Note: Covid extension and duration.
κ bugs in code.

1 Comparing Fixed Point with Newton-Krylov

1.1 1D Problems

The Dirichlet exact problem is solved with $N = 20$ and $n = 10$ in 3 seconds. The error in u and v decreases down to the order of 10^{-15}. The Neumann exact problem is solved with $N = 30$ and $n = 19$ in 15 seconds. The error in u and v decreases down to the order of 10^{-5}, which is very different from the Dirichlet problem. The same happens in the 2D case. The inner number of iterations goes up to 100 in the later outer iterations. However, this does not happen if I choose $N = 20$ and $n = 10$ and the error even decreases down to 10^{-12}. In general, the more points I use, the worse the error gets. This is also the case for the Dirichlet problem, but it remains reasonably accurate despite this, while the Neumann problem is only accurate to 10^{-2} or similar for high numbers of points. In particular, this happens when we increase n, increasing N does not seem to be the issue.

I then apply this observation to one of the paper examples (Example 2), with $\kappa = 1$. When I run the problem with $N = 50$ and $n = 10$, and then with the old code with $n = 30$ and interpolate to $n = 10$ in time, We get an error of 0.0362. If I run both versions with $n = 30$ we get an error of 3.8640×10^{-4}, which seems to contradict the above observation.
For $n = 30$, we get for $\kappa = -1$ and error of 0.0013 and for $\kappa = 0$ and error of 1.2519×10^{-8}. These problems take 56 seconds to solve.

For the paper example 1, we again choose $N = 50$ and $n = 30$. We get for $\kappa = 0$ an error of 1.5563×10^{-13}, for $\kappa = -1$ we have the error 0.0174 and for $\kappa = 1$, the error is 0.0037.

For the third example, we have Dirichlet BCs and again choose $N = 50$ and $n = 30$. We get for $\kappa = 0$ the eroor is 1.8015×10^{-7}, for $\kappa = -1$ the error is 1.6645×10^{-7} and for $\kappa = 1$ the error is 1.9485×10^{-7}.

1.2 2D Problems

The exact 2D Dirichlet problem (with and without external potential), for $N = 20$ and $n = 10$, has an error of 10^{-15} for both u and v and is solved in 40 and 60 seconds respectively. The exact Neumann problem reduces the error in u to 10^{-8} and in v to 10^{-7}. The inner iterations reach 100 and it takes 170 seconds. For the first paper problem, we have the following results.

For $\kappa = -1$ the error is 0.0032, for $\kappa = 0$, the error is 3.6813×10^{-9} and for $\kappa = 1$ the error is 0.0014. For the second paper problem, we have the following results. For $\kappa = -1$ the error is 0.0062, for $\kappa = 0$, the error is 3.1640×10^{-4} and for $\kappa = 1$ the error is 0.0017. All of these problems take 300 to 500 seconds.

2 H_1 control and similar

Below is a summary of some of the papers I found on regularization terms involving the gradient or divergence of the control. There are some references that suggest that no-flux boundary conditions on the control have to be imposed when solving the gradient equation. However, a lot of papers just defined the control boundary conditions as part of their modelling choice. I am not sure if that's allowed? Only the first reference mentioned below provides any kind of derivation.

In the book 'Perspectives in flow control and optimization', Chapter 2, by Gunzburger the author poses a problem with a control including a gradient, see Figure 1. The derivation shows the resulting no-flux boundary condition for the control.

The paper 'Optimal Control of Obstacle Problems by H1-Obstacles' by Ito is concerned with H_1 regularization, but as far as I can tell, neglects all boundary conditions. In the paper 'Parallel Multiscale Gauss Newton Krylov Methods for Inverse Wave Propagation', the control is a gradient. They do apply no-flux boundary conditions, but state that they 'assume' them, rather than derive them. The paper 'Image Sequence Interpolation Based on Optical Flow, Segmentation, and Optimal Control' has a gradient regularizer, for which Dirichlet boundary conditions are assumed in the optimality system.

In the paper 'Optimal Control Formulation for Determining Optical Flow', Borzi, Ito and Kunisch propose a cost functional with regularization terms for the different derivatives of the control. They impose Dirichlet boundary conditions by restricting the admissible controls to satisfy this condition and observe that the time derivative being zero at the temporal boundaries is a natural boundary condition for the control if a term of the form $\int \Phi \left(\frac{\partial \vec{w}}{\partial t}\right)^2$ is in the cost functional.

The above paper is refering to another paper for their model, called 'Image Sequence Interpolation Using Optimal Control'. However, there they don't have a gradient of the control b in the cost functional but $\nabla \Delta b$. This results in $b = 0$, $\nabla_n b = 0$, and $\Delta b = 0$ on the boundary.

