Boundary conditions and discrete consistency of adjoint operators

By T. Flint AND M. J. P. Hack

1. Motivation and objectives

The adjoint of a linear operator is related to the direct operator through Lagrange's identity and has classically been used as a tool for solving differential equations that are posed as boundary value problems (Morse & Feshbach 1953; Ince 1956). Giles & Pierce (1997) showed that the adjoint problem can be formed and is well posed for a number of partial differential equation (PDE) problems, including the compressible Navier-Stokes equations. Hill (1995) pioneered the application of the adjoint in fluid mechanics, using the adjoint eigenfunctions to determine where a flow field may be most receptive to external forcing. Here, the formulation of the adjoint operator and its boundary conditions for linear second-order systems of PDEs is discussed, including a comparison of the continuous and discrete formulation. The motivation is to connect the continuous and discrete adjoint approaches, with particular focus on their boundary conditions, and to see under what conditions they become equivalent.

2. Continuous and discrete forms of linear PDEs

A general linear, second-order, time-dependent system of PDEs can be written in the form

$$\frac{\partial \mathbf{q}(\mathbf{x},t)}{\partial t} = \mathcal{A}(\mathbf{x})\ \mathbf{q}(\mathbf{x},t), \tag{2.1}$$

where $\mathbf{q}$ is the vector of l unknowns that depends on both time, t, and space, $\mathbf{x} = (x_1, x_2, x_3)^{\mathsf{T}}$. The linear operator, $\mathcal{A}$, can also depend on $\mathbf{x}$ and can be written to expose the underlying spatial derivatives

$$\begin{aligned}\mathcal{A} =& \mathsf{A}_0 + \mathsf{A}_{x1}\frac{\partial}{\partial x_1} + \mathsf{A}_{x2}\frac{\partial}{\partial x_2} + \mathsf{A}_{x3}\frac{\partial}{\partial x_3} + \mathsf{A}_{x1x1}\frac{\partial^2}{\partial x_1^2} + \mathsf{A}_{x2x2}\frac{\partial^2}{\partial x_2^2} \\ &+ \mathsf{A}_{x3x3}\frac{\partial^2}{\partial x_3^2} + \mathsf{A}_{x1x2}\frac{\partial^2}{\partial x_1 \partial x_2} + \mathsf{A}_{x1x3}\frac{\partial^2}{\partial x_1 \partial x_3} + \mathsf{A}_{x2x3}\frac{\partial^2}{\partial x_2 \partial x_3},\end{aligned} \tag{2.2}$$

where $\mathsf{A}_{(\cdot)}$ denote square matrices of size l corresponding to the coefficients of each of the spatial derivatives of $\mathbf{q}$. The system can be solved numerically by reducing it to a system of ordinary differential equations through discretizing and discretely approximating the derivatives. The discrete system can then be written as a matrix-vector equation as

$$\frac{\partial \mathbf{s}(t)}{\partial t} = \mathsf{A}\ \mathbf{s}(t), \tag{2.3}$$

where the vector $\mathbf{s}$ now contains all of the unknowns in discretized form and is of size $n \times l$, where n is the number of discrete degrees of freedom chosen to represent each unknown. The matrix A describes a discrete approximation to $\mathcal{A}$ and is of size $(n \times l) \times (n \times l)$.

3. Continuous adjoint

The adjoint form of the continuous version of the equations (Eq. (2.1)) is constructed to satisfy the property

$$\langle \mathbf{q}^\dagger, \mathcal{A}\mathbf{q} \rangle = \langle \mathcal{A}^\dagger \mathbf{q}^\dagger, \mathbf{q} \rangle, \tag{3.1}$$

where $\mathcal{A}^\dagger$ refers to the adjoint operator and $\mathbf{q}^\dagger$ the adjoint field. In the continuous case, the inner product is defined as

$$\langle \mathbf{y}, \mathbf{z} \rangle = \int_\Omega \mathbf{y}^{\mathsf{H}} \mathbf{z} \, \mathrm{d}\Omega, \tag{3.2}$$

where the integral refers to an integral over the entire spatial domain, Ω, and the $(\cdot)^{\mathsf{H}}$ denotes conjugate transpose. Using this inner product, the adjoint is defined by

$$\int_\Omega \mathbf{q}^{\dagger\mathsf{H}} \mathcal{A}\mathbf{q} \, \mathrm{d}\Omega = \int_\Omega (\mathcal{A}^\dagger \mathbf{q}^\dagger)^{\mathsf{H}} \mathbf{q} \, \mathrm{d}\Omega. \tag{3.3}$$

Substituting the continuous operator (Eq. (2.2)) into the left-hand side of Eq. (3.3) and performing integration by parts yields

