Sensitivity Calculations for fun and profit

I. G. Abel [1, 2] G. W. Hammett [2]

[1] Princeton Center for Theoretical Science

[2] Princeton Plasma Physics Laboratory

Vienna Shindig 2016

Outline

Outline

Outline

Outline

What the ??

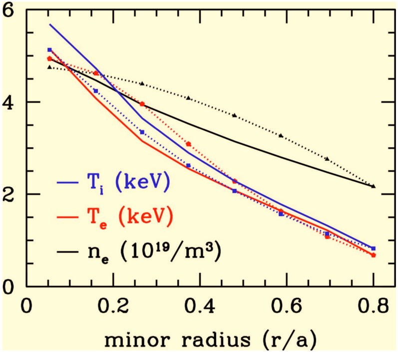
What are they?

Derivatives of functionals of our system, with respect to input parameters. Paradigmatically,

$$\frac{\partial Q_i}{\partial \left(R/\ L_{Ti} \right)} \tag{1}$$

Uses

- Stiffness
- Implicit Transport Algorithms (TRINITY)
- Dynamical Reduction
- Uncertainty Quantification
- Gradient-Driven optimization



Adjoint Methods for Sensitivity Calculations

- If your forward system is $F(g; m) = 0$
- Your functional is $Q[g, m]$, then

$$\frac{\partial Q}{\partial m} = \frac{\partial Q}{\partial g}\frac{\partial g}{\partial m} + \frac{\partial Q}{\partial m} = \left\langle \frac{\partial Q}{\partial g}, \frac{\partial g}{\partial m} \right\rangle + \frac{\partial Q}{\partial m} \tag{2}$$

- $\frac{\partial g}{\partial m}$ satisfies

$$\frac{\partial F}{\partial g}\frac{\partial g}{\partial m} = -\frac{\partial F}{\partial m} \tag{3}$$

- Use inner product structure

$$\left\langle \lambda, \frac{\partial F}{\partial g}\frac{\partial g}{\partial m} \right\rangle = -\left\langle \lambda, \frac{\partial F}{\partial m} \right\rangle \quad \text{implies} \quad \left\langle \frac{\partial F}{\partial g}^{\dagger}\lambda, \frac{\partial g}{\partial m} \right\rangle = -\left\langle \lambda, \frac{\partial F}{\partial m} \right\rangle \tag{4}$$

- $$\text{Hence, if} \quad \frac{\partial F}{\partial g}^{\dagger}\lambda = \frac{\partial Q}{\partial g} \quad \text{Then} \quad \frac{\partial Q}{\partial m} = -\left\langle \lambda, \frac{\partial F}{\partial m} \right\rangle + \frac{\partial Q}{\partial m} \tag{5}$$

Adjoint Methods for Sensitivity Calculations II

Direct Methods

- Just make a two-point approximation.

OR

- Differentiate with respect to paramter, solve tangent equtions.

Scales linearly with number of paramters to vary.

Adjoint Methods

- Solve adjoint state equation (backwards in time, around reference value)
- Take inner product of adjoint solution to get needed derivatives.

Scales linearly with number of functionals to differentiate.

The Gyrokinetic System

- Use $g_s = h_s - \frac{Z_s e}{T_s}\langle\delta\varphi\rangle_{\boldsymbol{R}} F_s$ to isolate the time derivative,
- Then the Gyrokinetic system is given by

$$\frac{\partial g_s}{\partial t} = (\mathbb{L}g)_s + (C_{\mathrm{GK}}g)_s - \frac{c}{B}\boldsymbol{b}\times\frac{\partial\langle\delta\varphi\rangle_{\boldsymbol{R}}}{\partial\boldsymbol{R}_s}\cdot\frac{\partial g_s}{\partial\boldsymbol{R}_s}. \tag{6}$$

$$\begin{aligned}(\mathbb{L}g)_s = &-\left(v_{\|}\boldsymbol{b} + \boldsymbol{V}_{\mathrm{D}s}\right)\cdot\frac{\partial}{\partial\boldsymbol{R}_s}\left(g_s + \frac{Z_s e}{T_s}\langle\delta\varphi[g]\rangle_{\boldsymbol{R}} F_{0s}\right)\\ &-\frac{c}{B}\frac{\partial F_{0s}}{\partial\psi}\boldsymbol{b}\times\frac{\partial\langle\delta\varphi[g]\rangle_{\boldsymbol{R}}}{\partial\boldsymbol{R}_s}\cdot\nabla\psi.\end{aligned} \tag{7}$$

- With

$$\delta\varphi = \left[\sum_s \frac{Z_s^2 e^2 n_s}{T_s}\left(1-\Gamma_{0s}\right)\right]^{-1}\sum_s Z_s e\int d^3\boldsymbol{w}\langle g_s\rangle_{\boldsymbol{r}}, \tag{8}$$

Free Energy as an Inner Product

- This conserves

$$W[g] = \sum_s \left\langle \left\langle \int d^3\mathbf{w} \frac{T_s g_s^2}{2F_{0s}} \right\rangle_\perp \right\rangle_\psi + \frac{1}{2} \langle \langle \delta\varphi[g] \, \Phi \delta\varphi[g] \rangle_\perp \rangle_\psi \,, \quad (9)$$

