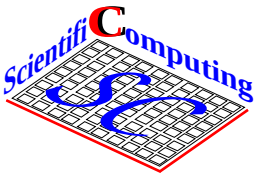
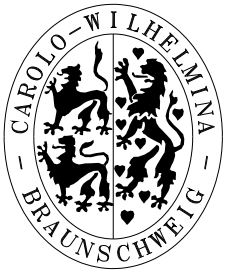


A Review of Petrov-Galerkin Stabilization Approaches and an Extension to Meshfree Methods

Thomas-Peter Fries, Hermann G. Matthies
Institute of Scientific Computing
Technical University Braunschweig
Brunswick, Germany

Informatikbericht Nr.: 2004-01

March 30, 2004



A Review of Petrov-Galerkin Stabilization Approaches and an Extension to Meshfree Methods

Thomas-Peter Fries, Hermann G. Matthies
Department of Computer Science
Technical University Braunschweig
Brunswick, Germany

Informatikbericht Nr.: 2004-01

March 30, 2004

Location
Institute of Scientific Computing
Technische Universität Braunschweig
Hans-Sommer-Strasse 65
D-38106 Braunschweig

Postal Address
Institut für Wissenschaftliches Rechnen
Technische Universtität Braunschweig
D-38092 Braunschweig
Germany

Contact

Phone:	+49-(0)531-391-3000
Fax:	+49-(0)531-391-3003
E–Mail:	wire@tu-bs.de

A Review of Petrov-Galerkin Stabilization Approaches and an Extension to Meshfree Methods

Thomas-Peter Fries, Hermann G. MatthiesMarch 30, 2004

Abstract

This paper gives a detailed review of popular stabilization approaches that have developed to be standard tools in the numerical world. The need for stabilization is outlined and stabilization ideas based on the Petrov-Galerkin concept are discussed. Stabilization methods are explained on the one hand in an illustrative approach with help of the artificial diffusion idea and on the other hand in a theoretical approach by outlining the mathematical background of stabilization.

A generalization to meshfree methods is investigated. We find that the structure of stabilizing perturbation terms can be applied in the same manner to meshfree methods. However, the weighting of the stabilizing terms, defined with the stabilization parameter, requires special attention. Using standard formulas for the stabilization parameter, raised in the meshbased finite element context, is only suitable for meshfree shape functions with small dilatation parameters. This is confirmed with numerical experiments for the advection-diffusion equation and incompressible Navier-Stokes equations.

Contents

1 Introduction

Numerical methods are indispensable for the successful simulation of physical problems due to approximation of the underlying partial differential equations. A huge number of methods have raised to accomplish this task. Most of these methods introduce a finite number of unknowns and can be based on the principles of weighted residual methods.

Let us separate these methods according to the aspect, whether they need a mesh for the approximation or not. Among the class of meshbased methods are famous members such as the finite volume and finite element methods. They enable fast and reliable approximation whenever the mesh can be maintained successfully during the calculation. We may label these methods "standard tools" in the numerical world. On the other hand, meshfree methods enable the solution of partial differential equations only based on a set of points without the need of an additional mesh. The price to pay for this is the relatively high computational burden associated with the usage of these methods. A large number of these comparatively new methods have been developed during the last three decades, among them the popular Element Free Galerkin method, Smoothed Particle Hydrodynamics etc. [3, 13] Rather than calling these methods an alternative for meshbased methods, we may label them "specific tools", because they are frequently used in problems, where meshes are impossible or only hardly to maintain. In approximations, where the mesh does not cause problems, standard meshbased methods are often preferable as they are, in general, considerably faster. We believe that an important step for the usage and acceptance of meshfree methods can be made, if meshbased and meshfree methods were easily couplable. Then, it would be possible, to use meshfree methods only in small regions of the domain, where a mesh is difficult to maintain and meshbased standard methods in all other parts.

Rather than separating numerical methods, one may differ formulations of the underlying differential equations. Here, most importantly Lagrangian and Eulerian viewpoints have to be mentioned which choose distinct coordinate systems for the description of the problem. The most important difference in the formulations is in the presence of an advection term in the Eulerian formulation, which is absent in the Lagrangian description. Advection terms are non-selfadjoint operators that often lead to problems in their numerical treatment [7]. This is particularly the case for Bubnov-Galerkin methods, where the test functions are chosen equal to the shape functions. Then, spurious oscillations may pollute the overall solution and stabilization is required.

There are also other problems where oscillations and other "unexpected" phenomena (locking, singular matrices etc.) occur such as with the so-called mixed problems [12]. Applying the same shape functions to all variables of the problems in a Bubnov-Galerkin setting (equal-order interpolation), which is from the computational viewpoint the most convenient way, leads to severe problems as a result from violating certain conditions. Then, again stabilization is required [12].

The need for stabilization is well studied in the meshbased context. A number of stabilization methods have been developed to overcome numerical problems. This also stems from the fact that for meshbased methods the Eulerian viewpoint is standard, because it seems impossible to maintain a conforming mesh for example in flow problems with the Lagrangian viewpoint. Then, stabilization is a crucial ingredient to obtain suitable solutions. Meshfree methods, however, are in general used for problems in Lagrangian formulation [13]. Then, there is no stabilization of the advection term necessary and the stabilization of mixed problems may be bypassed with other

ideas.

However, we believe that *Eulerian meshfree methods* are not only of interest in their own right, but can have a significant influence in the usage of meshfree methods in practice. This is mostly due to the aspect that it seems straightforward to couple meshbased and meshfree methods formulated with the Eulerian viewpoint, instead of combining Eulerian meshbased methods and Lagrangian meshfree methods. To successfully work with Eulerian meshfree methods stabilization is required. This is the most important aspect of this paper.

The paper is organized as follows: Section 2 is a problem statement. Different sources for oscillations and related numerical problems are considered, most notably convection-dominated problems and mixed problems. The historical development of stabilization ideas is briefly described, starting from introduction of artificial diffusion and ending with the established, most frequently used stabilization schemes, which are well-funded in theoretical analysis.

Section 3 gives a review of these popular stabilization schemes. The Streamline-Upwind/Petrov-Galerkin (SUPG), Pressure-stabilizing/Petrov-Galerkin (PSPG) and Galerkin/Least-Squares (GLS) stabilization are considered. Further on, the construction of discontinuity-capturing operators is discussed. The main idea of all these methods is to add products of suitable perturbation terms and the residuals, thereby maintaining consistency. The stabilization parameter τ weights these added stabilization terms. The different structures of each method are described briefly. This structure, i.e. the way these schemes modify the weak form of a problem, is independent on whether meshbased or meshfree shape functions are taken for the approximation. However, the stabilization parameter τ depends on the problem under consideration and the chosen numerical method.

In section 4 an illustrative approach to stabilization is described, considering the linear finite element method. Starting from introducing artificial diffusion, the SUPG and GLS stabilization are deduced, which are identical in case of linear interpolations. Special attention is given to the deduction of the stabilization parameter τ. With help of the one-dimensional advection-diffusion equation, the well-known coth-formula for τ is derived. Most of the stabilization schemes used nowadays still rely on alternative, similar versions of this formula to determine the stabilization parameter. We find that this formula is particularly useful also for the stabilization of other problems and show ways how to determine more suited stabilization parameters dependent on the element shape functions. Thereby, nodally exact solutions for the model problem not only for linear elements are obtained, but also for quadratic elements and it is shown how this can be generalized to other higher-order elements.

The theoretical background in the mathematical analysis is outlined in section 5. Standard techniques to prove important features, such as stability and convergence, of the stabilized finite element methods are briefly explained. Rather than copying these analyses, the interested reader is referred to the references given in this section. It can be shown that stabilized methods lead to higher-order but suboptimal convergence. Furthermore, design criteria for τ from mathematical analysis are compared with the results from the previous section and other proposals for the determination of τ. It is pointed out that the optimal choice of this parameter is still an open question.

In section 6 we turn to meshfree methods. In the previous sections it was our intention to develop an understanding of stabilization in the meshbased context. It is tried to apply the same

steps here for meshfree methods. We find that standard stabilization methods can be applied, however, the choice of the stabilization parameter τ requires special attention. The parameter τ, developed to obtain the nodally exact solution for meshbased and meshfree methods, has a very different character: obtained for the finite element methods it has a local character, i.e. τ can be determined with knowledge of the relative position of some neighboring nodes. In contrast, τ for meshfree methods depends on the position of all particles inside the domain. Despite this fact, it is shown that small dilatation parameters, i.e. small supports, of the meshfree shape functions justify the usage of standard formulas for τ —although derived in a meshbased context!— and show numerical results for the incompressible Navier-Stokes equations. These results agree well with the assumption that standard formulas for τ may be used for small dilatation parameters.

2 The Need for Stabilization

Using numerical methods in a straightforward way for the approximation of arbitrary differential equations may cause severe problems. Oscillations, locking, singular matrices and other problems may be the result of disregarding important basic rules related with a certain concrete problem. Then, stabilization is needed. In this section, it is described under which circumstances problems occur and stabilization may be needed to obtain satisfactory approximations.