$$\begin{aligned}
\int_\Omega \mathbf{q}^{\dagger\mathsf{H}} \mathcal{A}\mathbf{q} \, \mathrm{d}\Omega = \int_{\zeta_1}^{\zeta_2} \int_{\eta_1}^{\eta_2} \int_{\xi_1}^{\xi_2} \Bigg[& \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_0 \mathbf{q} - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1}}{\partial x_1} \mathbf{q} + \frac{\partial^2 \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1x1}}{\partial x_1^2} \mathbf{q} - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x2}}{\partial x_2} \mathbf{q} \\
& + \frac{\partial^2 \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x2x2}}{\partial x_2^2} \mathbf{q} + \frac{\partial^2 \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1x2}}{\partial x_1 \partial x_2} \mathbf{q} - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x3}}{\partial x_3} \mathbf{q} \\
& + \frac{\partial^2 \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x3x3}}{\partial x_3^2} \mathbf{q} + \frac{\partial^2 \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1x3}}{\partial x_1 \partial x_3} \mathbf{q} \\
& + \frac{\partial^2 \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x2x3}}{\partial x_2 \partial x_3} \mathbf{q} \Bigg] \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \\
+ \int_{\eta_1}^{\eta_2} \int_{\xi_1}^{\xi_2} \Bigg[& \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x3} \mathbf{q} + \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x3x3} \frac{\partial \mathbf{q}}{\partial x_3} - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x3x3}}{\partial x_3} \mathbf{q} \\
& - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1x3}}{\partial x_1} \mathbf{q} - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x2x3}}{\partial x_2} \mathbf{q} \Bigg]_{x_3=\zeta_1}^{x_3=\zeta_2} \mathrm{d}x_1 \mathrm{d}x_2 \\
+ \int_{\zeta_1}^{\zeta_2} \int_{\xi_1}^{\xi_2} \Bigg[& \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x2} \mathbf{q} + \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x2x2} \frac{\partial \mathbf{q}}{\partial x_2} - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x2x2}}{\partial x_2} \mathbf{q} \\
& - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1x2}}{\partial x_1} \mathbf{q} - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x2x3}}{\partial x_3} \mathbf{q} \Bigg]_{x_2=\eta_1}^{x_2=\eta_2} \mathrm{d}x_1 \mathrm{d}x_3 \\
+ \int_{\zeta_1}^{\zeta_2} \int_{\eta_1}^{\eta_2} \Bigg[& \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1} \mathbf{q} + \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1x1} \frac{\partial \mathbf{q}}{\partial x_1} - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1x1}}{\partial x_1} \mathbf{q} \\
& - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1x2}}{\partial x_2} \mathbf{q} - \frac{\partial \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1x3}}{\partial x_3} \mathbf{q} \Bigg]_{x_1=\xi_1}^{x_1=\xi_2} \mathrm{d}x_2 \mathrm{d}x_3 \\
+ \int_{\xi_1}^{\xi_2} \Bigg[& \Big[\mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x2x3} \mathbf{q} \Big]_{x_2=\eta_1}^{x_2=\eta_2} \Bigg]_{x_3=\zeta_1}^{x_3=\zeta_2} \mathrm{d}x_1
\end{aligned}$$

$$+\int_{\eta_1}^{\eta_2}\left[\left[\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x1x3}\mathbf{q}\right]_{x_1=\xi_1}^{x_1=\xi_2}\right]_{x_3=\zeta_1}^{x_3=\zeta_2}\mathrm{d}x_2$$
$$+\int_{\zeta_1}^{\zeta_2}\left[\left[\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x1x2}\mathbf{q}\right]_{x_2=\eta_1}^{x_2=\eta_2}\right]_{x_1=\xi_1}^{x_1=\xi_2}\mathrm{d}x_3, \tag{3.4}$$

where the domain is defined by $x_1 \times x_2 \times x_3 \in [\xi_1, \xi_2] \times [\eta_1, \eta_2] \times [\zeta_1, \zeta_2]$. The boundary conditions must be taken into account in order to determine the adjoint operators. Here, two cases are considered, one in which the boundary conditions on $\mathbf{q}$ are homogeneous Dirichlet and one in which one component of $\mathbf{q}$ is not specified, as is the case, for instance, at a wall in compressible flows (Poinsot & Lele 1992).

3.1. *Homogeneous Dirichlet*

In the case in which $\mathbf{q} = \mathbf{0}$ on the boundaries, all of the boundary terms involving $\mathbf{q}$ vanish and Eq. (3.4) reduces to

$$\begin{aligned}\int_\Omega \mathbf{q}^{\dagger\mathsf{H}}\mathcal{A}\mathbf{q}\,\mathrm{d}\Omega = \int_{\zeta_1}^{\zeta_2}\int_{\eta_1}^{\eta_2}\int_{\xi_1}^{\xi_2}\Bigg[&\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_0\mathbf{q} - \frac{\partial\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x1}}{\partial x_1}\mathbf{q} + \frac{\partial^2\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x1x1}}{\partial x_1^2}\mathbf{q} - \frac{\partial\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x2}}{\partial x_2}\mathbf{q}\\
&+ \frac{\partial^2\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x2x2}}{\partial x_2^2}\mathbf{q} + \frac{\partial^2\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x1x2}}{\partial x_1\partial x_2}\mathbf{q} - \frac{\partial\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x3}}{\partial x_3}\mathbf{q}\\
&+ \frac{\partial^2\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x3x3}}{\partial x_3^2}\mathbf{q} + \frac{\partial^2\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x1x3}}{\partial x_1\partial x_3}\mathbf{q}\\
&+ \frac{\partial^2\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x2x3}}{\partial x_2\partial x_3}\mathbf{q}\Bigg]\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3\\
+\int_{\eta_1}^{\eta_2}\int_{\xi_1}^{\xi_2}&\left[\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x3x3}\frac{\partial\mathbf{q}}{\partial x_3}\right]_{x_3=\zeta_1}^{x_3=\zeta_2}\mathrm{d}x_1\mathrm{d}x_2\\
+\int_{\zeta_1}^{\zeta_2}\int_{\xi_1}^{\xi_2}&\left[\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x2x2}\frac{\partial\mathbf{q}}{\partial x_2}\right]_{x_2=\eta_1}^{x_2=\eta_2}\mathrm{d}x_1\mathrm{d}x_3\\
+\int_{\zeta_1}^{\zeta_2}\int_{\eta_1}^{\eta_2}&\left[\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{x1x1}\frac{\partial\mathbf{q}}{\partial x_1}\right]_{x_1=\xi_1}^{x_1=\xi_2}\mathrm{d}x_2\mathrm{d}x_3.\end{aligned} \tag{3.5}$$