- We use this as a norm, and can show that it has a consistent inner product given by

$$\left\langle g^{(1)}, g^{(2)} \right\rangle = \sum_s \left\langle \left\langle \int d^3\mathbf{w} \frac{T_s g_s^{(1)} g_s^{(2)}}{2F_s} + \frac{1}{2} \delta\varphi\left[g^{(1)}\right] \Phi \delta\varphi\left[g^{(2)}\right] \right\rangle_\perp \right\rangle_\psi \,, \quad (10)$$

- This is consistent in that $W = \|g\|^2 = \langle g, g \rangle$.
- This is a good norm for subcritical studies, but we also need a time average. Thus we use

$$\left\langle g^{(1)}, g^{(2)} \right\rangle_{\mathrm{T}} = \sum_s \left\langle \left\langle \int d^3\mathbf{w} \frac{T_s g_s^{(1)} g_s^{(2)}}{2F_s} + \frac{1}{2} \delta\varphi\left[g^{(1)}\right] \Phi \delta\varphi\left[g^{(2)}\right] \right\rangle_{\mathrm{turb}} \right\rangle_\psi \,, \quad (11)$$

The Adjoint Gyrokinetic Operator

We then calculate the adjoint gyrokinetic operator in this norm:

$$(\mathbb{L}^\dagger g)_s = (v_\parallel \boldsymbol{b} + \boldsymbol{V}_{\mathrm{D}s}) \cdot \frac{\partial}{\partial \boldsymbol{R}_s}\left(g_s + \frac{Z_s e}{T_s}\langle \delta\varphi[g]\rangle_{\boldsymbol{R}} F_{0s}\right) - \frac{Z_s e}{T_s} F_{0s} \boldsymbol{b} \times \frac{\partial \langle \varsigma[g]\rangle_{\boldsymbol{R}}}{\partial \boldsymbol{R}_s} \cdot \nabla\psi. \quad (12)$$

The adjoint source field ζ is given by

$$\varsigma[g] = \left(\sum_s \frac{Z_s^2 e^2 n_s}{T_s}\right)^{-1} \frac{c}{B}\left\{\sum_s \left(\int d^3\boldsymbol{w}\, T_s g_s + Z_s e n_s \Gamma_{0s} \delta\varphi[g]\right) \frac{d \ln N_s}{d\psi} + \sum_s \left[\int d^3\boldsymbol{w}\, T_s \left(\frac{\varepsilon_s}{T_s} - \frac{3}{2}\right) g_s + Z_s e n_s \Gamma_{1s} \delta\varphi[g]\right] \frac{d \ln T_s}{d\psi}\right\} \quad (13)$$

The Adjoint Gyrokinetic State Equation

- Now we know $\mathbb{L}^\dagger$ we can define the adjoint state equation to be

$$\left(-\frac{\partial}{\partial t} + \mathbb{L}^\dagger - C_{\mathrm{GK}} + \mathbb{N}[g_0] - \frac{c}{B}\boldsymbol{b} \times \nabla \langle \delta\varphi[g_0] \rangle_{\boldsymbol{R}} \cdot \frac{\partial}{\partial \boldsymbol{R}_s} \right) g^\dagger = S_Q \qquad (14)$$

- The source depends on the quantitiy of interest, and is given by the functional derivative

$$\begin{aligned} S_Q = \frac{\partial Q_i}{\partial g} = &\frac{c}{B}\frac{\partial \langle \delta\varphi[g_0] \rangle_{\boldsymbol{R}}}{\partial y} \left(\frac{m_i v^2}{2} - \frac{3}{2} T_i \right) F_i \delta(s = i) \\ &+ \frac{c}{B} \frac{Z_s e}{\sum_{s'} Z_{s'}^2 e^2 / T_{s'}} \frac{\partial}{\partial y} \left\langle \int \mathrm{d}^3 \boldsymbol{v} \left(\frac{m_i v^2}{2} - \frac{3}{2} T_i \right) \langle g_{0i} \rangle_{\boldsymbol{r}} \right\rangle_{\boldsymbol{R}} \end{aligned} \qquad (15)$$

- This is formally solved backwards in time with final conditions $g^\dagger(T) = 0$. (see later)
- This equation is derived around a fixed background state g_0, and gives the derivative at that point.

Expressions for $\partial Q / \partial (\cdots)$

- Now we can compute

$$\frac{\partial Q_i}{\partial R/L_{Ti}} = -\left\langle g^\dagger, \frac{c}{B}\delta(s=i)\left(\frac{m_i v^2}{2} - \frac{3T_i}{2}\right)\frac{\partial \delta\varphi[g_0]}{\partial y}F_i\right\rangle \quad (16)$$

Or **any** other derivative, with appropriate RHS, e.g. in shifted circles,

$$\frac{\partial Q_i}{\partial q} = +\frac{1}{q}\left\langle g^\dagger, v_\parallel \nabla_\parallel \left(g_{0s} + \frac{Z_s e \delta\varphi[g_0]}{T_s}F_s\right)\right\rangle \quad (17)$$

- As above, we can do this all over again for perturbations around a transport solution
- ...

Obvious Distraction

What I didn't Tell You

Problems

- Adjoints grow without bound
- This is due to the butterfly effect, and interchange of averaging limits in our derivation.

This is why adjoints aren't used in fluid turbulence

Solutions

- Ensemble Adjoint Method – run backwards from noise for intermediate times
- Shadowing trajectories – pick a $g^\dagger(T)$ to prevent trajectory flying away
- Variations on the theme – perhaps a mixture of both

May Be Expensive

Summary & Future Developments

- Adjoint Methods may save the day
- But they may be expensive
- Test This!
- ...
- Profit