J Eng Math (2009) 65:1–24
DOI 10.1007/s10665-008-9258-7

Boundary and internal conditions for adjoint fluid-flow problems

Application to the quasi-1D Euler equations

E. V. Volpe · L. C. de Castro Santos

Received: 29 June 2007 / Accepted: 11 November 2008 / Published online: 30 November 2008

Abstract The ever-increasing robustness and reliability of flow-simulation methods have consolidated CFD as a major tool in virtually all branches of fluid mechanics. Traditionally, those methods have played a crucial role in the analysis of flow physics. In more recent years, though, the subject has broadened considerably, with the development of optimization and inverse design applications. Since then, the search for efficient ways to evaluate flow-sensitivity gradients has received the attention of numerous researchers. In this scenario, the adjoint method has emerged as, quite possibly, the most powerful tool for the job, which heightens the need for a clear understanding of its conceptual basis. Yet, some of its underlying aspects are still subject to debate in the literature, despite all the research that has been carried out on the method. Such is the case with the adjoint boundary and internal conditions, in particular. The present work aims to shed more light on that topic, with emphasis on the need for an internal shock condition. By following the path of previous authors, the quasi-1D Euler problem is used as a vehicle to explore those concepts. The results clearly indicate that the behavior of the adjoint solution through a shock wave ultimately depends upon the nature of the objective functional.

Keywords Adjoint method · Adjoint boundary conditions · Integral approach · Realizable flow variations

1 Introduction

The impact the use of CFD has made in so many branches of fluid dynamics is undeniable. In particular for the aerospace industry, it provides a very cost-effective means of analyzing different flow-geometry configurations. More recently, the development of inverse design and optimization methods has opened up new possibilities. On combining those methods with CFD codes, one can specify elaborate design goals and search for configurations that meet them.

The works by Lions [1], Pironneau [2–4], Jameson [5–7], Salas [8,9], Giles and Pierce [10,11] and many others [12–14] have established the adjoint method as a major contribution in the field. It offers an extremely attractive

E. V. Volpe (✉)
Department of Mechanical Engineering, University of São Paulo, Av. Prof. Mello Moraes, 2231, CEP 05508-900, São Paulo, SP, Brazil
e-mail: ernvolpe@usp.br

L. C. de Castro Santos
Department of Applied Mathematics, University of São Paulo, Rua do Matão, 1010, CEP 05508-900, São Paulo, SP, Brazil
e-mail: lsantos@ime.usp.br

capability, which is to compute sensitivity gradients at a cost that is largely independent of the number of design variables. It all hinges upon the solution of the equations governing the flow and the so-called adjoint equations.

It is worth noting that the solution of the adjoint problem incurs roughly the same computational costs as the flow simulation does. Even so, there are several works that aim at further reducing those costs. In [15], the one-shot method developed by Salas et al. [8,9] is further explored to reach a faster rate of convergence, which is achieved by taking advantage of a multi-grid scheme. On following a similar one-shot approach, Beux and Dervieux [16] analyze the influence of the geometry parameterization, by considering a different number of parameters in each cycle of the multi-grid scheme.

Other approaches have also been pursued successfully. Apparently, though, none of them seems to allow the same flexibility with regard to the flow-physics model and the measure of merit as the adjoint method does. Such is the case with automatic differentiation [17,18], which relies on software tools to *differentiate* a CFD code and, thus, to evaluate sensitivity gradients. Despite all the effort that has been put into developing appropriate tools for the job, the differentiation procedure remains a tedious, cumbersome, task.

Over the years, several application examples have been reported in the literature for the adjoint method. Ranging from nuclear-reactor thermo-hydraulics to atmospheric sciences [14,19], they span a wide variety of interests. Aerodynamics applications, in particular, range from the design of single [20] and multi-element profiles [21,22], to wing-shape optimization [23,24] and turbo-machinery design [25]. There have also been some successful applications of coupled aerodynamic-structure formulations, such as in [26]. More recently, several developments regarding unsteady problems have been reported in the literature as well [27–29].

Not only can the adjoint variables be used to evaluate functional gradients, as in an optimization procedure, but they also enable one to compute these functionals to a higher order of accuracy [30–33]. An entirely different class of applications has evolved around this idea for error analysis [34,35]. A closely related topic is grid adaptation based on functionals [36,37]. It can be shown that the usual spatial (or h) adaptation schemes, which are based on flow variables, have limited power to attain the levels of accuracy functionals are required to meet, in applications such as drag minimization [35,38].

The choices for a gradient-based optimization procedure can be summarized as follows: On taking the set of governing equations as a nonlinear operator R over a set of flow variables U, one can write [39]:

$$R(U) = 0. \tag{1}$$

By linearly perturbing the geometry, a linearized sensitivity equation can be constructed such that:

$$Lu = f. \tag{2}$$

Here L is the linearized operator, u the sensitivity of the flow variables and f accounts for the remaining terms resulting from the geometry perturbation. The nonlinear optimization problem can be represented as a functional $J(U)$ over the domain. When linearly perturbed, it can be written as an internal product:

$$I(u) = \langle g, u \rangle, \tag{3}$$

where $I(u) \equiv \delta J(U)$. At this point there are several choices:

- the classical brute-force approach, which is to perturb directly the nonlinear flow equations (1), and solve the problem N times, where N is the number of design variables, and recompute $J(U)$;
- develop a code to solve for the sensitivities (2), either directly or by using an automatic differentiation tool, and evaluating (3);
- follow an adjoint approach.

The latter option implies augmenting the functional by imposing (2) as a constraint, by means of continuous Lagrange multipliers v. It leads to:

$$I(u) = \langle g, u \rangle - \langle v, Lu - f \rangle. \tag{4}$$

As long as (2) holds, Eq. 4 is equivalent to (3). The adjoint operator L^* is then defined by the relation:

$$\langle v, Lu \rangle = \langle L^* v, u \rangle. \tag{5}$$

On assuming that u, v satisfy homogeneous boundary conditions, and replacing (5) in Eq. 4, one obtains:

$$I(u) = -\langle L^* v - g, u\rangle + \langle v, f\rangle. \tag{6}$$

As long as the adjoint equation is satisfied, one has:

$$L^* v = g \tag{7}$$

and the flow-variables sensitivities are removed from the functional sensitivity, which leads to:

$$I(u) = \langle v, f\rangle \tag{8}$$

Despite its theoretical simplicity, there are several implementation issues regarding the development of adjoint codes:

- it may be considerably difficult to extract the adjoint operator L^* analytically from the original equations;
- the boundary conditions for u, v are generally non-homogeneous, thus bringing additional integrals in Eq. 8;
- the boundary conditions for v have a non-intuitive mathematical behavior, and should be carefully considered, so as to fully remove the problem dependence on flow sensitivity.

These aspects have led researchers to follow two different alternatives: the discrete or the continuous formulation.

In the discrete approach, instead of handling the analytical development, the discretized operator L is explicitly constructed in matrix form. The discrete adjoint operator L^* is the matrix transpose of the discrete sensitivity operator L. In [40], one can find a comparison between the discrete and continuous approaches for Euler equations. In terms of optimization effectiveness there is not a sizable difference.

The generation of the adjoint discrete problem, although straightforward, can also be cumbersome. In order to facilitate such tasks, automatic differentiation techniques have also been applied in the development of discrete adjoint codes [41,42]. There is still the non-trivial computational work in solving the very large (fortunately sparse) linear system [43].

In the continuous approach, the adjoint operator is obtained analytically for each PDE. Similarly, the boundary conditions for the adjoint equations are derived for each choice of flow boundary conditions. An adjoint continuous solver for the Navier–Stokes equations was developed by Jameson et al. [40,44] and successfully used in inverse design and drag minimization. The only boundary conditions needed for external aerodynamics are far field and wall. For a wide variety of problems, other boundary conditions are needed. Therefore it is important to focus on their derivation.

A clear understanding of the boundary conditions is an important issue in the development of a continuous adjoint problem. The boundary conditions of the primal, (2), and dual, (4), problems are related in such a way that the characteristics of each are reversed. Depending on the nature of the flow, the number and specifications of each characteristic have to be matched, so as to retain consistency. The topic will be explored in detail below.

As noted by Giles and Pierce [39]: "*there is little discussion of the properties of the adjoint solutions themselves*", which is something that can only be achieved by following a continuous approach. Initially in [39], and much latter in [34], those authors state the need of an internal boundary condition, with the specific purpose of handling the shock discontinuity. While that condition is strictly required for the well-posedness of the problem, the adjoint-solution behavior through the shock wave depends on the nature of the objective functional. It will be shown below that such dependence is more general than what was anticipated by those authors, on the basis of their particular choice of functional.

To better illustrate the importance of that topic, it is instructive to quote a passage from [45], which reads:

> In Iollo et al. (1993), it is suggested that v = 0 could be imposed at the shock, but this over-constrains the adjoint problem, in addition to contradicting (2.5). Cliff et al. (1996, 1998) conclude that there is a 'shock' in the adjoint variables at the shock location, having proved that the adjoint variables undergo a change of sign across the shock. However, as this change of sign is entirely due to the non-standard coordinate system they employ in formulating the augmented Lagrange, the conclusion that the adjoint variables are discontinuous at the shock is misleading.

That interpretation of the adjoint internal condition seems to prevail in the literature. Venditti and Darmofal have developed a finite-element discrete adjoint implementation of the analytical work of Giles and Pierce and use it to explore both grid adaption and the super-convergence of functionals [36–38,46]. In their line of work they have made use of the same linear functional as Giles and Pierce [45], which does not reveal any discontinuity at the shock location. In his doctoral thesis, Xie [47] follows the same approach of Giles and Pierce. However, he has observed a shock discontinuity, and interpreted it as was denied in [45].

The contribution of this paper is to further explore and clarify some issues regarding the boundary and internal conditions for the adjoint equations, while focusing on the continuous formulation of the method. The relation between the objective functional and the adjoint-solution behavior through a shock wave is considered in detail. On following the steps of [37,39] and [47], the quasi-1D Euler equations are used as a vehicle for the analytical development. Yet, departing from previous references, the whole problem is formulated for a generic objective functional.

In this framework, the need for an internal condition at the shock wave is verified. The continuity of the adjoint variables is imposed at that location, just as Giles and Pierce had proposed [39]. Our findings reveal that, although the adjoint variables themselves meet the continuity condition as expected, their gradient exhibits a discontinuity. Furthermore, the strength of that discontinuity is directly related to the nature of the objective functional. The linear functional that has been picked by those authors is shown to be a particular case, for which the adjoint variables smoothly cross the shock wave. This fact seems to have been misinterpreted by some authors in the past, which led to contradictory conclusions of some references.

The test results clearly demonstrate the validity of the present theoretical claims. An alternative continuous theoretical development is pursued and a novel interpretation of some of the previous conclusions are drawn from its results.

2 Statement of the problem

Owing to its simplicity, the classical problem of Euler flow through a nozzle with a variable cross-section is chosen as an appropriate vehicle to explore the method's underlying concepts. In particular, for a slowly varying cross-sectional area, the flow can be conveniently modeled by the well-known quasi-1D Euler equations [48, Chap. 2], [49, Sect. 16.4],

$$\frac{\partial(S\mathbf{Q})}{\partial t} + \frac{\partial(S\mathbf{F})}{\partial x} = S'\mathbf{H}, \tag{9}$$

where $S = S(x)$ denotes the cross-sectional area of the nozzle, and $S' = \mathrm{d}S/\mathrm{d}x$ represents how the area varies lengthwise. The quantities $\mathbf{Q}$ and $\mathbf{F}$ represent the state variables and flux vector in conservative form, respectively, and the term $S'\mathbf{H}$ on the the RHS accounts for the effects of area change on the balance equations:

$$\mathbf{Q} \equiv \begin{pmatrix} \rho \\ \rho u \\ \rho e_t \end{pmatrix}; \quad \mathbf{F} \equiv \begin{pmatrix} \rho u \\ \rho u^2 + P \\ (\rho e_t + P)u \end{pmatrix}; \quad \mathbf{H} \equiv \begin{pmatrix} 0 \\ P \\ 0 \end{pmatrix}. \tag{10}$$

The ideal-gas relation for pressure and the total energy per unit of mass (e_t)[1] closes the set (9):

$$P = (\gamma - 1)\rho\left[e_t - \frac{u^2}{2}\right]. \tag{11}$$

As an illustrative example of a measure of merit, we define the following functional

$$I = \int_0^l g(\mathbf{V})\,\mathrm{d}x, \tag{12}$$

[1] $e_t = (e_i + u^2/2)$, where e_i represents the thermodynamic specific internal energy.

where $g(\mathbf{V})$ represents a generic scalar function of the primitive state variables $\mathbf{V} = (\rho, u, P)^T$. The integral is taken over the whole length of the nozzle, l, and the objective is to find the geometry $S(x)$ that minimizes the functional I. On computing the variation of (12), one must take into account the possible occurrence of a normal shock wave within the flow domain. Hence, it is convenient to split the integration into two parts, upstream and downstream of the shock location x_s [10,39,45]. This leads to

$$\delta I = \int_0^{x_s^-} \frac{\partial g}{\partial V}^T \cdot \frac{\partial V}{\partial Q} \cdot \delta \mathbf{Q}\, \mathrm{d}x - [g(V)]_{x_s^-}^{x_s^+} \delta x_s + \int_{x_s^+}^{l} \frac{\partial g}{\partial V}^T \cdot \frac{\partial V}{\partial Q} \cdot \delta \mathbf{Q}\, \mathrm{d}x, \tag{13}$$

where δx_s represents a shift in the shock-wave location. The rationale behind (13) is that any changes in the flow field, $\delta\mathbf{Q}$ and δx_s, are brought about by variations in a set of control parameters, which has yet to be defined.