In order to compare with the right-hand side of Eq. (3.3) and determine the adjoint operator, the remaining boundary terms must vanish. The only degree of freedom left is in the choice of the boundary conditions on the adjoint field, $\mathbf{q}^\dagger$, of which homogeneous Dirichlet conditions must be chosen: $\mathbf{q}^\dagger = \mathbf{0}$. Comparing coefficients of the remaining terms on the right-hand side of Eq. (3.5) with the right-hand side of Eq. (3.3) yields the adjoint operator, which is provided in Appendix A.

3.2. *Wall boundary conditions in compressible flow*

The boundary conditions at a solid isothermal wall in a compressible flow are that all three components of velocity and temperature are fixed, so in the linearized system the boundary conditions $u' = v' = w' = T' = 0$ on the fluctuations are specified. At the wall, only four boundary conditions are specified, despite transporting five state variables (Poinsot & Lele 1992); thus, to determine the corresponding adjoint boundary conditions the boundary terms in Eq. (3.4) must be revisited. Consider the case in which the transported variables of the linearized system are the density, three mass fluxes, and total energy, $\mathbf{q} = (\rho', (\rho u)', (\rho v)', (\rho w)', e')^{\mathsf{T}}$. There are three terms in Eq. (3.4) that cannot be put in a form comparable to Eq. (3.3) and so must vanish when the boundary conditions are considered,

$$\mathrm{I} = \int_{\eta_1}^{\eta_2} \int_{\xi_1}^{\xi_2} \left[\mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x3x3} \frac{\partial \mathbf{q}}{\partial x_3} \right]_{x_3=\zeta_1}^{x_3=\zeta_2} \mathrm{d}x_1 \mathrm{d}x_2, \tag{3.6}$$

$$\mathrm{II} = \int_{\zeta_1}^{\zeta_2} \int_{\xi_1}^{\xi_2} \left[\mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x2x2} \frac{\partial \mathbf{q}}{\partial x_2} \right]_{x_2=\eta_1}^{x_2=\eta_2} \mathrm{d}x_1 \mathrm{d}x_3, \tag{3.7}$$

$$\mathrm{III} = \int_{\zeta_1}^{\zeta_2} \int_{\eta_1}^{\eta_2} \left[\mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{x1x1} \frac{\partial \mathbf{q}}{\partial x_1} \right]_{x_1=\xi_1}^{x_1=\xi_2} \mathrm{d}x_2 \mathrm{d}x_3. \tag{3.8}$$

Enforcing five boundary conditions on the adjoint variables would make these terms vanish, though it is unclear whether or not this will be over-determined given that the direct equations required only four boundary conditions. From the definition of the matrices $\mathsf{A}_{(\cdot)}$ in Eq. (2.2), the ith row in the matrix $\mathsf{A}_{(\cdot)}$ corresponds to the coefficients of the associated derivative in the ith governing equation. The first equation governs the conservation of mass, and does not contain any second derivatives, meaning that the first rows of all matrices corresponding to the coefficients of second derivatives, A_{xixj}, must only contain zeros. The boundary terms in Eqs. (3.6)-(3.8) can then be written as

$$\begin{aligned} \mathbf{q}^{\dagger\mathsf{H}} \mathsf{A}_{xixi} \frac{\partial \mathbf{q}}{\partial x_i} &= [q_1^{\dagger *}, q_2^{\dagger *}, q_3^{\dagger *}, q_4^{\dagger *}, q_5^{\dagger *}] \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \end{bmatrix} \begin{bmatrix} \square \\ \square \\ \square \\ \square \\ \square \end{bmatrix}, \\ &= [q_1^{\dagger *}, q_2^{\dagger *}, q_3^{\dagger *}, q_4^{\dagger *}, q_5^{\dagger *}] \begin{bmatrix} 0 \\ \square \\ \square \\ \square \\ \square \end{bmatrix}, \end{aligned} \tag{3.9}$$

where a square denotes an unknown quantity and a star denotes a conjugate. The choice of $q_2^\dagger = 0$, $q_3^\dagger = 0$, $q_4^\dagger = 0$, and $q_5^\dagger = 0$ on the boundary guarantees that the boundary terms in Eqs. (3.6)-(3.8) vanish. In fact it can be shown that this choice of the adjoint boundary conditions causes all the boundary terms in Eq. (3.4) to vanish (see Appendix B). Comparing the remaining terms in Eq. (3.4) with Eq. (3.3) again recovers the adjoint operator shown in Appendix A.