2.1 Convection-dominated Problems

The phenomenon of *convection*, typically identified by first order terms in the differential equations of a model, divides the usability of numerical methods. Methods being successfully applied in structure problems, where no convection is present, may totally fail when they are applied to convection-dominated problems, as they occur frequently e.g. in fluid mechanics. This is particularly the case with Bubnov-Galerkin methods, which are weighted residual methods, where the test functions are set equal to the shape functions [7]. In practice, these test and shape functions are often provided with the FEM methodology, i.e. in a meshbased way. But the situation is the same with any other test and shape functions, such as those constructed with MMs. Figure 1 shows an oscillatory example for a meshbased and a meshfree approximation of a convection-dominated problem.

In structural analysis, where often the minimization of energy principles is the underlying idea, the application of Bubnov-Galerkin methods leads to symmetric matrices and "optimal" results. With "optimal" we refer to the fact that the solution often possesses the "best approximation" property, meaning that the difference between the approximate and the exact solution is minimized with respect to a certain norm [7].

The situation, however, is totally different in the presence of convective terms. Then, the matrix associated with the advective term is non-symmetric (non-self adjointness of the convective operator) and the "best approximation" property is lost [7]. As a result Bubnov-Galerkin methods applied to these problems are far from "optimal" and show spurious oscillations in the solutions, worsening with growing convection-domination. This does not only lead to qualitatively bad results but even violates basic physical principles like entropy [22] or the positive boundedness of concentrations etc. One finds that the pollution of the solution with oscillations is dependent on

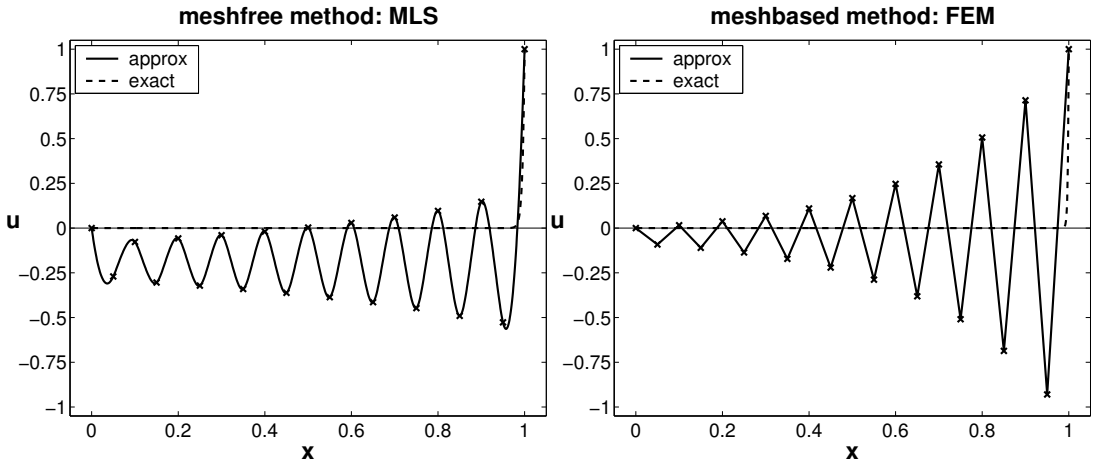


Figure 1: Typical oscillations in the approximation of an advection-dominated problem in one dimension. Standard Bubnov-Galerkin methods are applied, based on meshfree and meshbased shape functions.

the domination of the convection terms over other terms of the differential equation, in practice most often diffusion terms. The role of convection in differential equations is defined by well-known identification numbers such as the Peclet number and Reynolds number. The higher these numbers are, the more dominant is the convection term and the stronger is the pollution with oscillations.

The same situation can be found in the finite difference context. There, the same problem with oscillations occurs when using central differences for the advective operator. It can be easily shown that Bubnov-Galerkin treatment of the weak form and central differences applied to the strong form are closely related. The corresponding matrix line of node I of a one-dimensional advective operator $c\frac{\partial u}{\partial x}$ becomes for linear FEM and FDM in case of a regular node distribution:

$$\text{FEM}: \quad \frac{c}{2}\begin{pmatrix} & \ddots & & & \\ \dots & -1 & 0 & +1 & \dots \\ & & & \ddots & \end{pmatrix} \qquad \text{FDM}: \quad \frac{c}{2\Delta x}\begin{pmatrix} & \ddots & & & \\ \dots & -1 & 0 & +1 & \dots \\ & & & \ddots & \end{pmatrix}.$$

The only difference is in the constant term Δx, which cancels out for the FEM due to the integration of the weak form over the domain. It is thus not surprising that Bubnov-Galerkin FEM and central FDM show the same characteristics.

In the FDM context it is well-known that upwind differencing on the convective term does not show oscillatory solutions, but introduces over-diffusive results [7]. A simple Taylor series analysis proves that upwinding is only first order accurate, in contrast to the second order accurate — but oscillatory— central differences. This analysis also elucidates that upwinding can also be interpreted as central differences plus artificial diffusion. Thus, the "right" combination of central and upwind differences may introduce the optimal amount of artificial diffusion which leads to accurate and oscillation-free solutions [7].

Starting in the seventies a large number of FEMs raised with different ideas to include the up-

wind effect in finite elements. Often, the one-dimensional advection-diffusion equation served as a relevant model equation and generalization to other problems in multi-dimension was straightforward —and unsuccessful! The proposed methods all obtained nodally exact solutions for the 1D model equation, such that the resulting difference stencil of the FEM matches exactly the known nodally exact stencil from the FDM. This was realized in the "anisotropic balancing dissipation" approach by adding artificial diffusion in streamline direction and using standard Bubnov-Galerkin to discretize the modified problem [38]. Thereby, the consistency of the method is given up, i.e. the exact solution does not longer fulfill the modified weak form. Source terms require special treatment in this approach. Other approaches used a Petrov-Galerkin FEM, where the test functions are modified such that they weight the upwind node more than the downstream node, see e.g. [8, 20]. In [24], the advection term is integrated with only one integration point, which is placed inside the element in dependence of the convection-diffusion ratio, whereas all other terms are integrated in the standard way. All these approaches obtain the optimal difference stencils in the resulting system of equations leading to the nodally exact solution for the one-dimensional model problem. However, successful generalization to arbitrary, time-dependent problems and multi-dimensions failed —i.e. the results were either oscillatory or over-diffusive due to the crosswind diffusion effect— and a successful method was still outstanding.

The Streamline-Upwind/Petrov-Galerkin (SUPG) method, introduced from Brooks and Hughes in [7] (and [26]) can be considered as the first successful stabilization technique to prevent oscillations in convection-dominated problems in the FEM. The main steps are: introduce artificial diffusion in streamline direction only, interpret this as a modification of the test function of the advection terms and finally, enforce consistency, such that this modified test function is applied to *all* terms of the weak form. Then, the term "artificial diffusion" is not fully applicable any longer, because the stabilized weak form can not, in general, be manipulated such that only a diffusion term is extracted. The exact solution of the problem still satisfies the SUPG stabilized weak form.

In the following, SUPG was extended to coupled multi-dimensional advection-diffusion systems, where each equation has to be stabilized. The Euler and Navier-Stokes equations also fall into this class, the first being the governing equations of inviscid compressible flow, the latter of viscous compressible flow. (Incompressible flows can be handled very successfully without stabilizing each equation individually [7]). A first attempt to do this has been realized by Hughes and Tezduyar in [35]. However, it is later pointed out in [32] that this approach is not adequate, because it fails to appropriately treat the modal components of the system. A stabilization is needed, where a distinct stabilization for each component can be realized. Using the same stabilization for all components may lead to over-diffusive results for some components and oscillatory results for other.

With the purpose of enabling mathematical analysis and systematically generalizing SUPG such that each componential equation is treated in an optimal manner, Hughes *et al.* symmetrize the compressible Navier-Stokes equations in terms of entropy variables in [30], see also [19], instead of the conservation variables. Numerical solutions based upon this form fulfill entropy conditions, manifested in the Clausius-Duhem inequality or equivalently in the second law of thermodynamics, from the beginning, which is advantageous in mathematical stability proofs. Convergence of the generalized SUPG has been proven in [31]. As shown in [32], the symmetrized system can be transformed to an uncoupled system of scalar variational equations, for which an adequate

stabilization is known.

Later, Aliabadi, Ray and Tezduyar compare in [1] results for the different formulations of variables —conservative vs. entropy variables— without finding remarkable difference between the two.

The major part of the theoretical analysis of the SUPG has be done from Johnson, see [36, 46] and references therein. There, SUPG is often labeled with the term "streamline diffusion method".

Motivated from mathematical analysis, another type of stabilization scheme has been established, the Galerkin/Least-Squares (GLS) method. It is similar to the SUPG in certain aspects, and for purely hyperbolic equations and/or linear interpolation functions, the two become identical. In the GLS method, least-squares forms of the residuals are added to the Galerkin method, enhancing stability of the Bubnov-Galerkin method without giving up consistency or degrading accuracy [29]. There is no motivation from artificial diffusion as was the starting point for SUPG. The GLS method was introduced under this name as a method on its own in [29] by Hughes, Franca and Hulbert (although it has already been applied to other problems before, such as in [27]). They apply the GLS method for stationary and instationary advective-diffusive systems. In [50], Shakib uses the GLS for the solution of the compressible Euler and Navier-Stokes equations in the above mentioned symmetrized form.

