ON STABILIZED FINITE ELEMENT METHODS FOR THE STOKES PROBLEM IN THE SMALL TIME STEP LIMIT

Max Gunzburger

School of Computational Science

Florida State University

Joint work with

Pavel Bochev and Richard Lehoucq

Computational Mathematics and Algorithms Department

Sandia National Laboratories

Last updated in 2006

WHY IS THE SMALL TIME STEP LIMIT INTERESTING?

- Reacting flow simulations
  - the accurate temporal resolution of fast chemical reactions requires small time steps, even for implicit time stepping schemes
  - in fact, the time step often needs to be much smaller than that required for the accurate resolution of diffusion and convection phenomena
  - thus, the time step used is disparately small compared to the spatial grid size

- Low-order time stepping schemes used in conjunction with high-order spatial discretizations
  - for example, such a combination can yield errors proportional to

$$C_1 \Delta t + C_2 h^\beta \qquad \text{for } \beta > 1$$

  - to equilibrate the temporal and spatial contributions to the error, one has to take a small time step

- So, in these and other settings, the need to use small time steps is driven by questions of accuracy
  - normally, one does not expect any issue with stability if small enough time steps are used

WHY ARE STABILIZED METHOD INTERESTING?

- Stabilized methods for effecting spatial discretizations are in very common use in engineering practice

  - they allow for the use of equal-order interpolation

    - the same degree polynomials defined with respect to the same grid can be used for all variables

  - mixed-Galerkin finite element methods do not allow for equal-order interpolation

– equal-order interpolation is very useful in large-scale multi-physics codes

  - the codes deal with several dozen independent variables, e.g., several chemical species, velocity components, etc.

  - these codes require the transfer of information between its different components

  - the tasks of keeping track of all the variables and of information transfer are both greatly facilitated when one can use equal-order interpolation

- Thus, large, multi-physics codes typically use

  – consistently stabilized finite element methods for spatial discretization

  – implicit finite difference schemes for temporal discretization

    - projection schemes are also used

- Consistently stabilized finite element methods for the Stokes problem were developed in the steady-state setting; three well-known methods in this class are

  - pressure-Poisson stabilization

    T. Hughes, L. Franca, and M. Balestra; A new finite element formulation for computational fluid dynamics: Circumventing the Babuska-Brezzi condition: A stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations, *Comp. Meth. Appl. Mech. Engrg.* **59** 1986, 85-99

  - Galerkin least-squares stabilization

    T. Hughes and L. Franca; A new finite element formulation for computational fluid dynamics: VII. The Stokes problem with various well-posed boundary conditions: symmetric formulations that converge for all velocity pressure spaces, *Comp. Meth. Appl. Mech. Engrg.* **65** 1987, 85-96

  - Douglas-Wang stabilization

    J. Douglas and J. Wang; An absolutely stabilized finite element method for the Stokes problem, *Math. Comp.* **52** 1989, 495-508

- In the steady-state setting, these methods have proven to be very useful

- These methods were then used in time-dependent settings where they also have enjoyed much success

- However, when

  consistently stabilized finite element methods and

  implicit time integrators are used

  to effect spatial and temporal discretization

  instabilities in the pressure approximations are observed

  when small time steps are used

- The poor approximation of the pressure when using consistently stabilized finite element methods with small time steps was
  - was first pointed out to us by John Shadid of the Sandia National Laboratories
  - was subsequently studied in some recent papers

    P. Bochev, M.G., and J. Shadid; On stabilized finite element methods for transient problems with varying time scales, *Comp. Meth. Appl. Mech. Engrg.* **193** 2004, 1471-1489

    P. Bochev, M.G., and R. Lehoucq; On stabilized finite element methods for transient problems with varying time scales, *Proc. ECOMASS 2004, Jyväskylä, Findland, 2004*

  - these papers focused on
    - the fully discrete in space and time equations
    - demonstrating the sufficiency of a lower bound on the time step

- Here, we focus on
  - the semi-discrete in space equations
  - demonstrating the necessity of a lower bound on the time step

- Along the way, we show that
  - the semi-discrete pressure operator is unstable
    - this identifies the source of the instabilities in the pressure approximations in the small time step limit
  - a sufficiently large time step stabilizes the fully discrete pressure equation

- So, in the time-dependent case, consistently stabilized finite element methods find themselves in a curious and unusual situation
  - the semi-discrete in space equations are ill-posed
  - the fully-discrete in space and time equations are well-posed, provided the time step is large enough
  - thus, it is not surprising that something bad happens in the small time step limit if the spatial grid size does not also tend to zero sufficiently fast
- Because the cause of the instability is the semi-discrete pressure operator
  - the instability has nothing to do with reaction or convection terms terms
    - thus, although fast reaction terms provide a motivation for using small time steps, those terms do not cause the problems
  - thus, it suffices (very fortunately for us) to simply consider the Stokes equations

MIXED-GALERKIN FINITE ELEMENT METHODS FOR THE STOKES EQUATIONS

- Consider the Stokes system

$$\begin{cases} \dfrac{\partial \mathbf{u}}{\partial t} - \triangle \mathbf{u} + \nabla p = \mathbf{f} & \text{in} \quad \Omega \times (0,T) \\ \nabla \cdot \mathbf{u} = 0 & \text{in} \quad \Omega \times (0,T) \\ \mathbf{u} = \mathbf{0} & \text{on} \quad \Gamma \times (0,T) \\ \mathbf{u}\big|_{t=0} = \mathbf{u}_0 & \text{in} \quad \Omega \end{cases}$$

- Function spaces $\quad \mathcal{V} = H_0^1(\Omega) \qquad \mathcal{P} = L_0^2(\Omega)$

- Bilinear forms: with $(\cdot,\cdot)$ denoting the $L^2(\Omega)$ inner product

$$a(\mathbf{u},\mathbf{v}) = (\nabla \mathbf{v}, \nabla \mathbf{u}) \qquad\qquad b(\mathbf{v},p) = -(p, \nabla \cdot \mathbf{v})$$

- Weak formulation: seek $\mathbf{u} \in \mathcal{V}$ and $p \in \mathcal{P}$ such that for almost all $t \in (0, T]$

$$\begin{aligned} (\mathbf{u}_t, \mathbf{v}) + \ a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) \ &= (\mathbf{f}, \mathbf{v}) \qquad \forall\, \mathbf{v} \in \mathcal{V} \\ b(\mathbf{u}, q) \qquad\qquad &= 0 \qquad\qquad \forall\, q \in \mathcal{P} \end{aligned}$$

- Semi-discretization in space

  - choose conforming finite-dimensional subspaces $\mathcal{V}^h \subset \mathcal{V}$ and $\mathcal{P}^h \subset \mathcal{P}$

  - then, seek $\mathbf{u}^h \in \mathcal{V}^h$ and $p^h \in \mathcal{P}^h$ such that for all $t \in (0, T]$

$$\begin{aligned} ({\mathbf{u}^h}_t, \mathbf{v}^h) + \ a(\mathbf{u}^h, \mathbf{v}^h) + b(\mathbf{v}^h, p^h) \ &= (\mathbf{f}, \mathbf{v}^h) \qquad \forall\, \mathbf{v}^h \in \mathcal{V}^h \\ b(\mathbf{u}^h, q^h) \qquad\qquad &= 0 \qquad\qquad \forall\, q^h \in \mathcal{P}^h \end{aligned}$$

