ICOSAHOM 2018 – Oral Presentation

To Do:

- Timing (should be ~25 minutes).
- Practice until the text comes naturally!

1 Title

Good morning. I would like to start off by thanking the organizers of this session for their generous invitation and especially for their having put together such an exciting group of speakers.

My name Philip Zwanenburg, working under the supervision of professor Nadarajah at McGill University and I will be presenting on the selection of finite element test norms for hyperbolic PDEs based on their stabilization and convergence properties.

2 Motivation and Goals

The success of standard Galerkin finite element methods for the solution of elliptic partial differential equations (PDEs) arises from the relation between the variational formulations and the minimization of an energy functional. This is well established and many stable and optimally convergent finite element methods in both the continuous and discontinuous Galerkin settings have been proposed.

In the case of convection-dominated or purely hyperbolic PDEs, additional stabilization is required in regions of high local Peclet or Reynolds number. While the need for this stabilization is understood, the specific form which it should take in order to guarantee both stable and optimally convergent discretizations is not generally provided. For the purposes of this talk, we will restrict the discussion to the introduction of the stabilization either though suitably chosen numerical fluxes or through a modification of the test space, resulting in a Petrov-Galerkin method.

Taking the linear advection equation as a model problem, our goal here is to present our current understanding and results of our initial findings regarding the determination of the optimal form of stabilization, which we define as that resulting in the computed solution being the L2 projection of the exact solution. We are still at the beginning of our investigation of how these concepts may be applied to our ultimate aim of providing advantages compared to existing methods for the compressible Navier-Stokes equations and we are consequently very open to any suggestions or discussion. It is our hope that this talk will leave you with several interesting ideas and especially that you may find use for them in your own research.

3 Outline

We will first outline the methods under consideration with detailed discussion of the related theory. We have specifically considered: the discontinuous Petrov-Galerkin method of Demkowicz et al. with various test norms, the optimal trial Petrov-Galerkin method recently formulated by Brunken et al. and the discontinuous Galerkin method with upwind numerical fluxes.

We will then present several numerical results comparing the performance of each of the methods for one and two dimensional problems and end with a discussion on the successes and limitations which we have encountered and our expected future work.

4 Abstract Functional Setting – Continuous
We begin with the introduction of a linear abstract functional setting. Consider the abstract variational problem given here where U and V denote the trial and test spaces, which are assumed to be Hilbert spaces, and where the bilinear form 'b' and the linear form 'l' correspond to a particular variational formulation.

We enforce the standard requirements that the bilinear form is bounded with continuity constant M, stable with inf-sup constant \gamma greater than zero and that the linear form is continuous and satisfies the compatibility condition given here, namely that we do not have non-zero loads for test functions resulting in a vanishing bilinear form. Then by the Banach-Necas-Babuska theorem the problem is well-posed.

5 Abstract Functional Setting – Discrete
Letting variables with 'h' subscripts denote finite dimensional (or discrete) quantities and choosing discrete trial and test spaces, we obtain the discrete variational problem shown here. If a discrete version of the inf-sup condition holds, then the discrete problem is well-posed and satisfies the following error estimate for which we offer a quick proof as it emphasizes the importance of the continuity and inf-sup conditions.

6 Quick Proof
Beginning with the discrete inf-sup condition and substituting the following choice (w_h minus u_h) for u_h into all relevant terms we obtain the first line of the inequality below. The linearity of the bilinear form then allows us to add and subtract the exact solution. As a result of Galerkin orthogonality, the second term in the numerator is then equal to zero. Using the continuity condition and cancelling the term in the test space from the numerator and the denominator, we obtain the last inequality.

7 Quick Proof
We can then use the triangle inequality and that just obtained to recover the error estimate, noting that it has been demonstrated that the additional 1 can be removed in the Hilbert space setting.

