SENSITIVITIES FOR OPTIMIZATION

John Burkardt
Max Gunzburger

Mathematics Department
Iowa State University
Ames, Iowa

15 May 1996

Intro: Comments

Some control problems can be written as the minimization of some cost functional, which is defined in terms of a state constrained by governing equations, which is determined by some controllable parameters.

Discretization is typically applied to the continuous state problem to define a computable approximate state.

The relationship between the cost and the parameters is mediated by the state function, and involves partial derivatives of that state. We might choose to differentiate the state function, and then discretize it, or differentiate the already discretized state function.

In some cases, these processes may be interchanged without affecting the computation of the approximate state derivatives.

In other cases, there is a discrepancy between the two.

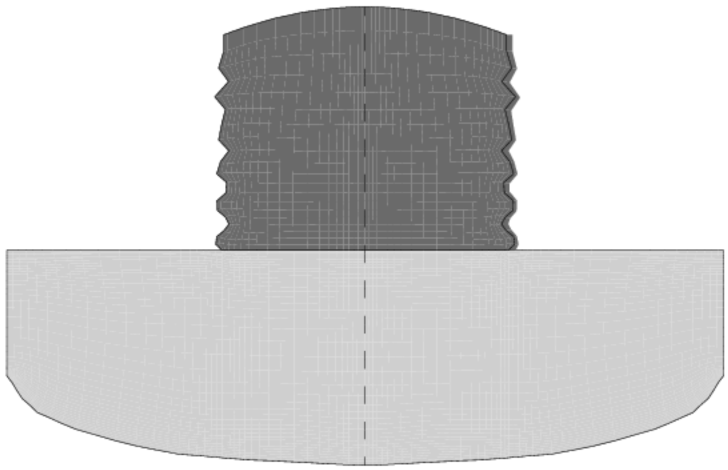
I'd like to examine some typical problems, look at the mathematical details behind the procedure, focus on a problem where the derivatives can be computed either way for one parameter, but not for others, and speculate about the causes of this problem.

OUTLINE

- **1: TYPICAL PROBLEMS**
- **2: SENSITIVITIES AND PDE CONTROL**
- **3: THE WIND TUNNEL PROBLEM**
- **4: OBSERVATIONS AND CONCLUSIONS**

1.0: TYPICAL PROBLEMS

- **1.1: Creation of Silicon Monocrystals**
- **1.2: Wind Tunnel Simulation of Flow Disturbances**
- **1.3: A Cylindrical Containment Vessel**



1.1: Creation of Silicon Monocrystals

Goal: Defect-free crystals.

Parameters: Heat, rotation, magnetic field, drawing rate.

Constraints: NUMEROUS.

Cost functionals: Crystal quality or process expense.

1.1: Comments

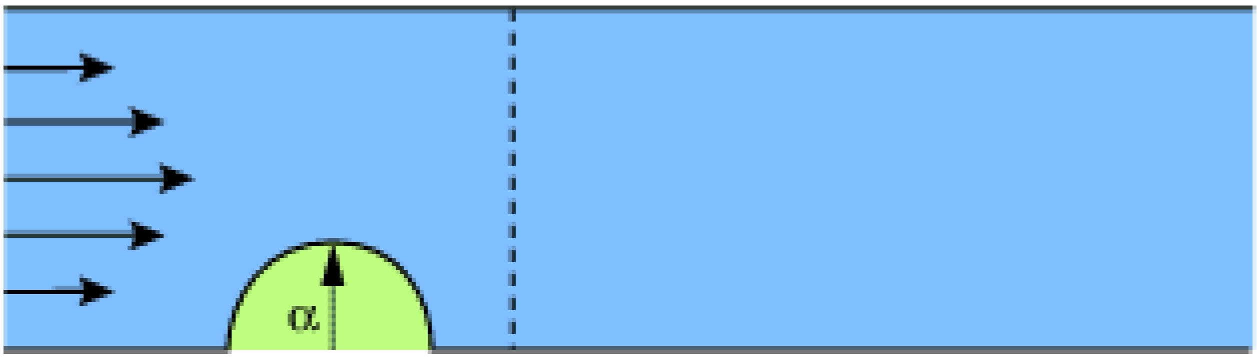
Goal: Large, cheap defect-free silicon crystals.

