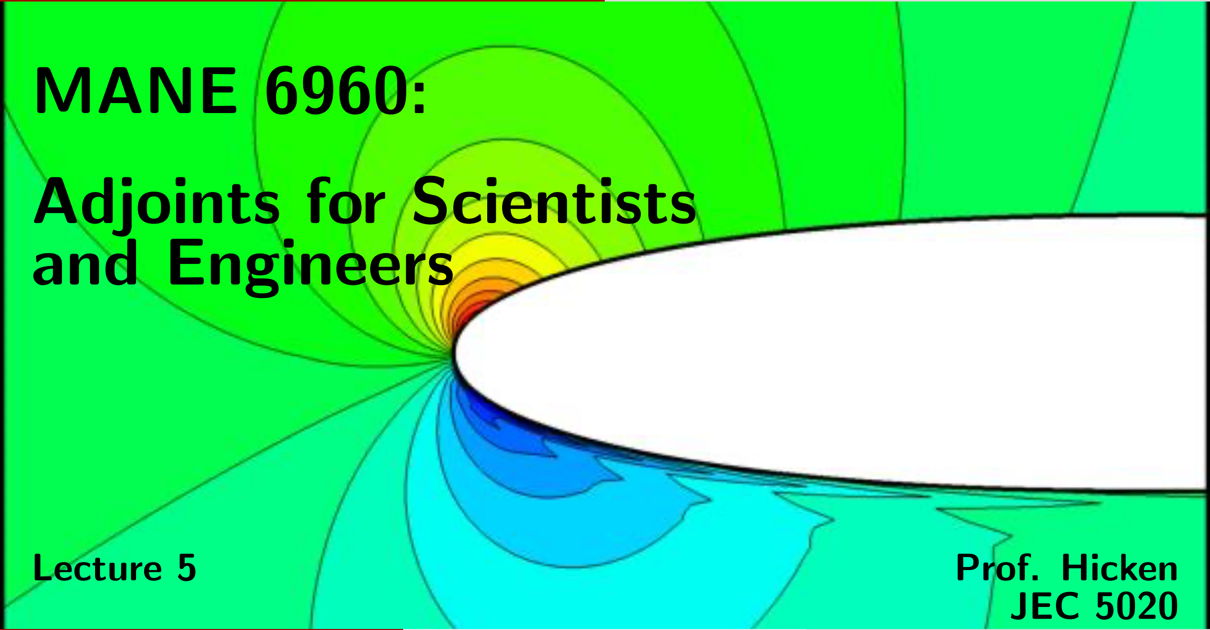


MANE 6960:

Adjoints for Scientists and Engineers

Lecture 5

Prof. Hicken
JEC 5020

Adjoint Consistent Discretizations

Recap and Lecture Objective

Last lecture we saw that there are two ways to derive/compute the adjoint variables:

- the continuous-adjoint approach; and
- the discrete-adjoint approach.

This lecture we will consider the case where these two approaches coincide and analyze one of the consequences of this.

Adjoint Consistency

As usual, we consider a generic primal linear BVP problem given by

$$\begin{aligned} Lu &= f, && \forall x \in \Omega, \\ Bu &= b, && \forall x \in \Gamma, \end{aligned} \tag{Pri}$$

and a generic linear functional given by

$$J(u) = (g, u)_\Omega + (c, Cu)_\Gamma. \tag{Fun}$$

Recall that the above define an adjoint BVP given by

$$\begin{aligned} L^*\psi &= g, && \forall x \in \Omega, \\ B^*\psi &= c, && \forall x \in \Gamma. \end{aligned} \tag{Adj}$$

Adjoint Consistency (cont.)

Definition: Adjoint Consistency [ABCM02, Lu05, HZ14]

The discretization of $(\textsc{Pri})$ and $(\textsc{Fun})$ given by

$$\begin{aligned} L_h u_h &= f_h, \\ \text{and} \qquad J_h(u_h) &= (g_h, u_h)_h, \end{aligned} \tag{$\star$}$$

is adjoint consistent of order $p > 0$ if

$$L_h^T [\psi]_h = g_h + \mathrm{O}(h^p),$$

where $[\psi]_h$ denotes the projection of the solution to the BVP $(\textsc{Adj})$ on to the discrete solution space.

