frac{df_i}{d\alpha_j}] dx. \tag{34}$$

5.2.1. Derivation of the RHS of the adjoint sensitivity equation

The RHS of Eq. (31) G_j , can be represented as follows:

$$G_j = -\frac{\partial L^*}{\partial \alpha_j} v - \frac{dU}{d\alpha_j} \frac{\partial L^*}{\partial U} v + \frac{dU}{d\alpha_j} \frac{\partial^2 p}{\partial U^2}, \quad (35)$$

where the adjoint operator L^* has been differentiated by parts:

$$\frac{dL^*}{d\alpha_j} = \frac{\partial L^*}{\partial \alpha_j} + \frac{dU}{d\alpha_j} \frac{\partial L^*}{\partial U}. \quad (36)$$

Recall that the application of the adjoint operator on the derivative term is defined by,

$$L^* \frac{dv}{d\alpha_j} \equiv -hA^T \frac{d}{dx} \frac{dv}{d\alpha_j} - \frac{dh}{dx} B^T \frac{dv}{d\alpha_j}.$$

The cost of method I is the solution of N sensitivity equations, N adjoint sensitivity equations, and 1 additional adjoint equation; all together $2N + 1$ linear equations. The terms L_U, L_α and $J_{UU}, J_{U\alpha}$ need to be computed as well. In particular the term containing the second order Jacobian tensor, J_{UU} , may be difficult to implement using modern CFD codes.

5.3. Analytical derivation of the Hessian method II: $O(N + 1)$ linear solutions

Our starting point is Eq. (29):

$$\Pi(i, j) \equiv \frac{d^2 J}{d\alpha_i d\alpha_j} = \int_{\Omega} \left[\left(\frac{dv}{d\alpha_j} \right)^T f_i + v^T \frac{df_i}{d\alpha_j} \right] dx. \quad (37)$$

5.3.1. The term $v^T (df_i/d\alpha_j)$

The derivative of f_i is computed explicitly from Eq. (26):

$$\frac{df_i}{d\alpha_j} = -\frac{d}{d\alpha_j} \frac{\partial R}{\partial \alpha_i} = -\frac{\partial R^2}{\partial \alpha_i \partial \alpha_j} - \frac{\partial}{\partial U} \frac{\partial R}{\partial \alpha_i} \frac{dU}{d\alpha_j} = -\frac{\partial L}{\partial \alpha_i} \frac{dU}{d\alpha_j}, \quad (38)$$

where the second derivative with respect to α can be eliminated, since the residual R is linear in the design variables:

$$\frac{\partial R^2}{\partial \alpha_i \partial \alpha_j} = 0.$$

5.3.2. The term $(dv/d\alpha_j)^T f_i$

The adjoint sensitivity term is derived by first considering the sensitivity equation:

$$\frac{\partial R}{\partial U} \frac{dU}{d\alpha_i} + \frac{\partial R}{\partial \alpha_i} = 0.$$

In terms of the sensitivities we can write

$$f_i = -\frac{\partial R}{\partial \alpha_i} = L \frac{dU}{d\alpha_i}.$$

Therefore,

$$\int_{\Omega} \left(\frac{dv}{d\alpha_j} \right)^T f_i dx = \int_{\Omega} \left(\frac{dv}{d\alpha_j} \right)^T L \frac{dU}{d\alpha_i} dx.$$

Taking the adjoint of the right hand side integrand we get (using the identity, $L^* \frac{dv}{d\alpha_j} = G_j$, and G_j is given by Eq. (35)),

$$\int_{\Omega} \left(\frac{dv}{d\alpha_j} \right)^T f_i dx = \int_{\Omega} \frac{dU^T}{d\alpha_i} L^* \frac{dv}{d\alpha_j} dx = \int_{\Omega} \frac{dU^T}{d\alpha_i} \left(\frac{\partial^2 p}{\partial U^2} \frac{dU}{d\alpha_j} - \left(\frac{\partial L^*}{\partial U} \frac{dU}{d\alpha_j} + \frac{\partial L^*}{\partial \alpha_j} \right) v \right) dx \quad (39)$$

5.3.3. The Hessian in matrix form

Substituting relations (38) and (39) in (37) and rearranging terms, the Hessian can be represented in the following matrix–vector form:

$$\Pi(i, j) = \int_{\Omega} \begin{bmatrix} \frac{dU}{d\alpha_i} \\ 1 \end{bmatrix}^T \left\{ \begin{bmatrix} \frac{\partial^2 p}{\partial U^2} & 0 \\ 0 & 0 \end{bmatrix} - v \otimes \begin{bmatrix} \frac{\partial L^*}{\partial U} & \frac{\partial L^*}{\partial \alpha_j} \\ \frac{\partial L}{\partial \alpha_i} & 0 \end{bmatrix} \right\} \begin{bmatrix} \frac{dU}{d\alpha_j} \\ 1 \end{bmatrix} dx \tag{40}$$