Today, the SUPG and GLS stabilizations are most frequently used. Both stabilization methods add products of perturbations and residuals to the weak form, weighted with a so-called stabilization parameter τ. The derivation of τ, leading to reliable approximations with both meshbased and meshfree methods is the most important aspect of this paper.

2.2 The Babuška-Brezzi Condition

Variational formulations associated with constraints lead to severe problems if standard numerical methods are used in a straightforward manner. One way to treat these problems is the usage of admissible functions satisfying the constraint *ab initio* [12]. The solution is then a member of a smaller space of functions than the space required from continuity conditions alone and suitable interpolations are not easy to find. Instead, the problem can be reformulated by introducing a second variable, the Lagrange multiplier [12]. The resulting variational formulation falls into an abstract class of "mixed" formulations. Lagrange multipliers and mixed formulations are thus intimately related. One of the most well-known examples of a mixed problem is Stokes flow in which the velocity-strain energy is minimized subject to the incompressibility constraint.

Figure 2 shows an example for Stokes flow with large oscillations in the pressure field as a consequence of violating the Babuška-Brezzi condition.

The approximation of mixed formulations requires careful choice of the combination of interpolation functions. In particular, equal-order interpolations, where the same ansatz is made for the primary and secondary (Lagrange multiplier) variables are *not* adequate in a Bubnov-Galerkin setting, although from an implementational viewpoint they are most desirable. Also, many other practically convenient interpolations fail to give satisfactory results, especially in three dimensions. The governing stability conditions for mixed problems are *K-ellipticity* and the *Babuška-Brezzi condition* [2, 5]. Violating them leads to pathologies such as spurious oscillations and locking [12], or the resulting system of equations may be singular not giving a solution at all.

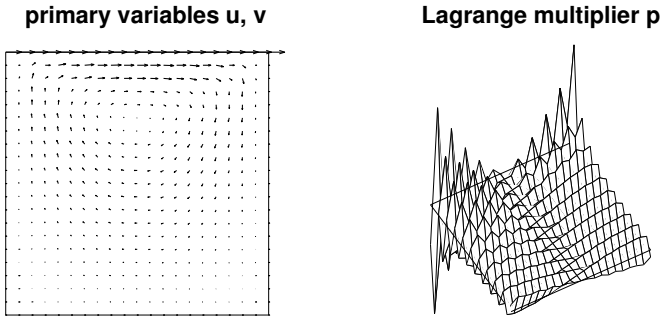


Figure 2: The solution for Stokes flow with P1/P1 FEM and a wrong stabilization parameter $\tau \approx 0$. Although the primary variables (velocity field) are reasonably approximated, the Lagrange multiplier (pressure field) shows large oscillations.

In [12] the authors claim that it depends on the concrete problem, which of the two criteria is more difficult to obtain. Lack of stability may come from the Lagrange multiplier or from the primary variable. For problems, in which K-ellipticity is difficult to satisfy —e.g. for linear isotropic incompressible elasticity emanating from the Hellinger-Reissner principle— , the problem comes from the primal variable and it is often easy to find interpolations *satisfying* the Babuška-Brezzi condition.

In contrast, for problems that fulfill the ellipticity requirement immediately —like Stokes flow— , stability problems arise from the Lagrange multiplier and it is difficult to fulfill the Babuška-Brezzi condition. Only very few combinations of interpolations are adequate. In this case, it is desirable to find ways to *circumvent* the condition. Motivated from theory this can be done by modifying the bilinear form such that it is coercive on the primal variable as well as the Lagrange multiplier. Then, there is no need to fulfill the Babuška-Brezzi condition for this method. This can be interpreted as some kind of stabilization which is realized by adding appropriate perturbation terms, without upsetting consistency. It will be shown later that this is realized —with the same fundamental idea as in other stabilizations— by a multiplication of perturbations with residual forms of the governing problem. In [12, 27, 28] such stabilizations with the aim to circumvent the Babuška-Brezzi conditions have been presented for Stokes flow. The formulations become stable for any combination of interpolations, but the methods have different requirements on the velocity/pressure spaces: some need continuous spaces, other become convergent for arbitrary spaces, hence they allow also discontinuous approximations.

For the mixed problems discussed throughout this paper, which are Stokes equations and the incompressible Navier-Stokes equations, we summarize that a stabilization has to be found such that the Babuška-Brezzi condition is circumvented. Finding interpolations satisfying this condition is difficult with the FEM and even more difficult for meshfree interpolations and will therefore not be further mentioned.

Stabilizations of the Stokes equations have first been presented in [28], later in [53] for the incompressible Navier-Stokes equations. Both methods are very similar in that they only perturb the test function of the Lagrange multiplier, i.e. the pressure, leading to unsymmetric systems of equations for Stokes flow. This kind of stabilization is called throughout this paper Pressure-Stabilizing/Petrov-Galerkin (PSPG) as proposed in [53]. In [27], Stokes flow has been stabilized

with GLS stabilization, leading to perturbations of all test functions but maintaining symmetry. Note that GLS was already mentioned in the previous subsection for the stabilization of convection-dominated problems and can also be used here to circumvent the Babuška-Brezzi condition. This is not the case for SUPG stabilization which is only successful in suppressing oscillations from convection-dominated problems.

2.3 Steep Solution Gradients

In subsection 2.1 it has been shown that convection-dominated problems require stabilization such that a pollution of the overall solution with oscillations is prevented. However, these stabilizations do not preclude "over- and undershooting" about sharp internal and boundary layers [34]. These somehow "localized" (in that they do not influence the whole domain) oscillations can be suppressed by getting control over the solution gradient. The aim is to obtain a *monotone* solution without any oscillations. These methods have also been called "maximum-principle satisfying methods" in the literature.

There is, however, a very severe restriction concerning the monotonicity of a numerical scheme, which is summarized in the theorem of Godunov. There, it is proven that no linear higher-order method can obtain monotone solutions [21]. Thus, there are only two ways to achieve monotonicity: Using first order accurate schemes such as upwind finite differences or using non-linear schemes. The first way is in fact no real alternative, as higher-order accuracy is essential in the reliable simulation of many problems, consequently non-linear schemes have to be developed.

In the resulting schemes, there is always some kind of analysis and control of an interim solution. In the finite difference and finite volume context this can for example be realized with the so-called slope-limiter methods, a subclass of the monotone Total Variation Diminishing (TVD) schemes [21]. The minmod-slope-limiter, Roe's superbee limiter, van-Leer-limiter are well-known examples of slope-limiters.

One of the first monotone methods in the finite element context for convection-diffusion problems is the one proposed in [49], where the non-linearity is introduced by detection of element downstream nodes and a specific element matrix for the advection term depending on that node. In [44], Mizukami and Hughes introduce the first consistent monotone *Petrov-Galerkin* FEM, valid for linear triangular elements with acute angles only. In Petrov-Galerkin FEMs the non-linearity lies in the dependence of the perturbed test function upon the solution gradient. The resulting discretized equations are non-linear even for a linear problem.

In [34] over- and undershoots are stabilized with a discontinuity-capturing term, being the first generalizable approach to complex multidimensional systems. This Petrov-Galerkin method contains test functions modified with the added discontinuity-capturing term, acting in the direction of the solution gradient. Note that, in contrast, the stabilizations of 2.1 act in the direction of the streamline. Having control in direction of the streamline and of the solution gradient enables higher-order monotone schemes with enhanced robustness with the price of non-linearity. In [55] the discontinuity operator is generalized to non-linear convection-diffusion-reaction equations and in [33] and later [50] to multidimensional advective-diffusive systems such as the Navier-Stokes equations.

The authors would like to annotate that a compromise has to be made whether the steepness

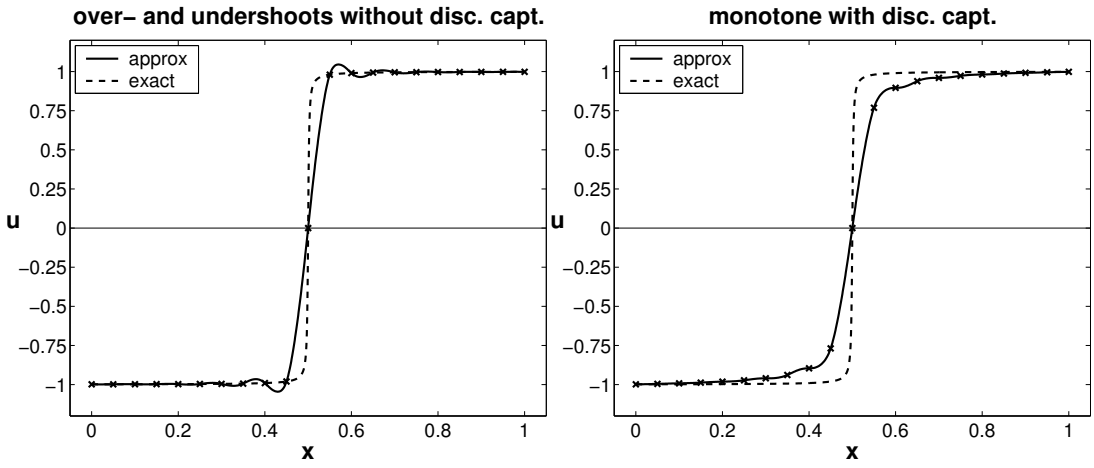


Figure 3: Approximation with over- and undershoots and monotone approximation.

of a solution or the monotonicity is of higher importance. It is an immanent feature of the shape functions of a numerical method —e.g. their supports and functional form— that only a certain gradient can be represented without over- and undershoots. This can be depicted from Figure 3. The only way to obtain a monotone solution is to smear out the steep gradients in the domain such that the method can represent it without over- and undershoots. Thus, for higher-order shape functions and meshfree shape functions a more accurate solution (in terms of approximation error and solution steepness) will be obtained in the presence of over- and undershoots, i.e. without or tuned influence of a discontinuity-capturing term.