Parameters: applied heat, rotation rates of crucible and seed crystal, applied magnetic field, drawing rate.

Constraints: Time dependent K-epsilon turbulence Navier Stokes, moving boundaries, heat conduction, crystal formation, surface behavior, ...

Cost functionals:

- Quality: melt vorticity, melt velocity near crystal, heat loss,
- Expense: time, energy usage.



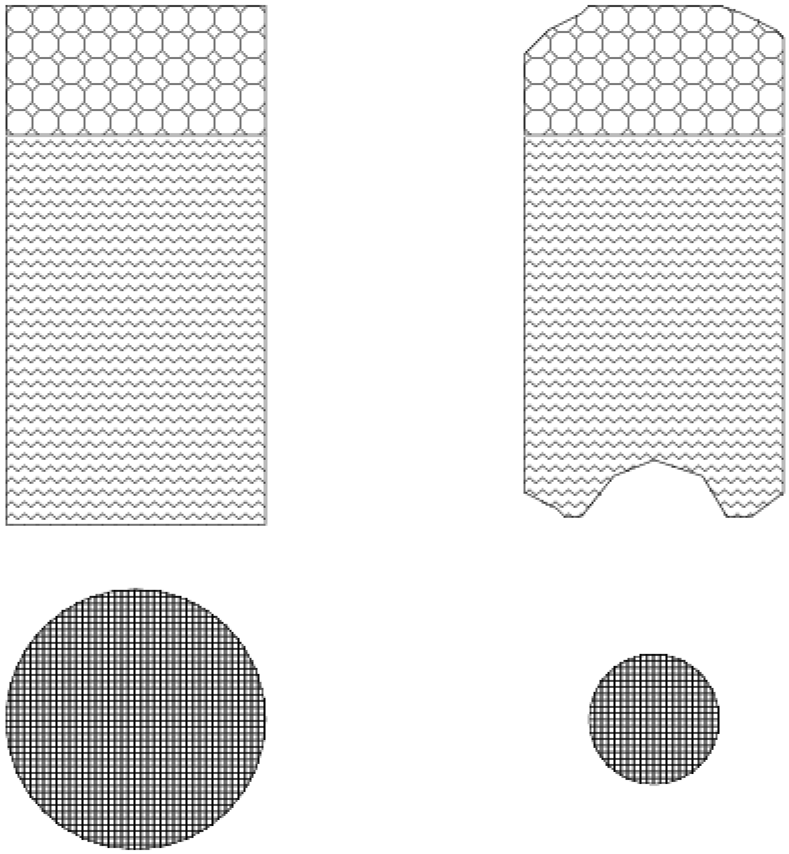
1.2: Wind Tunnel Simulation of Flow Disturbances

Goal: Small obstruction simulator for wind tunnel.

Parameters: obstruction shape, inflow conditions, roughness of surface.

Constraints: Navier Stokes and continuity.

Cost: match observed data, minimize surface area.



1.3: A Cylindrical Containment Vessel

Goal: Minimize materials for gas-liquid container.

Parameters: cross-section (may be polygonal); wall thickness; shape of vertical profile; size of top plate (must be flat); shape of bottom.

Constraints: Various, plus top must be flat, walls and bottom must resist deformation.

Cost: materials (top plate is expensive), processing.

2.0: SENSITIVITIES & PDE CONTROL

- **2.1: The Partial Differential Equation**
- **2.2: The State Partial Jacobian**
- **2.3: The Cost Functional**
- **2.4: An Equation for Discrete Sensitivities**
- **2.5: Discretize, then Differentiate**
- **2.6: Differentiate, then Discretize**
- **2.7: Commutivity?**

2.1: The Partial Differential Equation

We consider a PDE:

$$F(u, g) = 0 \text{ over } \bar{\Omega}. \tag{1}$$

where:

- Ω is bounded open domain with boundary $\partial\Omega$;
- u is unknown state functions;
- g is control parameters;
- F is partial differential equations plus conditions.