4. Discrete adjoint

The adjoint form of the discrete version of the equations (Eq. (2.3)) is constructed to satisfy the property

$$\langle \mathbf{s}^\dagger, \mathsf{A}\mathbf{s} \rangle = \langle \mathsf{A}^\dagger \mathbf{s}^\dagger, \mathbf{s} \rangle, \tag{4.1}$$

where, in the discrete case, the inner product is defined as

$$\langle \mathbf{y}, \mathbf{z} \rangle = \mathbf{y}^{\mathsf{H}} \mathbf{z}. \tag{4.2}$$

Taking the left-hand side of Eq. (4.1) we have

$$\begin{aligned} \langle \mathbf{s}^{\dagger}, \mathsf{A}\mathbf{s} \rangle &= \mathbf{s}^{\dagger\mathsf{H}} \mathsf{A}\mathbf{s} && (4.3)\\ &= (\mathsf{A}^{\mathsf{H}} \mathbf{s}^{\dagger})^{\mathsf{H}} \mathbf{s} && (4.4)\\ &= \langle \mathsf{A}^{\mathsf{H}} \mathbf{s}^{\dagger}, \mathbf{s} \rangle. && (4.5) \end{aligned}$$

Comparing with the right-hand side of Eq. (4.1) we see that $\mathsf{A}^{\dagger} \equiv \mathsf{A}^{\mathsf{H}}$.

We note that the discrete construction of the adjoint avoids the explicit treatment of the boundary conditions, as is necessary in the continuous case. The boundary treatment of the direct equations was embedded in the A matrix, and as such the adjoint, $\mathsf{A}^{\dagger}$, must now contain a consistent boundary treatment for the adjoint variables, $\mathbf{s}^{\dagger}$. While the discrete formulation may be conceptually simpler than the continuous one, it leaves interpretation of the adjoint equations opaque as their differential form is absent and the boundary conditions are implicitly defined.

5. Numerical boundary treatment

Solving both the continuous and the discrete version of the adjoint equations typically requires a numerical approach. In the continuous case, the adjoint operator must be discretized and boundary conditions implemented. In contrast, the discrete adjoint was defined for the discrete form of the equations and the boundary conditions were automatically embedded. It is of interest to investigate under what conditions, if any, these two approaches are consistent.

A simple example of a one-dimensional scalar PDE is considered that is discretized with a second-order central finite-difference scheme.

5.1. *Wave equation*

To begin, the conceptually simpler case of the wave equation is considered which only contains second spatial derivatives,

$$\begin{aligned} \frac{\partial^2 w}{\partial t^2} &= \alpha \frac{\partial^2 w}{\partial x^2} && (5.1)\\ &= \mathcal{A} w, && (5.2) \end{aligned}$$

where $w(x,t)$ is a scalar function of space and time, and α is some scalar constant. Discretization in space using five points in x gives

$$\mathcal{A} \approx \mathsf{A} = \frac{\alpha}{\Delta x} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{5.3}$$

where Δx is the grid spacing, which is assumed to be constant. Two homogeneous Dirichlet boundary conditions have been specified, one at each boundary. The boundary conditions have been implemented by removing the equations at the boundaries and replacing their coefficient in the interior of the domain with zeros. This results in a system in which the interior points evolve as though the boundary values are set to zero at all times. The discrete adjoint operator is then (from Eq. 4.5)

$$\mathsf{A}^\dagger = \mathsf{A}^\mathsf{H} = \frac{\alpha^*}{\Delta x}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{A}. \tag{5.4}$$

The discrete forward operator is the same as its adjoint, with exactly the same boundary conditions imposed.

The continuous adjoint operator is given by (Appendix A)

$$\mathcal{A}^\dagger = \alpha^* \frac{\partial^2}{\partial x^2}, \tag{5.5}$$

and when discretized, along with implementing the boundary conditions in the same way as was done for the direct equations, the discretized operator is identical to the discrete adjoint. Specifically, Eq. (3.4) indicates that homogeneous Dirichlet boundary conditions must also be specified on the adjoint field in order to use the adjoint operators specified in Appendix A. The discrete adjoint and the discretized continuous adjoint are discretely equivalent for this case, and both recover exact bi-orthogonality and conjugated eigenvalues when compared to the direct equations.

5.2. *Advection equation*

The advection equation is now considered. It contains only a single derivative in space

$$\frac{\partial w}{\partial t} = \alpha \frac{\partial w}{\partial x} \tag{5.6}$$

$$= \mathcal{A} w. \tag{5.7}$$

Discretization in space using five points in x yields

$$\mathcal{A} \approx \mathsf{A} = \frac{\alpha}{2\Delta x}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -2 & 2 \end{bmatrix}, \tag{5.8}$$

where, consistent with the order of the equation, a homogeneous Dirichlet boundary condition has been prescribed on one side of the domain. On the side of the domain on which no boundary condition is specified, a first-order one-sided derivative stencil has been introduced to evaluate the spatial gradient at the boundary point. Taking the conjugate transpose to recover the discrete adjoint gives

$$\mathsf{A}^\dagger = \mathsf{A}^\mathsf{H} = \frac{\alpha^*}{2\Delta x}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}. \tag{5.9}$$

In this case, the discrete adjoint is less obvious to interpret, as it imposes a homogeneous boundary condition on the same side as the direct equations, and the operators near the other boundary have been modified. In the interior of the domain, a negative first-order finite-difference stencil is recovered.