On assuming that the nozzle geometry can be accurately represented as a function of x with a set of parameters a_k, $S(x; a_k)$, one could attempt to minimize I by adjusting that set. Then the a_k would represent the control parameters, and the most important piece of information would be the sensitivity gradient $\partial I/\partial a_k$, which should be computed on the basis of (13). The problem with that approach lies in the dependence of the flow solution on those parameters, which is not usually known in closed form. As a result, an estimate of $\partial\mathbf{Q}/\partial a_k$ would require several solutions of the equations governing the flow (9), one for each variation δa_k, taken separately. Clearly, as the number of control parameters grows, the cost of such computations is bound to become prohibitive.

A common characteristic of all the simulations required to estimate $\partial\mathbf{Q}/\partial a_k$ is that, for each individual variation δa_k, there must correspond a realizable solution to the governing equations. Hence, if one could constrain the variations to the space of realizable solutions, *a priori*, then, perhaps, one could eliminate the need for expensive computations of $\partial\mathbf{Q}/\partial a_k$ to estimate the sensitivity gradient.

The adjoint method opens up such a possibility. Originally proposed by Jameson [5,6] for aerodynamic applications, it makes use of concepts from control theory to achieve that goal. In essence, it imposes the flow-governing equations as constraints on the optimization problem and, on doing so, it precludes unrealizable solutions. For the application in hand, one is mostly interested in steady flow conditions. Therefore, the steady form of (9) is imposed on (12) as a non-holonomic constraint, thus leading to the augmented functional [10,39,45]

$$I = \int_0^l g(\mathbf{V})\, \mathrm{d}x + \underbrace{\int_0^l \Psi^T \cdot \left(\frac{\partial (S\mathbf{F})}{\partial x} - S'\mathbf{H} \right) \mathrm{d}x + \Psi_s^T \cdot [S\mathbf{F}]_{x_s^-}^{x_s^+}}_{I_c}. \tag{14}$$

The third term on the RHS of (14) imposes the Rankine–Hugoniot (R–H) relations on the shock wave; the vectors Ψ^T and Ψ_s^T are the corresponding Lagrange multipliers. The symbol I_c is used to denote the whole set of constraints.

The variation of the first integral in (14) is obviously given by (13), whereas that of term I_c involves the variations of the constraint equation and of the R–H relations. The former yields

$$\frac{\partial}{\partial x}\delta(S\mathbf{F}) - \delta(S'\mathbf{H}) = 0, \tag{15}$$

where the variations $\delta(S\mathbf{F})$ and $\delta(S'\mathbf{H})$, are given by

$$\delta(S\mathbf{F}) = S\mathbf{A}\delta\mathbf{Q} + \mathbf{F}\delta S, \quad \delta(S'\mathbf{H}) = S'\mathbf{B}\delta\mathbf{Q} + \mathbf{H}\delta S'. \tag{16}$$

Both flow solution and geometry variations are implicitly assumed to be caused by control-parameter changes, but they are otherwise taken to be independent of each other. The symbol $\mathbf{A}$ represents the matrix of the flux Jacobian and $\mathbf{B}$ represents the Jacobian matrix of (9) RHS term:

$$\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{Q}} = \begin{pmatrix} 0 & 1 & 0 \\ (\gamma-3)\frac{u^2}{2} & (3-\gamma)u & (\gamma-1) \\ (\gamma-1)u^3 - \gamma e_t u & \gamma e_t - (\gamma-1)\frac{3u^2}{2} & \gamma u \end{pmatrix}, \tag{17}$$

$$\mathbf{B} = \frac{\partial \mathbf{H}}{\partial \mathbf{Q}} = (\gamma - 1) \begin{pmatrix} 0 & 0 & 0 \\ \frac{u^2}{2} & -u & 1 \\ 0 & 0 & 0 \end{pmatrix}. \tag{18}$$

As with the flux vector $\mathbf{F}$, one can easily show that $\mathbf{H}$ is homogeneous of degree one with respect to $\mathbf{Q}$. Hence, the relation $\mathbf{H} = \mathbf{B}.\mathbf{Q}$ holds as an exact result. The R–H term also depends on the shock-wave location x_s, in addition to the geometry S and the state variables $\mathbf{Q}$, thus its variation leads to

$$\delta\left[S\mathbf{F}\right]_{x_s^-}^{x_s^+} = \left[\frac{\partial (S\mathbf{F})}{\partial x}\right]_{x_s^-}^{x_s^+} \delta x_s + \left[S\mathbf{A}\delta\mathbf{Q}\right]_{x_s^-}^{x_s^+} + \left[\mathbf{F}\delta S\right]_{x_s^-}^{x_s^+} = \left[S'\mathbf{H}\right]_{x_s^-}^{x_s^+} \delta x_s + \left[\delta(S\mathbf{F})\right]_{x_s^-}^{x_s^+}. \tag{19}$$

Here, use was made of the steady form of the governing equation (9), to substitute for the flux derivative in the algebra. The first expression in (16) has also been used to collect the terms of $\delta(S\mathbf{F})$.

The first step to obtain the variation of I_c is to integrate the constraint functional by parts and to substitute (19) for the variation of the R–H term

$$\begin{aligned} \delta I_c = {} & \left[\Psi^T \cdot \delta(S\mathbf{F})\right]_0^{x_s^-} - \int_0^{x_s^-} \frac{\partial \Psi^T}{\partial x} \cdot \delta(S\mathbf{F}) + \Psi^T \cdot \delta(S'\mathbf{H})\,\mathrm{d}x + \Psi_s^T \cdot \left\{\left[S'\mathbf{H}\right]_{x_s^-}^{x_s^+} \delta x_s + \left[\delta(S\mathbf{F})\right]_{x_s^-}^{x_s^+}\right\} \\ & + \left[\Psi^T \cdot \delta(S\mathbf{F})\right]_{x_s^+}^{l} - \int_{x_s^+}^{l} \frac{\partial \Psi^T}{\partial x} \cdot \delta(S\mathbf{F}) + \Psi^T \cdot \delta(S'\mathbf{H})\,\mathrm{d}x. \end{aligned} \tag{20}$$

The terms that involve the shock-wave position can be regrouped as follows

$$\begin{aligned} \delta I_c = {} & \left[\Psi^T(x_s^-) - \Psi_s^T\right] \cdot \delta\left(S\mathbf{F}\right)_{x_s^-} - \left[\Psi^T(x_s^+) - \Psi_s^T\right] \cdot \delta\left(S\mathbf{F}\right)_{x_s^+} + \Psi_s^T \cdot \left[S'\mathbf{H}\right]_{x_s^-}^{x_s^+} \delta x_s + \left[\Psi^T \cdot \delta(S\mathbf{F})\right]_0^{l} \\ & - \int_0^{x_s^-} \left[\frac{\partial \Psi^T}{\partial x} \cdot \delta(S\mathbf{F}) + \Psi^T \cdot \delta(S'\mathbf{H})\right] \mathrm{d}x - \int_{x_s^+}^{l} \left[\frac{\partial \Psi^T}{\partial x} \cdot \delta(S\mathbf{F}) + \Psi^T \cdot \delta(S'\mathbf{H})\right] \mathrm{d}x. \end{aligned} \tag{21}$$

The first two terms on the RHS of (21) involve differences between the R–H Lagrange multipliers, Ψ_s, and the Ψ, where the latter are evaluated at the shock location. According to Giles and Pierce [10,39,45], these terms prompt the need for an internal boundary condition that should be imposed on the Ψ, at that location. Then, on making

$$\Psi(x_s^-) = \Psi_s = \Psi(x_s^+), \tag{22}$$

those two terms drop from (21), and one imposes the continuity of the Lagrange multipliers through the shock wave. Then, on introducing the variations (16) into the above equation and collecting like terms, one has

$$\begin{aligned} \delta I_c = {} & \Psi_s^T \cdot \left[S'\mathbf{H}\right]_{x_s^-}^{x_s^+} \delta x_s + \left[\delta\mathbf{Q}^T S\mathbf{A}^T \Psi\right]_0^{l} + \left[\delta S\mathbf{F}^T \cdot \Psi\right]_0^{l} \\ & - \int_0^{x_s^-} \left[\delta\mathbf{Q}^T \cdot \left(S\mathbf{A}^T \frac{\partial \Psi}{\partial x} + S'\mathbf{B}^T \Psi\right) + \delta S\mathbf{F}^T \cdot \frac{\partial \Psi}{\partial x} + \delta S'\mathbf{H}^T \cdot \Psi\right] \mathrm{d}x \\ & - \int_{x_s^+}^{l} \left[\delta\mathbf{Q}^T \cdot \left(S\mathbf{A}^T \frac{\partial \Psi}{\partial x} + S'\mathbf{B}^T \Psi\right) + \delta S\mathbf{F}^T \cdot \frac{\partial \Psi}{\partial x} + \delta S'\mathbf{H}^T \cdot \Psi\right] \mathrm{d}x. \end{aligned} \tag{23}$$

As a result of the continuity of Ψ and the fact that physical variables only undergo finite jump discontinuities through a shock wave, one can perform the integrations over the whole length of the nozzle, from 0 to l, as opposed to splitting them at the shock location, on the understanding that the equations governing the flow hold there in the sense of weak solutions, only. Hence, the only remaining term that is directly related to the R–H conditions is the one that multiplies the shock-wave shift δx_s.

Finally, the variation of the augmented functional (14) is obtained by adding the two equations (13) and (23) and collecting like terms,

$$\delta I = \underbrace{\left[\delta \mathbf{Q}^T S \mathbf{A}^T \Psi\right]_0^l}_{(a)} + \underbrace{\int_0^l \delta \mathbf{Q}^T \cdot \left[\left(\frac{\partial V}{\partial Q}\right)^T \frac{\partial g}{\partial V} - S\mathbf{A}^T \frac{\partial \Psi}{\partial x} - S'\mathbf{B}^T \Psi\right] \mathrm{d}x}_{(b)}$$

$$+ \underbrace{\left\{\Psi_s^T \cdot \left[S'\mathbf{H}\right]_{x_s^-}^{x_s^+} - \left[g(V)\right]_{x_s^-}^{x_s^+}\right\} \delta x_s}_{(c)} + \underbrace{\left[\mathbf{F}^T \cdot \Psi\, \delta S\right]_0^l}_{(d)} - \underbrace{\int_0^l \left[\mathbf{F}^T \cdot \frac{\partial \Psi}{\partial x}\, \delta S + \mathbf{H}^T \cdot \Psi\, \delta S'\right] \mathrm{d}x}_{(e)}. \tag{24}$$

The resulting equation represents the variation of the measure of merit subject to the chosen constraints. That is, it should satisfy the Euler equations in steady form, as well as the R–H relations. A split into geometric and physical variations can be recognized at once, since the terms (d) and (e) involve only the former, whereas (a), (b) and (c) involve only the latter.

The variation split can be used to advantage in eliminating the $\delta\mathbf{Q}$ from the above equation. In principle, one can make use of the expression within brackets in (b) as a means to solve for Ψ. That can be accomplished by requiring it to satisfy the PDE

$$-S\mathbf{A}^T \frac{\partial \Psi}{\partial x} - S'\mathbf{B}^T \Psi + \frac{\partial V}{\partial Q}^T \frac{\partial g}{\partial V} = 0 \tag{25}$$

on both sides of the shock wave, thereby suppressing the integral from (24). Boundary conditions for this PDE are obtained from term (a), along with the continuity condition (22). It will be shown below that these conditions can be derived in such a way as to ensure that the problem is well-posed and that the contribution of term (a) to the total variation is zero. The term (c) will also be shown to have no contribution to δI.