3 Review of Stabilization Methods

In this section some of the most important stabilization methods are described roughly. It is our aim to outline their different structures and for which kind of problems —referring to section 2— they are suited. All stabilization schemes described in the following are Petrov-Galerkin approaches. They all add perturbations to the original Bubnov-Galerkin weak form. These perturbations are formulated in terms of modifications of the Bubnov-Galerkin test functions. They are multiplied with the residuals of the differential equations and thereby ensure consistency. Additionally, a stabilization parameter τ weights the influence of the added stabilization terms.

3.1 Streamline-Upwind/Petrov-Galerkin (SUPG)

In subsection 2.1 it was claimed that, in the FDM, introducing artificial diffusion in a smart way smoothes out the oscillations in convection-dominated problems. This motivation is the starting point of the Streamline-Upwind/Petrov-Galerkin (SUPG) method, first published by Hughes *et al.* in [26, 7]. It introduces a certain amount of artificial diffusion in streamline direction only. The latter aspect ensures that no diffusion perpendicular to the flow direction is introduced, which was the reason for excessive over diffusion in other methods. The details, how the SUPG introduces artificial diffusion in streamline direction and the determination of the "right" amount —controlled

with the stabilization parameter τ— is intensively discussed in section 4.

Here, only the main structure of the SUPG method is shown. Starting point is a PDE of the general form

$$\mathcal{L}u = f,$$

where $\mathcal{L}$ is any differential operator. Introducing an ansatz of the kind $\widetilde{u}(\mathbf{x}) = \mathbf{N}^T(\mathbf{x})\,\mathbf{u} = \sum N_I(\mathbf{x})\,u_I$, the weak form of the problem is

$$\int_\Omega \mathbf{w}^\star\,(\mathcal{L}\widetilde{u} - f)\,d\Omega = \mathbf{0}.$$

Choosing $\mathbf{w}^\star = \mathbf{N}$ leads to a Bubnov-Galerkin method, any $\mathbf{w}^\star \neq \mathbf{N}$ is called Petrov-Galerkin method. In the SUPG, as the name already implies, $\mathbf{w}^\star$ is chosen differently from $\mathbf{N}$. The standard Bubnov-Galerkin test functions $\mathbf{w} = \mathbf{N}$ are modified by a streamline upwind perturbation [7] of the kind

$$\mathbf{w}^\star = \mathbf{w} + \tau\mathcal{L}_{adv}\mathbf{w},$$

where $\mathcal{L}_{adv}$ is the advective part of the whole operator $\mathcal{L}$, and τ is the stabilization parameter that weights the perturbation. It is later shown, why this particular modification ban be interpreted as an introduction of artificial diffusion to the problem.

Note that the perturbation is multiplied with the residual form of the differential equation. Thereby, consistency is fulfilled from the beginning in that the exact solution also fulfills the stabilized weak form exactly. *Stabilization through a product of a perturbation and the residual* is a fundamental aspect of successful stabilization schemes and is realized in all stabilization method described herein.

There is one important aspect to mention at this point: In the FEM, piecewise polynomials are particularly useful shape functions, often having C^0 continuity in the domain Ω (and C^∞ inside an element). Then, the first derivatives include jumps at the element boundaries and second derivatives are Dirac-δ functions at the element boundaries. Integration in Ω over the product of two functions, where e.g. a jump and a Dirac-δ function coincide is not allowed (this occurs in terms such as $\int_\Omega w_{,x}N_{,xx}d\Omega$). This problem is well-known in the context of the least-squares FEM, see subsection 5.5, and may be handled there by using C^1-continuous shape functions, which are comparatively expansive. However, in the context of stabilization, where very similar terms as in the least-squares FEM occur, this problem is circumvented by defining the stabilization contributions *only inside element interiors*, where the shape functions are C^∞;

$$\int_\Omega \mathbf{w}\,(\mathcal{L}\widetilde{u} - f)\,d\Omega + \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau\mathcal{L}_{adv}\mathbf{w}\,(\mathcal{L}\widetilde{u} - f)\,d\Omega = \mathbf{0}.$$

Thereby, the stabilization does not upset higher continuity requirements as needed for the Bubnov-Galerkin weak form of the same problem. Note that *meshfree* shape functions used in practice are always at least C^1 continuous —they can be constructed to have arbitrary continuity— and that therefore no summation over subdomains of Ω has to be considered. Therefore, through-

out this paper, we do not overemphasize this aspect and write

$$\int_\Omega \left(\mathbf{w} + \tau \mathcal{L}_{adv}\mathbf{w}\right)\left(\mathcal{L}\widetilde{u} - f\right) d\Omega = \mathbf{0}$$

for simplicity whenever the continuity consideration is of less importance.

3.2 Pressure-Stabilizing/Petrov-Galerkin (PSPG)

Considering stabilization of mixed problems as described in subsection 2.2, the PSPG stabilization is a common technique. It has been introduced for the stabilization of the Stokes equations [28] and incompressible Navier-Stokes equations [53]. To avoid unnecessary confusion related to an introduction of abstract universal mixed formulations, the PSPG method is described for the case of Stokes equations only:

$$\begin{aligned} \text{momentum equations:} \quad \nabla \cdot \boldsymbol{\sigma} &= \mathbf{f}, \text{ with } \boldsymbol{\sigma} = -p\mathbf{I} + 2\mu\boldsymbol{\varepsilon} \\ \text{continuity equation:} \quad \nabla \cdot \mathbf{u} &= 0, \end{aligned}$$

$$\int_\Omega \mathbf{w} \cdot \left(\nabla \cdot \boldsymbol{\sigma} - \mathbf{f}\right) d\Omega + \int_\Omega q \left(\nabla \cdot \mathbf{u}\right) + \int_\Omega \tau \nabla q \cdot \left(\nabla \cdot \boldsymbol{\sigma} - \mathbf{f}\right) d\Omega = 0.$$

$\mathbf{u}$ is the velocity vector, p the pressure, $\mathbf{I}$ is the identity tensor, μ the dynamic viscosity and $\boldsymbol{\varepsilon} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$. The continuity condition is also called incompressibility constraint, underlining the mixed character of this formulation. Note that the equations governing Stokes flow are identical to the equations of classical isotropic incompressible elasticity, where $\mathbf{u}$ stands for displacement and μ for the shear modulus. The third term of the weak form is the PSPG stabilization term, which consists of a perturbation $\tau\nabla q$ multiplied with the residual of the momentum equation. The existence of a time-dependent term (instationary Stokes equations) or the existence of additional advective terms (Navier-Stokes equations) does not influence the structure of the PSPG stabilization, only the residual is modified then.

In mixed convection-dominated problems, such as the incompressible Navier-Stokes equations with high Reynolds-numbers, SUPG *and* PSPG (called herein SUPG/PSPG) stabilization have to be applied to obtain satisfactory results. It should also be mentioned that the PSPG stabilization parameter τ does not necessarily have to be identical with the SUPG stabilization parameter [53]. However, this is almost always the case in practice and can be explained with theoretical analysis, see section 5.

3.3 Galerkin/Least-Squares (GLS)

The GLS stabilization has been introduced in 1988 by Hughes and Franca in [29]. Before, it has been applied to a large number of separate problems which has then be summarized to a method on its own under the name GLS. It can be interpreted as a generalization of the SUPG method and was motivated from mathematical analysis rather than the artificial diffusion aspect. In the GLS method, the operator over the test functions is the differential operator of the original problem.

Hence, for the same weak form of

$$\int_\Omega \mathbf{w}^\star \left(\mathcal{L}\widetilde{u} - f\right) d\Omega = \mathbf{0}$$

from subsection 3.1, there is a modification of the test function of

$$\mathbf{w}^\star = \mathbf{w} + \tau \mathcal{L}\mathbf{w},$$

i.e. the whole operator $\mathcal{L}$ of a differential equation is taken to perturb the test function. One can thus see that the difference to the SUPG is in the modification of $\tau\mathcal{L}\mathbf{w}$ for the GLS instead of $\tau\mathcal{L}_{adv}\mathbf{w}$ for the SUPG. For hyperbolic systems (no diffusion, i.e. second order terms) and/or linear test and shape functions, the GLS stabilization reduces to the SUPG stabilization [29].