Several points are now of note. First, it can be seen that the computed solution has an error proportional to the infimum of that of all possible discrete solutions __in the trial space norm__ set according to the continuity and inf-sup conditions. If the trial norm is given by the L2 norm and the continuity and inf-sup constants are independent of the mesh size and solution regime, it is thus observed that the computed solution will converge optimally in the L2 norm. This will be a focus later in the presentation.

8 Selection of Norms

Second, a case of particular interest occurs when noting that the computed solution will be given by the __best__ approximation in the trial norm when the continuity and inf-sup constants are equal. It has been proven that the inf-sup constant being equal to one is implied by the continuity constant being equal to one and the existence of corresponding trial and test spaces and solution-dependent test functions, such that equality in the continuity condition is attainable. In what follows, it will be shown that the optimal trial Petrov-Galerkin method has the choice of norms for the two spaces dictated by this exact mechanism. We characterize this choice with the description of the norms being naturally induced by the problem.

It is thus clear that a trend towards using a stronger norm for the test space and a weaker norm for the trial space used in the discrete inf-sup condition and having the continuity and inf-sup constants take values of 1 leads to an optimal method.

9 Selection of Norms

Once a test norm is selected, we can define the map from trial to test space, T, using the test space inner product, such that the optimal test space, achieving the supremum in the discrete inf-sup constant, is given by the solution to the following linear system for each of the trial functions. The discrete version of this equation is used to compute test functions for discrete trial functions as shown at the bottom of the page. Note that the continuous test space is still used above and that the specific choice of discrete test space must be made according to the __practical__ realization of the method.

This presentation corresponds to the stabilization introduced through the variation of the test space. The alternative which we consider, based on the introduction of stabilization through a suitably chosen numerical flux, can also be interpreted in a similar manner. We consider this approach in the context of the discontinous Galerkin method where the trial and test space are equal, and it is very interesting to assess the well-posedness of the upwind method in terms of the corresponding trial and test norms required for boundedness and stability.

With the above abstract theory in place, we now move on to the description of the model problem and specific methods considered in this work.

10 Model Problem

The model problem considered here is the steady linear advection equation. We denote the solution by 'u', the advection velocity by 'b', the source term by 's', and outward pointing unit normal vectors by 'n'.

The weak formulation is obtained by partitioning the domain into volumes, denote by the caligraphic 'V', and integrating the above with respect to a test function 'v'. Derivatives are moved to the test functions by integrating by parts once. The single-valued normal numerical fluxes, denoted by f*, have been introduced to enforce the coupling between adjacent elements and will either be explicitly specified or become trace unknowns in the methods to be subsequently presented.

In preparation for the specification of the norms, we reformulate the above in terms of contributions over mesh faces as opposed to over volume boundaries with double square brackets denoting the standard jump operator.

11 The Discontinuous Petrov-Galerkin Method

The discontinuous Petrov-Galerkin method was originally proposed by Demkowicz et al. in 2010. From the previous discussion regarding the motivation of using the weakest possible norm for the trial space, we note that it is desirable to move derivatives in the bilinear form to the test space. From a reinterpretation of the weak form from the previous slide, we obtain the formal L2 adjoint along with additional trace terms and subsequently define the adjoint graph space as that where all terms in the operator form of the adjoint problem are bounded in L2. We say that the test space is H_b conforming when it is equal to the adjoint graph space.

12 The DPG Method – Broken Test Space

We note from the abstract analysis above that no assumptions were yet made on the conformity of the trial and test spaces. As the eventual goal of the methodology is to solve for the solution over a tessellation of the discretized domain, we note that using an H_b-conforming test space results in each of the optimal test functions potentially having global support.

In order to make the method practical, the discontinous Petrov-Galerkin method employs broken energy spaces such that the required inversion of the Riesz operator for the computation of optimal test functions can be done elementwise.

Discrete test spaces are thus most commonly chosen from piecewise polynomial spaces.

13 Investigated norms + Addtional

For the purposes of this study, we have investigated the following test norms, specifically the H1 along characteristics and graph norms. It is generally difficult to ascertain the specific form of the trial norm for each of the chosen test norms, however.