Under these circumstances, the total variation reduces to the sum of (d) and (e), both of which involve only area variations. In essence, then, on solving (25) for Ψ and substituting the result in the expression of δI, one is implicitly constraining that variation to the space of realizable solutions. As a result, one should be able to estimate the sensitivity gradient without having to resort to expensive computations of individual design-parameter variations. This could be done on the basis of a simplified expression, instead.

$$\delta I = \left[\mathbf{F}^T \cdot \Psi \delta S\right]_0^l - \int_0^l \left[\mathbf{F}^T \cdot \frac{\partial \Psi}{\partial x}\, \delta S + \mathbf{H}^T \cdot \Psi\, \delta S'\right] \mathrm{d}x. \tag{26}$$

Equation 25 corresponds to the adjoint equation for the problem. However, it can be cast in a more instructive form, which stresses its similarity to the equation governing the flow (9). The importance of such a similarity is twofold: not only does it indicate an approach to solving the adjoint equation; but, more importantly, it gives us guidance in deriving appropriate boundary conditions to impose on it.

3 The adjoint equations

The similarity between the adjoint PDE (25) and the Euler equation (9) becomes apparent when one writes the latter in terms of the Jacobian matrix $\mathbf{A}$ and makes use of the identity $\mathbf{H} = \mathbf{B}.\mathbf{Q}$. On introducing these results into the steady form of (9) and comparing it to (25) one gets

$$-S\mathbf{A}^T \frac{\partial \Psi}{\partial x} - S'\mathbf{B}^T \Psi = -\frac{\partial V}{\partial Q}^T \frac{\partial g}{\partial V} \equiv -\frac{\partial g}{\partial Q}, \tag{27}$$

$$S\mathbf{A}\frac{\partial \mathbf{Q}}{\partial x} - S'\left(\mathbf{B} - \mathbf{A}\right)\mathbf{Q} = 0. \tag{28}$$

As can be seen, both equations share a similar structure. The only differences are in the sign of the flux term and the non-homogeneous term on the RHS of (27). The origin of the latter can be traced back to the variation of the objective functional in (13). Whereas the terms on the LHS of (27) all come from the variation of the governing equation (23) and bear no influence of the objective functional.

The apparent similarity between the two PDEs opens up the possibility of using the same body of knowledge that has been developed for the Euler equation to tackle the adjoint equation. The hyperbolic character the Euler equation exhibits in the time–space domain is what ultimately determines flow boundary conditions. Given the structure of (27), the same feature could be shared by the adjoint equation if only one could postulate a time dependence for the variable Ψ similar to $\mathbf{Q}$. In principle, that could be accomplished by simply adding a time derivative to the LHS of (27). This step leads to the following PDE

$$\frac{\partial (S\Psi)}{\partial t} - S\mathbf{A}^T \frac{\partial \Psi}{\partial x} - S'\mathbf{B}^T \Psi = -\frac{\partial g}{\partial Q}, \tag{29}$$

$$\frac{\partial (S\mathbf{Q})}{\partial t} + S\mathbf{A} \frac{\partial \mathbf{Q}}{\partial x} - S' (\mathbf{B} - \mathbf{A}) \mathbf{Q} = 0, \tag{30}$$

which is paired with the full Euler equation for comparison. In view of the relation between the two, the flow and the optimization are usually referred to as primal and dual problems, respectively [31]. For its role in the latter, Ψ is often referred to as the co-state vector in the literature. The same terminology shall be adopted henceforth.

Clearly, one is only interested in the steady solution to (29), which actually recovers (27) as the time derivative vanishes. Equation 29 is just a model equation, but it does fulfill its role regarding the hyperbolic character. In its final form, the adjoint equation (29) shares the same characteristics as the Euler equation (30), but for a reversal of sign of the characteristic velocities. Quite different from the latter, though, the adjoint equation is linear. For neither the flux Jacobian $\mathbf{A}$ nor the matrix $\mathbf{B}$ depend on Ψ.

An immediate consequence of the adjoint equation linearity is the fact that singularities can only travel along characteristics, and they must be borne by the Cauchy data [50,51, Sect. 3.8]. Then, there is no reason for the Ψ to develop jump discontinuities anywhere in the solution domain, unless they are brought in by the boundary or initial conditions. That is the case of the sonic line singularity $(u - a) = 0$, which is reported in the literature [52].

In a strict sense, the shock wave divides the flow field into two solution domains for the adjoint equation, and they are connected by the continuity condition (22). A numerical method, however, may handle that connection in either of two ways: by enforcing (22) explicitly at the shock location, as in a shock-fitting approach; or, alternatively, by treating the shock as a discontinuity within the domain, in a fashion that is quite similar to the shock-capturing methods of flow simulation. The second option is usually preferred over the first one because it should, in principle, involve a simpler implementation of the method.

In any case, the behavior of the co-state variables Ψ through the shock wave can be analyzed with the aid of the adjoint equation (29) and of the term (c) in the total variation (24). That term imposes a condition on a control volume of infinitesimal thickness that encloses the shock wave, and it implies

$$\left(g^+ - g^-\right) = \Psi_s^T \cdot \left[\left(S'\mathbf{H}\right)^+ - \left(S'\mathbf{H}\right)^-\right], \tag{31}$$

where the symbols $(\,)^-$ and $(\,)^+$ stand for conditions on the upstream and downstream sides of the shock wave, x_s^- and x_s^+ respectively. On substituting the identity $\mathbf{H} = \mathbf{BQ}$ in (31), it yields

$$\left(g^+ - g^-\right) = \Psi_s^T \cdot \left[\left(S'\mathbf{BQ}\right)^+ - \left(S'\mathbf{BQ}\right)^-\right]. \tag{32}$$

On accounting for the fact that Ψ is assumed to be continuous through the shock wave, and on making use of the stationary form of the adjoint equation (29), the RHS of (32) can be written as

$$\left(g^+ - g^-\right) = -\left(S\frac{\partial \Psi}{\partial x}^T \mathbf{AQ}\right)^+ + \left.\frac{\partial g}{\partial Q}\right|_+^T \cdot \mathbf{Q}^+ + \left(S\frac{\partial \Psi}{\partial x}^T \mathbf{AQ}\right)^- - \left.\frac{\partial g}{\partial Q}\right|_-^T \cdot \mathbf{Q}^-. \tag{33}$$

Finally, by making use of the identity $\mathbf{F} = \mathbf{AQ}$ and of the R–H condition across the shock wave $(S\mathbf{F})^+ = (S\mathbf{F})^-$, one gets

$$\left(g^+ - g^-\right) = + \left.\frac{\partial g}{\partial Q}\right|_+^T \cdot \mathbf{Q}^+ - \left.\frac{\partial g}{\partial Q}\right|_-^T \cdot \mathbf{Q}^- - \left(\left.\frac{\partial \Psi}{\partial x}\right|_+^T - \left.\frac{\partial \Psi}{\partial x}\right|_-^T\right) \cdot (S\mathbf{F})^+ . \tag{34}$$

The above equation relates the co-state variables to the nature of the objective function $g(V)$: If this function is homogeneous of degree one with respect to $\mathbf{Q}$, then the first derivative of Ψ must be continuous through the shock wave, which, in turn, implies that the co-state variables are continuous of class $\mathcal{C}^1$. If that is not the case, then the first derivative of Ψ undergoes a jump discontinuity through the shock, the intensity of which is given by (34).

3.1 Boundary conditions for the adjoint equation

As was discussed above, the steady solution to the adjoint equation (29) should allow one to eliminate the integral (b) from the the total variation δI (24). Boundary conditions for that equation can be derived from (a), which would then be the only remaining term in the variation that involves $\delta\mathbf{Q}$:

$$\left[\delta\mathbf{Q}^T S\mathbf{A}^T \Psi\right]_0^l = \left(\delta\mathbf{Q}^T S\mathbf{A}^T \Psi\right)_l - \left(\delta\mathbf{Q}^T S\mathbf{A}^T \Psi\right)_0 . \tag{35}$$

Besides being consistent with the PDE (29), these boundary conditions should ideally eliminate the above terms from the total variation δI, thus suppressing its dependence on the state vector variation.

The terms in (35) involve both state and co-state vectors at the boundaries of the flow domain. The variation of the state vector is governed by flow boundary conditions, which impose relations among its components, δq_i. Boundary conditions for the ψ_i should follow the same reasoning, given the similarity between the PDEs. Only the sign reversal of the characteristics leads to boundary conditions that are complementary to those of the primal problem. As will be shown below, it is this relation between primal and dual boundary conditions that ultimately enables one to eliminate all terms in (35) from δI.

The rationale behind the adjoint boundary conditions is best conveyed on replacing the terms $\mathbf{A}$ and $\delta\mathbf{Q}$ by their counterpart in (16), $\mathbf{A}\delta\mathbf{Q} = \delta\mathbf{F}$. That leads to an alternative form for (35)

$$\left[S\delta\mathbf{F}^T \cdot \Psi\right]_0^l = \left(S\delta\mathbf{F}^T \cdot \Psi\right)_l - \left(S\delta\mathbf{F}^T \cdot \Psi\right)_0 . \tag{36}$$

Both terms on the RHS of (36) involve scalar products between the co-state variables Ψ and flux variations $\delta\mathbf{F}$ at the boundaries of the flow domain. The nullity of these products implies the Ψ should be orthogonal to all realizable flux variations $\delta\mathbf{F}$—the term *realizable*, here, refers to those that satisfy the equations governing the flow. This way, the co-state variables may be thought of as *generalized constraint forces* that impose mass, momentum and energy conservation on all flux variations at the boundaries. Therein lies a possible interpretation of the Ψ and their boundary conditions.

Given the way the variational problem is constructed, the same interpretation of the co-state variables should actually hold throughout the flow domain, as opposed to just the boundaries. It suffices to recall that $\delta\mathbf{Q}$ (term b in (24)) must be realizable, but it is otherwise arbitrary. Hence the only way to make that integral null is to take Ψ as the solution to the adjoint equation. Within this framework, the rationale behind the method is akin to minimizing the virtual work of the generalized constraint forces Ψ, thereby ensuring that the system trajectory in state space is fully realizable. Although these ideas do not have any direct bearing on the algebra, they can guide us through the derivation that follows.

As was mentioned above, the flux Jacobian of the adjoint equation (29) is the same as that of the Euler equation, except for the transposition and sign reversal. Therefore, it has the same characteristic velocities, but with opposite

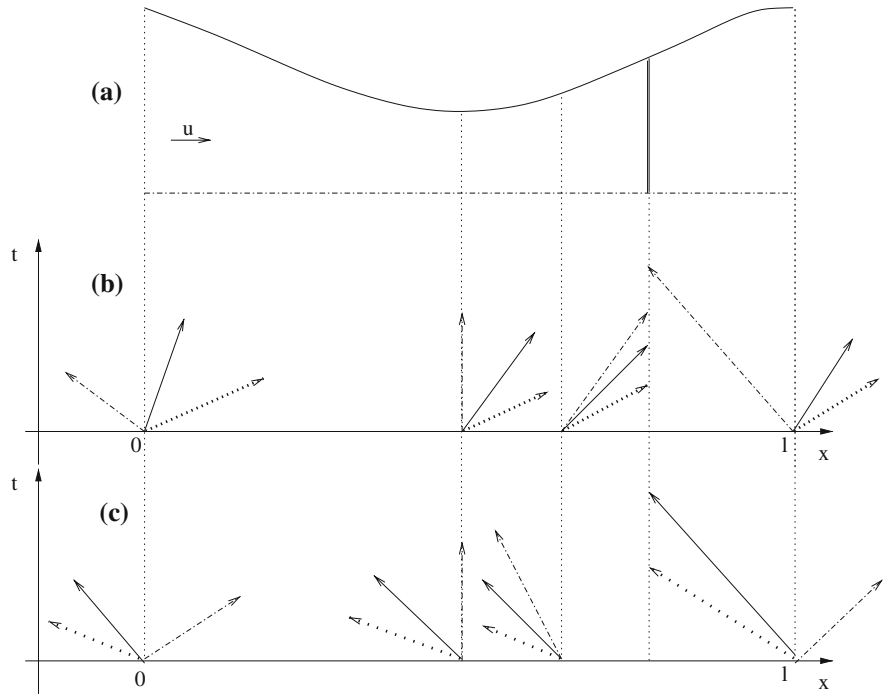

Fig. 1 Quasi-1D nozzle flow. **a** Physical space, with sonic conditions at the throat and a normal shock wave in the divergent portion—represented by a *double solid line*. **b** Flow characteristics: *solid line* , u; *dotted line*, $(u+c)$; *dash–dot line*, $(u-c)$. **c** Adjoint characteristics: *solid line*, $-u$; *dotted line*, $-(u+c)$; *dash–dot line*, $-(u-c)$

signs: $-u$, $-(u+c)$, $-(u-c)$. Figure 1 illustrates the situation in an arbitrary quasi-1D nozzle. It depicts both primal and dual characteristics under different flow conditions.