The continuous adjoint operator is given by (Appendix A)

$$\mathcal{A}^\dagger = -\alpha^* \frac{\partial}{\partial x}. \tag{5.10}$$

To eliminate the boundary terms in Eq. (3.4) in the recovery of the adjoint operators, a homogeneous Dirichlet condition must be imposed on the side of the domain where no boundary condition was prescribed in the direct problem. Discretization of the continuous adjoint (Eq. (5.10)) gives

$$\mathcal{A}^\dagger \approx \frac{\alpha^*}{2\Delta x} \begin{bmatrix} 2 & -2 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{5.11}$$

While this operator has a negative finite-difference stencil in the center of the domain like the discrete adjoint, the operator near the boundaries differs from the discrete adjoint. In fact, the discrete adjoint operator implies boundary conditions on the adjoint field which are not consistent with the continuous derivation.

Vectors that remain unchanged apart from a re-scaling of their magnitude upon application of a linear operator, A, are called eigenvectors, $\mathbf{w}$, and the corresponding scaling factor is the eigenvalue, λ,

$$\mathsf{A}\mathbf{w} = \lambda \mathbf{w}. \tag{5.12}$$

Eigenvalue problems can be useful, for example, in understanding how a time-evolving system may behave. In fluid mechanics, Hill (1995) used the eigenvectors of the adjoint operator to study the receptivity of a boundary layer to forcing. Here, we study the properties of the eigenvalues and eigenvectors of the aforementioned operators in order to identify differences between the discrete and continuous adjoint approaches.

The eigenvalues for the discrete direct system (Eq. (5.8)), the discrete adjoint (Eq. (5.9)), and the discretized continuous adjoint (Eq. (5.11)) are plotted in Figure 1. Both methods of constructing the adjoint operator produce the complex conjugate of the spectrum of the direct system, in this case to machine precision.

Assuming that they are diagonalizable, the operators must have different eigenvectors if they have the same eigenvalues. The eigenvector with the largest magnitude eigenvalue is plotted in Figure 2, using both five and 20 grid points. The eigenvectors of the discrete adjoint and the discretized continuous adjoint satisfy different boundary conditions, and differ in the interior of the domain. However, they converge as the number of grid points is increased, indicating that they are approximating the same operator, though their discrete forms differ.

The difference between the eigenvectors produced by the two operators varies for different eigenvectors depending on how well they are resolved. For example, in Figure 3 the eigenvector of the discrete adjoint and the discretized continuous adjoint with the tenth largest magnitude eigenvalue resolved with 20 points are compared. While the eigenvector shown in Figure 2 compared well, this eigenvector differs more.

A rigorous test of the accuracy of an adjoint eigenvector computation is to evaluate the bi-orthogonality property of the direct and adjoint eigenvectors, which states that the adjoint eigenvectors are orthogonal to all direct eigenvectors except for the one with

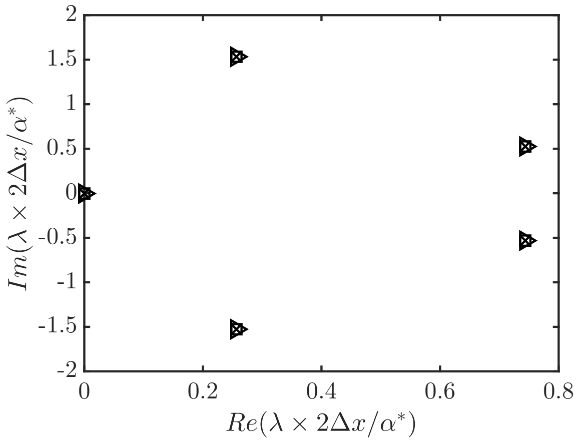

FIGURE 1. Eigenvalues of the discretized first derivative operators on five grid points. Crosses, complex conjugate of the eigenvalues of the direct equations (Eq. (5.8)); squares, discrete adjoint (Eq. (5.9)); triangles, discretized continuous adjoint (Eq. (5.11)).

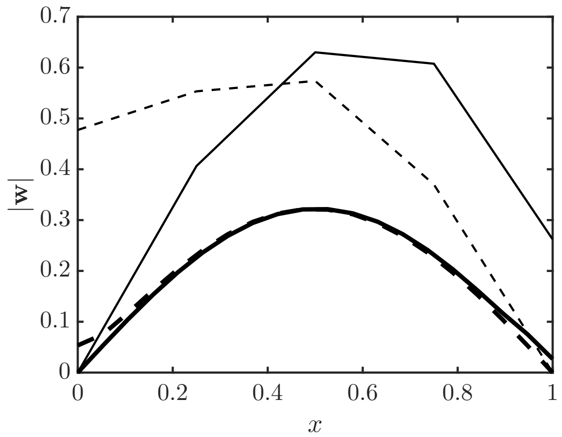

FIGURE 2. Eigenvector, $\mathbf{w}$, with the largest magnitude eigenvalue of the discretized adjoint operators at two resolutions. Thin lines, five points; thick lines, 20 points; solid lines, discrete adjoint (Eq. (5.9)); dashed lines, discretized continuous adjoint (Eq. (5.11)).

the corresponding conjugate eigenvalue, and vice versa. Mathematically

$$\langle \mathbf{w}_i, \mathbf{w}_j^\dagger \rangle = \beta \delta_{ij}, \tag{5.13}$$

where β is a normalization factor. The vectors $\mathbf{w}_i$ and $\mathbf{w}_i^\dagger$ are the direct and adjoint eigenvectors respectively, with the label i corresponding to the index of their corresponding eigenvalue. By construction, the eigenvectors computed using the discrete adjoint satisfy bi-orthogonality exactly. The bi-orthogonality of the eigenvectors computed using the discretized continuous adjoint is plotted in Figure 4. To examine how well Eq. (5.13) is satisfied, the inner product of an adjoint eigenvector with all of the direct eigenvectors is computed. The difference in magnitude between the inner product when $i = j$ and that of the largest other inner product is computed and normalized by the inner product