  - the discretized initial data is determined from

$$\left(\mathbf{u}^h|_{t=0}, \mathbf{v}^h\right) = \left(\mathbf{u}_0, \mathbf{v}^h\right) \qquad \forall\, \mathbf{v}^h \in \mathcal{V}^h$$

- Let

$\{\boldsymbol{\xi}_i^h\}_{i=1}^N$ denote a basis for $\mathcal{V}^h$ $\qquad$ $\{\chi_k^h\}_{k=1}^M$ denote a basis for $\mathcal{P}^h$

so that

$$\mathbf{u}^h(\mathbf{x},t) = \sum_{j=1}^N U_j(t)\boldsymbol{\xi}_j^h(\mathbf{x}) \qquad p^h = \sum_{m=1}^M P_m(t)\chi_m^h(\mathbf{x})$$

for some functions $\{U_j(t)\}_{j=1}^N$ and $\{P_m(t)\}_{m=1}^M$

- Then, the semi-discrete equations are equivalent to

$$\begin{pmatrix} \mathbb{M}\dot{\vec{U}} \\ \vec{0} \end{pmatrix} + \begin{pmatrix} \mathbb{A} & \mathbb{B}^\top \\ \mathbb{B} & 0 \end{pmatrix} \begin{pmatrix} \vec{U} \\ \vec{P} \end{pmatrix} = \begin{pmatrix} \vec{F} \\ \vec{0} \end{pmatrix} \qquad \text{and} \qquad \mathbb{M}\vec{U}|_{t=0} = \vec{U}_0$$

where

$$\mathbb{A}_{ij} = a(\boldsymbol{\xi}_i^h, \boldsymbol{\xi}_j^h) = (\nabla\boldsymbol{\xi}_i^h, \nabla\boldsymbol{\xi}_j^h) \qquad \mathbb{B}_{im} = b(\chi_m^h, \boldsymbol{\xi}_i^h) = -(\chi_m^h, \nabla\cdot\boldsymbol{\xi}_j^h)$$

$$\mathbb{M}_{ij} = (\boldsymbol{\xi}_i^h, \boldsymbol{\xi}_j^h) \qquad (\vec{F})_i = (\mathbf{f}, \boldsymbol{\xi}_i^h) \qquad (\vec{U}_0)_j = (\mathbf{u}_0, \boldsymbol{\xi}_j^h)$$

- The conformity $\mathcal{V}^h \subset \mathcal{V}$ and $\mathcal{P}^h \subset \mathcal{P}$ of the approximation spaces is not sufficient to guarantee the stability of mixed-Galerkin finite element methods for the Stokes problem

- In addition, these spaces are required to satisfy the LBB or inf-sup compatibility condition

$$\inf_{q^h \in \mathcal{P}^h,\, q^h \neq 0} \; \sup_{\mathbf{v}^h \in \mathcal{V}^h,\, \mathbf{v}^h \neq \mathbf{0}} \frac{(q^h, \nabla \cdot \mathbf{v}^h)}{\|\mathbf{v}^h\|_1 \|q^h\|_0} = \kappa_h \geq \kappa_h^{\min} > 0$$

- The inf-sup constant κ_h has the matrix characterization

$$\kappa_h^2 = \min_{\vec{Z} \in \Re^M} \frac{\vec{Z}^\top \mathbb{B} \mathbb{A}^{-1} \mathbb{B}^\top \vec{Z}}{\vec{Z}^\top \mathbb{M}_p \vec{Z}} \qquad \text{where} \qquad (\mathbb{M}_p)_{ij} = (\chi_i^h, \chi_j^h)$$

- The LBB condition rules out the use of equal-order interpolation with respect to the same mesh for both $\mathcal{V}^h$ and $\mathcal{P}^h$

- Note that the semi-discrete mixed-Galerkin finite element equations can be written in the form

$$(\mathbf{u}^h_t, \mathbf{v}^h) + a(\mathbf{u}^h, \mathbf{v}^h) + b(\mathbf{v}^h, p^h) - b(\mathbf{u}^h, q^h) - (\mathbf{f}, \mathbf{v}^h) = 0$$

$$\forall \{\mathbf{v}^h, q^h\} \in \mathcal{V}^h \times \mathcal{P}^h, \quad \forall\, t \in (0, T]$$

with initial data determined from

$$\big(\mathbf{u}^h|_{t=0}, \mathbf{v}^h\big) = \big(\mathbf{u}_0, \mathbf{v}^h\big) \qquad \forall\, \mathbf{v}^h \in \mathcal{V}^h$$

CONSISTENTLY STABILIZED FINITE ELEMENT METHODS FOR THE STOKES EQUATIONS

- Let $\mathcal{T}_h$ denote a subdivision of the domain Ω into finite elements
  - then $\mathcal{K}$ denotes a finite element in $\mathcal{T}_h$
  - then $(\cdot,\cdot)_{\mathcal{K}}$ denotes the $L^2(\mathcal{K})$ inner product

- The discretized initial data is again determined from

$$\big(\mathbf{u}^h|_{t=0}, \mathbf{v}^h\big) = \big(\mathbf{u}_0, \mathbf{v}^h\big) \qquad \forall\, \mathbf{v}^h \in \mathcal{V}^h$$

- Then, consistently stabilized methods for the Stokes problem take the form

$$(\mathbf{u}_t^h, \mathbf{v}^h) + a(\mathbf{u}^h, \mathbf{v}^h) + b(\mathbf{v}^h, p^h) - b(\mathbf{u}^h, q^h) - (\mathbf{f}, \mathbf{v}^h) \quad \Leftarrow \text{mixed Galerkin terms}$$

$$- \sum_{\mathcal{K}\in\mathcal{T}_h} \tau \Big(\mathbf{u}_t^h - \triangle \mathbf{u}^h + \nabla p^h - \mathbf{f} \;,\; -\gamma \triangle \mathbf{v}^h + \nabla q^h \Big)_{\mathcal{K}} \quad \Leftarrow \text{stabilization terms}$$

$$= 0 \qquad \forall \, \{\mathbf{v}^h, q^h\} \in \mathcal{V}^h \times \mathcal{P}^h, \quad \forall \, t \in (0, T]$$

where

$$\tau \qquad \text{is a stabilization parameter}$$

and γ is used to define the three most popular stabilized methods

γ	$=$	0	pressure-Poisson stabilization
γ	$=$	1	Galerkin least-squares stabilization
γ	$=$	-1	Douglas-Wang stabilization

- Let's dissect the stabilization terms

$$-\underbrace{\sum_{\mathcal{K}\in\mathcal{T}_h}}_{\text{sum over elements}} \underbrace{\tau}_{\text{stabilization parameter}} \Big(\underbrace{\mathbf{u}_t^h - \triangle \mathbf{u}^h + \nabla p^h - \mathbf{f}}_{\text{residual of momentum equation}} \;,\; \underbrace{-\gamma \triangle \mathbf{v}^h + \nabla q^h}_{\text{test function terms}} \Big)_{\mathcal{K}}$$

- The crucial term that provides stabilization is the pressure-Poisson term

$$-\sum_{\mathcal{K}\in\mathcal{T}_h} \tau \Big(\mathbf{u}_t^h - \triangle \mathbf{u}^h + \nabla p^h - \mathbf{f} \;,\; -\gamma \triangle \mathbf{v}^h + \nabla q^h \Big)_{\mathcal{K}}$$