It is important to note that GLS stabilization automatically allows arbitrary combinations of interpolations, which is realized by circumventing the Babuška-Brezzi conditions from the beginning, see e.g. [27]. Hence oscillations and other problems described in subsections 2.1 and 2.2 can be stabilized with GLS stabilization. It is an interesting fact that SUPG/PSPG stabilization, e.g. for the incompressible Navier-Stokes equations, can be motivated from the GLS stabilization with only a few reductions [53]; in case of linear FEM, they fully agree.

3.4 Discontinuity Capturing

As pointed out in subsection 2.3 over- and undershoots in the solution can be prevented by getting control in the direction of the solution gradient. Using a Petrov-Galerkin approach, this is done with the following modification of the test functions [34]

$$\mathbf{w}^\star = \mathbf{w} + \tau \mathbf{c}_{\|} \cdot \mathbf{w},$$

where $\tau \mathbf{c}_{\|} \cdot \mathbf{w}$ is the discontinuity-capturing term. Note that additionally a stabilization with SUPG or GLS is necessary to get control in the direction of the streamline. The parameter $\mathbf{c}_{\|}$ is a projection of the advection direction $\mathbf{c}$ onto the solution gradient $\nabla\widetilde{u}$ as shown in Figure 4. It is defined as

$$\mathbf{c}_{\|} = \begin{cases} \frac{\mathbf{c}\cdot\nabla\tilde{u}}{|\nabla\tilde{u}|_2^2}\nabla\tilde{u}, & \text{if } \nabla\tilde{u} \neq 0 \\ 0, & \text{if } \nabla\tilde{u} = 0 \end{cases}.$$

The parameter τ is defined differently from the stabilization parameters for SUPG, PSPG and GLS.

4 Illustrative Approach: Linear FEM and Artificial Diffusion

In the previous section several stabilization schemes are described roughly to emphasize their different structures. In this section the focus is on motivation and understanding rather than only describing how the methods are defined. We restrict ourselves to the stabilization of convection-dominated problems as described in subsection 2.1. There it is pointed out that SUPG and GLS

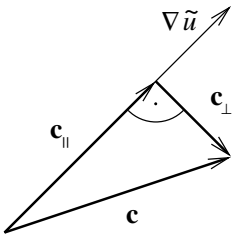


Figure 4: Projection of the advection direction $\mathbf{c}$ onto the solution gradient $\nabla\tilde{u}$.

stabilization are suited to smooth out oscillations in the solution. Our considerations are further restricted to the linear FEM, because then, SUPG and GLS become equal. Then, they can both be motivated with the same underlying idea of introducing artificial diffusion in a smart way.

It is our aim to explain in an illustrative way how the methods work, rather than from mathematical analysis. Therefore, the stabilization of simple model equations is carefully analyzed and conclusions are made for more general cases. Special attention is given to the choice of the stabilization parameter τ.

4.1 One-dimensional Advection-Diffusion Equation

A particularly simple differential equation which shows the typical oscillations in case of convection domination is the linear, one-dimensional advection-diffusion equation which reads in strong form

$$c\frac{\partial u}{\partial x} - K\frac{\partial^2 u}{\partial x^2} = 0.$$

A scalar quantity $u\,(x)$ is advected with the velocity c and thereby undergoes a diffusion dependent on K. The exact solution of this problem is known as

$$u^{\mathrm{ex}}\,(x) = C_1 e^{\frac{c}{K}x} + C_2,$$

where the constants C_1 and C_2 can be determined with help of the boundary conditions.

4.1.1 Finite Difference Method

The FDM is a good starting point to demonstrate oscillations in dependence of the convection-diffusion ratio and to motivate artificial diffusion as a help for stabilization. It is probably the simplest numerical method to solve this problem and it is still closely related to the Bubnov-Galerkin handling with the linear FEM. Assume $c > 0$, i.e. flow from left to right, and a regular node distribution with $\Delta x = \mathrm{const}$ as shown in Figure 5.

Two different difference formulas for the advection term are compared, while the diffusion term is approximated with the same stencil in both cases:

$$\begin{aligned}
\text{advection term: } \frac{\partial u}{\partial x} &\approx \frac{u_{i+1} - u_{i-1}}{2\Delta x} \text{ (central) or } \frac{u_i - u_{i-1}}{\Delta x} \text{ (upwind),} \\
\text{diffusion term: } \frac{\partial^2 u}{\partial x^2} &\approx \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2} \text{ (central).}
\end{aligned}$$

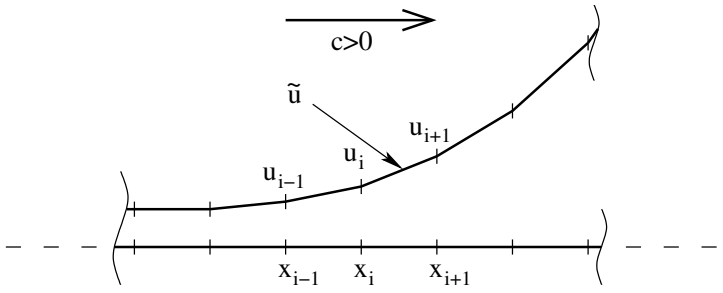


Figure 5: Regular point distribution and approximate solution $\tilde{u}$ in one dimension.

Results are displayed in Figure 6, each showing —for varying advection-diffusion ratios— the exact solution and the two approximated solutions with central and upwind differencing on the advective term. The element Peclet number $Pe = \frac{c\Delta x}{2K}$ is an important characteristic number for the oscillations; for linear elements oscillations occur if $Pe > 1$.

It can be clearly seen that the upwind solutions do not show oscillations but introduce more diffusion (the solutions are less steep than the exact solutions), whereas the central solutions show severe oscillations with growing advection-diffusion ratios. Note that in the upper left figure, where no oscillations occur, the center solution is even steeper than the exact solution. This fact gives rise to the interpretation that center differences are *under*-diffusive whereas upwind differences are *over*-diffusive. The exact solution is in-between the two approximations as long as no oscillations occur. It can thus be concluded that the exact solution can be better approximated by adding artificial diffusion to the center solution.

4.1.2 Modification of the Weak Form

The aim is now to introduce artificial diffusion in Bubnov-Galerkin methods, where one deals with the weak form of a differential equation. The weak form of the advection-diffusion equation, after an ansatz of the kind $\widetilde{u}(x) = \mathbf{N}^T(x)\,\mathbf{u} = \sum N_I(x)\,u_I$ becomes

$$\int_\Omega \mathbf{w}\left(c\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega = \mathbf{0},$$

where $\mathbf{w} = \mathbf{N}$. Artificial diffusion $\widetilde{K}$ is now introduced additionally to the physically existent diffusion K:

$$\begin{aligned}
\int_\Omega \mathbf{w}\left(c\frac{\partial \widetilde{u}}{\partial x} - \left(K+\widetilde{K}\right)\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega &= \mathbf{0} \\
\left[\int_\Omega c\mathbf{w}\frac{\partial \mathbf{N}^T}{\partial x} - \widetilde{K}\mathbf{w}\frac{\partial^2 \mathbf{N}^T}{\partial x^2} - K\mathbf{w}\frac{\partial^2 \mathbf{N}^T}{\partial x^2} d\Omega\right]\mathbf{u} &= \mathbf{0} \\
\left[c\int_\Omega \mathbf{w}\frac{\partial \mathbf{N}^T}{\partial x} d\Omega + \widetilde{K}\int_\Omega \frac{\partial \mathbf{w}}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} d\Omega + K\int_\Omega \frac{\partial \mathbf{w}}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} d\Omega - K\oint_\Gamma \mathbf{w}\frac{\partial \mathbf{N}^T}{\partial x} d\Gamma\right]\mathbf{u} &= \mathbf{0}.
\end{aligned}$$

Note that in the last step, the divergence theorem (or: partial integration) has been applied, leading to a boundary term *only* for the physical diffusion term and *not* for the artificial diffusion term. This certain aspect becomes mathematically justified, by considering the stabilization