In 'Control Problems for the Navier Stokes Equations' by Gunzburger et al, a H_1 control for the boundary is introduced, and no-flux boundary conditions are imposed. However, it is unclear whether it is an assumption or results from the derivation of the optimality conditions.

The paper by Mang et al 'Constrained H1 Regularization Schemes for Diffeomorphic Image

Registration' proposes a different way of enforcing a divergence free flow, see Figure 2. They set periodic boundary conditions, which I cannot see to be derived in the process of getting the optimality system. They mention ill-posedness issues for $\nabla \cdot v = 0$, but I don't know what they would be/ if that applies to us. The paper 'On Some Control Problems in Fluid Mechanics', a curl is introduced in the cost functional, however, it is the curl of the state variable, therefore entering the optimality system in the adjoint. We have Dirichlet boundary conditions, so it doesn't seem like the curl is entering the boundary conditions.

The objective, or cost, or performance functional is given by

$$\mathcal{J}(\phi, f, g) = \frac{\sigma_3}{2}\int_{\Omega_2} |\nabla\phi|^2 \, d\Omega + \frac{\sigma_4}{2}\int_{\Gamma_2}\left(\frac{\partial\phi}{\partial n} - \Psi\right)^2 d\Gamma$$
$$+\frac{\sigma_1}{2}\int_{\Omega_1}(\sigma_5|\nabla f|^2 + f^2)\, d\Omega + \frac{\sigma_2}{2}\int_{\Gamma_1}(\sigma_6|\nabla_s g|^2 + g^2)\, d\Gamma\,, \tag{2.76}$$

(infinite-dimensional) scalar-valued control functions. The constraints are the nonlinear, second-order, elliptic partial differential equation

$$-\Delta\phi + \phi^3 = \begin{cases} f & \text{in } \Omega_1, \\ 0 & \text{in } \Omega\backslash\Omega_1, \end{cases} \tag{2.74}$$

along with the boundary condition

$$\phi = \begin{cases} g & \text{on } \Gamma_1, \\ 0 & \text{on } \Gamma\backslash\Gamma_1\,. \end{cases} \tag{2.75}$$

Setting the first variation of $\mathcal{L}$ with respect to the distributed control f to zero results in

$$\int_{\Omega_1}\Big(\sigma_1(\sigma_5\nabla f \cdot \nabla\widetilde{f} + f\widetilde{f}) + \xi\widetilde{f}\Big)\, d\Omega = 0\,,$$

[39] At first glance, the adjoint system (2.78) seems very complicated due to the appearance of the jump conditions along $\partial\Omega_2$. However, in a finite element discretization of (2.78), the first jump condition can be enforced by using continuous test and trial spaces and the second is a natural interface condition; thus, the jump conditions pose no great difficulty.

where the variation $\widetilde{f}$ is arbitrary. Integrating by parts to remove derivatives from $\widetilde{f}$, one obtains

$$\int_{\Omega_1}\widetilde{f}\Big(\sigma_1(-\sigma_5\Delta f + f) + \xi\Big)\, d\Omega + \sigma_1\sigma_5\int_{\partial\Omega_1}\widetilde{f}\frac{\partial f}{\partial n}\, d\Gamma = 0$$

from which it follows that

$$-\sigma_5\Delta f + f = -\frac{1}{\sigma_1}\xi \quad \text{in } \Omega_1 \qquad \text{and} \qquad \sigma_5\frac{\partial f}{\partial n} = 0 \quad \text{on } \partial\Omega_1\,.$$

Figure 1: Gunzburger Book Result

1.1. Outline of the method. We introduce a pseudo-time variable $t > 0$ and solve for a *stationary velocity field* $\boldsymbol{v} \in \mathcal{V}$, $\mathcal{V} \subset L^2(\Omega)^d$, and a *mass source* $w \in \mathcal{W}$, $\mathcal{W} \subset L^2(\Omega)$, as follows:

$$\min_{\boldsymbol{v},w} \; \frac{1}{2}\|m_R - m_1\|^2_{L^2(\Omega)} + \frac{\beta_v}{2}\|\boldsymbol{v}\|^q_{\mathcal{V}} + \frac{\beta_w}{2}\|w\|^2_{\mathcal{W}} \tag{2a}$$

subject to

$$\partial_t m + \nabla m \cdot \boldsymbol{v} = 0 \qquad \text{in } \Omega \times (0,1], \tag{2b}$$

$$m - m_T = 0 \qquad \text{in } \Omega \times \{0\}, \tag{2c}$$

$$\nabla \cdot \boldsymbol{v} - w = 0 \qquad \text{in } \Omega, \tag{2d}$$

Figure 2: Mang