Boundary conditions for supersonic flow at an entrance impose constant values on all state vector components q_i. That, in turn, implies that all of their variations are zero: $\delta q_i = 0$. Since the adjoint characteristics run in the opposite direction, the co-state variables should not be imposed there. As a result, the corresponding ψ_i boundary conditions are free, but the nullity of $\delta\mathbf{Q}$ ensures that $\left(\delta\mathbf{Q}^T S\mathbf{A}^T\Psi\right)_0 = 0$ at that boundary. To put it another way, $\delta\mathbf{Q} = 0 \Rightarrow \delta\mathbf{F} = 0$. Therefore, the orthogonality condition holds for any vector Ψ.

At a supersonic exit, no boundary conditions are imposed on $\mathbf{Q}$ and, thus, all variations δq_i are left unspecified. However, the reverse sign of the adjoint characteristics implies that all ψ_i must be imposed boundary conditions there. In principle, these could be picked so as to ensure that $\left(\delta\mathbf{Q}^T S\mathbf{A}^T\Psi\right)_l = 0$ at that boundary. This can be accomplished by simply imposing homogeneous boundary conditions on the co-state variables: $\psi_i = 0$. Here again, the unspecified δq_i imply that all $\delta\mathbf{F}$ are allowed, hence the need for imposing homogeneous boundary conditions on Ψ.

Boundary conditions for subsonic flow need more careful consideration. In that case there are characteristics that run both ways, upstream and downstream, and flux variations at the boundaries are constrained. The idea is to evaluate the realizable $\delta\mathbf{F}$ in each case, and then to find out the Ψ that are orthogonal to these variations. For simplicity, the algebra will be done in terms of $\delta\mathbf{Q}$, but the correspondence between the two forms is immediate, since it is given by a simple relation, $\delta\mathbf{F} = \mathbf{A}\delta\mathbf{Q}$.

At a subsonic entrance the boundary conditions impose the flow direction, along with constant values for the stagnation pressure P_o and temperature T_o, thereby implying that their variations are zero. Owing to the quasi-1D character of this particular application, only the latter two quantities are considered. On writing them in terms of the state variables q_i, one gets

$$T_o = \frac{(\gamma-1)}{2\gamma R q_1^2}\left[(1-\gamma)q_2^2 + 2\gamma q_1 q_3\right], \tag{37}$$

$$P_o = \frac{(\gamma-1)\left(2q_1q_3-q_2^2\right)}{2q_1}\left[1+\frac{q_2^2}{\gamma(2q_1q_3-q_2^2)}\right]^{\frac{\gamma}{\gamma-1}}, \tag{38}$$

and, on computing their first variations and collecting the δq_i, they yield

$$\delta T_o = \frac{\gamma-1}{\gamma R\rho}\left\{\gamma\delta q_3 - u(\gamma-1)\delta q_2 - \left[e_t\gamma-(\gamma-1)u^2\right]\delta q_1\right\}, \tag{39}$$

$$\delta P_o = \frac{1}{2(\gamma-1)u^2-4\gamma e_t}\left\{\left[1+\frac{u^2}{\gamma(2e_t-u^2)}\right]^{\frac{\gamma}{\gamma-1}}\left[u(u\delta q_1-2\delta q_2)\left(-2e_t(\gamma-2)\gamma\right.\right.\right.$$
$$\left.\left.\left.+(\gamma-1)^2u^2\right)+2\left(-2e_t(\gamma-1)\gamma+\left(1+(\gamma-1)\gamma\right)u^2\right)\delta q_3\right]\right\}. \tag{40}$$

Finally, on imposing the condition of zero variation on T_o and P_o, in (39) and (40), and on solving them as a set for two out of the three components δq_i, one obtains two linear expressions involving these variations.

$$\delta q_2 = \left[\frac{(2+\gamma(\gamma-1))u}{2}-\frac{e_t\gamma(\gamma-1)}{u}\right]\delta q_1, \quad \delta q_3 = \left[\frac{(\gamma-1)^2u^2}{2}-e_t\gamma(\gamma-2)\right]\delta q_1. \tag{41}$$

The two equations of (41) correspond to the physical boundary conditions that are imposed on T_o and P_o. They give the variations of two state variables in terms of the third, which is left unspecified. For simplicity, we have chosen to write δq_2 and δq_3 in terms of δq_1, but the relations can be expressed differently. In any case, these equations represent the realizable $\delta\mathbf{Q}$ at a subsonic entrance, and the $\delta\mathbf{F}$ thereof.

The corresponding values of Ψ at such a boundary are obtained by the orthogonality condition (36). On computing the scalar product $\left(\delta\mathbf{Q}^T S\mathbf{A}^T\Psi\right)_0$, substituting (41) for δq_2 and δq_3 and collecting like terms, one obtains the following equation:

$$\left(\delta\mathbf{Q}^T S\mathbf{A}^T\Psi\right)_0 = -\left[2\gamma(\gamma-1)e_t+\left(-2+\gamma-\gamma^2\right)u^2\right]\left[2\psi_1+2e_t\gamma\psi_3+u\left(2\psi_2-u(\gamma-1)\psi_3\right)\right]\frac{\delta q_1}{4u}. \tag{42}$$

Since the variation δq_1 must be left unspecified, the only way one can make the product vanish is to assume its coefficient goes to zero. That assumption gives rise to the following boundary condition, to be imposed on the adjoint PDE at a subsonic entrance

$$\psi_1 = \left[\frac{(\gamma-1)u^2}{2}-e_t\gamma\right]\psi_3 - u\psi_2. \tag{43}$$

It is worth noting that, on imposing $\psi_1 = \psi_1(\psi_2, \psi_3)$, one is effectively constraining one degree of freedom (DOF), which corresponds to the adjoint characteristic that runs into the flow domain. Two remaining DOFs are preserved, and they correspond to the two characteristics that convey information from the domain to the entrance boundary.

Adjoint boundary conditions for the subsonic exit are derived in a similar way. The physical boundary condition implies that the static pressure, alone, be specified at the subsonic exit. On casting the static pressure (11) in terms of the q_i, it reads

$$P = (\gamma-1)\left(q_3-\frac{q_2^2}{2q_1}\right). \tag{44}$$

The corresponding variation is given by

$$\delta P = \frac{(\gamma-1)}{2}\left(2\delta q_3 - 2u\delta q_2 + u^2\delta q_1\right). \tag{45}$$

Then, on imposing $\delta P = 0$ on the above equation and on solving it for δq_3 in terms of δq_1 and δq_2, one gets

$$\delta q_3 = u\delta q_2 - \frac{u^2}{2}\delta q_1, \tag{46}$$

Table 1 Summary of boundary conditions for the adjoint equation (29)

Flow direction	Regime	State vector variation $\delta\mathbf{Q}$	Co-state vector Ψ boundary conditions
Entrance	Subsonic	$\delta q_2(\delta q_1)$, $\delta q_3(\delta q_1)$ Eqs. 41	$\psi_1(\psi_2, \psi_3)$ Eq. 43
	Supersonic	$\delta q_i = 0$	no b.c. on ψ_i
Exit	Subsonic	$\delta q_3(\delta q_1, \delta q_2)$ Eq. 46	$\psi_1(\psi_3)$, $\psi_2(\psi_3)$ Eqs. 48
	Supersonic	δq_i are free	$\psi_i = 0$

where both δq_1 and δq_2 are not specified. Equation (46) gives the realizable $\delta\mathbf{Q}$ and the corresponding $\delta\mathbf{F}$ that represent the subsonic exit boundary condition. On computing the scalar product between the latter and Ψ, substituting (46) for δq_3 and collecting like terms, it yields

$$\left(\delta\mathbf{Q}^T S\mathbf{A}^T\Psi\right)_l = \left[2\left(\psi_1 + e_t\gamma\psi_3 + 2u\psi_2\right) + (3-\gamma)u^2\psi_3\right]\frac{\delta q_2}{2} + \left[-2\left(e_t\gamma\psi_3 + u\psi_2\right) + (\gamma-2)u\psi_3\right]\frac{u\delta q_1}{2}. \tag{47}$$

Since δq_1 and δq_2 are free, the only way to meet the orthogonality condition is to assume their coefficients are zero. On doing so, one gets a set of two equations on the ψ_i. On solving the set for ψ_1 and ψ_2 in terms of ψ_3, it yields

$$\psi_1 = \left[\gamma e_t - \frac{(\gamma-1)u^2}{2}\right]\psi_3, \quad \psi_2 = \left[\frac{(\gamma-2)u}{2} - \frac{\gamma e_t}{u}\right]\psi_3. \tag{48}$$

Similar to (43), the above relations represent the adjoint boundary conditions to be imposed at the subsonic exit boundary. They constrain two DOFs, which correspond to the two adjoint characteristics that run into the flow domain. One remaining DOF is preserved, and it corresponds to the single characteristic that conveys information from the domain to the exit boundary. Table 1 summarizes the above results, and presents the adjoint boundary conditions that correspond to each flow regime.

4 Inverse design application

Given the general structure of the adjoint equation (29), one may illustrate the above results with a simple inverse-design application. The measure of merit is defined as the mean square error of the actual pressure distribution P with respect to a target distribution P_D

$$I = \frac{1}{2}\int_0^l (P - P_D)^2 \, \mathrm{d}x. \tag{49}$$

The integrand of functional (49) yields the following gradient with respect to the state variables $\mathbf{Q}$:

$$\frac{\partial g}{\partial Q} = (P - P_D)(\gamma - 1)\begin{pmatrix} \frac{u^2}{2} \\ -u \\ 1 \end{pmatrix}. \tag{50}$$

This expression gives rise to the non-homogeneous term of the adjoint PDE (29). Moreover, it can be shown that g is not a homogeneous function of degree one on $\mathbf{Q}$. According to (34), it implies that the ψ_i derivatives should undergo jump discontinuities through a shock wave, and these should meet the relation

$$\left(\frac{\partial\Psi}{\partial x}\bigg|_+^T - \frac{\partial\Psi}{\partial x}\bigg|_-^T\right)\cdot(S\mathbf{F})^+ = \frac{1}{2}\left[P^2 - P_D^2\right]_-^+. \tag{51}$$

If the geometry of the nozzle can be cast in the generic form $S = S(x; a_k)$, for $k = 1, \ldots, n$, and Ψ satisfies (29). Then the sensitivity gradient is obtained directly from (26),

$$\frac{\partial I}{\partial a_k} = \left[\mathbf{F}^T\cdot\Psi\frac{\partial S}{\partial a_k}\right]_0^l - \int_0^l\left[\mathbf{F}^T\cdot\frac{\partial\Psi}{\partial x}\frac{\partial S}{\partial a_k} + \mathbf{H}^T\cdot\Psi\frac{\partial S'}{\partial a_k}\right]\mathrm{d}x. \tag{52}$$

On picking a particular family of functions to represent $S(x;a_k)$, say polynomials, one is able to evaluate the derivatives $\partial S/\partial a_k$ in closed form.

In any case, the adjoint PDE is solved on the basis of a steady converged solution to the Euler equations, so that all of its coefficients are known beforehand and remain constant throughout the solution of the adjoint equation. The inverse-design application comprises a sequence of: flow simulation for a given geometry, adjoint solution, evaluation of the sensitivity gradient and the procedure of gradient-based optimization. The last step leads to a new geometry and, thus, closes the cycle. The cycles are repeated until a local extremum of the measure of merit is reached, within a prescribed accuracy level.

4.1 Analytic solutions

Before turning our attention to the actual applications, it is instructive to consider analytic solutions to a few particular cases, since that would provide a necessary validation test. Giles and Pierce [39] have proposed a very elegant approach to the analytic adjoint solutions. It is based on the Green's function and relies on an extensive investigation those authors have conducted on the properties of the adjoint equations, their duality and solution behavior [11,45,52]. For the sake of space, this section brings only a brief summary of the relevant material. The reader is referred to the original paper [39] for further details.

The co-state variables are computed directly by means of Eq. 3.2 from [39, p. 332]. In our notation and sign convention that equation becomes:

$$\Psi^T = -\big(\delta I_1(\xi)|\delta I_2(\xi)|\delta I_3(\xi)\big)\cdot\big(f_1(\xi)|f_2(\xi)|f_3(\xi)\big)^{-1}. \tag{53}$$

Here the $f_i(\xi)$ represent three linearly independent source vectors, each corresponding to a different mode of perturbation to the flow solution. Together, these modes comprise a set of small disturbances: in the specific-mass flow rate ($m \equiv \rho u$) at constant stagnation enthalpy (h_o) and stagnation pressure (P_o); in h_o at fixed P_o and Mach number (M); and in P_o at constant h_o and M—in short, $\delta m|_{h_o,P_o}$, $\delta h_o|_{P_o,M}$ and $\delta P_o|_{h_o,M}$, respectively. The $\delta I_i(\xi)$ represent the effect of each perturbation mode on the objective functional, and the inverse matrix in (53) is given by:

$$\big(f_1(\xi)|f_2(\xi)|f_3(\xi)\big)^{-1} = \begin{pmatrix} \dfrac{P+\rho u^2}{2P} & \dfrac{-\rho u}{P} & \dfrac{P+\rho u^2}{2Ph_o} \\ \dfrac{-h_o}{S\rho u} & 0 & \dfrac{1}{S\rho u} \\ \dfrac{-uP_o}{2PS} & \dfrac{P_o}{PS} & \dfrac{-uP_o}{2PSh_o} \end{pmatrix}. \tag{54}$$

Two cases are considered here, for the purpose of verifying the adjoint boundary and internal conditions: the *supersonic flow*, and the *shocked flow*. The former was specifically designed to test the corresponding adjoint boundary conditions (Table 1). Whereas the latter allows for checking two sets of conditions: those pertaining to the subsonic boundaries, and the ones that are imposed on the shock-wave.