2.1: Comments

- F and u depend on space, but this is suppressed;
- dependence on time t could easily be included;
- Ω may depend on the controls g.

2.2: The State Partial Jacobian

Newton's method applied to an estimated solution (u_0, g_0):

$$F_u(u_0, g_0)\ \Delta u_0 = -F(u_0, g_0). \tag{2}$$

Differentiate the state equation with respect to parameter g, we get a *sensitivity* equation for u_g:

$$F_u(u_0, g_0)\ u_g = -F_g(u_0, g_0). \tag{3}$$

2.2: Comments

Repeated differentiation yields similar equations for higher order sensitivities, always with the same linear operator on the left. This means that, after a Newton iteration, the sensitivities can be found very cheaply.

2.3: The Cost Functional

A cost functional $\mathcal{J}(u, g)$ is to be minimized:

Since u is an implicit function of g, we can define:

$$J(g) \equiv \mathcal{J}(u(g), g). \tag{4}$$

The controlled PDE problem may be reformulated as an *unconstrained minimization*:

Find controls g that minimize the cost $J(g)$.

We may need first derivatives of $J(g)$:

$$\frac{\partial J}{\partial g} = \frac{\partial \mathcal{J}}{\partial u}\frac{\partial u}{\partial g} + \frac{\partial \mathcal{J}}{\partial g} \tag{5}$$

2.3: Comments

$\mathcal{J}$ is usually an integral of some quantity over the domain, or some portion thereof. For the matching problem, it might have the form

$$\mathcal{J}(u, \lambda, \alpha, R) = \int_{x=x^*} (u - u*)\, ds. \tag{6}$$

$\mathcal{J}$ could measure flow matching along a profile line, or total vorticity, or the spatial variation from a desired temperature, and so on.

$\frac{\partial \mathcal{J}}{\partial u}$ should be easily derivable, and the second factor is just u_g, the state sensitivity.

2.4: An Equation for Discrete Sensitivities

The continuous Navier Stokes horizontal momentum equation is:

$$-(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}) + R\ (u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + \frac{\partial p}{\partial x}) = 0 \qquad (7)$$

We discretize Equation (7) to solve for u^h:

$$\int_\Omega \frac{\partial u^h}{\partial x}\frac{\partial \psi}{\partial x} + \frac{\partial u^h}{\partial y}\frac{\partial \psi}{\partial y} + R(u^h\frac{\partial u^h}{\partial x} + v^h\frac{\partial u^h}{\partial y} + \frac{\partial p^h}{\partial x})\psi\ d\Omega = 0 \qquad (8)$$

For computation of derivatives of J^h, we need equations that determine derivatives of u^h.

There are two possible routes to this information.

2.5: Discretize, then Differentiate

We have already discretized Equation (7), to define u^h. Therefore, simply differentiate Equation (8) (and move into the integral, if possible) to get:

$$\begin{aligned} & \int_\Omega \frac{\partial u_R^h}{\partial x}\frac{\partial \psi}{\partial x} + \frac{\partial u_R^h}{\partial y}\frac{\partial \psi}{\partial y} \\ +R & \ (u_R^h \frac{\partial u^h}{\partial x} + u^h \frac{\partial u_R^h}{\partial x} + v_R^h \frac{\partial u^h}{\partial y} + v^h \frac{\partial u_R^h}{\partial y} + \frac{\partial p_R^h}{\partial x}) \psi \ d\Omega \\ = & \ - \int_\Omega (u^h \frac{\partial u^h}{\partial x} + v^h \frac{\partial u^h}{\partial y} + \frac{\partial p^h}{\partial x}) \psi \ d\Omega \end{aligned} \tag{9}$$