The notation \otimes is introduced in order to clarify the application of the operators in the integrand on the sensitivities and adjoint variables. It should be interpreted as follows:

$$\begin{aligned} v \otimes \frac{\partial L^*}{\partial U} \frac{dU}{d\alpha_j} &= \frac{\partial L^*}{\partial U} \left(\frac{dU}{d\alpha_j} \right) \cdot v \\ v \otimes \frac{\partial L^*}{\partial \alpha_j} &= \frac{\partial L^*}{\partial \alpha_j} (v) \\ v \otimes \frac{\partial L}{\partial \alpha_i} \frac{dU}{d\alpha_j} &= v \cdot \frac{\partial L}{\partial \alpha_i} \left(\frac{dU}{d\alpha_j} \right) \end{aligned}$$

where the parenthesis $A(x)$ denotes the action of an operator A on x , and the dot $v \cdot x$ denotes a dot product between v and x . The cost of method II consists of solving N linear sensitivity equations, and 1 additional adjoint equation; all together $N + 1$ linear equations. As in method I, the terms L_U, L_x and J_{UU}, J_{Ux} need to be computed as well.

Although method II seems to have an advantage over method I with regards to computational cost, we suspect that method II is likely to be more delicate in practice for flows with shocks, based on the following argument. In the Hessian representation given by Eq. (40) the terms $\frac{dU}{d\alpha_i}$, and $\frac{dU}{d\alpha_j}$ contain a delta function at the shock, while the terms $\frac{\partial L}{\partial \alpha_j}, \frac{\partial L^*}{\partial \alpha_j}, \frac{\partial L^*}{\partial U}$ and $\frac{\partial^2 p}{\partial U^2}$ contain a Heaviside function. As mentioned before the adjoint variable, v , is continuous at the shock. The numerical implementation of Eq. (40) may be more delicate compared to that of Eq. (29) since some of the terms involve multiplication of delta functions at the shock. This is not the case for the Hessian representation of method I (recall the observation we have made following Eq. (29)).

In the case of shocked flow for which the domain is not splitted to shockless subdomains, a more rigorous mathematical treatment is required to show that Eq. (40) can be interpreted in the sense of generalized functions, similar to the analysis done for the gradient in [11] and [12]. This is beyond the scope of this paper and is left for future investigation.

6. Numerical verification of the adjoint based methods for supersonic flow

We demonstrate the two methods on the model problem for shockless flow. In the quasi-one-dimensional approximation the domain can always be split into sub-regions where the solution is non-singular. We show here that the two methods yield the same results as the direct approach for a shockless case. The case considered is supersonic. The nozzle geometry is the same as that in previous sections. At the inlet the total pressure is given, $p^0 = 2$, as well as total enthalpy $H = 4$ and Mach number $M = 3$. We consider a pressure minimization objective function over $\Omega = [-1, 1]$. The shape functions are 10 B-splines with support evenly distributed in the interval $[-1/2, 1/2]$. All the calculations are done using the Mathematica[®] symbolic manipulation software.

The Hessian is first determined by the direct method, Eq. (20), on the unperturbed nozzle, to obtain a reference point. The result is shown in Fig. 9 and later compared with results obtained using the adjoint based methods.

6.1. Adjoint method I: $O(2N + 1)$ linear solutions

The Hessian representation by Eq. (29) requires the adjoint solution and its sensitivity derivatives. The adjoint is calculated following the Green’s function method presented in [5]. In the following, the adjoint sensitivity derivatives are calculated in a similar manner for supersonic flow condition.