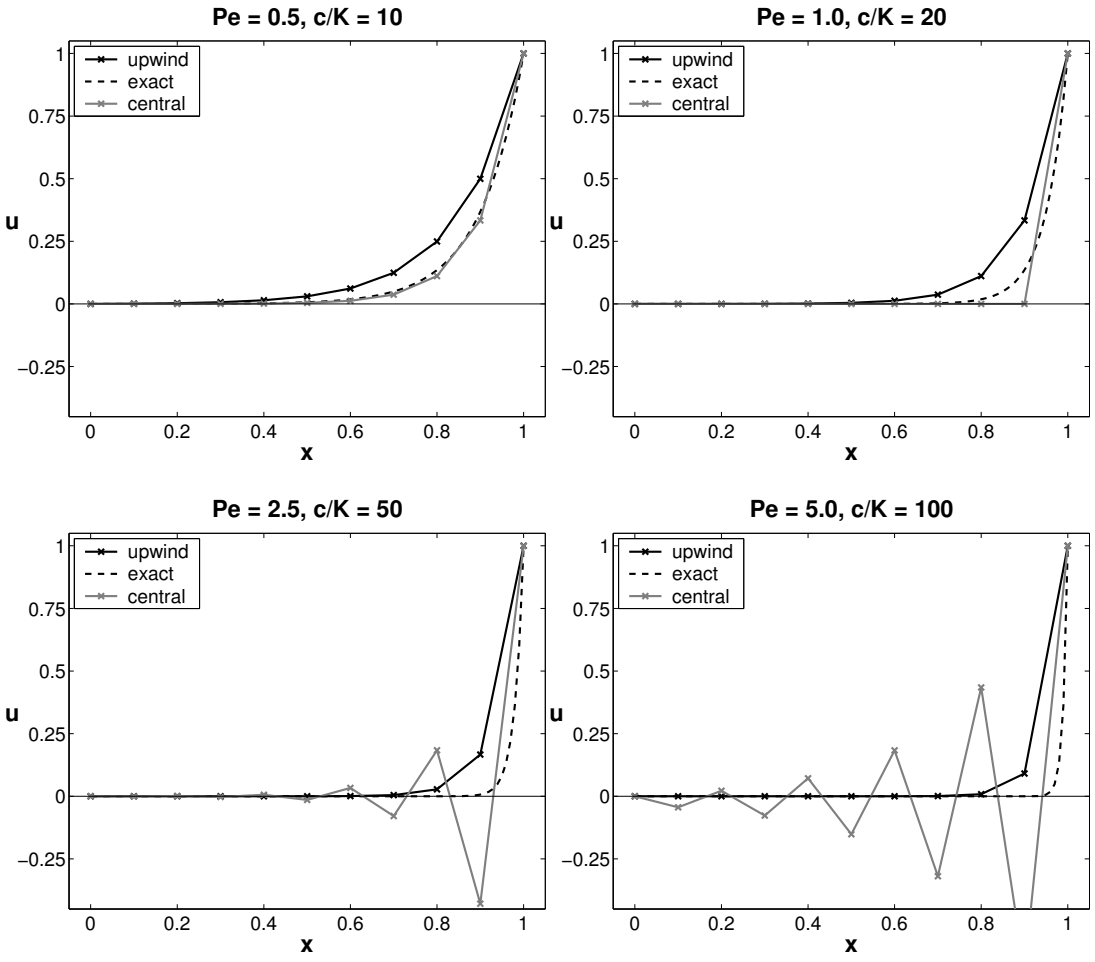


Figure 6: Comparison of the monotone over-diffusive upwind solution with the under-diffusive and —in convection dominated cases— oscillatory central solution and the exact solution.

contribution only in element interiors, i.e.

$$\int_\Omega \mathbf{w}\left(c\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega - \sum_{e=1}^{n_{el}} \int_{\Omega_e} \mathbf{w}\left(\widetilde{K}\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega = \mathbf{0}$$

instead of the above formulation, but then, the clear motivation as artificial diffusion becomes less obvious. Rearranging the above multiplied out equation gives

$$\left[c\int_\Omega \left(\mathbf{w} + \frac{\widetilde{K}}{c}\frac{\partial \mathbf{w}}{\partial x}\right)\frac{\partial \mathbf{N}^T}{\partial x} d\Omega + K\int_\Omega \frac{\partial \mathbf{w}}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} d\Omega - K\oint_\Gamma \mathbf{w}\frac{\partial \mathbf{N}^T}{\partial x} d\Gamma\right]\mathbf{u} = \mathbf{0}.$$

Comparing this with the weak form of the original problem without artificial diffusion,

$$\left[c\int_\Omega \mathbf{w}\frac{\partial \mathbf{N}^T}{\partial x} d\Omega + K\int_\Omega \frac{\partial \mathbf{w}}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} d\Omega - K\oint_\Gamma \mathbf{w}\frac{\partial \mathbf{N}^T}{\partial x} d\Gamma\right]\mathbf{u} = \mathbf{0}$$

one can be interpret artificial diffusion as a modification of the weak form, such that the test functions of the advective term becomes $\mathbf{w} + \frac{\widetilde{K}}{c}\frac{\partial \mathbf{w}}{\partial x}$ instead of only $\mathbf{w}$. There are many ways to construct numerical methods for this problem having "optimal" solutions, i.e. having solutions that are neither over- nor under-diffusive and approximate the exact solution very accurate without any oscillations [20, 24, 38]. However, the appearance of source terms, generalization to instationary and/or multi-dimensional problems gives very bad results [7]. The reason for this is the *inconsistency* of the above equation: the exact solution will not satisfy the weak form of the problem with artificial diffusion.

The SUPG, however, is a consistent method, which is part of the success of this method. The modified test function of the advection term is applied to *all* terms of the weak form, leading to

$$\begin{aligned}\Bigg[c\int_\Omega \left(\mathbf{w} + \frac{\widetilde{K}}{c}\frac{\partial \mathbf{w}}{\partial x}\right)\frac{\partial \mathbf{N}^T}{\partial x} d\Omega + K\int_\Omega \frac{\partial}{\partial x}\left(\mathbf{w} + \frac{\widetilde{K}}{c}\frac{\partial \mathbf{w}}{\partial x}\right)\frac{\partial \mathbf{N}^T}{\partial x} d\Omega \\ -K\oint_\Gamma \left(\mathbf{w} + \frac{\widetilde{K}}{c}\frac{\partial \mathbf{w}}{\partial x}\right)\frac{\partial \mathbf{N}^T}{\partial x} d\Gamma\Bigg]\mathbf{u} &= \mathbf{0}.\end{aligned}$$

Re-application of partial integration gives

$$\int_\Omega \left(\mathbf{w} + \frac{\widetilde{K}}{c}\frac{\partial \mathbf{w}}{\partial x}\right)\left(c\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega = \mathbf{0},$$

where the consistency of the method becomes obvious. Note that instead of a Bubnov-Galerkin method a Petrov-Galerkin method results with the test function $\mathbf{w}^\star = \mathbf{w} + \frac{\widetilde{K}}{c}\frac{\partial \mathbf{w}}{\partial x} = \mathbf{w} + \tau c\frac{\partial \mathbf{w}}{\partial x}$, where $\tau = \frac{\widetilde{K}}{c^2}$ and $\mathbf{w} = \mathbf{N}$. Mathematically more correct one should again consider that in the FEM, the stabilization is defined only in element interiors which gives a weak form of

$$\int_\Omega \mathbf{w}\left(c\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega + \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau c\frac{\partial \mathbf{w}}{\partial x}\left(c\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega = \mathbf{0}.$$

The question of how to choose the right amount of artificial diffusion $\widetilde{K}$, which is totally equivalent to find the right stabilization parameter τ, to obtain "optimal" results is still open and

is discussed in the following. Note that the interpretation of $\widetilde{K}$ as artificial diffusion is not fully correct after application of the modified test function of the advection term to all other terms, however, it remains to be the main effect of the stabilization.

4.1.3 Weighting the Modification: Stabilization Parameter τ and the coth-Formula

With help of the exact solution, which is known for this simple model problem, it is possible to determine the stabilization parameter τ in any way the "optimal" approximation is desired. A particularly useful choice for the "optimality" of the approximation is to obtain the *nodally exact solution* (which is in this case equivalent to minimizing the H^1 seminorm of the solution error $\widetilde{u} - u$ [50], see subsection 4.1.9). In the following this "optimal" choice for τ is determined, such that the approximation coincides with the exact solutions at all nodes. Starting point is

$$\int_\Omega \mathbf{w} \left(c\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2} \right) d\Omega + \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau c \frac{\partial \mathbf{w}}{\partial x} \left(c\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2} \right) d\Omega = \mathbf{0}.$$

From this system of equations a certain equation for node I is extracted and the ansatz is applied

$$\left[\int_\Omega w_I \left(c\frac{\partial \mathbf{N}^T}{\partial x} - K\frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) d\Omega + \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau_I c \frac{\partial w_I}{\partial x} \left(c\frac{\partial \mathbf{N}^T}{\partial x} - K\frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) d\Omega \right] \mathbf{u} = \mathbf{0}.$$

The vector of "unknown" values $\mathbf{u}$ is in fact known for this problem due to the exact solution

$$\mathbf{u}^{\text{ex}} = C_1 e^{\frac{c}{K}\mathbf{x}} + C_2 = \begin{pmatrix} C_1 e^{\frac{c}{K}x_1} + C_2 \\ C_1 e^{\frac{c}{K}x_2} + C_2 \\ \vdots \\ C_1 e^{\frac{c}{K}x_n} + C_2 \end{pmatrix},$$

where x_i are the node positions. Thus, equation I for τ_I can be rearranged:

$$\left[\sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau_I c \frac{\partial w_I}{\partial x} \left(c\frac{\partial \mathbf{N}^T}{\partial x} - K\frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) \right] \mathbf{u}^{\text{ex}} = - \left[\int_\Omega w_I \left(c\frac{\partial \mathbf{N}^T}{\partial x} - K\frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) \right] \mathbf{u}^{\text{ex}}$$

$$\begin{aligned} \Rightarrow \tau_I &= -\frac{\left[\int_\Omega w_I \left(c\frac{\partial \mathbf{N}^T}{\partial x} - K\frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) \right] \mathbf{u}^{\text{ex}}}{\left[\sum_{e=1}^{n_{el}} \int_{\Omega_e} c\frac{\partial w_I}{\partial x} \left(c\frac{\partial \mathbf{N}^T}{\partial x} - K\frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) \right] \mathbf{u}^{\text{ex}}} \\ &= -\frac{\left[c\int_\Omega w_I \frac{\partial \mathbf{N}^T}{\partial x} + K\int_\Omega \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} - K\oint_\Gamma w_I \frac{\partial \mathbf{N}^T}{\partial x} \right] \mathbf{u}^{\text{ex}}}{\left[\sum_{e=1}^{n_{el}} c^2 \int_{\Omega_e} \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} - cK\int_{\Omega_e} \frac{\partial w_I}{\partial x}\frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right] \mathbf{u}^{\text{ex}}}. \end{aligned}$$