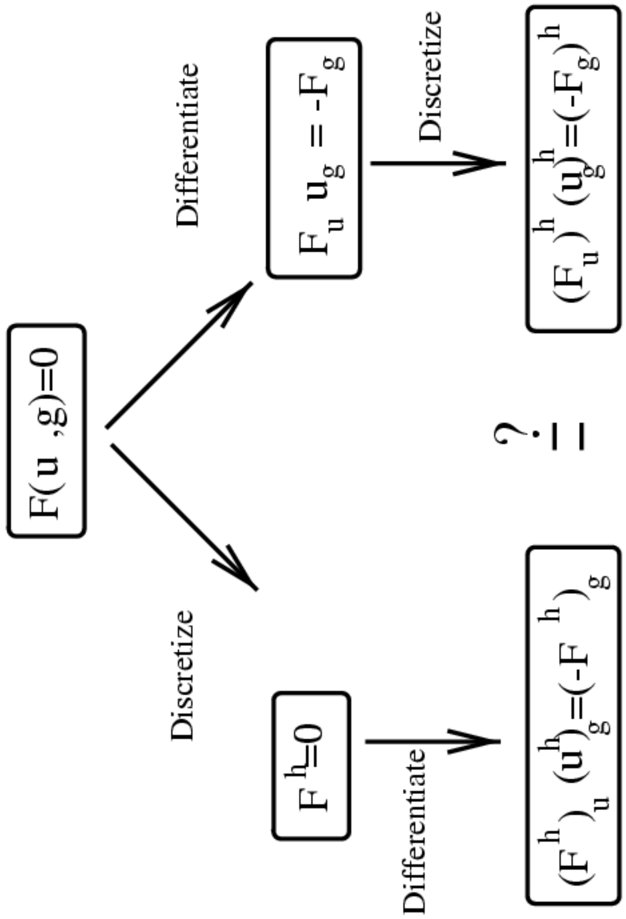
2.6: Differentiate, then Discretize

Differentiate Equation (7) with respect to R:

$$\begin{aligned} & -(\frac{\partial^2 u_R}{\partial x^2} + \frac{\partial^2 u_R}{\partial y^2}) \\ + & \; R(u_R \frac{\partial u}{\partial x} + u \frac{\partial u_R}{\partial x} + v_R \frac{\partial u}{\partial y} + v \frac{\partial u_R}{\partial y} + \frac{\partial p_R}{\partial x}) \\ = & \; -(u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + \frac{\partial p}{\partial x}) \end{aligned} \tag{10}$$

Then discretize Equation (10) to get:

$$\begin{aligned} & \int_\Omega \frac{\partial u_R}{\partial x}\frac{\partial \psi}{\partial x} + \frac{\partial u_R}{\partial y}\frac{\partial \psi}{\partial y} \\ + & \; R(u_R \frac{\partial u}{\partial x} + u \frac{\partial u_R}{\partial x} + v_R \frac{\partial u}{\partial y} + v \frac{\partial u_R}{\partial y} + \frac{\partial p_R}{\partial x}\psi) \; d\Omega \\ = & \; -\int_\Omega (u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + \frac{\partial p}{\partial x}\psi) \; d\Omega \end{aligned} \tag{11}$$