We make use of the Lagrange identity for this case ($j = 1, \dots, N$ and $k = 1, 2, 3$):

$$(G_j, u_k) = \left(\frac{dv}{d\alpha_j}, \tilde{d}_k \right). \tag{41}$$

Eq. (41) is a direct consequence of the adjoint definition:

0.123894	0.0769523	0.0104256	0.000123508	0.	0.	0.	0.	0.	0.
0.0769523	0.216507	0.155398	0.0233237	0.000291494	0.	0.	0.	0.	0.
0.0104256	0.155398	0.483357	0.357767	0.0514667	0.000563739	0.	0.	0.	0.
0.000123508	0.0233237	0.357767	1.03299	0.64799	0.0726091	0.000563739	0.	0.	0.
0.	0.000291494	0.0514667	0.64799	1.42087	0.64799	0.0514667	0.000291494	0.	0.
0.	0.	0.000563739	0.0726091	0.64799	1.03299	0.357767	0.0233237	0.000123508	0.
0.	0.	0.	0.000563739	0.0514667	0.357767	0.483357	0.155398	0.0104256	0.0000619795
0.	0.	0.	0.	0.000291494	0.0233237	0.155398	0.216507	0.0769523	0.00602804
0.	0.	0.	0.	0.	0.000123508	0.0104256	0.0769523	0.123894	0.0524921
0.	0.	0.	0.	0.	0.	0.0000619795	0.00602804	0.0524921	0.100228

Fig. 9. Hessian by the direct method for supersonic flow condition (explicit).

$$\left(L^* \frac{dv}{d\alpha_j}, u_k \right) = \left(\frac{dv}{d\alpha_j}, Lu_k \right).$$

The left hand side of Eq. (41) is defined explicitly by,

$$(G_j, u_k) \equiv \int_{-1}^1 G_j(x) u_k(x) dx.$$

The right hand side is defined by,

$$\left(\frac{dv}{d\alpha_j}, \tilde{d}_k \right) \equiv \int_{-1}^1 \frac{dv}{d\alpha_j}(x) d_k(x) \delta(x - \xi) dx,$$

where $\delta(x)$ denotes the delta function at x .

The functions u_k are assumed to be the solutions of the following linearized equation corresponding to three independent right hand sides d_k :

$$Lu_k = d_k \delta(x - \xi). \tag{42}$$

In the supersonic case d_k is given by the columns of the matrix $\mathcal{D}(x)$, and \mathcal{D} is defined by:

$$\mathcal{D} = \begin{pmatrix} 1 & -h\rho q/(2H) & h\rho q/p^0 \\ q & 0 & h(\rho q^2 + p)/p^0 \\ H & h\rho q/2 & h\rho qH/p^0. \end{pmatrix}$$

The solutions of Eq. (42) are found in terms of the fundamental components (see [5])

$$\left(\frac{1}{h} \frac{\partial U}{\partial m} \Big|_{H,p^0}, \frac{\partial U}{\partial H} \Big|_{p^0,M}, \frac{\partial U}{\partial p^0} \Big|_{H,M} \right). \tag{43}$$

For supersonic flow the Mach number, M , the enthalpy, H , and the total pressure, p^0 , are fixed at the inlet and there are no boundary conditions at the outlet, resulting in

$$\begin{aligned} u_1(x, \xi) &= \mathcal{H}(x - \xi) \frac{1}{h(x)} \frac{\partial U}{\partial m}(x) \Big|_{H,p^0} \\ u_2(x, \xi) &= \mathcal{H}(x - \xi) \frac{\partial U}{\partial H}(x) \Big|_{p^0,M} \\ u_3(x, \xi) &= \mathcal{H}(x - \xi) \frac{\partial U}{\partial p^0}(x) \Big|_{H,M}, \end{aligned}$$

where $\mathcal{H}(x)$ is the Heaviside function of x . Each of the above three partial derivatives, (43), has to be evaluated for the three fundamental variables, $U = (\rho, \rho q, \rho E)^T$.

The following matrix contains the key information required:

$$\chi = \begin{bmatrix} \frac{1}{h} \frac{\partial \rho}{\partial m} \Big|_{H,p^0} & \frac{\partial \rho}{\partial H} \Big|_{p^0,M} & \frac{\partial \rho}{\partial p^0} \Big|_{H,M} \\ \frac{1}{h} & \frac{\partial m}{\partial H} \Big|_{p^0,M} & \frac{\partial m}{\partial p^0} \Big|_{H,M} \\ \frac{1}{h} \frac{\partial \rho E}{\partial m} \Big|_{H,p^0} & \frac{\partial \rho E}{\partial H} \Big|_{p^0,M} & \frac{\partial \rho E}{\partial p^0} \Big|_{H,M} \end{bmatrix} \tag{44}$$

In terms of χ the solutions u_k can be written as follows:

$$\mathcal{U} = \chi \mathcal{A},$$

where u_k form the columns of the matrix \mathcal{U} , and where the columns of the matrix \mathcal{A} are the three coefficient vectors

$$\mathcal{A} = \begin{bmatrix} \mathcal{H}(x - \xi) & 0 & 0 \\ 0 & \mathcal{H}(x - \xi) & 0 \\ 0 & 0 & \mathcal{H}(x - \xi) \end{bmatrix}. \tag{45}$$