Remarks on Adjoint Consistency

In other words, a discretization is adjoint consistent if the discrete adjoint equation,

$$L_h^T \psi_h = g_h,$$

is a consistent discretization of the adjoint BVP.

- Adjoint consistency is a property of the discretization, both the primal BVP discretization and functional discretization
- Adjoint consistency is also known as dual consistency

Remarks on Adjoint Consistency (cont.)

Another way to put it: a discretization is adjoint consistent if the discrete-adjoint approach is also a continuous-adjoint approach

- not all continuous-adjoint approaches will be a discrete-adjoint approach, since there is only one discrete-adjoint equation for a given primal discretization, but there are many consistent ways to discretize the adjoint BVP.

Implications

What are the implications of having a scheme that is, or is not, adjoint consistent?

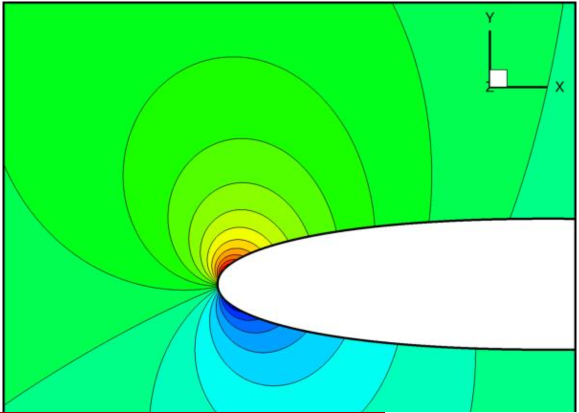
Let's first look at the qualitative implications.

- consider a high-order summation-by-parts discretization of the Euler equations
- two schemes, one adjoint consistent and one adjoint inconsistent
- only difference is how the slip-wall boundary condition is treated

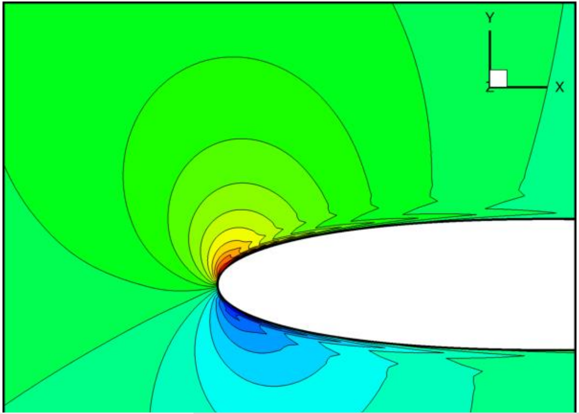
Implications (cont.)

Contours of the drag-functional adjoint associated with ρv.

Adjoint Consistent



Adjoint Inconsistent



Implications (cont.)

So, obviously there is a qualitative difference, but adjoint consistency also has important quantitative effects on

- functional accuracy;
- functional error estimation; and
- error localization (for mesh adaptation).

We will examine the first of these in more detail now, and discuss the other two later in the course.

Superconvergent Functional Estimates

Superconvergence

In the next few slides, we will prove that adjoint-consistent discretizations produce superconvergent functional estimates.

What does superconvergent mean?

Suppose that the discrete solution is order h^p accurate in some norm:

$$\|u_h - [u]_h\| \leq c_u h^p,$$

where c_u is independent of h, and $[u]_h$ denotes the projection of the exact solution to the primal BVP on to the discrete solution space.

Superconvergence (cont.)

Based on the above bound on the solution error, we might expect the functional error to behave as

$$\begin{aligned} |J(u) - J_h(u_h)| &\leq |J(u) - J_h(u)| + |J_h(u - u_h)| \\ &\leq c_J h^p, \end{aligned}$$

for some constant c_J independent of h.

However, when the discretization is adjoint consistent and the solution is sufficiently accurate, we will find that the functional can be significantly more accurate than this.

Some Assumptions

We will make the following assumptions.

Assumption 1: The discrete functional, when evaluated using $[u]_h$, is such that $|J(u) - J_h([u]_h)| = \mathrm{O}(h^{2p})$.

Assumption 2: The solution error is bounded as $\|u_h - [u]_h\| = \mathrm{O}(h^p)$.