The *supersonic flow* case, with (49) for the measure of merit, implies that the perturbation modes $\delta I_i(\xi)$ are given by simple integrals.

$$\begin{cases} \delta I_1(\xi) = \int_\xi^l \dfrac{1}{S}\dfrac{\partial g}{\partial m}\bigg|_{h_o,P_o} \mathrm{d}x = -\int_\xi^l \dfrac{(P-P_D)u}{S(1-M^2)}\mathrm{d}x \\ \delta I_2(\xi) = \int_\xi^l \dfrac{\partial g}{\partial h_o}\bigg|_{P_o,M} \mathrm{d}x = 0 \\ \delta I_3(\xi) = \int_\xi^l \dfrac{\partial g}{\partial P_o}\bigg|_{h_o,M} \mathrm{d}x = \int_\xi^l \dfrac{(P-P_D)P}{P_o}\mathrm{d}x \end{cases}, \tag{55}$$

where, as is pointed out by Giles and Pierce [39], the zero result for δI_2 reflects the fact that the pressure is constant at fixed M and P_o.

The *shocked flow* case also involves Eqs. 53 and 54, and the $\delta I_i(\xi)$ involve integrals that are similar to those in (55). However, the limits of integration change and the perturbation modes are evaluated in a quite different fashion. For the sonic condition at the throat and the presence of a shock wave must be accounted for. The latter entails a displacement of the shock front, δx_s, which is given by Eq. 7.1 from [39, p. 337].

$$c_2 = c_1 f(M_1) + P_{o1} f'(M_1) \left[\frac{\mathrm{d}M}{\mathrm{d}x}\delta x_s + \frac{a_1}{S} \left.\frac{\partial M}{\partial m}\right|_{h_o,P_o} \right]_{x_s^-}, \tag{56}$$

where the quantity $f(M_1)$ represents the ratio between stagnation pressures on the upstream (x_s^-) and downstream (x_s^+) sides of a normal shock wave: $f(M_1) = P_{o2}/P_{o1}$. Expressions for the derivatives $\partial M/\partial m|_{h_o,P_o}$ and $f'(M_1)$ are also obtained from normal shock relations:

$$\left.\frac{\partial M}{\partial m}\right|_{h_o,P_o} = \frac{M}{m}\left\{\frac{1+\left[(\gamma-1)/2\right]M^2}{1-M^2}\right\}, \tag{57}$$

$$f'(M_1) = \frac{-4\gamma\left(M_1^2-1\right)^2}{2(\gamma+1)M_1+\left(\gamma^2-1\right)M_1^3}\left\{\frac{(\gamma+1)^2M_1^2}{\left[2+(\gamma-1)M_1^2\right]\left[1+\gamma\left(2M_1^2-1\right)\right]}\right\}^{\frac{\gamma}{\gamma-1}}. \tag{58}$$

The flow domain is divided into three portions for estimating the first perturbation mode $\delta I_1(\xi)$. In each one of them, the functional variation is given by a different expression:

a. From the entrance section to the nozzle throat (x_t), $0 \le \xi < x_t$, $\delta I_1(\xi)$ is given by Eq. (6.1) from [39, p. 338],

$$\delta I_1(\xi) = -\int_0^{\xi} \left.\frac{1}{S}\frac{\partial g}{\partial m}\right|_{h_o,P_o} \mathrm{d}x. \tag{59}$$

b. From the throat section to the normal shock position (x_s), $x_t \le \xi < x_s$, $\delta I_1(\xi)$ is given by the integral [39, Sect. 7.2.1, p. 338]

$$\delta I_1(\xi) = \int_{\xi}^{x_s} \left.\frac{1}{S}\frac{\partial g}{\partial m}\right|_{h_o,P_o} \mathrm{d}x + \int_{x_s}^{1} \left.\frac{a_2}{S}\frac{\partial g}{\partial m}\right|_{h_o,P_o} \mathrm{d}x + \int_{x_s}^{1} c_2 \left.\frac{\partial g}{\partial P_o}\right|_{h_o,M} \mathrm{d}x - [g(V)]_{x_s^-}^{x_s^+}\delta x_s, \tag{60}$$

and the coefficients a_2 and c_2 are obtained from the set

$$\begin{cases} a_2 + S\left.\frac{m}{Po}\right|_{x_s^+} c_2 = 1 \\ \left.\frac{-u}{S(1-M^2)}\right|_1 a_2 + \left.\frac{P}{P_o}\right|_1 c_2 = 0 \end{cases}, \tag{61}$$

where use was made of the fact that $(\partial m/\partial P_o|_{h_o,M}) = m/P_o$. The shock-wave displacement, δx_s, is estimated by means of (56)–(58) with $a_1 = 1$ and $c_1 = 0$.

c. From the shock-wave position to the exit section, $x_s \le \xi \le 1$, $\delta I_1(\xi)$ is given by [39, Sect. 7.2.2, p. 339]

$$\delta I_1(\xi) = \int_{x_s}^{\xi} \left.\frac{a_2}{S}\frac{\partial g}{\partial m}\right|_{h_o,P_o} \mathrm{d}x + \int_{\xi}^{1} \left.\frac{a_3}{S}\frac{\partial g}{\partial m}\right|_{h_o,P_o} \mathrm{d}x + \int_{x_s}^{1} c \left.\frac{\partial g}{\partial P_o}\right|_{h_o,M} \mathrm{d}x - [g(V)]_{x_s^-}^{x_s^+}\delta x_s, \tag{62}$$

where the coefficients a_2, a_3 and c satisfy the set

$$\begin{cases} a_3 - a_2 = 1 \\ + a_2 + S\left.\frac{m}{Po}\right|_{x_s^+} c = 0 \\ \left.\frac{-u}{S(1-M^2)}\right|_1 a_3 + \left.\frac{P}{P_o}\right|_1 c = 0 \end{cases}, \tag{63}$$

and δx_s is given by (56)–(58) as usual, but now with $a_1 = c_1 = 0$.

As in the supersonic flow case, the second perturbation mode is identically zero throughout the whole domain, $\delta I_2(\xi) = 0$, [39, Sect. 7.3, p. 339]. The third perturbation mode δI_3, on the other hand, requires that the nozzle be divided into two portions:

a. From the entrance section to the shock wave position, $0 \leq \xi < x_s$, δI_3 is given by [39, Sect. 7.4.1, p. 340]

$$\delta I_3(\xi) = \int_{x_s}^{1} \frac{a_2}{S}\frac{\partial g}{\partial m}\bigg|_{h_o,P_o} \mathrm{d}x + \int_{\xi}^{x_s} \frac{\partial g}{\partial P_o}\bigg|_{h_o,M} \mathrm{d}x + \int_{x_s}^{1} c_2 \frac{\partial g}{\partial P_o}\bigg|_{h_o,M} \mathrm{d}x - [g(V)]_{x_s^-}^{x_s^+}\,\delta x_s, \tag{64}$$

with the coefficients a_2 and c_2 as given by the set

$$\begin{cases} a_2 + S\,\frac{m}{Po}\big|_{x_s^+}\, c_2 = S\,\frac{m}{Po}\big|_{x_s^-} \\ \frac{-u}{S(1-M^2)}\Big|_1 a_2 + \quad \frac{P}{P_o}\Big|_1 c_2 = \quad 0 \end{cases}, \tag{65}$$

and δx_s is obtained from (56)–(58) with $a_1 = 0$ and $c_1 = 1$.

b. Finally, from the normal shock position to the exit section, $x_s \leq \xi < 1$, δI_3 is given by [39, Sect. 7.4.2, p. 340]

$$\delta I_3(\xi) = \int_{x_s}^{1} \frac{a}{S}\frac{\partial g}{\partial m}\bigg|_{h_o,P_o} \mathrm{d}x + \int_{x_s}^{\xi} c_2 \frac{\partial g}{\partial P_o}\bigg|_{h_o,M} \mathrm{d}x + \int_{\xi}^{1} c_3 \frac{\partial g}{\partial P_o}\bigg|_{h_o,M} \mathrm{d}x - [g(V)]_{x_s^-}^{x_s^+}\,\delta x_s, \tag{66}$$

where a, c_2 and c_3 are given by the set

$$\begin{cases} + c_3 \quad - c_2 = 1 \\ a \quad + S\,\frac{m}{Po}\big|_{x_s^+}\, c_2 = 0 \\ \frac{-u}{S(1-M^2)}\Big|_1 a + \frac{P}{P_o}\Big|_1 c_3 \quad = 0 \end{cases}, \tag{67}$$

and the shock-wave displacement is determined by (56)–(58) with $a_1 = c_1 = 0$.

For our purposes, the most important feature of the rationale outlined above is that it does not formulate adjoint boundary conditions explicitly. Instead, it imposes boundary and realizability conditions on the flow-field perturbations [39]. Then the behavior of the adjoint variables at the boundaries becomes an implicit result of those conditions, and the same holds true for their behavior across the shock wave. That is an essential difference between Giles and Pierce's approach and ours. Yet, in principle, the adjoint solutions are expected to agree. For both approaches rely on the central concept of realizable flow-field variations.

In what follows, the analytical solutions to the *supersonic* and the *shocked* transonic cases will be presented along with the corresponding numerical ones, for comparison. They should enable one to verify the adjoint boundary and internal conditions, and also provide a means of assessing the accuracy of the numerical method.

4.2 Numerical solutions

A single piece of code was written to tackle the whole procedure. The Euler equations are explicitly integrated by a split-flux algorithm that is based on the Modified Steger–Warming method. The method was chosen for providing the right amount of numerical dissipation at shock waves, so as not to interfere significantly with the convergence of the adjoint solution. Owing to the linearity of the adjoint problem and the lack of a physical flux definition, the method of choice for solving that PDE was the implicit Beam and Warming three-point backward scheme. Artificial dissipation was later added to that scheme, so as to give it greater stability [49, Sect. 18.1], [53]. The conjugate-gradient algorithm was picked as the optimization method [54, Chap. 3].

It must be noted that the code does not enforce continuity of the ψ_i, strictly, across shock waves. Instead, condition (22) comes out as a result of the numerical scheme and, in particular, its artificial dissipation, as in a "shock-capturing" approach. The behavior of the first derivatives of ψ_i through the wave front—Eqs. (34) and (51)—is likewise determined by the numerical solution alone.

For the purpose of comparison, a program was dedicated to computing the sensitivity gradient by brute force, over the same search path that had been pursued by the adjoint. To that end, each control parameter is perturbed separately about its original value in the baseline configuration. The gradient is evaluated by a central

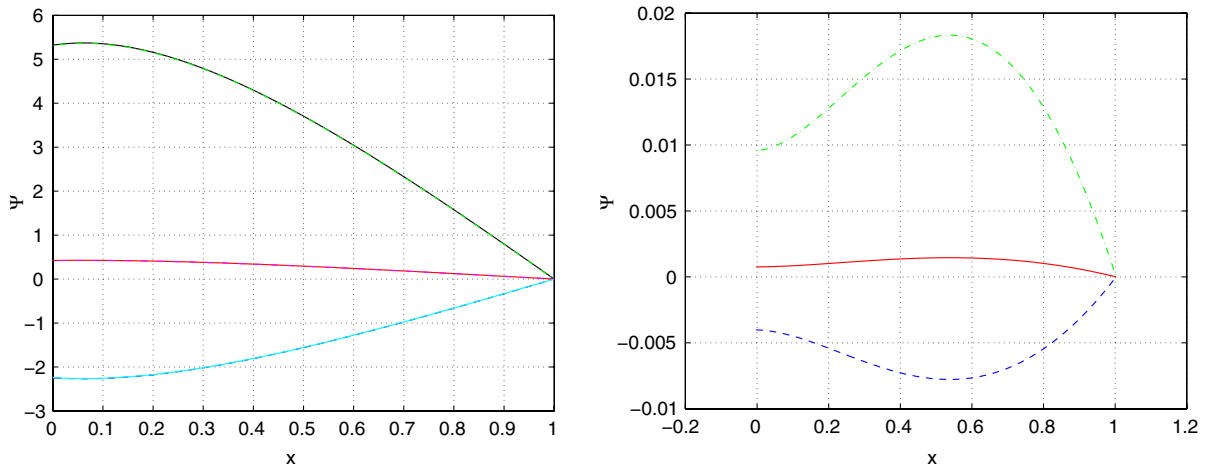

Fig. 2 Supersonic flow through a conical nozzle. Adjoint-equation solutions: *Left*, first cycle solution, analytic, $\psi_{i(a)}$, and numerical, ψ_i. *Right*, last cycle numerical solution, ψ_i. Numerical solutions on both sides: *dash–dot green line*, ψ_1; *dashed blue line* ψ_2; *solid red line* ψ_3. Analytic solutions on the left: $\psi_{1(a)}$, *black line*; $\psi_{2(a)}$, *cyan line*; $\psi_{3(a)}$, *magenta dash–dot line* (see online version for the colors)

finite-difference scheme to second-order accuracy. Two perturbed flow simulations are needed for each parameter, in addition to the baseline solution.