  – this accounts for the minus sign in front of the stabilization term

- The stabilization terms vanish for the exact solution of the Stokes equations

$$-\sum_{\mathcal{K}\in\mathcal{T}_h} \tau\Big(\mathbf{u}_t - \triangle\mathbf{u} + \nabla p - \mathbf{f}\ ,\ -\gamma\triangle\mathbf{v}^h + \nabla q^h\Big)_{\mathcal{K}} = 0$$

  – the exact solution of the Stokes equations also satisfies the mixed-Galerkin formulation

$$(\mathbf{u}_t, \mathbf{v}^h) + a(\mathbf{u}, \mathbf{v}^h) + b(\mathbf{v}^h, p) - b(\mathbf{u}, q^h) - (\mathbf{f}, \mathbf{v}^h) = 0$$

  – as a result, the exact solution of the Stokes equations satisfies the stabilized equations

  – for this reason, this class of methods is referred to as consistently stabilized methods

- The parameter γ in the test function terms serves to define different methods

  – pressure-Poisson stabilization $\gamma = 0$

$$-\sum_{\mathcal{K}\in\mathcal{T}_h} \tau\Big(\mathbf{u}_t^h - \triangle\mathbf{u}^h + \nabla p^h - \mathbf{f}\ ,\ +\nabla q^h\Big)_{\mathcal{K}}$$

  - provides the simplest means for effecting stabilization and consistency
  - in the stationary case, it is seemingly unconditionally stable i.e., it is stable for all values for the stability parameter

  – Galerkin least-squares stabilization $\gamma = 1$

$$-\sum_{\mathcal{K}\in\mathcal{T}_h} \tau\Big(\mathbf{u}_t^h - \triangle\mathbf{u}^h + \nabla p^h - \mathbf{f}\ ,\ -\triangle\mathbf{v}^h + \nabla q^h\Big)_{\mathcal{K}}$$

  - in the stationary case, it provides a symmetric stabilization term
  - in the stationary case, it is only conditionally stable i.e., it is stable for only a limited range of values for the stability parameter

– Douglas-Wang stabilization $\gamma = -1$

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \tau \Big(\mathbf{u}_t^h - \triangle \mathbf{u}^h + \nabla p^h - \mathbf{f} \ , \ -\triangle \mathbf{v}^h - \nabla q^h \Big)_{\mathcal{K}}$$

- in the stationary case, the test function terms involve the adjoint of the momentum equation operator

- in the stationary case, it is provably unconditionally stable, i.e., it is stable for all values for the stability parameter

- Why have different consistently stabilized finite element methods been introduced?

  - pressure-Poisson stabilization was the first consistently stabilized finite element method

  - Galerkin least-squares stabilization was introduced because of its symmetry property

  - Douglas-Wang stabilization was introduced because it can be rigorously proved to be unconditionally stable

    - no such proof exists for pressure-Poisson stabilization

    - however, computational evidence indicates that pressure-Poisson stabilization does lead to unconditional stability

    - the Galerkin least-squares method is only conditionally stable

- an unconditional stability proof does exists for an optimally convergent, modified (inconsistent) pressure-Poisson stabilization method

  P. Bochev and M.G.; An absolutely stable pressure-Poisson stabilized finite element method for the Stokes equations; *SIAM J. Numer. Anal.* **42** 2004, 1189-1207

- a detailed study of consistently stabilized finite element methods for the stationary Stokes equations can be found in

  T. Barth, P. Bochev, M. G., and J. Shadid; A Taxonomy of consistently stabilized finite element methods for the Stokes problem, *SIAM J. Sci. Comput.* **25** 2004, 1585–1607

• If one is going to use a consistent stabilization method, there is no compelling justification for using anything but pressure-Poisson stabilization methods

- thus, for the time being, we focus on that method

- we return to the Galerkin least-squares and Douglas-Wang methods later

• Since a main motivation for using stabilized finite element methods is that they allow for equal-order interpolation, we will use them in that manner

- In the paper

  P. Bochev, M.G., and J. Shadid; On stabilized finite element methods for transient problems with varying time scales, *Comp. Meth. Appl. Mech. Engrg.* **193** 2004, 1471-1489

  the stability parameter τ was chosen as

  $$\tau = \delta h^2$$

  for some constant δ and the sufficient condition for stability

  $$\Delta t > \delta h^2$$

  was demonstrated

  – computational evidence indicates that the optimal choice is $\delta = 0.05$; see

    T. Barth, P. Bochev, M. G., and J. Shadid; A Taxonomy of consistently stabilized finite element methods for the Stokes problem, *SIAM J. Sci. Comput.* **25** 2004, 1585–1607

  – we also choose $\tau = \delta h^2$

  – we will also discuss another choice for τ

SOME MOTIVATIONAL COMPUTATIONAL EXAMPLES

- A uniform grid of 200 triangles on the unit square is used $h = 0.1$

- The smooth exact solution is chosen as

$$\mathbf{u} = \begin{pmatrix} \dfrac{\partial \psi}{\partial x} \\ -\dfrac{\partial \psi}{\partial y} \end{pmatrix} \qquad \text{with} \qquad \psi(x, y) = x^2 (1-x)^2 \sin^2(\pi y)$$

$$p(x, y) = \sin(x)\cos(y) + (\cos(1) - 1)\sin(1)$$

- We first look at pressure plots determined using a mixed-Galerkin finite element method

  – of course, we cannot use equal-order interpolation

  – so we use the Taylor-Hood element pair

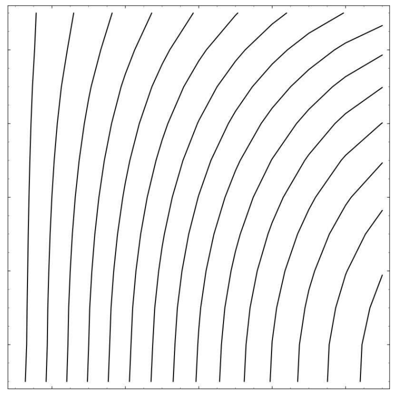
    - continuous piecewise quadratic velocity approximation

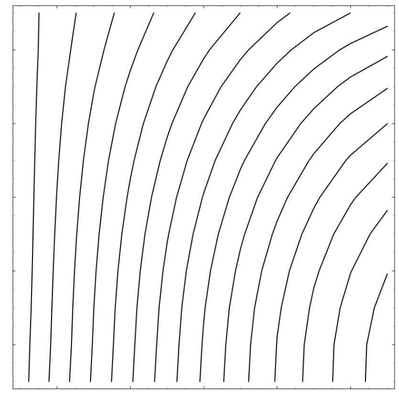
    - continuous piecewise linear pressure approximation

    with respect to the same grid

  – this element pair is known to define a stable mixed-Galerkin finite element method

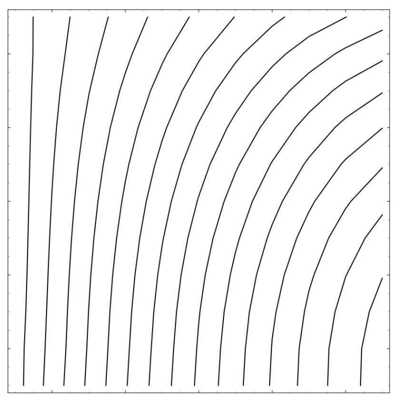
$\Delta t = 0.001$

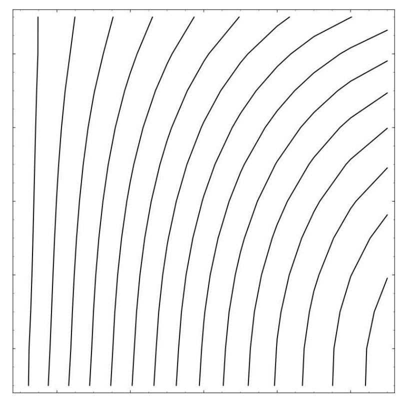




$\Delta t = 0.0001$

$\Delta t = 0.00001$





$\Delta t = 0.000001$

Approximate pressures determined using the Taylor-Hood element pair for different time steps