In terms of the above definitions, the adjoint sensitivity can be written in matrix notation as follows:

$$\frac{dv}{d\alpha_m}(\xi) = \mathcal{D}^{-1}(\xi) \left(\int_{\xi}^1 G_m(x) \chi(x) dx \right). \tag{46}$$

The results are presented in Fig. 10. The numerical results for the Hessian, as computed by method I, coincide with the reference solution, computed by the direct method, up to a numerical error. Let Π be the Hessian, we have $\|\Pi_{direct} - \Pi_{method I}\| / \|\Pi_{direct}\| = 1.5 \times 10^{-4}$. The source of numerical errors is primarily the numerical integration procedure, that is applied multiple times.

0.123856	0.0769981	0.0104029	0.00012806	0.	0.	0.	0.	0.	0.
0.0769981	0.216441	0.155464	0.0232916	0.000295957	0.	0.	0.	0.	0.
0.0104029	0.155464	0.483259	0.357865	0.0514269	0.000566326	0.	0.	0.	0.
0.00012806	0.0232916	0.357865	1.03285	0.648104	0.0725744	0.000566126	0.	0.	0.
0.	0.000295957	0.0514269	0.648104	1.42073	0.648072	0.0514451	0.000294136	0.	0.
0.	0.	0.000566326	0.0725744	0.648072	1.03291	0.357807	0.0233116	0.000124958	0.
0.	0.	0.	0.000566126	0.0514451	0.357807	0.483324	0.155417	0.0104191	0.0000628964
0.	0.	0.	0.	0.000294136	0.0233116	0.155417	0.216491	0.076963	0.00602333
0.	0.	0.	0.	0.	0.000124958	0.0104191	0.076963	0.123881	0.0524997
0.	0.	0.	0.	0.	0.	0.0000628964	0.00602333	0.0524997	0.100231

Fig. 10. Hessian by the adjoint method I for supersonic flow condition ($O(2N + 1)$ linear solutions).

0.123889	0.0769521	0.0104256	0.000123507	0.	0.	0.	0.	0.	0.
0.0769521	0.216507	0.155399	0.0233238	0.000291495	0.	0.	0.	0.	0.
0.0104256	0.155399	0.483359	0.357768	0.0514667	0.000563737	0.	0.	0.	0.
0.000123507	0.0233238	0.357768	1.03299	0.647989	0.072609	0.000563739	0.	0.	0.
0.	0.000291495	0.0514667	0.647989	1.42086	0.64799	0.0514667	0.000291494	0.	0.
0.	0.	0.000563737	0.072609	0.64799	1.03299	0.357768	0.0233237	0.000123507	0.
0.	0.	0.	0.000563739	0.0514667	0.357768	0.483357	0.155398	0.0104257	0.0000619797
0.	0.	0.	0.	0.000291494	0.0233237	0.155398	0.216507	0.0769523	0.00602789
0.	0.	0.	0.	0.	0.000123507	0.0104257	0.0769523	0.123893	0.0524929
0.	0.	0.	0.	0.	0.	0.0000619797	0.00602789	0.0524929	0.100231

Fig. 11. Hessian by adjoint method II for supersonic flow condition ($O(N + 1)$ linear solutions).

6.2. Adjoint method II: $O(N + 1)$ linear solutions

The Hessian representation by Eq. (40) does not require the adjoint sensitivity derivatives. In that sense method II is simpler to implement, it requires only the sensitivities and adjoint solutions. The result is demonstrated in Fig. 11. In this case the numerical results have a smaller numerical integration error compared with method I: $\|\Pi_{\text{direct}} - \Pi_{\text{method II}}\| / \|\Pi_{\text{direct}}\| = 2.4 \times 10^{-6}$. Method I is less accurate since there is one more integration involved to compute the adjoint sensitivities, as opposed to method II where we integrate only once to compute the adjoint variables. This result indicates that in practice method I will require a tighter convergence tolerance than method II in order to achieve a similar accuracy.

7. Discussion

Industrial aerodynamic optimization problems are usually defined by hundreds of design variables and millions of state variables. The state-of-the-art practice is to use the adjoint method to compute the gradient, and quasi-Newton method to accelerate the convergence. Quasi-Newton approximates the Hessian (or its inverse) by a low rank update method taking the identity matrix to be the initial guess. That choice corresponds to having the gradient as the initial search direction in the optimization process. Fig. 6 demonstrates how the gradient can serve as a poor search direction; the solid line (circles) represents the gradient and the dot-dashed line (diamonds) represents the error. The gradient is both quantitatively and qualitatively pointing in the wrong direction. For practical industrial aerodynamic optimization problems only a few iterations can be afforded, resulting in poor convergence. Therefore, we think that direct approximation of the Hessian is essential to achieve convergence in a relatively small number of iterations. Such an approximation can serve as the initial guess for a quasi-Newton method.