A number of test results is shown below, with the aim of illustrating the adjoint boundary and internal conditions. In all cases the target P_D is taken as the actual pressure distribution of a specific geometry, under given flow conditions. Although unusual in actual applications, the procedure ensures two essential features of the validation tests: not only is the target realizable, but its geometry is known beforehand. For simplicity, $S(x; a_k)$ is chosen as a second-order polynomial with three control parameters (a_0, a_1, a_2).

For all the inverse-design applications, accuracy levels have been set to the same value, for flow and adjoint solutions alike: convergence required the discretized equations residue to fall below 10^{-5} in magnitude. Both solvers have also made use of the same uniform mesh with 403 points—including ghost cells at the boundaries. On the other hand, the validation of our approach hinges on comparing numerical with analytical results. To that end, numerical adjoint solutions to the supersonic and transonic cases are plotted along with the corresponding analytical solutions in Figs. 2–left and 5–left, respectively.

A preliminary convergence study was also conducted. It has involved additional simulations of those cases on a finer uniform mesh with 603 points, while the accuracy levels were raised to 10^{-7}. Further tests have also been performed on two coarser meshes, one with 103 points and the other with 203. Accuracy levels were set to 10^{-5} for both cases. In view of the large variations in magnitude the ψ_i undergo, besides the occurrence of zeros and singularities, it seems more appropriate to estimate the numerical solutions relative error as

$$\varepsilon_i = \left| \frac{\psi_i - \psi_{i(a)}}{\max\left(\psi_{i(a)}\right) - \min\left(\psi_{i(a)}\right)} \right|; \tag{68}$$

here the $\psi_{i(a)}$ represent the analytical solutions for each adjoint variable, and ψ_i stand for the corresponding numerical results. The ε_i profiles are plotted in Figs. 3–left, 6–left and 7–left.

4.2.1 Supersonic case

The first set of results concerns the case of a diverging nozzle, under the following entrance boundary conditions[2]: $P_o = 15.65$, $T_o = 3.6$ and $M = 2.0$. The flow is supersonic throughout its length and, hence, the adjoint PDE

[2] Dimensionless stagnation properties are scaled to a reference state.

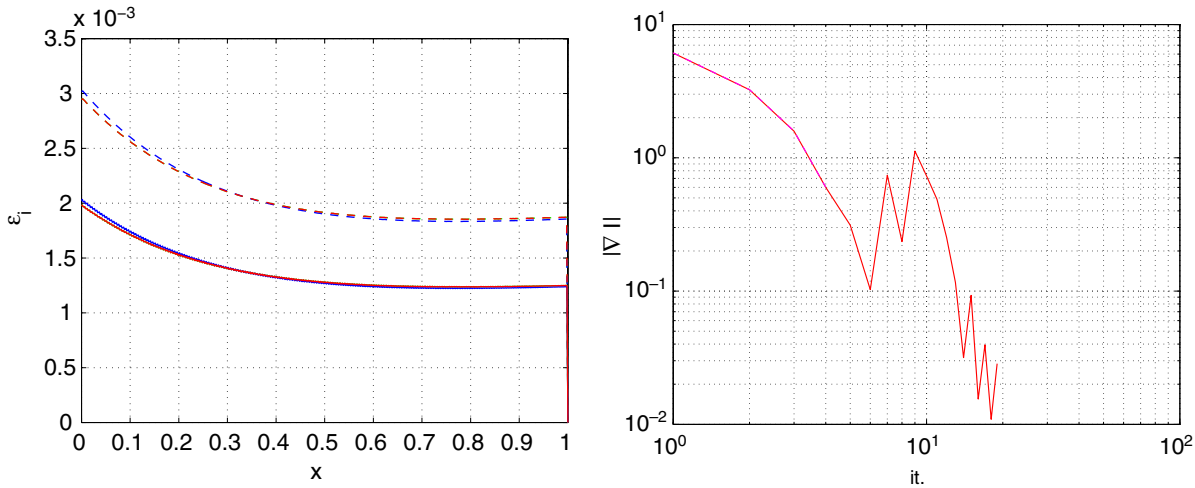

Fig. 3 Supersonic flow through a conical nozzle. *Left*, numerical adjoint solution relative error ε_i: *dashed lines*, coarse mesh (403 pts.); *solid lines*, fine mesh (603 pts.); ψ_1, *green*; ψ_2, *blue*; ψ_3 red. *Right*, gradient magnitude evolution: *solid line*, adjoint method; *dash–dot line*, brute force

is only subject to homogeneous boundary conditions at the nozzle exit (Table 1). The results were achieved in 18 inverse-design cycles.

Figure 2 shows the first and last solutions to the adjoint PDE—left and right sides, respectively. All Ψ components are given by smooth continuous functions that meet the homogeneous boundary condition at the domain exit. Besides, a comparison of both sides indicates that the ψ_i magnitude decrease as the minimum of I is approached. That is an intrinsic feature of the inverse-design applications, where a complete convergence should lead to a trivial solution to the adjoint PDE.

In particular, Fig. 2–left compares the analytic and numerical solutions on the 603 points mesh. It presents an excellent agreement between them. Indeed, Fig. 3–left shows that the relative error profiles ε_i of the latter remain below 2×10^{-3} throughout the domain (solid lines). On the other hand, the coarser-mesh solutions incur consistently larger errors: the 403 points mesh maximum error (dashed lines) is about 3×10^{-3}, the 203 and the 103 points meshes (not shown here) lead to maximum errors of 6×10^{-3} and 1.2×10^{-2}, respectively. Hence, the results seem to indicate the numerical solutions should recover the analytical ones, as the accuracy level is increased.

The evolution of the sensitivity gradient is depicted in Fig. 3–right, which presents its magnitude. For the first four cycles, the gradient was also evaluated by brute force. A comparison between these and the adjoint results shows a maximum error of less than 0.004%—which is also an important part of the validation procedure. Figure 4 presents the geometry changes and the corresponding pressure distributions. For clarity, though, only the first, the last and the target profiles are shown.

Although the pressure distribution approaches P_D and the final geometry is parallel to that of the target, they are not coincident. The result implies the method has reached a local minimum of Eq. 49, which corresponds to a different mass-flow rate. However, such behavior could be anticipated, since the latter quantity has not been constrained.

4.2.2 Transonic case

Next, results are shown of the second case, which concerns a converging–diverging nozzle under the following boundary conditions: $P_o = 20.0$ and $T_o = 2.08$ at the entrance, and the static back pressure $P_b = 15.0$. The results were obtained after 50 inverse design cycles. However, in this case the nozzle-throat area has been constrained to a fixed value, so as to limit the maximum mass-flow rate. The minimum area was imposed as an external constraint, by a relation among control parameters. The procedure was designed so as not to interfere with the gradient

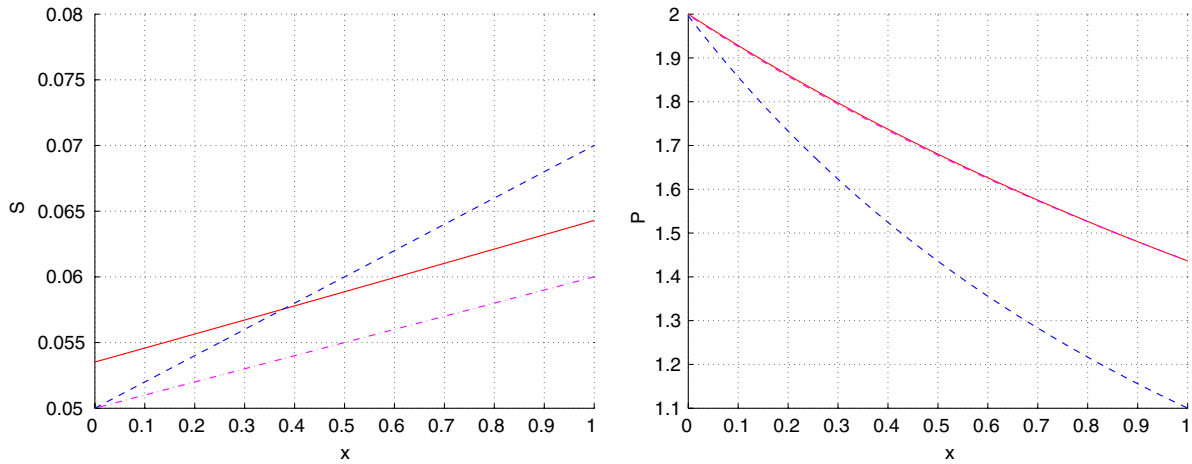

Fig. 4 Supersonic flow through a diverging nozzle. *Left*, geometry changes. *Right*, pressure evolution. *Dashed lines*, original curves; *solid lines*, final results; *dash–dot lines*, target profiles

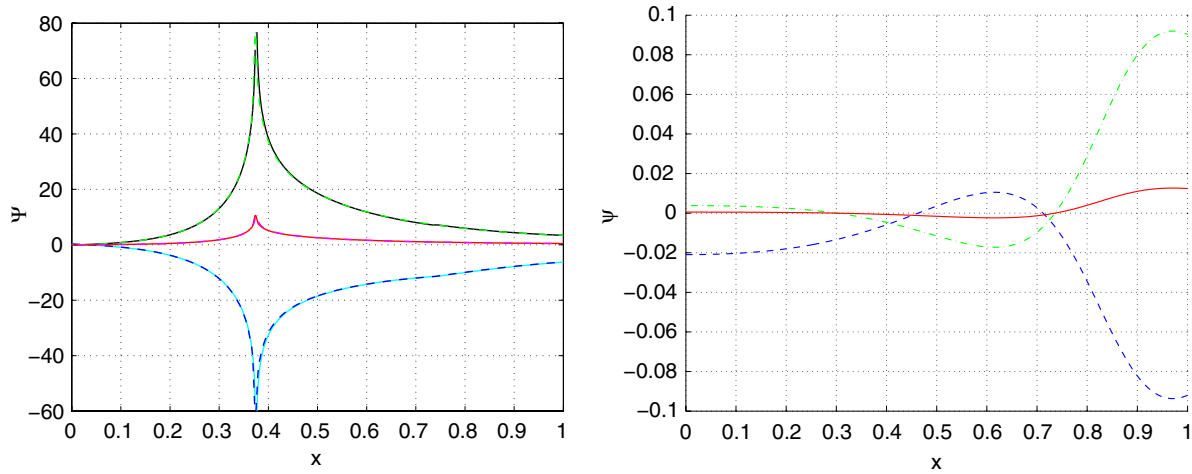

Fig. 5 Transonic flow through a convergent–divergent nozzle. Adjoint-equation solutions: *Left*, first-cycle solution, analytic, $\psi_{i(a)}$, and numerical, ψ_i. *Right*, last-cycle numerical solution, ψ_i. Numerical solutions on both sides: *dash–dot green line*, ψ_1; *dashed blue line* ψ_2; *solid red line* ψ_3. Analytic solutions on the left: $\psi_{1(a)}$, *black line*; $\psi_{2(a)}$, *cyan line*; $\psi_{3(a)}$, magenta *dash–dot line*

computation, but to change the search direction, only. The target pressure distribution P_D has been set for subsonic isentropic flow, which is feasible within the active constraints.

Figure 5 depicts the adjoint solutions in the first and last cycles—left and right sides, respectively. In this case, the first-cycle flow solution has a shock wave in the divergent portion of the nozzle (see Fig. 8–right), which implies that the throat flow is choked. Indeed, the first adjoint solution (Fig. 5–left) exhibits the characteristic singularity at the throat, as reported in the literature [39,52] and mentioned above.

A comparison between analytic and numerical adjoint solutions of the first inverse-design cycle is presented in Fig. 5–left, and it corresponds to the shocked-transonic flow. Here again, both solutions refer to the 603 points mesh. The relative error profiles ε_i are shown in Fig. 6–left for the domain as a whole, and in Fig. 6–right for the region between the shock front and the nozzle exit.