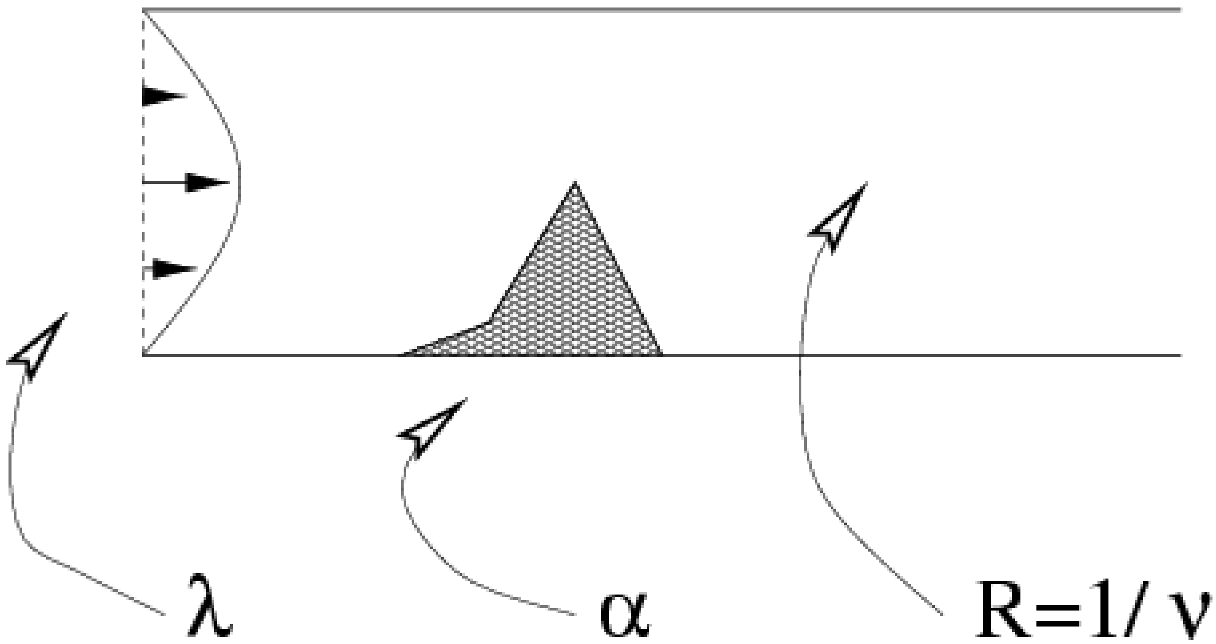
2.7: Commutivity?

2.7: Comments

- If the equations are the same, the solutions are the same.
- If $(u_g)^h$ is equal to, or suitably close to, $(u^h)_g$ we say it is a *consistent* approximation.
- Since $(u_g)^h$ is a finite element approximation to u_g, it is *convergent* to it.

3.0: THE WIND TUNNEL PROBLEM

- **3.1: Problem Parameters**
- **3.2: Optimization on** R
- **3.3: Numerical Results for** R
- **3.4: The Equation for** $(u^h)_R$
- **3.5: The Equation for** $(u_R)^h$
- **3.6: Optimization on** α
- **3.7: Numerical Results for** α
- **3.8: The Equation for** $(u^h)_\alpha$
- **3.9: The Equation for** $(u_\alpha)^h$
- **3.10: The Discrepancy Between** $(u^h)_\alpha$ **and** $(u_\alpha)^h$



3.1: Problem Parameters

- λ determines inflow;
- α sets bump shape;
- R is inverse viscosity, almost the Reynolds number.

3.2: Optimization on R

Given R, the current estimate of minimizer, compute:

- discretized $u^h(R)$;
- solution sensitivity $(u_R)^h$ or $(u^h)_R$;
- cost $J^h(R)$;
- cost sensitivity $(J_R)^h$ or $(J^h)_R$;

Then call minimization code for next estimate.

Note that:

- $(u^h)_R$ estimated by finite differences;
- $(u_R)^h$ computed by discretized sensitivity equation;
- $(J^h)_R$ uses $(u^h)_R$ in chain rule;
- $(J_R)^h$ uses $(u_R)^h$ in chain rule.

3.3: Numerical Results for R

Step	R	$J^h \times 10^{-4}$
1	1.00000	6.134345
2	1.00018	6.134015
3	1.00091	6.132
4	1.00200	6.130
5	1.00927	6.117
6	1.02	6.097
7	1.09	5.966
8	1.20	5.774
9	1.92	4.581
10	2.95	3.167
11	6.80	0.439
12	8.19	0.124
13	9.39	0.012
14	9.87	0.0004
15	9.99146	0.0000000002
16	9.99988	0.0

3.3: Comments

- The optimization proceeds smoothly. The cost sensitivity $(J_R)^h$ is estimated using the discretized sensitivities $(u_R)^h$.
- We need higher accuracy at beginning and end of iteration, when stepsize is small. Also need more accurate derivatives there.
- If finite differences are used to estimate $(u^h)_R$, from which we calculate $(J^h)_R$, essentially the same results are found. This suggests that, at least for R, *the order of differentiation and discretization doesn't matter!*.

