INFORMATION GEOMETRIC REGULARIZATION NOTES

DAVID WINTERS, FLORIAN SCHÄFER, AND QI TANG

Abstract. [ADD!]

Key words. [ADD!]

MSC codes. [ADD!]

1. Introduction. [1]

2. Problem Set Up.

The Barotropic Euler equation describes an iviscid and barotropic gas.

$$\partial_t \begin{pmatrix} \rho \boldsymbol{u} \\ \rho \end{pmatrix} + \overrightarrow{\operatorname{div}} \begin{pmatrix} \rho \boldsymbol{u} \otimes \boldsymbol{u} + P(\rho)\mathbf{I} \\ \rho \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ 0 \end{pmatrix}$$

Unfortunately, this equation sometimes produces shocks as is most easily seen in the case of the Burgers equation having $P, f \equiv 0$. Solutions to this equation can be thought of as **dual geodesics** on the (infinite-dimensional) diffeomorphism manifold, and a shock is achieved upon reaching the boundary of this manifold. To fix this, we will modify the geometry of the manifold so that dual geodesics are kept away from the boundary using information-geometric forces. The case of an Eulerian description of Euler's equations has been solved for in [1] and is given below:

$$\begin{cases} \partial_t \begin{pmatrix} \rho \boldsymbol{u} \\ \rho \end{pmatrix} + \overrightarrow{\operatorname{div}} \begin{pmatrix} \rho \boldsymbol{u} \otimes \boldsymbol{u} + (P(\rho) + \Sigma)\mathbf{I} \\ \rho \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ 0 \end{pmatrix} \\ \rho^{-1}\Sigma - \alpha \overrightarrow{\operatorname{div}} \left(\rho^{-1} \nabla \Sigma \right) = \alpha \left(\operatorname{tr}([\mathrm{D}\boldsymbol{u}])^2 + \operatorname{tr}\left([\mathrm{D}\boldsymbol{u}]^2\right) \right). \end{cases}$$

The term α controls the magnitude of the IG force, and in the limit $\alpha \to 0$, we expect to recover the solution to the original PDE. This is an alternative method to constructing solutions to the standard method of adding a vanishing viscosity term.

3. Burgers' Equation.

The article continues by working the case of Burgers equation in detail:

$$\partial_t u + u\,\partial_x u = \partial_t + \frac{1}{2}\partial_x u^2 = 0.$$

This equation can be solved by the method of characteristics until the time at which two characteristics intersect. At this point a shock forms and a differentiable solution no longer exists. When doing this by hand, we may create a weak solution by merging the two characteristics into a single line.

If we are evolving two characteristics on a computer, we may use interior point methods to prevent them from intersecting. This is done by adding a barrier term, $-\log(\det)$, that makes each characteristic approach the average of the two lines asymptotically. So, while we start with characteristic lines for the burgers' equation, when two lines approach one another, they asymptotically approach the average of the two slopes. To improve the approximation of a weak solution, we can rerun the simulation with a weaker barrier term and thus the two characteristics approach more quickly.

This manuscript is for review purposes only.

In the context of information geometry, the negative log barrier function produces a dually flat (information) geometry. That is, the barrier function $\psi = \frac{1}{2}x^2 - \alpha\log(x')$ induces a dual set of coordinates $\eta = \nabla\psi$, whose inverse exists by the strict convexity of ψ. This also induces a Riemannian metric via its Hessian. When we say that the Burgers' equation has a dually flat information geometry, we mean its solutions are straight lines in the coordinates of η. We may map back to coordinates in terms of x via $x(t) = \nabla\psi^{-1}\eta(t)$. By adding this log term, the "straight lines" avoid the boundary and allow the simulation to proceed without shocks. There are some details I am missing here such as the coordinates $(\overline{\Phi}, \Phi')$ to define ψ, but I will instead jump to the infinite dimensional case.