As can be seen, all ε_i spike at the throat region, $0.3 \leq x \leq 0.45$, for both solutions alike. In the remainder of the domain, though, the error in the coarser mesh (403 pts.) is bounded below 2×10^{-2}, while in the finer mesh (603 pts.) the error remains below 2.5×10^{-3}. At the boundaries, $\varepsilon_i \leq 8 \times 10^{-3}$ in the coarser mesh, while

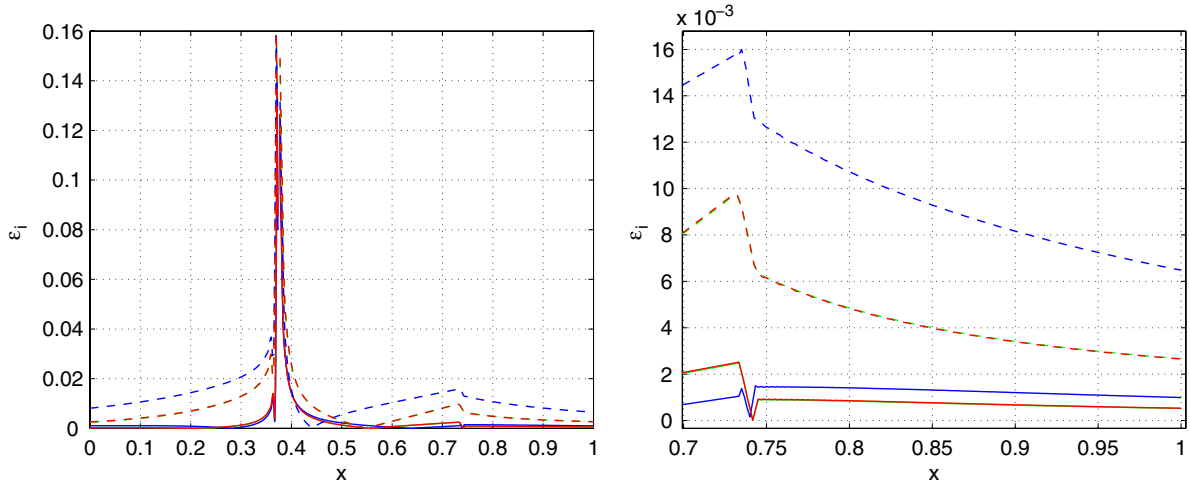

Fig. 6 Transonic flow through a convergent–divergent nozzle. *Left*, numerical adjoint-solution relative error ε_i: *dashed lines*, coarse mesh (403 points); *solid lines*, fine mesh (603 points); ε_1, *green*; ε_2, *blue*; ε_3, *red*. *Right*, detail of ε_i at the shock wave, curves are represented by the same line types and color code

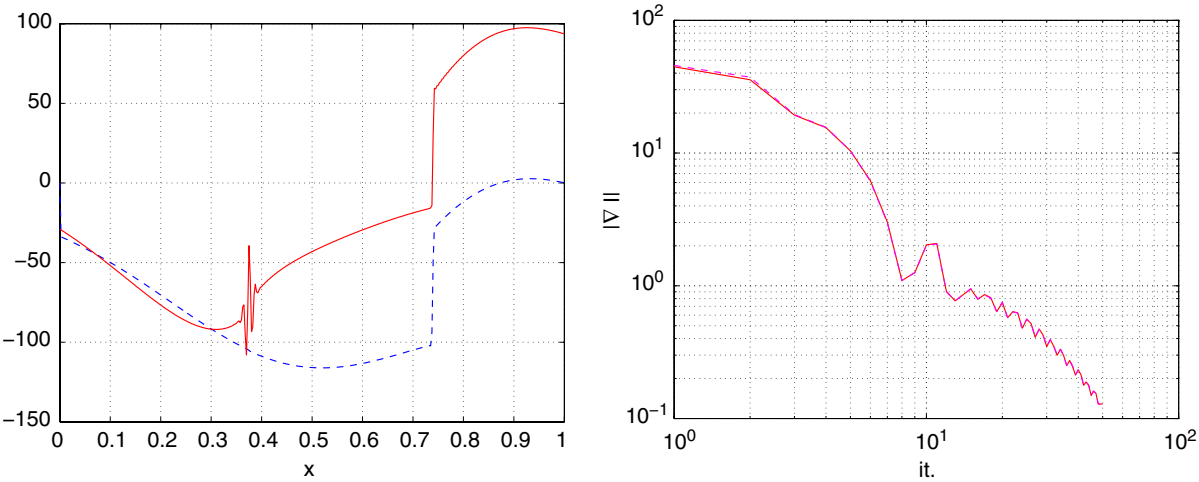

Fig. 7 Transonic flow through a convergent–divergent nozzle. *Left*, discontinuities in $\partial\Psi/\partial x$ at the first inverse-design cycle. In reference to Eq. 51: *solid line*, $(\partial\Psi/\partial x)^T \cdot (S\mathbf{F})$; *dashed line*, $g(\mathbf{V}) = (P^2 - P_D^2)/2$. *Right*, gradient magnitude evolution: *solid line*, adjoint method; *dash–dot line*, brute force

$\varepsilon_i \leq 1.2 \times 10^{-3}$ in the finer mesh. In particular at the shock wave, Fig. 6–left, the former shows $\varepsilon_i \leq 1.6 \times 10^{-2}$, whereas in the latter $\varepsilon_i \leq 2.5 \times 10^{-3}$. As for the two coarsest meshes (103 and 203 pts.), within the intervals $0 \leq x < 0.3$ and $0.45 < x \leq 1$, the error remains below 0.35 for the former and 0.30 for the latter. Such a consistent accuracy improvement with mesh refinement seems to indicate that the numerical solutions should indeed recover the analytical $\psi_{i(a)}$. Furthermore, it fully verifies the present formulation of the adjoint boundary and internal conditions.

On the other hand, Ψ does not seem to undergo any noticeable slope discontinuity at the shock location. The whole picture changes when one evaluates the terms in relation (51), which are plotted in Fig. 7–left. There one clearly sees that both the derivative of Ψ (solid line) and the pressure square difference (dashed line) undergo finite jumps at the location of the shock wave, precisely. Furthermore, the jumps appear to have the same strength, within the accuracy of these estimates: the former is about 73.58, whereas the latter is about 74.43. It must be noted that

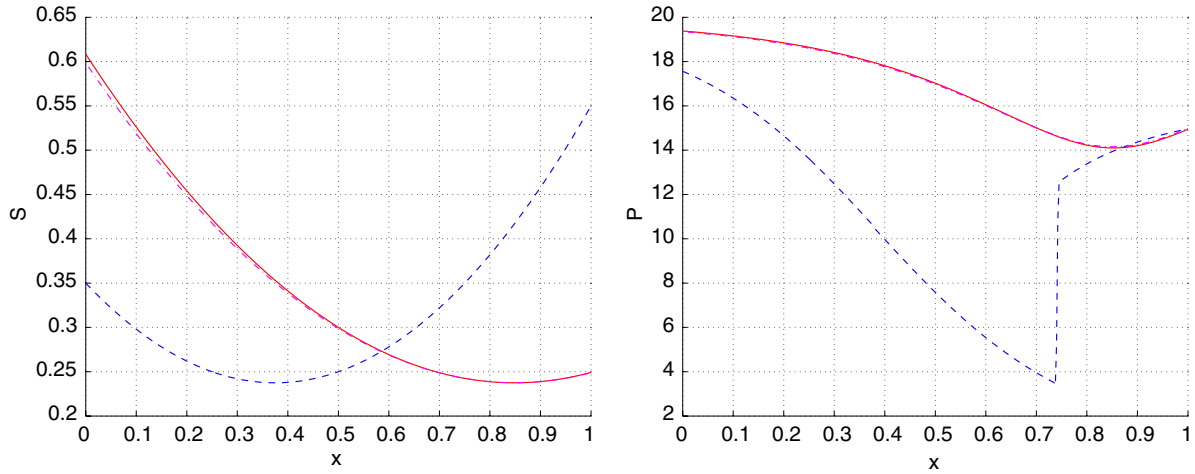

Fig. 8 Transonic flow through a convergent–divergent nozzle. *Left*, geometry changes. *Right*, pressure evolution. *Dashed lines*, original curves; *solid lines*, final results; *dash–dot lines*, target profiles

the term $\partial\Psi/\partial x$ shows another discontinuity in Fig. 7–left, but this one is a direct result of the Ψ-singularity at the nozzle throat.

As the cycles progress and the solution approaches P_D, the flow becomes fully subsonic, the mass-flow rate drops from its limiting value, the ψ_i become smooth functions (Fig. 5–right) and their magnitude also drops off. As with the previous case, the adjoint equations are subject to the same boundary conditions throughout the process. Only, in this case they concern subsonic as opposed to supersonic flow.

The evolution of the sensitivity gradient is shown in Fig. 7–right. Here again, the gradient has also been estimated by brute force, at the same points as the adjoint method. A comparison between the two reveals the error of the latter with respect to the former method remains below 4.5%, which certainly is much larger than, the previous case. However, that may be attributed to the fact that the subsonic adjoint boundary conditions involve solving characteristic equations which incur numerical error. The supersonic ones reduce to simple homogeneous Dirichlet conditions at the domain exit.

Lastly, geometry changes and the corresponding pressure distributions are depicted in Fig. 8 left and right sides, respectively. There, it can be seen that the final geometry gets closer to the target, as the pressure distribution approaches P_D.

4.3 A homogeneous functional of degree one

Still in regard to the behavior of the adjoint variables through a shock wave, it is instructive to consider an alternative measure of merit. Equation 34 shows that any objective functional, which is homogeneous of degree one with respect to the state variables $\mathbf{Q}$, should lead to smooth ψ_i of class $\mathcal{C}^1$. Such a functional would make for an illustrative comparison with the previous case of (51) and Fig. 7–right. Then, it should be interesting to pick the same pressure integral as used by Giles and Pierce [39,45].

$$I = \int_0^l P \, \mathrm{d}x, \tag{69}$$

for which the integrand yields the gradient

$$\frac{\partial g}{\partial Q} = (\gamma - 1) \begin{pmatrix} \frac{u^2}{2} \\ -u \\ 1 \end{pmatrix}. \tag{70}$$

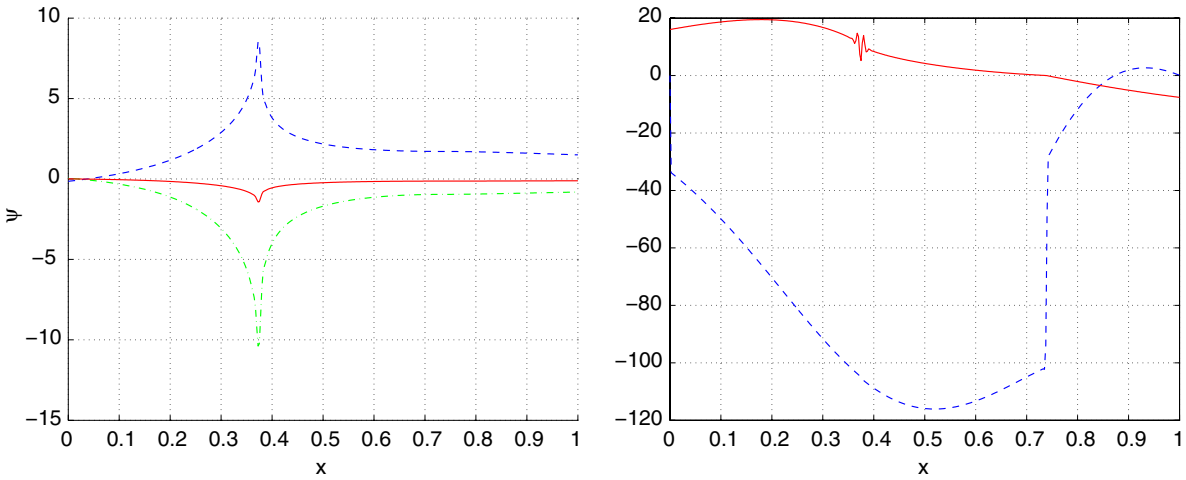

Fig. 9 Transonic flow through a convergent–divergent nozzle. *Left*, adjoint solutions to Giles' objective functional: *dash-dot lines*, ψ_1; *dashed lines* ψ_2; *solid lines* ψ_3. Right, Eq. 71: *solid line*, $(\partial\Psi/\partial x)^T \cdot (S\mathbf{F})$; *dashed line*, pressure distribution $g(\mathbf{V}) = P$

As before, the expression corresponds to the non-homogeneous term of the adjoint PDE (29). However, in this case it can be seen that $g(\mathbf{Q})$ is an homogeneous function of degree one and, as a result, the ψ_i-derivatives should undergo no jumps through the shock wave. That is, they should meet the condition

$$\left(\frac{\partial \Psi}{\partial x}\bigg|_+^T - \frac{\partial \Psi}{\partial x}\bigg|_-^T \right) \cdot (S\mathbf{F})^+ = 0, \tag{71}$$

which implies continuity, given that the flux $(S\mathbf{F})$ through the shock wave is obviously non-zero.