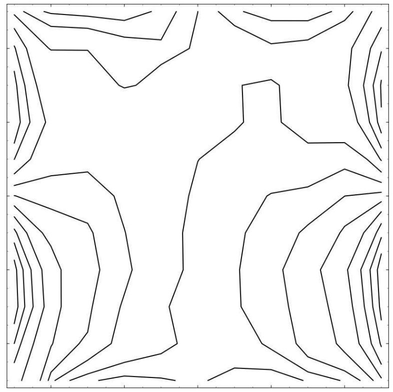
- This is an illustration of the fact that mixed Galerkin finite element methods do not encounter any difficulties, e.g., pressure instabilities, in the small time step limit

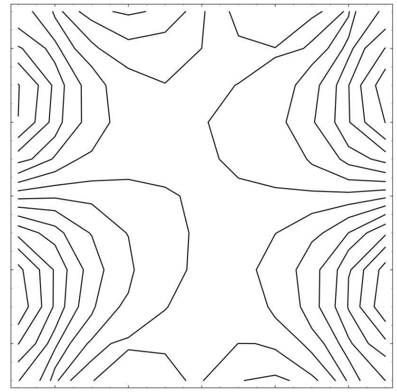
---

- We next look at pressure plots determined using a consistently stabilized finite element method
  - pressure-Poisson stabilization is used $\gamma = 0$
  - we still have $h = 0.1$ so that $\tau = \delta h^2 = \delta/100$
  - standard continuous piecewise
    linear,
    quadratic, and
    cubic

    nodal finite elements spaces are used for both the velocity and pressure

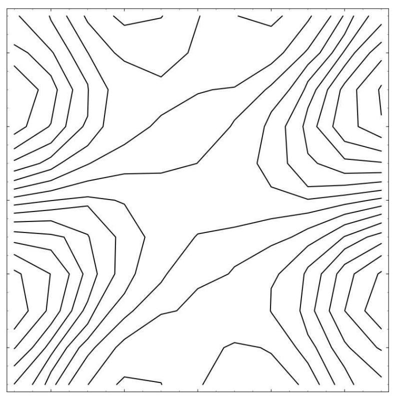
$\Delta t = 0.001$

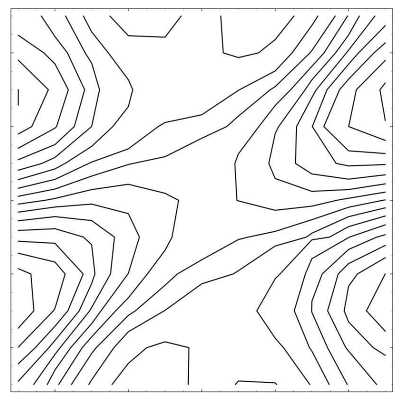




$\Delta t = 0.0001$

$\Delta t = 0.00001$

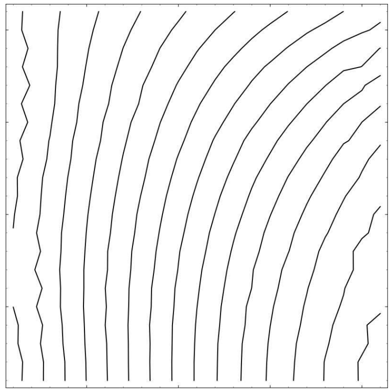


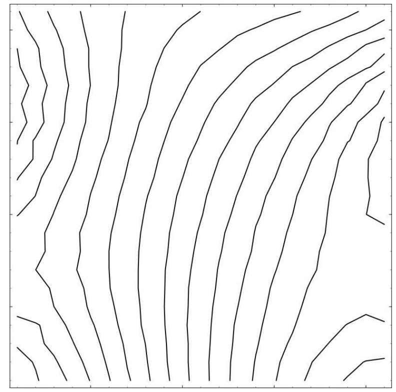


$\Delta t = 0.000001$

Approximate pressures determined using pressure-Poisson stabilization with $\delta = 0.05$ for different time steps and for equal-order piecewise linear interpolation

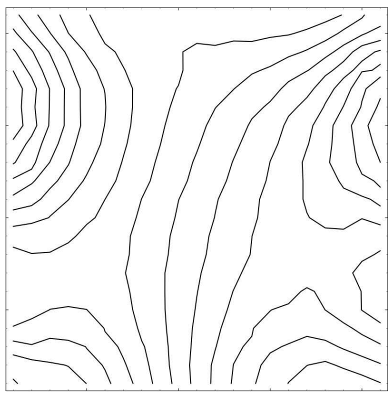
$\Delta t = 0.001$

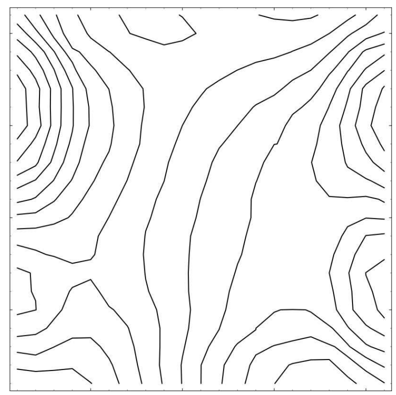




$\Delta t = 0.0001$

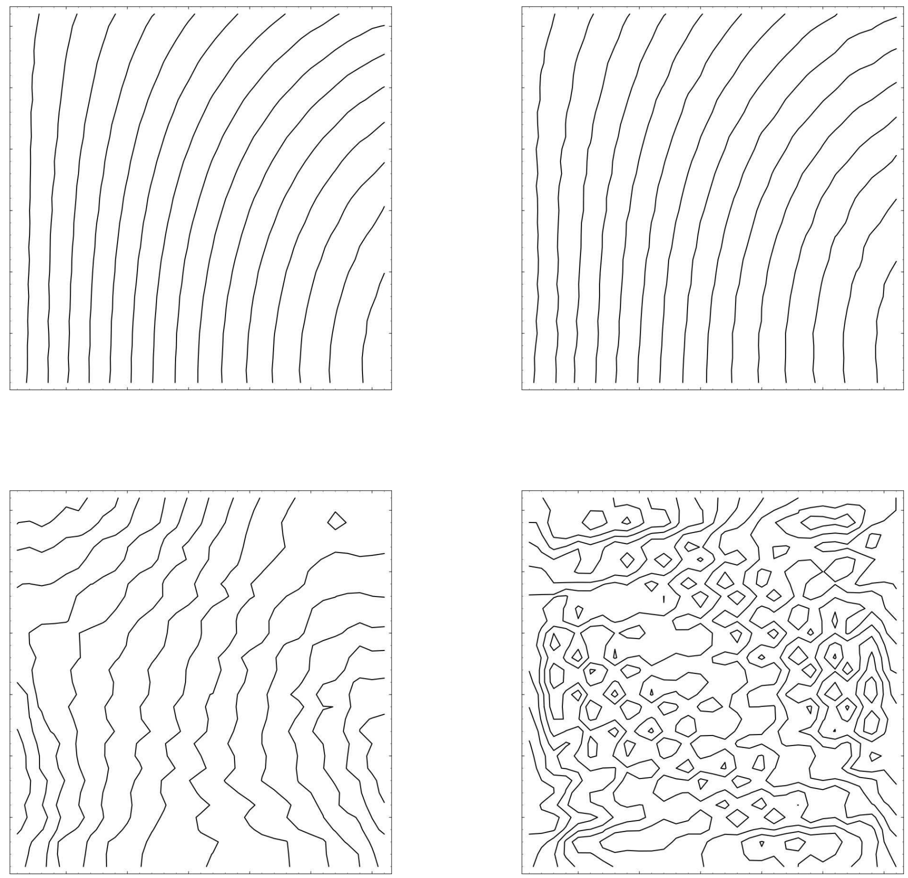
$\Delta t = 0.00001$





$\Delta t = 0.000001$

Approximate pressures determined using pressure-Poisson stabilization with $\delta = 0.05$ for different time steps and for equal-order piecewise quadratic interpolation