Assumption 3: The adjoint error is bounded as $\|\psi_h - [\psi]_h\| = \mathrm{O}(h^p)$.

Assumption 4: The discrete bilinear form $B_h(u_h, v_h) \equiv (v_h, L_h u_h)_h$ satisfies

$$([v]_h, L_h[u]_h)_h = (v, Lu)_\Omega + (C^* v, Bu)_\Gamma + \mathrm{O}(h^{2p-m}),$$

for all $u, v \in \mathbb{V}$, where m is the order of the PDE.

Some Assumptions (cont.)

Assumption 1 basically says that the functional uses a sufficiently accurate quadrature rule to evaluate integrals (note that the exact solution is used in this assumption, not the discrete solution u_h).

Assumption 2 requires that the discrete error, measured in some norm, is order h^p. This is reasonable provided the discretization is sufficiently high-order and well-conditioned.

$$L_h u_h = f_h \quad , \quad L_h [u]_h = f_h + O(h^p)$$

$$\Rightarrow \quad \| L_h (u_h - [u]_h) \| \leq M h^p, \text{ some } M$$

$$\Rightarrow \quad \| u_h - [u]_h \| \leq \| L_h^{-1} \| M h^p$$

Some Assumptions (cont.)

Assumption 3 is where we require adjoint consistency. If we have adjoint consistency of order h^p, and $\|L_h^{-T}\|$ is bounded — L_h^{-T} has to be bounded if the discretization is well-posed — then

$$L_h^T \psi_h = g_h, \quad \overbrace{L_h^T [\psi]_h = g_h + O(h^p)}^{\text{adj. consistency}}$$

$$\implies \quad L_h^T (\psi_h - [\psi]_h) = O(h^p)$$

$$\implies \quad \| \psi_h - [\psi]_h \| \leq \| L_h^{-T} \| M_\psi h^p$$

Some Assumptions (cont.)

Assumption 4 is the least obvious (at least for me). It says that when any two functions, $u, v \in \mathbb{V}$, are projected on to the discrete solution space, the discrete bilinear form is an $\mathrm{O}(h^{2p-m})$ approximation to the continuous bilinear form, where m is the order of the PDE.

Example: degree $p-1$ FEM discretization of $\frac{d^2u}{dx^2}$

error in L^2 norm; $\|u - \mathcal{I}u\|_{\Omega,q} \leq C h^{p-q}$

where $\mathcal{I}u$ is polynomial interpolation of degree $p-1$.

Functional Superconvergence

$$\left([v]_h, L_h [u]_h\right)_h = -\int_\Omega \frac{d}{dx}(\mathcal{I}v)\frac{d}{dx}(\mathcal{I}u)\,dx + \int_\Gamma (\mathcal{I}v)\frac{d}{dx}(\mathcal{I}u)\,dx$$

$$= -\int_\Omega \frac{dv}{dx}\frac{du}{dx}\,dx + \int_\Gamma v\frac{du}{dx}\,dx \underbrace{+ O(h^{p-1})\,O(h^{p-1}) + O(h^{p})\,O(h^{p-1})}_{O(h^{2p-2})}$$

Functional Superconvergence (cont.)

Theorem: Functional Superconvergence [PG00]

Let $L_h u_h = f_h$ be a discretization of the primal BVP,

$$Lu = f, \quad \forall x \in \Omega, \qquad Bu = b, \quad \forall x \in \Gamma,$$

and let $J_h(u_h) = (g_h, u_h)_h$ be a discretization of

$$J(u) = (g, u)_\Omega + (c, Cu)_\Gamma$$

Then, under Assumptions 1–4, we have the bound

$$|J(u) - J_h(u_h)| \leq c_J h^{2p-m}.$$

For proof see last 2 pages

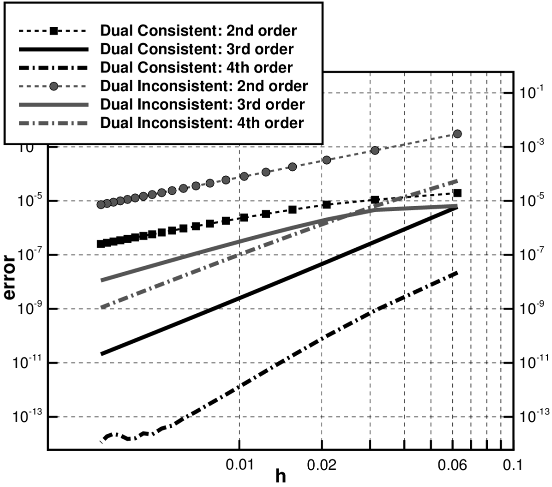
Remarks

This result really hinges on adjoint consistency, since the other assumptions are satisfied by many discretizations.