3.4: The Equation for $(u^h)_R$

Pretend there's just one state equation, with one term in u:

$$Ru\frac{\partial u}{\partial x} = 0 \qquad (12)$$

The discretized state equation for u^h is:

$$\int_\Omega R\, u^h\, \frac{\partial u^h}{\partial x}\, u^h\, \psi\, d\Omega = 0 \qquad (13)$$

Differentiate Equation (13) and interchange orders:

$$\int_\Omega R(u^h)_R\frac{\partial u^h}{\partial x} + Ru^h\frac{\partial (u^h)_R}{\partial x}\psi\, d\Omega = \int_\Omega -u^h\frac{\partial u^h}{\partial x}\psi\, d\Omega \qquad (14)$$

3.5: The Equation for $(u_R)^h$

The continuous sensitivity equation for u_R:

$$Ru_R \frac{\partial u}{\partial x} + Ru \frac{\partial u_R}{\partial x} = -u \frac{\partial u}{\partial x} \tag{15}$$

Now discretize Equation (15) to get the equation for $(u_R)^h$, the "discretized sensitivity":

$$\int_\Omega R(u_R)^h \frac{\partial u^h}{\partial x} + Ru^h \frac{\partial (u_R)^h}{\partial x} \psi \, d\Omega = \int_\Omega -u^h \frac{\partial u^h}{\partial x} \psi \, d\Omega \tag{16}$$

This equation has the same form as that for $(u^h)_R$, therefore:

$$(u_R)^h = (u^h)_R. \tag{17}$$

3.6: Optimization on α

Given R and α, the current estimates of minimizing parameters,

- compute discretized $u^h(R, \alpha)$;
- compute cost $J^h(R, \alpha)$;
- compute solution sensitivity $(u_\alpha)^h$ or $(u^h)_\alpha$;;
- compute cost sensitivity $(J_\alpha)^h$ or $(J^h)_\alpha$;
- call minimization code for new estimate.

Note that:

- α is a vector of 3 shape parameters;

3.7: Numerical Results For α

Step	$(J^h)_\alpha$	$(J_\alpha)^h$
1	0.429	6.134345
2	0.389	6.134015
3	0.00922	6.132
4	0.00901	6.130
5	0.00813	6.117
6	0.00658	6.097
7	0.00067	5.966
8	0.00056	5.774
9	0.00001	4.581
10	0.000009	3.167
11	0.000009	0.439
12	0.000008	0.124
13	0.000008	0.012
14	0.000007	0.0004
15	0.000006	0.0000000002
20	0.000002	0.012
–	———	——
30	0.000001	
39	2.5×10^{-13}	

3.7: Comments:

On the third step, we got the inflow about right. After step 10, iterates enter a long thin valley.

3.8: The Equation for $(u^h)_\alpha$

Differentiate Equation (13) and interchange orders:

$$\int_\Omega R(u^h)_\alpha \frac{\partial u^h}{\partial x} + Ru^h \frac{\partial (u^h)_\alpha}{\partial x} \psi \; d\Omega \tag{18}$$

$$\int_\Gamma Ru^h \frac{\partial u^h}{\partial x} \psi \frac{dn}{ds} \; ds = 0 \tag{19}$$

3.8: Comments:

- Situation much messier when replace integral by finite sum.
- Consider a single element on the boundary.