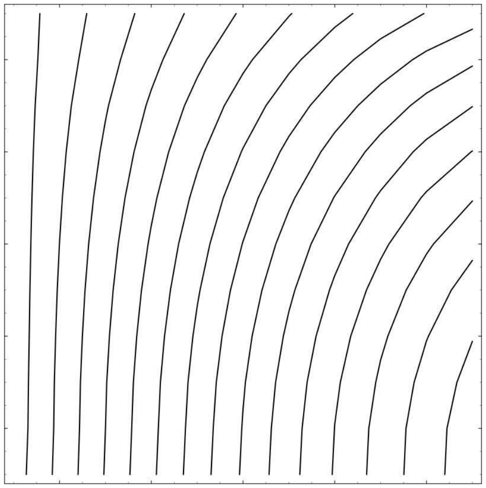
$\Delta t = 0.001$ $\Delta t = 0.0001$

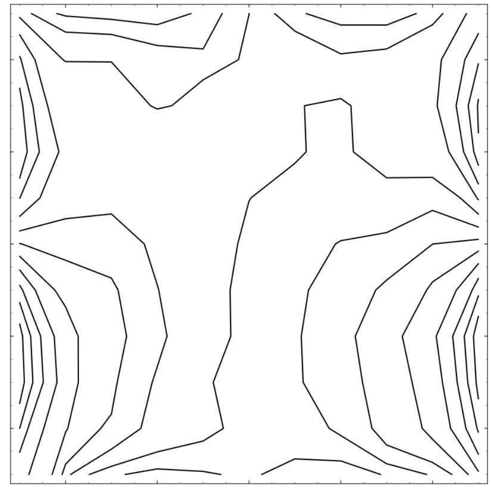
$\Delta t = 0.00001$ $\Delta t = 0.000001$

Approximate pressures determined using pressure-Poisson stabilization with $\delta = 0.05$ for different time steps and for equal-order piecewise cubic interpolation

- Why are the piecewise linear results wrong even for large time steps?

Taylor Hood





PP linear

$$\Delta t = 0.001 \qquad \delta = 0.05$$

– for piecewise linear elements, the stabilization term

$$-\sum_{\mathcal{K}\in\mathcal{T}_h} \tau\Big(\mathbf{u}_t^h - \triangle\mathbf{u}^h + \nabla p^h - \mathbf{f}\ ,\ \nabla q^h\Big)_{\mathcal{K}}$$

reduces to

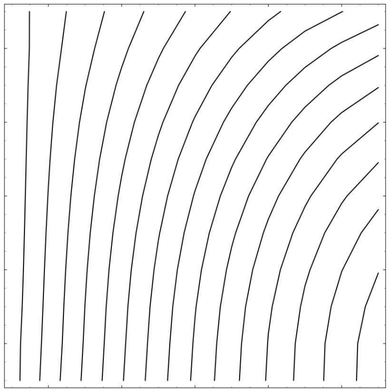
$$-\sum_{\mathcal{K}\in\mathcal{T}_h} \tau\Big(\mathbf{u}_t^h + \nabla p^h - \mathbf{f}\ ,\ \nabla q^h\Big)_{\mathcal{K}}$$

which, as it happens, induces an incorrect boundary condition on the pressure

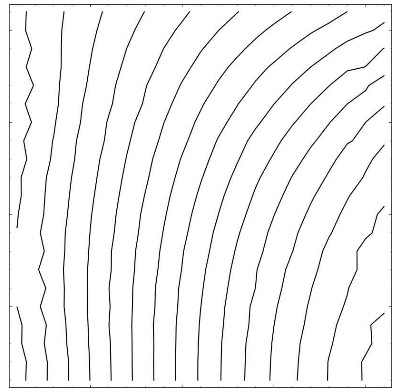
– but note the smooth apparent "convergence" (to the wrong answer) of the pressure approximation

- it would be difficult to detect that something is wrong with the pressure approximation

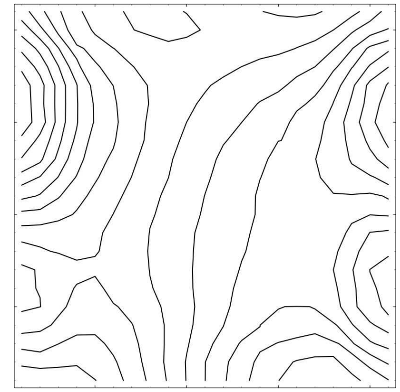
- What happens to the piecewise quadratic results as Δt decreases?



Taylor-Hood



PP $\Delta t = 0.001$



PP $\Delta t = 0.000001$

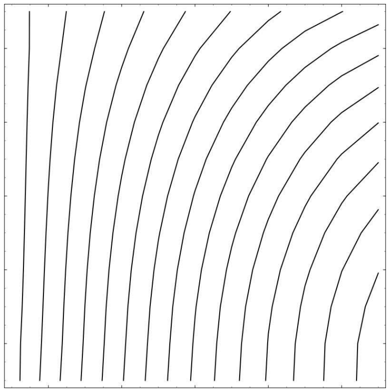
  – for "large" Δt, the results appear to be sort of correct

  – but as Δt gets small, the pressure approximation deviates from the correct solution

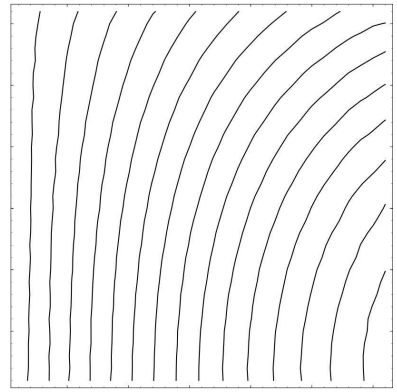
  – this transition seem to occur smoothly

    - again, it would be difficult to detect that something is wrong with the pressure approximation

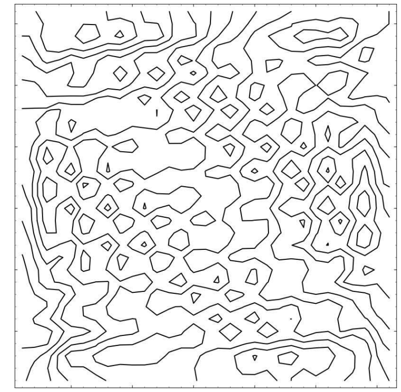
- What happens to the piecewise cubic results as Δt decreases is even stranger