We conclude the introduction to the DPG method with a few of its noteworthy characteristics.

First, it has been shown that the test space for the conforming Petrov-Galerkin methodology outlined in the abstract setting is a subset of that of the practical DPG methodology. This implies, when both formulations are uniquely solvable, that their solutions coincide.

We also note, in comparison to the discontinous Galerkin method, that the required selection of a suitable numerical flux has been replaced with that of the selection of a test norm. However, we do note the test norm approach may offer additional flexibility.

Finally, regarding the practical implementation, when also choosing the trial space to be discontinuous, additional trace unknowns are introduced on internal faces allowing for static condensation of volume degrees of freedom in the global system solve resulting in a hybridized formulation.

14 The Optimal Trial Petrov-Galerkin Method

The original formulation of the optimal trial Petrov-Galerkin method considered here was performed by Brunken et al. quite recently. It is largely based on similar motivations to those of the DPG method but with a stricter constraint on the choice of test norm.

(SHOW EXCITEMENT!) In fact, this is the method for which the test norm is chosen as that which is naturally induced by the problem such that the trial norm corresponds to the L2 norm, and the continuity and inf-sup constants are both equal to one!

15 OPG Formulation

Proceeding, we apply the Cauchy-Schwarz inequality to the bilinear form and, after separating terms, obtain the naturally induced test and trial space norms. For the unity inf-sup constant, we require equality and thus define the test space according to the following equations so that this is achieved.

Note that the formulation is simplified in the case of using a continuous test space where the jump terms are no longer present.

16 OPG Formulation
Finally, imposing sufficient conditions on the test space, notably regarding the selection of the discrete basis and the imposition of adjoint boundary conditions (on the outflow portion of the domain), an equivalent problem is obtained from the original where now it is the test functions __having global support__ which are solved for and from which the solution is then computed through the application of the adjoint operator.

Upon first inspection, the method appears to be truly fantastic, with the computed solution being given by the L2 projection of the exact solution. However, we will highlight, when discussing the results, two major drawbacks which we have encountered and for which we have yet to find general remedies.

17 The DG Method
The discontinuous Galerkin method is obtain from the general problem statement by specifying the single-valued normal flux based on the upwind solution value. For convenience regarding the comparison with the test space stabilization, we present the alternative formulation based on the penalization of interface jumps.

The first interface term, which can be interpreted in the sense of using a central numerical flux when taken in isolation, is added to ensure the satisfaction of a discrete coercivity constraint in the norm shown here. While stability is achieved, additional terms in the norm required for continuity result in the h-convergence being sub-optimal by a full order. The second interface term is thus added to strengthen the stability of the bilinear form to arrive at an error estimate which is closer to being optimal; note that the addition of the combination of these two face terms is identical to using the upwind numerical flux.

18 DG – Coercivity norms
The following norms have been used to establish stability and boundedness for the upwind DG method. Leading to a sharp error estimate for L2 solution convergence in the first upwind norm at a rate which is sub-optimal by one half order with respect to mesh refinement.

For the DG formulation, coercivity is used in place of the inf-sup stability condition because the trial and test spaces are the same. What we would like you to take away from this slide is that as the stabilization is strengthened in the test space, so is the norm being used to measure the solution. With the previous discussion of the Petrov-Galerkin methods in mind, this can be interpreted as a counteracting of the trend towards using stronger norms in the test space and __weaker__ norms in the trial space to achieve optimal solution convergence in L2.

19 Results – Test Cases
In order to elaborate on the implementation details of the methods, we now present results for two test

cases. First we consider the simplest case of a manufactured solution for a one dimensional problem with the exact solution shown. Second, we consider results for a particularly intersting test case which was conceived by Peterson to demonstrate that the sub-optimality in the L2 convergence of the solution is sharp for the DG method using upwind numerical fluxes.