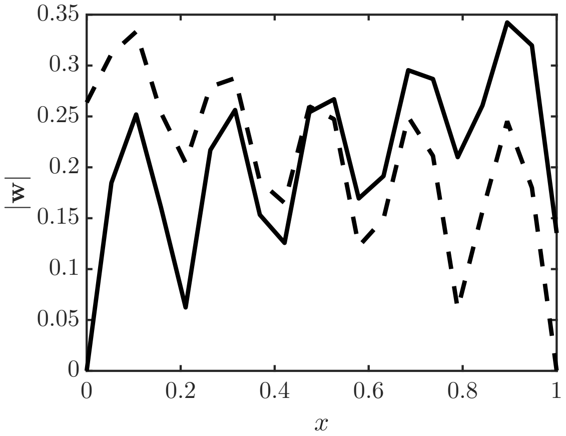

FIGURE 3. Eigenvector, $\mathbf{w}$, with the tenth largest magnitude eigenvalue of the discretized adjoint operators resolved with 20 points. Solid line, discrete adjoint (Eq. (5.9)); dashed line, discretized continuous adjoint (Eq. (5.11)).

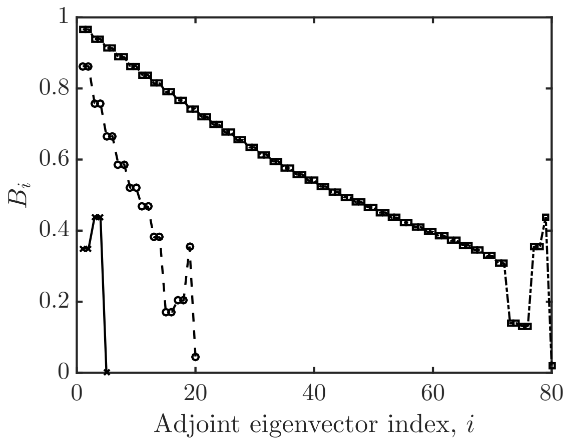

FIGURE 4. Bi-orthogonality of each adjoint eigenvector computed using the discretized continuous adjoint formulation (Eq. (5.11)) at three resolutions. Solid line, five grid points; dashed line, 20 grid points; dot-dashed line, 80 grid points. The adjoint eigenvectors are sorted from largest to smallest absolute value of their corresponding eigenvalue.

when $i = j$,

$$B_i = \frac{|\mathbf{w}_i^H \mathbf{w}_i^\dagger| - \max_{j, j \neq i} |\mathbf{w}_j^H \mathbf{w}_i^\dagger|}{|\mathbf{w}_i^H \mathbf{w}_i^\dagger|}, \tag{5.14}$$

where summation over like indices is not implied. If bi-orthogonality is exactly satisfied, this measure will evaluate to unity. Figure 4 indicates that as the resolution increases, the bi-orthogonality property is better satisfied for the adjoint eigenvectors with larger-magnitude eigenvalues. However, owing to the resolution restrictions becoming more stringent for more oscillatory eigenfunctions, lower adjoint eigenvectors fail to satisfy the bi-orthogonality property.

6. Conclusions

The construction of the adjoint operator for a general linear system of PDEs has been discussed with a focus on application to the linearized compressible Navier-Stokes equations. The continuous operator has been constructed from its direct counterpart, and the choice of the adjoint boundary conditions for various boundary types has been discussed. In the case of homogeneous Dirichlet conditions on the direct system, the same boundary conditions must be enforced on the adjoint problem. In the case of an isothermal wall in a compressible flow, a density perturbation is not prescribed for the direct system. When transporting density, mass flux, and total energy perturbations the correct adjoint boundary conditions to impose are that of homogeneous Dirichlet on all but the first adjoint variable.

Formulation of the discrete adjoint is conceptually simple, though it does not lend itself well to interpretation and its equivalence to the continuous formulation is unclear. For the one-dimensional wave equation, with a second-order derivative in space, the discrete and the continuous formulations are equivalent. However, for the advection equation, with a first-order spatial derivative, using a finite-difference discretization to compute the discrete adjoint results in different boundary conditions than the continuous derivation predicts.

It is possible to construct an adjoint operator that has the same eigenvalues as the conjugate eigenvalues of the direct system to machine precision but with eigenvectors that fail to satisfy the bi-orthogonality property. The eigenvectors of the discrete and the discretized continuous adjoint are, in general, different but converge if the eigenvectors are sufficiently resolved.

For the advection equation, the discrete adjoint and the discretized continuous adjoint operators differed only at the boundaries. This suggests that there may be alternative boundary conditions that could be applied to the discretized continuous adjoint and would make it discretely equivalent to the discrete adjoint, satisfying bi-orthogonality exactly.