Taylor-Hood



PP $\Delta t = 0.001$



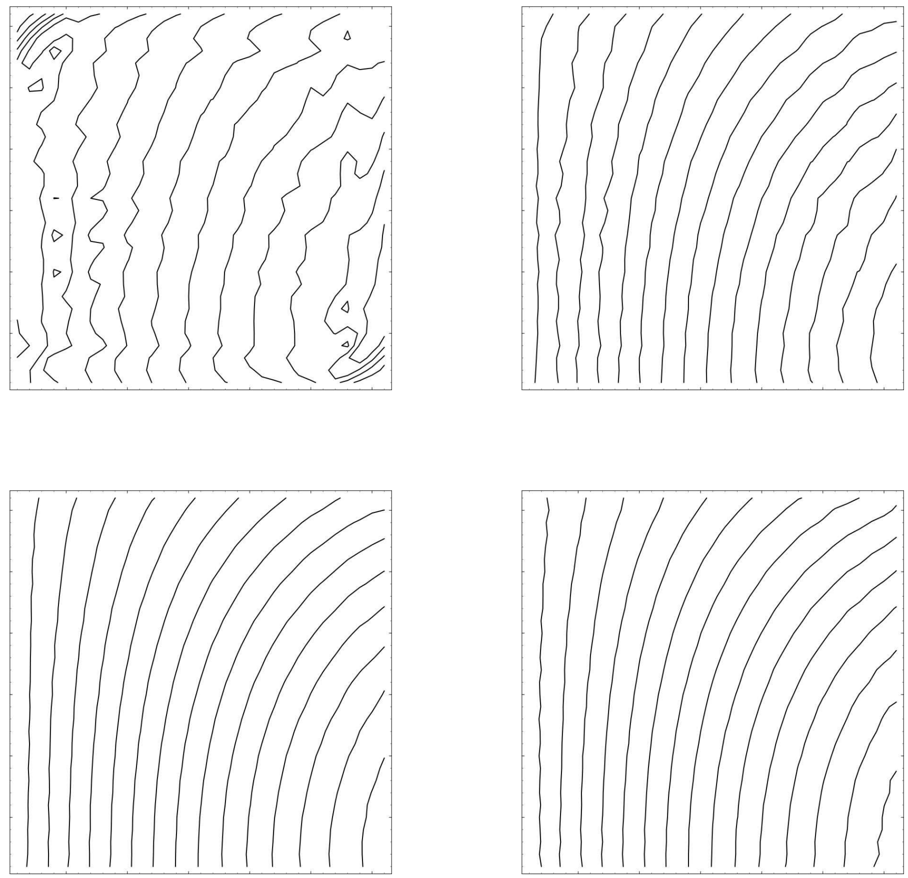
PP $\Delta t = 0.000001$

  - now as Δt gets small, spurious oscillations appear in the pressure approximation
  - the oscillations get worse as Δt decreases further

- Let's investigate the cubic case some more
  - we examine what happens as we vary δ

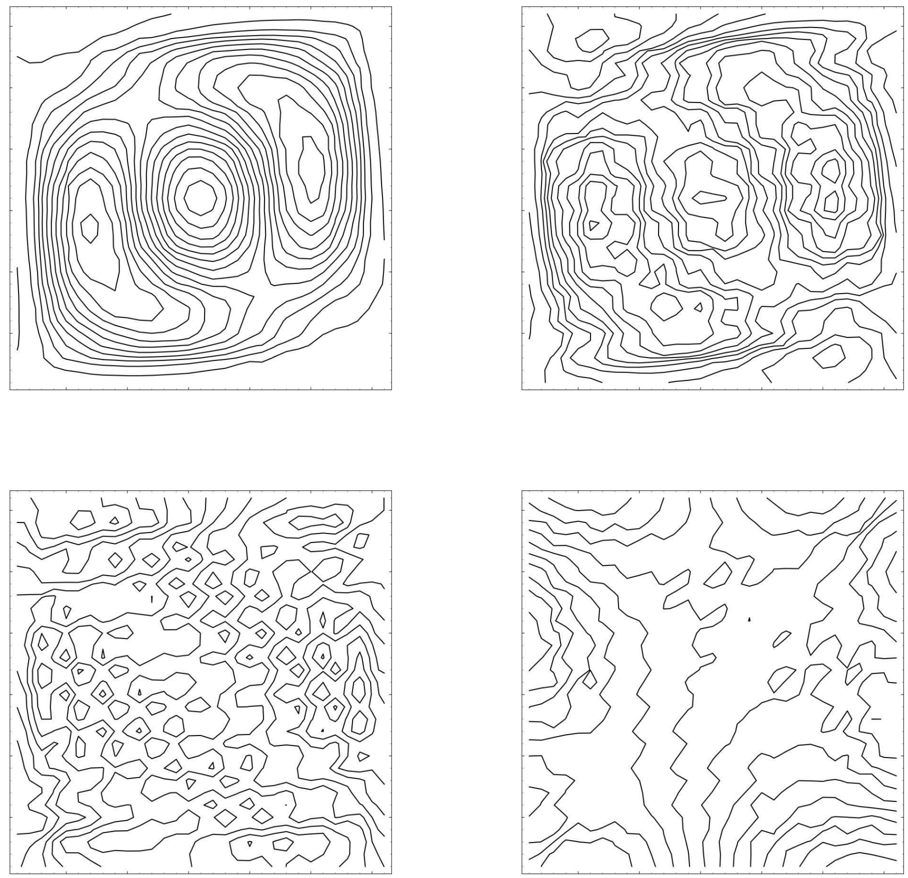
$\delta = 5$ $\delta = 0.5$

$\delta = 0.05$ $\delta = 0.005$



$\Delta t = 0.1 \quad \Leftarrow$ large time step

Good answers for a wide range of δ



$\delta = 5$ $\delta = 0.5$

$\delta = 0.05$ $\delta = 0.005$

$\Delta t = 0.000001 \quad \Leftarrow$ small time step

Garbage for all values of δ

ANALYSIS OF SEMI-DISCRETE EQUATIONS

- For the pressure-Poisson stabilized case, the semi-discrete equations are given by

$$\begin{pmatrix} \mathbb{M}\,\dot{\vec{U}} \\ \tau\mathbb{B}\,\dot{\vec{U}} \end{pmatrix} + \begin{pmatrix} \mathbb{A} & \mathbb{B}^\top \\ -\mathbb{B} - \tau\mathbb{S} & \tau\mathbb{K} \end{pmatrix} \begin{pmatrix} \vec{U} \\ \vec{P} \end{pmatrix} = \begin{pmatrix} \vec{F} \\ \tau\vec{G} \end{pmatrix}$$

where

$$\mathbb{K}_{km} = (\nabla\chi_m^h, \nabla\chi_k^h) \qquad \mathbb{S}_{kj} = (\Delta\boldsymbol{\xi}_j^h, \nabla\chi_k^h) \qquad (\vec{G})_k = (\mathbf{f}, \nabla\xi_k^h)$$

- Eliminating $\dot{\vec{U}}$, one obtains the semi-discrete pressure equation

$$\left(\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top\right)\vec{P} = \vec{G} - \mathbb{B}\mathbb{M}^{-1}\vec{F} + \left(\frac{1}{\tau}\mathbb{B} + \mathbb{S} + \mathbb{B}\mathbb{M}^{-1}\mathbb{A}\right)\vec{U}$$