20 Results – Discrete Spaces

We would like to draw particular attention to the discrete spaces used for the study of the two Petrov-Galerkin type methods. We note in the cases of both the DPG and OPG implementations that the degree of the globally coupled variables was one higher than that of the volume solution and that this was __required__ for the optimal convergence results to be presented shortly (The normal flux trace for the DPG method and the C0 continuous volume test functions for the 1D OPG method). In two dimensions, the test space used for the OPG method had to be of the same degree as that of the solution in order to guarantee an empty NULL space.

21 Results – Manufactured (Conv. Rates)

The convergence rates are shown simply to highlight that all of the methods are converging optimally in this simple 1D case.

22 Results – Manufactured (Errors)

Of greater interest are the magnitudes of the errors for each of the methods. We note in particular that the OPG solution has errors __exactly__ equal to those of the L2 projection, confirming that the continuity and inf-sup constants are taking values of one. While convergence rates remain optimal, we note that the error magnitudes increased for the other methods as stronger norms were used for the trial space.

23 Results – Peterson (Conv. Rates)

We now show the convergence rates for the solution computed on the Peterson meshes. We first note, as expected, that the convergence of the DG method is suboptimal by half of an order and that convergence for the other methods appears to be optimal. Recalling the discussion of the degree of the globally coupled trace unknowns, we note that both the OPG and DPG methods were optimal, but __only__ with global trace having a polynomial degree one higher than the solution! Thus, despite the static condensation, reducing the dimension of the global problem by one, the problem for the globally coupled solve corresponds to a trace problem with degree one higher than the solution; a comparison with a hybridized DG method would be informative but unfortunately did not perform this test yet.

24 Results – Peterson (Errors)

While we no longer have the exact correspondence between the OPG and L2 projection errors, we note the same trend in the increasing errors from the OPG to DPG to DG methods, similar to the 1D results.

25 Results – Overconstrained OPG

Finally, we note the second major problem with the OPG method which was encounted. Motivated by the physics of the adjoint problem, Dirichlet adjoint boundary conditions for the test function (equal to zero) were specified along all outflow boundaries. In the case of a quadrilateral element, this results in the value of the solution forced to be zero on the outflow corner of the mesh shown here for a diagonal advection velocity. Thus, the piecewise polynomial discrete test space is overly restrictive even for such a simple problem and this poses a major challenge for the adoption of the method. Rather than address the problem using a test space more in line with the adjoint graph space, we plan to revisit the issue once a diffusion term is introduced and see if it persists.

26 Conclusion

The relationship between the trial and test norms and the solution convergence in L2 offers very interesting possibilities for the formulation of stable and optimally convergent methods.

However, despite the advantageous optimal convergence and static condensation properties, the Petrov-Galerkin methods considered here had several serious drawbacks in relation to required/standard discrete spaces.

Finding situations in which the improved stability justifies the additional cost is of primary importance if these methods are to be seen as practically valuable.

27 Future

Our future work will focus on the extension of the ideas presented here to nonlinear problems attempting to both:
1) formulate test norms which have better properties than those previously proposed;
2) draw motivation from numerical fluxes used for DG-type methods through a similar interpretation of stabilization in terms of penalization of face terms in the norm used to establish coercivity.

Of particular relevance based on the budding interest and results for nonlinear stability, we would also like to investigate if any relation can be found between nonlinear stability (perhaps in terms of kinetic energy or entropy) and optimally stable Petrov-Galerkin formulations.

Thank you very much for your attention!

OPG:
- Discuss limitation of convergence degree for OPG (sub-optimal). Discussion of non-trivial NULL space when choosing test to have higher degree. Discuss briefly the alternative formulation of Dahmen2012 where the problem is circumvented by not enforcing the that the system matrix results from the exact linearization of the residual, resulting in an iterative Uzawa algorithm to iteratively converge (not iterative global solves for a linear PDE and potentially slow convergence!).

- Discussion of optimal test functions being discontinuous with multiple inflow faces => Motivation for Brunken to compute solution based on test space which is local polynomials.