Appendix A. Continuous adjoint operator

The continuous adjoint operator can be written as

$$\begin{aligned}\mathcal{A}^\dagger =&\mathsf{A}_0^\dagger + \mathsf{A}_{x1}^\dagger \frac{\partial}{\partial x_1} + \mathsf{A}_{x2}^\dagger \frac{\partial}{\partial x_2} + \mathsf{A}_{x3}^\dagger \frac{\partial}{\partial x_3} + \mathsf{A}_{x1x1}^\dagger \frac{\partial^2}{\partial x_1^2} + \mathsf{A}_{x2x2}^\dagger \frac{\partial^2}{\partial x_2^2} \\ &+ \mathsf{A}_{x3x3}^\dagger \frac{\partial^2}{\partial x_3^2} + \mathsf{A}_{x1x2}^\dagger \frac{\partial^2}{\partial x_1 \partial x_2} + \mathsf{A}_{x1x3}^\dagger \frac{\partial^2}{\partial x_1 \partial x_3} + \mathsf{A}_{x2x3}^\dagger \frac{\partial^2}{\partial x_2 \partial x_3}, \end{aligned} \tag{A 1}$$

where

$$\begin{aligned}\mathsf{A}_0^\dagger =&\mathsf{A}_0^\mathsf{H} - \frac{\partial \mathsf{A}_{x1}^\mathsf{H}}{\partial x_1} - \frac{\partial \mathsf{A}_{x2}^\mathsf{H}}{\partial x_2} - \frac{\partial \mathsf{A}_{x3}^\mathsf{H}}{\partial x_3} + \frac{\partial^2 \mathsf{A}_{x1x1}^\mathsf{H}}{\partial x_1^2} + \frac{\partial^2 \mathsf{A}_{x2x2}^\mathsf{H}}{\partial x_2^2} + \frac{\partial^2 \mathsf{A}_{x3x3}^\mathsf{H}}{\partial x_3^2} \\ &+ \frac{\partial^2 \mathsf{A}_{x1x2}^\mathsf{H}}{\partial x_1 \partial x_2} + \frac{\partial^2 \mathsf{A}_{x1x3}^\mathsf{H}}{\partial x_1 \partial x_3} + \frac{\partial^2 \mathsf{A}_{x2x3}^\mathsf{H}}{\partial x_2 \partial x_3}, \end{aligned} \tag{A 2}$$

$$\mathsf{A}_{x1}^\dagger = - \mathsf{A}_{x1}^\mathsf{H} + 2\frac{\partial \mathsf{A}_{x1x1}^\mathsf{H}}{\partial x_1} + \frac{\partial \mathsf{A}_{x1x2}^\mathsf{H}}{\partial x_2} + \frac{\partial \mathsf{A}_{x1x3}^\mathsf{H}}{\partial x_3}, \tag{A 3}$$

$$\mathsf{A}_{x2}^\dagger = - \mathsf{A}_{x2}^\mathsf{H} + 2\frac{\partial \mathsf{A}_{x2x2}^\mathsf{H}}{\partial x_2} + \frac{\partial \mathsf{A}_{x1x2}^\mathsf{H}}{\partial x_1} + \frac{\partial \mathsf{A}_{x2x3}^\mathsf{H}}{\partial x_3}, \tag{A 4}$$

$$\mathsf{A}^{\dagger}_{x3} = -\,\mathsf{A}^{\mathsf{H}}_{x3} + 2\frac{\partial \mathsf{A}^{\mathsf{H}}_{x3x3}}{\partial x_3} + \frac{\partial \mathsf{A}^{\mathsf{H}}_{x1x3}}{\partial x_1} + \frac{\partial \mathsf{A}^{\mathsf{H}}_{x2x3}}{\partial x_2}, \tag{A 5}$$

$$\mathsf{A}^{\dagger}_{x1x1} = \mathsf{A}^{\mathsf{H}}_{x1x1}, \tag{A 6}$$

$$\mathsf{A}^{\dagger}_{x2x2} = \mathsf{A}^{\mathsf{H}}_{x2x2}, \tag{A 7}$$

$$\mathsf{A}^{\dagger}_{x3x3} = \mathsf{A}^{\mathsf{H}}_{x3x3}, \tag{A 8}$$

$$\mathsf{A}^{\dagger}_{x1x2} = \mathsf{A}^{\mathsf{H}}_{x1x2}, \tag{A 9}$$

$$\mathsf{A}^{\dagger}_{x1x3} = \mathsf{A}^{\mathsf{H}}_{x1x3}, \tag{A 10}$$

$$\mathsf{A}^{\dagger}_{x2x3} = \mathsf{A}^{\mathsf{H}}_{x2x3}. \tag{A 11}$$

These coefficients were found by comparing the non-boundary terms on the right-hand side of Eq. (3.4) with the desired form on the right-hand side of Eq. (3.3).

Appendix B. Elimination of the boundary terms at an isothermal wall in compressible flow

The equation resulting from the continuous integration by parts (Eq. (3.4)) includes boundary terms of the form

$$\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{xi}\mathbf{q}, \tag{B 1}$$

$$\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{xixj}\mathbf{q}, \tag{B 2}$$

$$\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{xixi}\frac{\partial \mathbf{q}}{\partial x_i}, \tag{B 3}$$

$$\frac{\partial \mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{xixj}}{\partial x_i}\mathbf{q} = \frac{\partial \mathbf{q}^{\dagger\mathsf{H}}}{\partial x_i}\mathsf{A}_{xixj}\mathbf{q} + \mathbf{q}^{\dagger\mathsf{H}}\frac{\partial \mathsf{A}_{xixj}}{\partial x_i}\mathbf{q}. \tag{B 4}$$

It was shown in Section 3.2 that the term in Eq. (B 3) vanishes when the boundary conditions of $u' = v' = w' = T' = 0$ and $q_2^{\dagger} = q_3^{\dagger} = q_4^{\dagger} = q_5^{\dagger} = 0$ are applied. The same argument used for the term in Eq. (B 3) applies to the term in Eq. (B 2) and to the right half of the right-hand side of the term in Eq. (B 4).