A test was set up where the adjoint solution is computed for the objective functional (69), in the same original geometry and under the same conditions as the converging–diverging nozzle above. The results thus obtained are presented in Fig. 9, and they should be compared to those of the first cycle in the previous case.

On the left, Fig. 9 presents the solution for ψ_i, which should be compared to Fig. 5–left. On the right it depicts the term $(\partial\Psi/\partial x)^T \cdot (S\mathbf{F})$ and the pressure distribution—the latter picture should be compared to Fig. 7–right.

At a first glance, Figs. 5–left and 9–left show both solutions share similar features, such as the throat singularity at the same position—although the ψ_i signs have changed. However, a closer look reveals the solutions behave differently through the shock wave, which is at the same position for both. In the latter case (Fig. 9–left) the ψ_i cross the shock wave smoothly with zero slope, as opposed to the non-zero angles that appear in former.

As expected, the results on Fig. 9–left corroborate the findings that are presented in [39,45] for the same objective functional (69). The assertion becomes clearer on Fig. 9–right. There it can be seen the first derivatives of Ψ are continuous, and reach zero at the shock-wave location (solid line). Whereas the integrand of the objective functional $g(\mathbf{V})$ undergoes a finite jump at that point (dashed line). It is quite a different picture from Fig. 7–right, where both the derivatives of Ψ and $g(\mathbf{V})$ experience jumps of the same magnitude, within the results accuracy. These findings, in turn, seem to confirm the analysis of (34), which relates the nature of the objective functional to the behavior of the adjoint variables through a shock wave.

5 Conclusions

The primary objective has been to investigate the adjoint boundary and internal conditions. Two main aspects of the topic have been given special attention, namely: (1) the relation between the primal and dual problems,

and (2) the well-posedness of the latter. In particular, with respect to internal conditions, the idea was to address a question that is still subject to debate in the literature.

A novel approach to the problem has been proposed, which is geared toward achieving those specific purposes. It consists in seeing the adjoint variables as generalized constraint forces, which impose realizability conditions on flow-physics variations. Within that framework, the terms that give rise to boundary conditions amount to inner products, which only vanish when the adjoint variables are orthogonal to all allowable variations—the latter are viewed as virtual displacements in the system's state-space. In principle, this rationale ensures that the adjoint problem should be well-posed, as long as the primal problem is well-posed.

As for the internal condition, the need for imposing the adjoint-variables' continuity through a shock-wave is verified. That result certainly confirms Giles and Pierce's assertions on the subject. However, our findings also make it clear that the adjoint-solution behavior through the shock front depends on the nature of the objective functional. There, the solution can only remain continuous up to the first derivative for a specific class of functionals, i.e., those that are homogeneous of degree one with respect to the state variables. Objective functionals that lack this property will give rise to a discontinuity in the gradient of the adjoint variables at the shock position, precisely. The discontinuity strength is determined by the non-homogeneous term of the adjoint equation. That term, in turn, comes from the variation of the objective functional kernel with respect to state variables.

These claims have been verified by comparing the results, thus obtained, with corresponding analytical solutions to the adjoint equations. The comparison was made possible, because the latter solutions are currently available in the literature for the quasi-1D Euler flow. Complete inverse-design applications have also been developed, so as to validate the procedure as a whole.

Further research into the topic seems to indicate this approach can be extended to a more general context. That involves 2 and 3–D applications, besides more complex fluid-dynamics models—such as the time-dependent Euler flows. The full Navier–Stokes equations could also be investigated in the light of these ideas. However, in that case one must account for the changes the diffusion mechanism brings about in the momentum and energy contour problems. In addition, it would be of great interest to consider how the adjoint-variables behavior through shock waves could be used to advantage in mesh refinement and error-control applications.

Acknowledgements The second author would like to acknowledge the support of FAPESP grants 97/01229–7 and 99/03105–9 which were pivotal in the development of this research paper.

References

1. Lions JL (1971) Optimal control of systems governed by partial differential equations. Number 170 in Die Grundelehren der mathematischen Wissenschaften. 1st edn. Springer–Verlag, Berlin
2. Mohammadi B, Pirroneau O (2001) Applied shape optimization for fluids. 1st edn. Oxford University Press, Oxford
3. Pirroneau O (1973) On optimal profiles in stokes flow. J Fluid Mech 59(1st edn):117–128
4. Pirroneau O (1974) On optimal design in fluid dynamics. J Fluid Mech 64(1):97–110
5. Jameson A (1988) Aerodynamic design via control theory. In: 12th IMACS world congress on scientific computation, MAE Report 1824, Paris, July 1988
6. Jameson A (1988) Aerodynamic design via control theory. J Sci Comput 3:233–260
7. Jameson A (1995) Optimum aerodynamic design using cfd and control theory. In: AIAA 95-1729-CP, 1995
8. Kuruvila G, Taasan S, Salas MD (1994) Airfoil optimization by the one-shot method, optimum design methods in aerodynamics. AGARD-FDP-VKI Special Course, 1994
9. Taasan S, Kuruvila G, Salas MD (1992) Aerodynamic design and optimization in one shot. In: 30th Aerospace Sciences Meeting and Exhibit, Reno, NV, AIAA 92-0025, January 1992
10. Giles MB, Pierce NA (1998) On the properties of solutions of the adjoint Euler equations. In: Baines MJ (ed) Numerical methods for fluid dynamics VI, ICFD, June 1998
11. Giles MB, Pierce NA (2000) An introduction to the adjoint approach to design. Flow Turbul Control 65(3):393–415
12. Angrand F (1983) Optimum design for potential flows. Int J Numer Methods Fluids 3:265–282
13. Cabuk H, Sung C-H, Modi V (1991) Adjoint operator approach to shape design for incompressible flows. In: 3rd international conference on inverse design concepts and optimization in engineering sciences (ICIDES), College Park, PA, pp. 391–404
14. Cacuci DG, Weber CF, Oblow EM, Marable JH (1980) Sensitivity theory for general systems of non-linear equations. Nucl Sci Eng 75:88–110

15. Marco N, Beux F (1993) Multilevel optimization: application to one-shot shape optimum design. Rapport de recherche No. 2068, INRIA, 1993
16. Beux F, Dervieux A (1993) A hierarchical approach for shape optimization. Rapport de recherche No. 1868, INRIA, 1993
17. Cusdin P, Mueller JD (2003) Automatic differentiation and sensitivity analysis methods for computational fluid dynamics. Internal Report QUB-SAE-03-old, Queen's University Belfast, May 2003
18. Taylor AC III, Putko MM (2003) Some advanced concepts in discrete aerodynamic sensitivity analysis. AIAA J 41(7):1224–1229
19. Hall MCG, Cacuci D (1983) Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models. J Atmos Sci 40:2537–2546
20. Nadarajah SK (2003) The discrete adjoint approach to aerodynamic shape optimization. PhD thesis, Stanford University
21. Kim H, Nakahashi K (2005) Unstructured adjoint method for Navier-Stokes equations. JSME Int J 48(2):202–207
22. Kim S, Alonso JJ, Jameson A (2004) Multi-element high-lift configuration design optimization using viscous continuous adjoint method. J Aircr 41(5):September–October
23. Kim HJ, Sasaki D, Obayashi S, Nakahashi K (2001) Aerodynamic optimization of supersonic transport wing using unstructured adjoint method. AIAA J 39(6):1011–1020
24. Qiao Z, Qin X, Yang X (2002) Wing design by solving adjoint equations. In: 40th aerospace sciences meeting & exhibit, Reno, NV, January 2002. American Institute of Aeronautics and Astronautics, AIAA
25. Duta MC, Giles MB, Campobasso MS (2002) The harmonic adjoint approach to unsteady turbomachinery design. Int J Numer Methods Fluids 40:323–332
26. Martins JRRA, Alonso JJ, Reuthers JJ (2005) A coupled-adjoint sensitivity analysis method for high-fidelity aero-structural design. Optim Eng 6:33–62
27. Mani K, Mavriplis D (2008) Unsteady discrete adjoint formulation for two–dimensional flow problems with deforming meshes. AIAA J 46(6):1351–1364
28. Nadarajah SK, Jameson A (2006) Optimum shape design for unsteady three-dimensional viscous flows using a non-linear frequency domain method. In: 24th applied aerodynamics conference, San Francisco, CA, June 2006. American Institute of Aeronautics and Astronautics, AIAA
29. Thomas JP, Hall KC, Dowell EH (2005) Discrete adjoint approach for modeling unsteady aerodynamic design sensitivities. AIAA J 43(9):1931–1936
30. Giles MB, Pierce NA (1998) Superconvergent lift estimates through adjoint error analysis. http://citeseerx.ist.psu.edu/viewdoc/summary?doi:10.1.1.27.152
31. Giles MB, Pierce NA (1999) Adjoint recovery of superconvergent functionals from approximate solutions of partial differential equations. Report 98/18, Oxford University Computing Laboratory, Oxford, August 1999
32. Giles MB, Pierce NA (1999) Improved lift and drag estimates using adjoint Euler equations. AIAA Paper 99–3293, 1999
33. Giles MB, Pierce NA (2004) Adjoint and defect error bounding and correction for functional estimates. J Comput Phys 200:769–794
34. Giles MB, Pierce NA, Siili E (2004) Progress in adjoint error correction for integral functionals. Comput Visual Sci 6: 113–121
35. Giles MB, Siili E (2002) Adjoint methods for pdes: a posteriori error analysis and postprocessing by duality. Acta Numerica 11:145–236
36. Venditti DA, Darmofal DL (1999) A multilevel error estimation and grid adaptive strategy for improving the accuracy of integral outputs. AIAA Paper 99–3292, 1999.
37. Venditti DA, Darmofal DL (2000) Adjoint error estimation and grid adaptation for functional outputs: Application to quasi-one-dimensional flow. J Comput Phys 164:204–227
38. Venditti DA, Darmofal DL (2003) Anisotropic grid adaptation for functional outputs: application to two-dimensional viscous flows. J Comput Phys 187:22–46
39. Giles MB, Pierce NA (2001) Analytic adjoint solutions for the quasi-1D Euler equations. J Fluid Mech 426:327–345
40. Nadarajah SK, Jameson A (2000) A comparison of the continuous and discrete adjoint approach to automatic aerodynamic optimization. In: 38th AIAA aerospace sciences meeting and exhibit, Reno, NV, AIAA–2000–667, January 10–13, 2000
41. Giles MB, Ghate D, Duta MC (2005) Using automatic differentiation for adjoint cfd code development. Post SAROD Workshop, 2005. Bangalore, India
42. Mueller JD, Cusdin P (2005) On the performance of discrete adjoint cfd codes using automatic differentiation.. Int J Numer Methods Fluids 47:939–945
43. Mavriplis DJ (2006) Multigrid solution of the discrete adjoint for optimization problems on unstructured meshes. AIAA J 44(1): 42–50
44. Nadarajah SK, Jameson A (2001) Studies of the continuous and discrete adjoint approaches to viscous automatic aerodynamic shape optimization. In: *Computational fluid dynamics conference*, Anaheim, CA, AIAA–2001–2530, June 2001
45. Giles MB, Pierce NA (2000) Analytic adjoint solutions for the quasi-1D Euler equations. Report 00103, Oxford University Computing Laboratory, Oxford, March 2000
46. Venditti DA, Darmofal DL (2002) Grid adaptation for functional outputs: application to two-dimensional inviscid flows. J Comput Phys176:40–69
47. Xie L (2002) Gradient-based optimum aerodynamic design using adjoint methods. PhD thesis, Virginia Polytechnic Institute and State University

48. Hirsch C (1994) Numerical computation of internal and external flows, vol I of Wiley series in numerical methods in engineering, 1st edn. Fundamentals of Numerical Discretization. Wiley, NY
49. Hirsch C (1994) Numerical computation of internal and external flows, vol II of Wiley series in numerical methods in engineering, 1st edn. Computational methods for inviscid and viscous flows. Wiley, NY
50. Lax PD (1973) Hyperbolic systems of conservation laws and the mathematical theory of shock waves 1st edn. Society for Industrial and Applied Mathematics–SIAM, Philadelphia, PA
51. Leveque RJ (2002) Finite volume methods for hyperbolic systems. 1st edn. Cambridge University Press, Cambridge texts in applied mathematics.
52. Giles MB, Pierce NA (1997) Adjoint equations in cfd: duality, boundary conditions and solution behavior. AIAA Paper 97–1850, 1997
53. Pulliam TH (1986) Artificial dissipation models for the euler equations. AIAA J 24(12):1931–1940
54. Vanderplaats GN (1984) Numerical optimization techniques for engineering design: with applications. Series in mechanical engineering, 1st edn. McGraw-Hill