This expression is still independent of the chosen FEM shape functions. In here, *linear* shape functions are considered, consequently, the term $\int_{\Omega_e} \frac{\partial w_I}{\partial x}\frac{\partial^2 \mathbf{N}^T}{\partial x^2}$ cancels out as $\frac{\partial^2 \mathbf{N}^T}{\partial x^2} = 0$ in the element interior and there is no point in shifting the derivative onto the test function. Note that at this point it becomes important to have the stabilization in element interiors only, because at the element boundaries, the second derivative is *not* zero, but the Dirac delta function! This becomes clear from Figure 7.

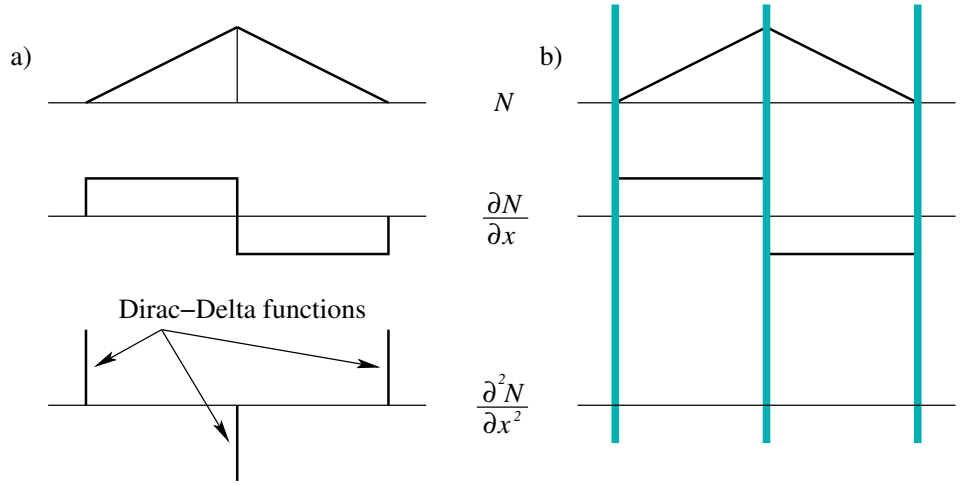


Figure 7: Difference between the integration over the whole linear shape function support and the element interiors only. Note that $\int_\Omega N d\Omega = \sum \int_{\Omega_e} N d\Omega$ and $\int_\Omega N_{,x} d\Omega = \sum \int_{\Omega_e} N_{,x} d\Omega$, but $\int_\Omega N_{,xx} d\Omega \neq \sum \int_{\Omega_e} N_{,xx} d\Omega$.

Note also that $w_I = 0$ at the element boundaries and consequently the boundary term in the numerator cancels out, too:

$$\begin{aligned}\tau_I &= -\frac{\left[c\int_\Omega w_I \frac{\partial \mathbf{N}^T}{\partial x} + K\int_\Omega \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x}\right]\mathbf{u}^{\text{ex}}}{\left[\sum_{e=1}^{n_{el}} c^2 \int_{\Omega_e} \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x}\right]\mathbf{u}^{\text{ex}}} \\ &= -\frac{\left[c\int_\Omega w_I \frac{\partial \mathbf{N}^T}{\partial x}\right]\mathbf{u}^{\text{ex}}}{\left[\sum_{e=1}^{n_{el}} c^2 \int_{\Omega_e} \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x}\right]\mathbf{u}^{\text{ex}}} - \frac{\left[K\int_\Omega \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x}\right]\mathbf{u}^{\text{ex}}}{\left[\sum_{e=1}^{n_{el}} c^2 \int_{\Omega_e} \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x}\right]\mathbf{u}^{\text{ex}}}.\end{aligned}$$

With $\int_\Omega w_I \frac{\partial \mathbf{N}^T}{\partial x} = \sum_{e=1}^{n_{el}} \int_{\Omega_e} w_I \frac{\partial \mathbf{N}^T}{\partial x}$ and $\int_\Omega \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} = \sum_{e=1}^{n_{el}} \int_{\Omega_e} \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x}$ (but $\int_\Omega \left(\frac{\partial w_I}{\partial x}\frac{\partial^2 \mathbf{N}^T}{\partial x^2}\right) \neq \sum_{e=1}^{n_{el}} \int_{\Omega_e} \frac{\partial w_I}{\partial x}\frac{\partial^2 \mathbf{N}^T}{\partial x^2}$, due to the negligence of the Dirac delta function contributions on the element boundaries in case of the right hand side term, see Figure 7) follows

$$\tau_I = -\frac{\left[\int_\Omega w_I \frac{\partial \mathbf{N}^T}{\partial x}\right]\mathbf{u}^{\text{ex}}}{\left[c\int_\Omega \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x}\right]\mathbf{u}^{\text{ex}}} - \frac{K}{c^2}.$$

One can evaluate these integrals explicitely for linear shape and test functions and a regular point

distribution, which is exactly the situation plotted in Figure 5. The result is

$$\tau_I = -\frac{\int_\Omega \begin{bmatrix} \cdots & 0 & w_I\frac{\partial N_{I-1}}{\partial x} & w_I\frac{\partial N_I}{\partial x} & w_I\frac{\partial N_{I+1}}{\partial x} & 0 & \cdots \end{bmatrix} \begin{bmatrix} \vdots \\ u^{\text{ex}}_{I-1} \\ u^{\text{ex}}_I \\ u^{\text{ex}}_{I+1} \\ \vdots \end{bmatrix}}{c\int_\Omega \begin{bmatrix} \cdots & 0 & \frac{\partial w_I}{\partial x}\frac{\partial N_{I-1}}{\partial x} & \frac{\partial w_I}{\partial x}\frac{\partial N_I}{\partial x} & \frac{\partial w_I}{\partial x}\frac{\partial N_{I+1}}{\partial x} & 0 & \cdots \end{bmatrix} \begin{bmatrix} \vdots \\ u^{\text{ex}}_{I-1} \\ u^{\text{ex}}_I \\ u^{\text{ex}}_{I+1} \\ \vdots \end{bmatrix}} - \frac{K}{c^2}$$

$$= -\frac{\begin{bmatrix} \cdots & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & \cdots \end{bmatrix} \begin{bmatrix} \vdots \\ C_1e^{\frac{c}{K}x_{I-1}}+C_2 \\ C_1e^{\frac{c}{K}x_I}+C_2 \\ C_1e^{\frac{c}{K}x_{I+1}}+C_2 \\ \vdots \end{bmatrix}}{c\begin{bmatrix} \cdots & 0 & -\frac{1}{\Delta x} & \frac{2}{\Delta x} & -\frac{1}{\Delta x} & 0 & \cdots \end{bmatrix} \begin{bmatrix} \vdots \\ C_1e^{\frac{c}{K}x_{I-1}}+C_2 \\ C_1e^{\frac{c}{K}x_I}+C_2 \\ C_1e^{\frac{c}{K}x_{I+1}}+C_2 \\ \vdots \end{bmatrix}} - \frac{K}{c^2}$$

$$= -\frac{-\frac{1}{2}\left(C_1e^{\frac{c}{K}x_{I-1}}+C_2\right)+0\left(C_1e^{\frac{c}{K}x_I}+C_2\right)+\frac{1}{2}\left(C_1e^{\frac{c}{K}x_{I+1}}+C_2\right)}{-c\frac{1}{\Delta x}\left(C_1e^{\frac{c}{K}x_{I-1}}+C_2\right)+c\frac{2}{\Delta x}\left(C_1e^{\frac{c}{K}x_I}+C_2\right)-c\frac{1}{\Delta x}\left(C_1e^{\frac{c}{K}x_{I+1}}+C_2\right)} - \frac{K}{c^2}.$$