Instead of considering a finite number of geodesics, we can imagine that the solution Φ_t is instead a diffeomorphism which tells us the net deformation of the initial data at time zero. As the initial data evolves, we record this as a new diffeomorphism which evolves differentiably in time. The correct notion of coordinates are

$$\overline{\Phi} = \int_{\mathbb{R}} \Phi(x) - x\, dx \quad \text{and} \quad \Phi' = \partial_x \Phi.$$

So, $\overline{\Phi}$ is a one-dimensional coordinate system in the space of diffeomorphisms, while Φ' is function-valued (infinite collection of coordinates). Then the barrier in the infinite dimensional setting is:

$$\psi(\Phi) = \psi_E(\Phi) - \alpha \int_{\mathbb{R}} \log(\partial_x \Phi)\, dx \qquad \psi_E(\Phi) = \frac{1}{2}\|\Phi(\cdot) - (\cdot)\|_{L^2}^2.$$

If we extend from Burgers to Euler where Φ takes values in $\mathbb{R}^n$:

$$\hat{\psi}(\Phi) = \psi(\Phi, \det D\Phi) \coloneqq \psi_E(\Phi) - \alpha \int_{\mathbb{R}^d} \log(\det D\Phi)\, dx.$$

Finally, we want to be able to recover the PDE from the deformation maps $\Phi(x)$ (change back to Eulerian coordinates). This is done via:

$$u(x) = \dot{\Phi}(x) \;\text{ and }\; \rho(x) = (\det(D\Phi))^{-1} \implies \int_{\mathbb{R}^d} \log(\det D\Phi)\, dx = \int_{\mathbb{R}^d} \log(\rho)\rho\, dx.$$

4. Derivation for Flat Geometry. They derive in $\mathbb{R}^d$ for a general barrier ψ. Ideas:

- We can extend a tangent vectors beginning at a particular point into a "line" via the exponential map given by the Riemannian metric.
- The dual exponential map maps a tangent vector to the dual space direction, extends this to a path via the exponential map, and changes back coordinates $\eta \to x$.
- These geodesics tell us how things will evolve in the absence of forces.

With forces, we have a dual Newton's Second Law (derived later) for force K:

$$\ddot{\Phi}_t + \left[\mathrm{D}^2\psi\left(\Phi_t\right)\right]^{-1}\left[\mathrm{D}^3\psi\left(\Phi_t\right)\right]\left(\dot{\Phi}_t, \dot{\Phi}_t\right) = K\left(t, \Phi_t, \dot{\Phi}_t\right).$$

In the univariate case, the barrier function is:

$$\psi(\Phi) = \frac{1}{2}\int_{\mathbb{R}} |\Phi(x) - x|^2 \, \mathrm{d}x + \alpha \int_{\mathbb{R}} -\log \partial_x \Phi(x) \mathrm{d}x.$$

The article goes on to explicitly solve for Newton's Second Law in terms of the barrier function given above. Because ψ is a linear functional, we need to take variations of the functional and are derived in 3.2.2. The conclusion is the equation:

$$\left((\cdot) - \alpha\partial_x\left(\left[\partial_x\Phi_t\right]^{-2}\left[\partial_x(\cdot)\right]\right)\right)\left(\ddot{\Phi}_t - K\left(t, \Phi_t, \dot{\Phi}_t\right)\right) = -2\alpha\partial_x\left(\left[\partial_x\Phi_t\right]^{-3}\left[\partial_x\dot{\Phi}_t\right]^2\right)$$

Next, they derive the equation in Eulerian coordinates using the equations given at the end of section 2. The final part of section 3 is to derive a new conservation law.

Questions to Answer:. 1. How did we derive Newton's Second Law?

We will perform the case of no external forces. Using the equation for the dual geodesic:

$$\Phi_t = Exp^{\psi}_{(\Phi_0)}(t\dot{\Phi}_0) = \nabla\psi^{-1} \circ (\nabla\psi(\Phi_0) + t[D^2\psi(\Phi_0)]\dot{\Phi}_0)$$