- Clearly, the stability of the pressure approximation depends on the properties of the semi-discrete pressure matrix

$$\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^{\top}$$

  – this matrix is the difference of two discrete Laplacians

    - the pressure-Poisson matrix $\mathbb{K}$

    - the composed matrix $\mathbb{B}\mathbb{M}^{-1}\mathbb{B}^{\top}$ involving discrete divergence and gradient matrices $\mathbb{B}$ and $\mathbb{B}^{\top}$, respectively

- In the paper

  R. Codina and M. Vázquez, and O. Zienkiewicz; A general algorithm for compressible and incompressible flows. Part III: the semi-implicit form, *Int. J. Numer. Meth. Fluids.* **27** 1998, 13-32

  it was shown that the semi-discrete pressure matrix is symmetric and positive semi-definite but no estimates for the smallest eigenvalue were provided

- We have derived several refinements of this result; in particular, we have derived grid dependent bounds for the smallest eigenvalue

PROPOSITION

- Let

$$\mu_{max}^2 \equiv \max_{\vec{Q}\in\Re^M} \frac{\vec{Q}^\top \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top \vec{Q}}{\vec{Q}^\top \mathbb{K}\vec{Q}}$$

- Then

$$(1-\mu_{max}^2)\vec{Q}^\top \mathbb{K}\vec{Q} \le \vec{Q}^\top \left(\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top\right)\vec{Q} \qquad \forall\,\vec{Q}\in\Re^M$$

- Moreover, there exists a $\vec{Q}_{max}$ such that equality holds

---

- We also have obtained an upper bound, but that is not of interest to our discussion

- We next give a variational characterization of of μ_{max}

PROPOSITION

- Let

$$\|\vec{Q}\|_{\mathbb{K}}^2 = \vec{Q}^\top \mathbb{K} \vec{Q} \qquad \text{and} \qquad \|\vec{V}\|_{\mathbb{M}}^2 = \vec{V}^\top \mathbb{M} \vec{V}$$

- Then

$$\mu_{max} = \sup_{q^h \in \mathcal{P}^h,\, q^h \neq 0} \; \sup_{\mathbf{v}^h \in \mathcal{V}^h,\, \mathbf{v}^h \neq \mathbf{0}} \frac{(q^h, \nabla \cdot \mathbf{v}^h)}{|q^h|_{\color{red}1} \|\mathbf{v}^h\|_{\color{red}0}} = \max_{\vec{Q} \in \Re^M} \; \max_{\vec{V} \in \Re^N} \frac{\vec{Q}^\top \mathbb{B} \vec{V}}{\|\vec{V}\|_{\mathbb{M}} \|\vec{Q}\|_{\mathbb{K}}}$$

- Note that the norms are "reversed" with respect to the inf-sup stability condition associated with Stokes equations!

$$\inf_{q^h \in \mathcal{P}^h,\, q^h \neq 0} \; \sup_{\mathbf{v}^h \in \mathcal{V}^h,\, \mathbf{v}^h \neq \mathbf{0}} \frac{(q^h, \nabla \cdot \mathbf{v}^h)}{\|\mathbf{v}^h\|_{\color{red}1} \|q^h\|_{\color{red}0}} = \kappa_h > 0$$

- To answer (in the negative) the question

  does there exists a $\widehat{\mu}_{max} < 1$ such that $\mu_{max} \le \widehat{\mu}_{max}$ uniformly in $h > 0$?

  we specialize to the unit square and to equal-order bilinear, biquadratic, and bicubic interpolation

THEOREM

- Let h denote a characteristic mesh size
- Then,

$$1 - \omega h + O(h^2) \leq \mu_{max} \leq 1$$

where

$$\omega = \begin{cases} \dfrac{2}{3} & \text{for bilinear elements} \\ \dfrac{4}{15} & \text{for biquadratic elements} \\ \dfrac{16}{105} & \text{for bicubic elements} \end{cases}$$

- The results obtained lead to the first result we are looking for

COROLLARY

- There exists $\vec{Q} \in \Re^M$ such that

$$0 \leq \vec{Q}^\top \big(\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top\big)\vec{Q} \leq \big(2\omega h + O(h^2)\big)\, \vec{Q}^\top \mathbb{K} \vec{Q}$$

- Thus, the semi-discrete pressure matrix $\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top$ is not uniformly invertible

- Note that $\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top$ does not depend on the stabilization parameter τ
  - thus the non-uniform invertibility of semi-discrete pressure matrix is unconditional
    - for all values of the stabilization parameter τ, the semi-discrete pressure equation is ill posed

THE STABILIZING ROLE OF TIME DISCRETIZATION

- Curiously, time discretization has a stabilizing effect on the ill-posed semi-discrete equations

- The fully discrete equations obtained using the backward Euler method are given by

$$\begin{pmatrix} \mathbb{M} + \Delta t \mathbb{A} & \Delta t \mathbb{B}^\top \\ (\tau - \Delta t)\mathbb{B} - \tau \Delta t \mathbb{S} & \tau \Delta t \mathbb{K} \end{pmatrix} \begin{pmatrix} \vec{U}^{(n+1)} \\ \vec{P}^{(n+1)} \end{pmatrix} = \begin{pmatrix} \mathbb{M}\vec{U}^{(n)} \\ \tau \mathbb{B}\vec{U}^{(n)} \end{pmatrix} + \Delta t \begin{pmatrix} \vec{F}^{(n+1)} \\ \tau \vec{G}^{(n+1)} \end{pmatrix}$$

where

$$\vec{U}^{(0)} = \vec{U}_0$$

and

$$(\cdot)^{(n)} \quad \text{denotes the approximation at time level } t_n = n\Delta t$$

- Eliminating $\vec{U}^{(n+1)}$, one obtains

$$(\mathbb{K} + \mathbb{N}\mathbb{B}^\top)\vec{P}^{(n+1)} = \frac{1}{\Delta t}\Big(\mathbb{N}\mathbb{M} + \mathbb{B}\Big)\vec{U}^{(k)} + \mathbb{N}\vec{F}^{(n+1)} + \vec{G}^{(n+1)}$$

where

$$\mathbb{N} = -\mathbb{B}(\mathbb{M} + \Delta t\mathbb{A})^{-1} + \Delta t\left(\mathbb{S} + \frac{1}{\tau}\mathbb{B}\right)(\mathbb{M} + \Delta t\mathbb{A})^{-1}$$

THEOREM

- As $\Delta t \longrightarrow 0$, we have that

$$\mathbb{K} + \mathbb{N}\mathbb{B}^\top = \mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top + \frac{\Delta t}{\tau}\Big(\mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top + O(\Delta t)\Big) + O(\Delta t)$$

- If

$$\frac{\Delta t}{\tau} \geq \alpha > 0$$

for any $\alpha > 0$, then

$$\vec{Q}^\top\big(\mathbb{K} + \mathbb{N}\mathbb{B}^\top\big)\vec{Q} \geq \frac{1}{2}\min\{1, \alpha\}\|\vec{Q}\|^2_{\mathbb{K}}$$

- A consequence of this theorem, we reach our ultimate goal: for fixed τ, showing the necessity of a uniform lower bound on the time step

COROLLARY

- If

$$\frac{\Delta t}{\tau} \geq \alpha > 0$$

where the value of $\alpha > 0$ is independent of h and Δt, then the fully discrete pressure matrix $\mathbb{K} + \mathbb{N}\mathbb{B}^\top$ is uniformly positive definite

- On the other hand, if

$$\frac{\Delta t}{\tau} \rightarrow 0 \quad \text{as} \quad \Delta t \rightarrow 0$$

$$\mathbb{K} + \mathbb{N}\mathbb{B}^\top \longrightarrow \mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top$$

i.e., the fully-discrete pressure matrix reduces to the unstable semi-discrete pressure matrix $\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top$