Consider the term in Eq. (B 1). Continuity is the first equation so the coefficient matrix of the first derivatives, A_{xi}, will contain a one on the first row in the column corresponding to the associated derivative direction. So the term in Eq. (B 1) will be

$$\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{xi}\mathbf{q} = [q_1^{\dagger *}, q_2^{\dagger *}, q_3^{\dagger *}, q_4^{\dagger *}, q_5^{\dagger *}]\begin{bmatrix} 0 & \delta_{i1} & \delta_{i2} & \delta_{i3} & 0 \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \end{bmatrix}\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \end{bmatrix}. \tag{B 5}$$

On the boundaries $q_2 = q_3 = q_4 = 0$, yielding

$$\mathbf{q}^{\dagger\mathsf{H}}\mathsf{A}_{xi}\mathbf{q} = [q_1^{\dagger *}, q_2^{\dagger *}, q_3^{\dagger *}, q_4^{\dagger *}, q_5^{\dagger *}]\begin{bmatrix} 0 & \delta_{i1} & \delta_{i2} & \delta_{i3} & 0 \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square \end{bmatrix}\begin{bmatrix} q_1 \\ 0 \\ 0 \\ 0 \\ q_5 \end{bmatrix}, \tag{B 6}$$

$$=[q_1^{\dagger *}, q_2^{\dagger *}, q_3^{\dagger *}, q_4^{\dagger *}, q_5^{\dagger *}]\begin{bmatrix} 0 \\ \square \\ \square \\ \square \\ \square \end{bmatrix}, \tag{B 7}$$

so with the adjoint boundary conditions of $q_2^\dagger = q_3^\dagger = q_4^\dagger = q_4^\dagger = 0$ this term vanishes at the boundaries.

Now consider the left half of the right-hand side of the term in Eq. (B 4),

$$\frac{\partial \mathbf{q}^{\dagger \mathsf{H}}}{\partial x_i}\mathsf{A}_{xixj}\mathbf{q}. \tag{B 8}$$

The case for the coefficients of mixed derivatives and that of the double derivatives must be treated separately here. When evaluated at the boundary (enforcing no-slip on the base state velocity), Eq. (B 8) with the mixed derivative coefficients becomes

$$\frac{\partial \mathbf{q}^{\dagger \mathsf{H}}}{\partial x_i}\mathsf{A}_{xixj}\mathbf{q} = [\square,\square,\square,\square,\square]\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \square & \square & \square & 0 \\ 0 & \square & \square & \square & 0 \\ 0 & \square & \square & \square & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} q_1 \\ 0 \\ 0 \\ 0 \\ q_5 \end{bmatrix} \tag{B 9}$$

$$= \mathbf{0}. \tag{B 10}$$

For the double derivative case, evaluated at the boundary, we have

$$\frac{\partial \mathbf{q}^{\dagger \mathsf{H}}}{\partial x_i}\mathsf{A}_{xixi}\mathbf{q} = [\square,\square,\square,\square,\square]\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \square & 0 & 0 & 0 \\ 0 & 0 & \square & 0 & 0 \\ 0 & 0 & 0 & \square & 0 \\ -\frac{\overline{T}}{Pr\overline{\rho}} & 0 & 0 & 0 & \frac{\gamma}{Pr\overline{\rho}} \end{bmatrix}\begin{bmatrix} q_1 \\ 0 \\ 0 \\ 0 \\ q_5 \end{bmatrix} \tag{B 11}$$

$$= [\square,\square,\square,\square,\square]\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -\frac{\rho'\overline{T}}{Pr\overline{\rho}} + \frac{\gamma}{Pr\overline{\rho}}\frac{p'}{\gamma-1} \end{bmatrix} \tag{B 12}$$

$$= [\square,\square,\square,\square,\square]\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -\frac{\rho'\overline{T}}{Pr\overline{\rho}} + \frac{\rho'\overline{T}}{Pr\overline{\rho}} \end{bmatrix} \tag{B 13}$$

$$= \mathbf{0}, \tag{B 14}$$

where we have used the linearized ideal gas law, $p'\gamma/(\gamma-1) = \rho'\overline{T} + \overline{\rho}T'$, evaluated at the wall. Note that if instead of using the boundary condition $T' = 0$ we had used $q_5 = 0$ for the direct system, this boundary term would remain and have to be accounted for in the adjoint operator.

Acknowledgments

This investigation was funded by the Office of Naval Research under Grant No. N00014-17-1-2341.

REFERENCES

Giles, M. B. & Pierce, N. A. 1997 Adjoint equations in CFD: Duality, boundary conditions and solution behaviour. *13th Comput. Fluid Dyn. Conf.*, pp. 182–198.

Hill, D. C. 1995 Adjoint systems and their role in the receptivity problem for boundary layers. *J. Fluid Mech.* **292**, 183–204.

Ince, E. L. 1956 *Ordinary Differential Equations*. Dover Publications, Inc.

Morse, P. M. & Feshbach, H. 1953 *Methods of Theroetical Physics*. McGraw-Hill Book Company, Inc.

Poinsot, T. J. & Lele, S. K. 1992 Boundary conditions for direct simulations of compressible viscous flows. *J. Comput. Phys.* **101**, 104–129.