In which we can clearly see that we are transforming a straight line in the dual space. Next, we take a derivative:

$$\dot{\Phi}_t = \frac{d}{dt}\nabla\psi^{-1}\circ(\nabla\psi(\Phi_0)+t[D^2\psi(\Phi_0)]\dot{\Phi}_0) = (\psi^{-1})'(\nabla\psi(\Phi_0)+t[D^2\psi(\Phi_0)]\dot{\Phi}_0)\cdot[D^2\psi(\Phi_0)]\dot{\Phi}_0$$

$$= D[D\psi^{-1}](\nabla\psi(\Phi_t)) \cdot [D^2\psi(\Phi_0)]\dot{\Phi}_0 = [D^2\psi(\Phi_t)]^{-1} \cdot [D^2\psi(\Phi_0)]\dot{\Phi}_0$$

because

$$\nabla(\nabla\psi)^{-1}(y) = \left(\nabla^2\psi\left((\nabla\psi)^{-1}(y)\right)\right)^{-1}.$$

Similarly, we can take the second derivative:

$$\ddot{\Phi}_t = \frac{d}{dt}[D^2\psi(\Phi_t)]^{-1}\cdot[D^2\psi(\Phi_0)]\dot{\Phi}_0 = -[D^2\psi(\Phi_t)]^{-1}\frac{d[D^2\psi(\Phi_t)]}{dt}[D^2\psi(\Phi_t)]^{-1}\cdot[D^2\psi(\Phi_0)]\dot{\Phi}_0$$

$$= -[D^2\psi(\Phi_t)]^{-1}[D^3\psi(\Phi_t)](\dot{\Phi}_t)[D^2\psi(\Phi_t)]^{-1}\cdot[D^2\psi(\Phi_0)]\dot{\Phi}_0 = -[D^2\psi(\Phi_t)]^{-1}[D^3\psi(\Phi_t)](\dot{\Phi}_t, \dot{\Phi}_t).$$

Here we used the fact that:

$$\frac{d}{dt}[D^2\psi(\Phi_t)][D^2\psi(\Phi_t)]^{-1} = \frac{d}{dt}I = 0 = [D^2\psi(\Phi_t)]\frac{d[D^2\psi(\Phi_t)]^{-1}}{dt} + \frac{d[D^2\psi(\Phi_t)]}{dt}[D^2\psi(\Phi_t)]^{-1}.$$

Thus we have derived Newton's Second Law:

$$\ddot{\Phi}_t + [D^2\psi(\Phi_t)]^{-1}[D^3\psi(\Phi_t)](\dot{\Phi}_t, \dot{\Phi}_t) = 0.$$

2. How to derive the Euler-Lagrange equations? This method is ultimately not what we are looking for, but this shows an interesting derivation.

Here, we will be using the Lagrangian $\mathcal{L} = \frac{1}{2}\dot{\Phi}^T[D^2\psi(\Phi)]\dot{\Phi}$. The Euler-Lagrange equations:

$$\frac{d}{dt}\frac{\mathcal{L}}{\partial\dot{\Phi}} = \frac{\partial\mathcal{L}}{\partial\Phi} \quad \text{becomes} \quad \frac{d}{dt}[D^2\psi(\Phi)]\dot{\Phi} = \frac{1}{2}[D^3\psi](\dot{\Phi}, \dot{\Phi}, \cdot)$$

or alternatively:

$$[D^3\psi](\dot{\Phi}, \dot{\Phi}, \cdot) + [D^2\psi(\Phi)]\ddot{\Phi} = \frac{1}{2}[D^3\psi](\dot{\Phi}, \dot{\Phi}, \cdot) \quad \Longrightarrow \quad \ddot{\Phi} + \frac{1}{2}[D^2\psi(\Phi)]^{-1}[D^3\psi](\dot{\Phi}, \dot{\Phi}, \cdot) = 0.$$

So, if we have an external force $F(\Phi, \dot{\Phi})$, the equation becomes:

$$\ddot{\Phi} + \frac{1}{2}[D^2\psi(\Phi)]^{-1}[D^3\psi](\dot{\Phi}, \dot{\Phi}, \cdot) = F(\Phi, \dot{\Phi}).$$

Instead, we would like to do the same for the dual equations.

3. The case of the dual equations:

5. Multivariate Case.

6. Numerical experiments.

7. Comparison, conclusion, and outlook.

Acknowledgments.

REFERENCES

[1] R. Cao and F. Schäfer, *Information geometric regularization of unidimensional pressureless Euler equations yields global strong solutions*, arXiv preprint arXiv:2411.15121, (2024).