3.9: The Equation for $(u_\alpha)^h$

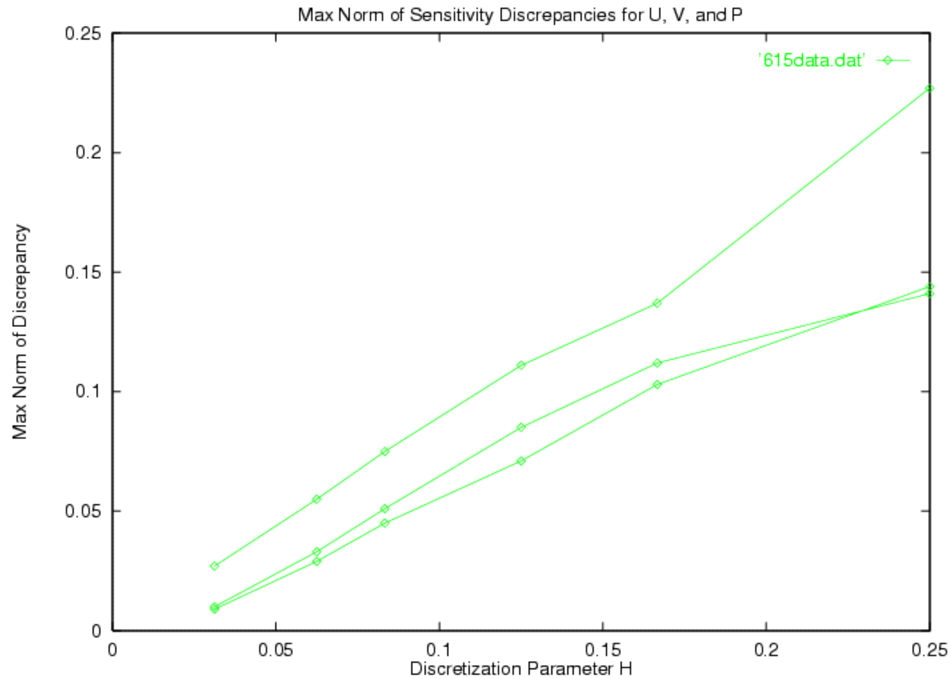
If we simply differentiate the continuous equation:

$$R(u_\alpha \frac{\partial u}{\partial x} + u \frac{\partial u_\alpha}{\partial x}) = 0 \quad (20)$$

$$u_\alpha = -u_y y_\alpha \quad (21)$$

...and then discretize, we get:

$$\int_\Omega R(u^h)_\alpha \frac{\partial u^h}{\partial x} + R u^h \frac{\partial (u^h)_\alpha}{\partial x} \psi \; d\Omega = 0 \quad (22)$$



3.10: The Discrepancy Between $(u^h)_\alpha$ and $(u_\alpha)^h$

3.10: Comments

- Discrepancy dropping like $O(h)$.
- Suggests that discrepancy problems can be reduced by reducing h.
- Approximation error of BC slows down convergence of $(u_\alpha)^h$ to u_α.

4.0: OBSERVATIONS AND CONCLUSIONS

Discretized Sensitivities $(u_R)^h$:

- Easy to derive sensitivity equation;
- Easy to discretize the sensitivity equation (avoids differentiation of the discretization method);
- Cheap to solve for;
- $(u_R)^h$ converges to continuous u_R at a known rate, because it is a finite element approximation.

Discretized Sensitivities in Numerical Optimization:

- $(u_R)^h$ is not necessarily a derivative of anything.
- Differentation and discretization need not commute.
- Sensitivity-approximated cost gradients can be unreliable.
- Errors in cost gradients hurt most at the beginning and end.

Differentiation of Discrete Solution, $(u^h)_R$:

- May be difficult to calculate explicitly; finite differences or ADIFOR may be necessary

- Guaranteed to be correct derivative of u^h, and to produce correct derivatives J_R.
- The convergence properties of $(u^h)_R$ are not understood. They are not even *guaranteed*!
- Differentiation of the discretization method may introduce extra complexitities, nonphysical local minima, and poor convergence to u_R.

4.0 Comments:

Discretized Sensitivities:

- Note that the boundary conditions may be hard to properly evaluate. In my case, I needed the value of , which was known to an order of $O(h)$ rather than $O(h^2)$, degrading my approximation.

Discretized Sensitivities in Numerical Optimization:

- Specificially, differentiation with respect to a geometric parameter, and discretization using finite elements.