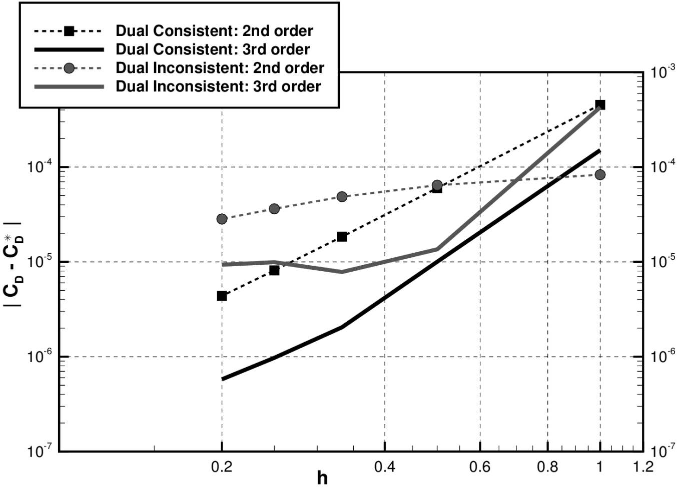
The moment you loose adjoint consistency, you loose functional superconvergence.

- superconvergence is mostly a concern for high-order $(p > 2)$ discretizations
- however, for high-order methods, adjoint consistency can significantly improve the efficiency of the scheme from the perspective of functional accuracy

Example: Inviscid Vortex Flow [HZ14]



Example: Onera M6 wing [HZ14]



References

[ABCM02] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and L. Donatella Marini, *Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems*, SIAM Journal on Numerical Analysis **39** (2002), no. 5, 1749–1779.

[HZ14] Jason E. Hicken and David W. Zingg, *Dual consistency and functional accuracy: a finite-difference perspective*, Journal of Computational Physics **256** (2014), 161–182.

[Lu05] James C. Lu, *An a posteriori error control framework for adaptive precision optimization using discontinuous Galerkin finite element method*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 2005.

[PG00] Niles A. Pierce and Michael B. Giles, *Adjoint recovery of superconvergent functionals from PDE approximations*, SIAM Review **42** (2000), no. 2, 247–264.

Proof:

$$J(u) = (g, u)_\Omega + (c, Cu)_\Gamma$$

$$= (g_h, [u]_h)_h + O(h^{2p}) \quad , \text{ Assump \#1}$$

$$= (g_h, u_h)_h - (g_h, u_h - [u]_h)_h + O(h^{2p})$$

$$= (g_h, u_h)_h - (\underbrace{L_h^T \psi_h}_{\equiv g_h}, u_h - [u]_h)_h + O(h^{2p})$$

$$= (g_h, u_h)_h - (\psi_h, L_h (u_h - [u]_h))_h + O(h^{2p})$$

(con't)

$$J(u) = J_h(u_h) - (\psi_h, \underset{\uparrow}{f_h} - L_h[u]_h)_h + O(h^{2p})$$

$$\because \; L_h u_h = f_h$$

$$= J_h(u_h) - ([\psi]_h, f_h - L_h[u_h])_h$$

$$+ ([\psi]_h - \psi_h, f_h - L_h[u]_h)_h + O(h^{2p})$$

$$= J_h(u_h) - \underbrace{([\psi]_h, f_h)_h + ([\psi_h], L_h[u_h])_h}_{O(h^{2p-m}) \text{ by assum. \#4}}$$

$$+ (\underbrace{[\psi]_h - \psi_h}_{O(h^p), \text{ assum. \#3}}, \underbrace{f_h - L_h[u]_h}_{O(h^{p-m}) \text{ assum. \#2}})_h + O(h^{2p})$$

$$= J_h(u_h) + O(h^{2p-m})$$