In this work we compute the Hessian analytically for an optimal shape design problem governed by the quasi-one-dimensional Euler equations, modeling inviscid flow in a channel. For that ideal model problem we can study the Hessian directly by representing the flow, sensitivity, and adjoint solutions as explicit functions of the design variables. That allows us to explore the question of Hessian existence in aerodynamic optimization problems under transonic flow with shocks. The Hessian exists in the strong sense if the objective function is twice differentiable at every design point. Fig. 8 serves as an example to the lack of smoothness that may occur in such problems. The objective function as a function of one of the design variables (α_g) is not smooth at $\alpha_g = 0$, thus both the gradient and the Hessian do not exist at that point. We think that more research is required to further understand the existence of the gradient and Hessian in the presence of shocks, particularly at the discrete level where the choice of numerical scheme may make a difference.

We also propose two formulations to approximate the Hessian that can be applied to general problems for which the direct approach is not practical. These formulations are based on the adjoint and sensitivity solutions. The first method is more expensive. It requires solving the adjoint sensitivity equations in addition to the standard sensitivity equations and adjoint equation ($O(2N + 1)$ linear solutions), while the second method requires only the sensitivity and adjoint equations ($O(N + 1)$ linear solutions). In practice the Hessian is used to obtain an improved search direction by preconditioning the gradient and therefore only an approximated Hessian is required. The computational effort can be greatly reduced by a number of ways,

for example by converging the sensitivity equations only to a low tolerance value. Also, the sensitivity equations system is easily parallelizable as it consists of a single operator with many right hand sides.

The preliminary analysis presented in this paper suggests that the first method, although more expansive, may be more stable, since the second method involves multiplications of delta functions (δ^2) at the shock location suggesting large numbers in the numerical computation (see discussion at the end of Section 5). We demonstrate the application of the proposed two adjoint based methods on the model problem for supersonic flow. More research is required to better understand, from a numerical analysis view point, method I and method II to compute the Hessian. This is left for future study.

References

- [1] A. Jameson, Aerodynamic design via control theory, ICASE report NO. 88-64, November, *Journal of Scientific Computing* 3 (1988) 233–260.
- [2] F. Beux, A. Dervieux, Exact-gradient shape optimization of a 2-D Euler flow, *Finite Elements in Analysis and Design* 12 (1992) 281–302.
- [3] P.D. Frank, G.R. Shubin, A comparison of optimization-based approaches for a model computational aerodynamics design problem, *Journal of Computational Physics* 98 (1992) 74–89.
- [4] A. Iollo, M.D. Salas, Contribution to the optimal shape design of two-dimensional internal flows with embedded shocks, *Journal of Computational Physics* 125 (1996) 124–134.
- [5] M. Giles, N. Pierce, Analytic adjoint solutions for the quasi-1D Euler equations, *Journal of Fluid Mechanics* 426 (2001) 327–345.
- [6] B. Mohammadi, O. Pironneau, Shape optimization in fluid mechanics, *Annual Review of Fluid Mechanics*. 36 (2004) 255–279.
- [7] E. Arian, S. Ta'asan, Analysis of the Hessian for aerodynamic shape optimization: inviscid flow, *Computers and Fluids* 28 (1999) 853–877.
- [8] E. Arian, V. Vatsa, A preconditioning method for shape optimization governed by the Euler equations, *International Journal of Computational Fluid Dynamics* 12 (1999) 17–27.
- [9] L. Sherman, A. Taylor III, L. Green, P. Newman, First- and second-order aerodynamic sensitivity derivatives via automatic differentiation with incremental iterative methods, *Journal of Computational Physics* 129 (1996) 307–331.
- [10] D. Ghatge, M. Giles, Efficient Hessian calculation using automatic differentiation, in: *AIAA Paper 2007-4059*, 25th AIAA Applied Aerodynamics Conference 25–28 June 2007, Miami, FL.
- [11] C. Bardos, O. Pironneau, A formalism for the differentiation of conservation laws, *Comptes Rendus Mathematique* 335 (2002) 839–845.
- [12] C. Bardos, O. Pironneau, Data assimilation for conservation laws, *Methods Applied Analysis* 12 (2005) 103–134.