- In other words, we see from

$$\mathbb{K} + \mathbb{N}\mathbb{B}^\top = \mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top + \frac{\Delta t}{\tau}\Big(\mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top + O(\Delta t)\Big) + O(\Delta t)$$

that implicit time discretization contributes the stabilizing term

$$\frac{\Delta t}{\tau}\mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top$$

  - if $\Delta t/\tau > \alpha$ with α fixed, this term is sufficient to overcome the destabilizing term

$$-\mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top$$

  - however, if $\Delta t/\tau \to 0$, the stabilization term disappears and we are left with the unstable semi-discrete pressure matrix $\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^\top$

CLEANING UP SOME LOSE ENDS

- What about other choices for the stabilization parameter τ?
  - we have chosen to use the spatial $\tau = \delta h^2$
  - another possibility is the transient choice

$$\tau = \frac{\Delta t}{2}\left(1 + \left(\frac{\Delta t}{\delta h^2}\right)^2\right)^{-1/2}$$

  - the good news: h fixed and $\Delta t \to 0 \quad \Rightarrow \quad \Delta t/\tau > 2$
  - the bad news: for h fixed and $\Delta t \to 0 \quad \Rightarrow \quad \tau \to 0$

    $\Rightarrow$ stabilization is lost!

– therefore, we see that the small time-step limit definition of τ is subject to two conflicting constraints

- to stabilize, τ must scale as $O(h^2)$

- to ensure that $\Delta t/\tau > \alpha$, it must scale as $O(\Delta t)$

– these constraints on τ are impossible to satisfy if Δt and h are allowed to vary independently

- in particular, in the small time-step limit,

  the spatial discretization step h must necessarily decrease as $\Delta t \to 0$

- What about the velocity approximation?

  - it can be shown that (using backward Euler time stepping)

    $$U^{(n+1)} = U^{(n)} + O(\Delta t)$$

    independently of the stabilization parameter

  - hence, there are no stability problems for velocity approximations

- What about Galerkin least-squares and Douglas-Wang stabilization?

  – for any $\gamma \in \{-1, 0, 1\}$, the semi-discrete equations are given by

$$\begin{pmatrix} (\mathbb{M} + \gamma\tau\mathbb{C})\,\dot{\vec{U}} \\ \tau\mathbb{B}\,\dot{\vec{U}} \end{pmatrix} + \begin{pmatrix} \mathbb{A} - \gamma\tau\mathbb{D} & \mathbb{B}^\top + \gamma\tau\mathbb{S}^\top \\ -\mathbb{B} - \tau\mathbb{S} & \tau\mathbb{K} \end{pmatrix} \begin{pmatrix} \vec{U} \\ P \end{pmatrix} = \begin{pmatrix} \vec{F} + \gamma\tau\vec{H} \\ \tau\vec{G} \end{pmatrix}$$

  where

$$\mathbb{C}_{ij} = \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\boldsymbol{\xi}_j^h, \Delta\boldsymbol{\xi}_i^h\right)_{\mathcal{K}} \qquad \mathbb{D}_{ij} = \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\Delta\boldsymbol{\xi}_j^h, \Delta\boldsymbol{\xi}_i^h\right)_{\mathcal{K}}$$

$$(\vec{H})_i = \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\mathbf{f}, \triangle\boldsymbol{\xi}_i^h\right)_{\mathcal{K}}$$

– the associated semi-discrete pressure matrix is given by

$$\mathbb{K} - (\mathbb{B} + \gamma\tau\mathbb{S})(\mathbb{M} + \gamma\tau\mathbb{C})^{-1}\mathbb{B}^\top$$

– the invertibility of the matrix $\mathbb{M} + \gamma\tau\mathbb{C}$ may place a further restriction on the stabilization parameter τ

– in any case, with $\tau = \delta h^2$, we have that

$$\mathbb{K} - (\mathbb{B} + \gamma\tau\mathbb{S})(\mathbb{M} + \gamma\tau\mathbb{C})^{-1}\mathbb{B}^\top \sim \mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}$$

i.e., for $\gamma = -1$ or $\gamma = 1$, the semi-discrete pressure matrix is spectrally equivalent to the corresponding matrix for $\gamma = 0$

- we have made this notion precise by providing the comparability constants

– thus, for $\gamma = -1$ and $\gamma = 1$, the semi-discrete pressure matrix is again not uniformly invertible

$\Rightarrow$ end up with the same conclusions as for $\gamma = 0$

- What does this all mean in practice?
  - we have $\tau = \delta h^2$ and we need to have

    $$\Delta t > \alpha h^2 \qquad \text{for fixed } \alpha$$

  - since
    - the time step may be governed by features not under our control, e.g., fast reactions,

    satisfaction of the lower bound requires that
    - the spatial grid size be sufficiently small, i.e.,

    $$h < \sqrt{\frac{\Delta t}{\alpha}}$$

  - in this case,
    - the spatial grid size is not determined by spatial accuracy considerations

    so that
    - the spatial grid size may end up being prohibitively small

- What about non-uniform grids?

  – for non-uniform grids, it may be desirable to choose the stabilization term in the form

$$-\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_{\mathcal{K}}\Big(\mathbf{u}_t^h - \triangle\mathbf{u}^h + \nabla p^h - \mathbf{f}\ ,\ -\gamma\triangle\mathbf{v}^h + \nabla q^h\Big)_{\mathcal{K}}$$

  where

$$\tau_{\mathcal{K}} = \delta h_{\mathcal{K}}^2$$

  and $h_{\mathcal{K}}$ is the diameter of the element $\mathcal{K}$

- What's the connection with pressure-projection methods?
  - interestingly, the operator $\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^T$ also arises in pressure-projection methods
  - in a pressure-projection method, this matrix effectively relaxes the discretized continuity equation to

$$-\mathbb{B}\vec{U} + \Delta t(\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^T)\vec{P} = 0$$

  - here, all that is needed for stabilization is that the matrix $\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^T$ be positive semi-definite
    - it does not have to be uniformly positive definite

– a pressure-projection method implemented with equal-order finite element spaces will also become unstable when Δt is small relative to h^2

  - the cause of this instability is fundamentally different from the one in stabilized methods

  - in a pressure-projection method, the instability arises from the insufficient amount of stabilization provided by $\Delta t(\mathbb{K} - \mathbb{B}\mathbb{M}^{-1}\mathbb{B}^T)$ for small Δt

THE MORALS OF THE STORY

- A lower bound on the time step is necessary if one uses a fully-discrete, consistent stabilization method
  - if one has to use small time steps, instabilities arise
  - however, if one does not have to use small time steps, consistently stabilized finite element methods are good methods to use

- If small time steps are needed, one should use mixed Galerkin finite element methods, if possible
  - in fact, mixed Galerkin finite element methods are always a good choice, unless they are ruled out by practical considerations

- If practical issues, e.g., complex, multiphysics problems, prevent the use of standard mixed Galerkin finite element methods, possible alternatives to consistently stabilized finite element methods include
  - coupled, space-time discretizations
  - inconsistently stabilized finite element methods
  - least-squares finite element methods

DETAILS

- Details about today's talk can be found in the paper

  P. Bochev, M. G., and R. Lehoucq; On stabilized finite element for the Stokes problem in the small time step limit, to appear in *Int. J. Numer. Meth. Fluids*