A number of algebraic manipulations is possible to simplify this expression. The constant C_1 can be canceled out from the fraction, C_2 cancels out after a simple factorization due to $-\frac{1}{2}+\frac{1}{2}=0$ and $-\frac{1}{\Delta x}+\frac{2}{\Delta x}-\frac{1}{\Delta x}=0$. It remains

$$\begin{aligned}
\tau_I &= -\frac{1}{2c}\frac{-e^{\frac{c}{K}x_{I-1}}+e^{\frac{c}{K}x_{I+1}}}{-\frac{1}{\Delta x}e^{\frac{c}{K}x_{I-1}}+\frac{2}{\Delta x}e^{\frac{c}{K}x_I}-\frac{1}{\Delta x}e^{\frac{c}{K}x_{I+1}}} - \frac{K}{c^2} \\
&= -\frac{1}{2c}\frac{-e^{\frac{c}{K}(x_I-\Delta x)}+e^{\frac{c}{K}(x_I+\Delta x)}}{-\frac{1}{\Delta x}e^{\frac{c}{K}(x_I-\Delta x)}+\frac{2}{\Delta x}e^{\frac{c}{K}x_I}-\frac{1}{\Delta x}e^{\frac{c}{K}(x_I+\Delta x)}} - \frac{K}{c^2} \\
&= -\frac{\Delta x}{2c}\frac{e^{\frac{c}{K}x_I}\left(-e^{-\frac{c}{K}\Delta x}+e^{\frac{c}{K}\Delta x}\right)}{e^{\frac{c}{K}x_I}\left(-e^{-\frac{c}{K}\Delta x}+2-e^{\frac{c}{K}\Delta x}\right)} - \frac{K}{c^2} \\
&= \frac{\Delta x}{2c}\frac{\frac{1}{2}\left(e^{\frac{c}{K}\Delta x}-e^{-\frac{c}{K}\Delta x}\right)}{\frac{1}{2}\left(e^{\frac{c}{K}\Delta x}+e^{-\frac{c}{K}\Delta x}\right)-1} - \frac{K}{c^2} \\
&= \frac{\Delta x}{2c}\frac{\sinh\left(\frac{c}{K}\Delta x\right)}{\cosh\left(\frac{c}{K}\Delta x\right)-1} - \frac{K}{c^2}.
\end{aligned}$$

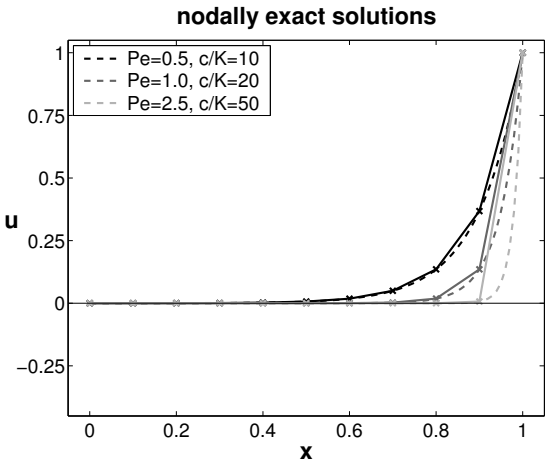


Figure 8: Approximate solutions obtained with help of the coth-solution: The solution is free of oscillations and nodally exact.

In the last step, the knowledge of $\sinh(a) = \frac{1}{2}(e^a - e^{-a})$ and $\cosh(a) = \frac{1}{2}(e^a + e^{-a})$ has been used. It requires a number of modifications to further reduce the remaining fraction, which is done in the following auxiliary calculation:

$$\begin{aligned}
\frac{\sinh(a)}{\cosh(a)-1} &= \frac{\frac{1}{2}(e^a - e^{-a})}{\frac{1}{2}(e^a + e^{-a}) - 1} = \frac{e^a - \frac{1}{e^a}}{e^a + \frac{1}{e^a} - 2} = \frac{\frac{e^{2a}}{e^a} - \frac{1}{e^a}}{\frac{e^{2a}}{e^a} + \frac{1}{e^a} - \frac{2e^a}{e^a}} \\
&= \frac{e^{2a} - 1}{e^{2a} - 2e^a + 1} = \frac{(e^a - 1)(e^a + 1)}{(e^a - 1)(e^a - 1)} = \frac{(e^a + 1)}{(e^a - 1)} \\
&= \frac{e^{\frac{a}{2}}\left(e^{\frac{a}{2}} + e^{-\frac{a}{2}}\right)}{e^{\frac{a}{2}}\left(e^{\frac{a}{2}} - e^{-\frac{a}{2}}\right)} = \frac{\frac{1}{2}\left(e^{\frac{a}{2}} + e^{-\frac{a}{2}}\right)}{\frac{1}{2}\left(e^{\frac{a}{2}} - e^{-\frac{a}{2}}\right)} = \frac{\cosh\left(\frac{a}{2}\right)}{\sinh\left(\frac{a}{2}\right)} \\
&= \coth\left(\frac{a}{2}\right).
\end{aligned}$$

With this knowledge the final equation for τ_I becomes

$$\begin{aligned}
\tau_I &= \frac{\Delta x}{2c}\coth\left(\frac{c\Delta x}{2K}\right) - \frac{K}{c^2} \\
&= \frac{\Delta x}{2c}\left(\coth\left(\frac{c\Delta x}{2K}\right) - \frac{2K}{c\Delta x}\right).
\end{aligned}$$

The term $\frac{c\Delta x}{2K}$ can be identified with the well-known Peclet number, $Pe = \frac{c\Delta x}{2K}$, consequently

$$\tau_I = \frac{\Delta x}{2c}\left(\coth(Pe) - \frac{1}{Pe}\right).$$

With this definition of the stabilization parameter one obtains the nodally exact solution for the one-dimensional advection-diffusion equation, approximated with linear FEM and a regular node distribution. The success of this formulation can be seen from Figure 8. This definition of τ has often be called "optimal" in the literature [7, 20, 24, 38, ...].

One finds that this formula for τ has the following important properties:

- It is totally independent of C_1 and C_2, i.e. of the boundary conditions.
- It depends on the *relative* positions of the two neighboring nodes only. Neither are the absolute positions of any nodes of importance nor the relative positions of any other than the neighboring nodes.

In this paper these properties of a stabilization criterion are called *local*. Any formula depending on boundary conditions and/or on the positions of all nodes is called *global*.

4.1.4 Different Ways to Obtain the coth-Formula

There are also other possibilities to obtain the coth-formula of the previous subsection.

Analytical solution of a difference equation One way is to solve the resulting difference equation analytically and then setting it equal to the analytical solution of the differential equation, i.e. the advection-diffusion equation. This approach is the origin of the coth-formula and is mentioned in [8, 20, 24, 38]. The deduction is outlined in the following.

The resulting difference equation for all nodes of the stabilized weak form

$$\int_\Omega \mathbf{w}\left(c\frac{\partial\widetilde{u}}{\partial x} - K\frac{\partial^2\widetilde{u}}{\partial x^2}\right)d\Omega + \sum_{e=1}^{n_{el}}\int_{\Omega_e}\tau c\frac{\partial\mathbf{w}}{\partial x}\left(c\frac{\partial\widetilde{u}}{\partial x} - K\frac{\partial^2\widetilde{u}}{\partial x^2}\right)d\Omega = \mathbf{0}$$

is

$$\begin{aligned}\left(-\frac{1}{2}c - \frac{K}{\Delta x} - \tau_I\frac{c^2}{\Delta x}\right)u_{i-1} + \left(2\frac{K}{\Delta x} + 2\tau_I\frac{c^2}{\Delta x}\right)u_i + \\ \left(\frac{1}{2}c - \frac{K}{\Delta x} - \tau_I\frac{c^2}{\Delta x}\right)u_{i+1} \quad &= \quad 0\end{aligned}$$

in case of linear FEM with a regular node distribution and evaluated integrals (note that $\sum_{e=1}^{n_{el}}\int_{\Omega_e}\tau cK\frac{\partial\mathbf{w}}{\partial x}\frac{\partial^2\widetilde{u}}{\partial x^2} = 0$). This difference equation can be solved analytically by choosing an ansatz of $u_n = C\rho^n$. Then, $u_{n-1} = C\rho^{n-1}$ and $u_{n+1} = C\rho^{n+1}$ and a quadratic characteristic polynomial arises. Solving this polynomial for the eigenvalues ρ_1 and ρ_2 gives

$$\rho_1 = \frac{-\frac{1}{2}c - \frac{K}{\Delta x} - \tau_I\frac{c^2}{\Delta x}}{\frac{1}{2}c - \frac{K}{\Delta x} - \tau_I\frac{c^2}{\Delta x}} \qquad\qquad \rho_2 = 1$$

and the analytical solution for the difference equation immediately follows as

$$u_n = C_1\rho_1^n + C_2\rho_2^n = C_1\left(\frac{-\frac{1}{2}c - \frac{K}{\Delta x} - \tau_I\frac{c^2}{\Delta x}}{\frac{1}{2}c - \frac{K}{\Delta x} - \tau_I\frac{c^2}{\Delta x}}\right)^n + C_2.$$

Equating this with the exact solution $u_n = C_1e^{\frac{c}{K}x_n} + C_2 = C_1e^{\frac{c}{K}(n\cdot\Delta x)} + C_2$, canceling the constants C_1 and C_2, and taking the n-th root of the equation gives

$$\frac{-\frac{1}{2}c - \frac{K}{\Delta x} - \tau_I\frac{c^2}{\Delta x}}{\frac{1}{2}c - \frac{K}{\Delta x} - \tau_I\frac{c^2}{\Delta x}} = e^{\frac{c}{K}\Delta x}.$$

Isolating τ_I leads to

$$\begin{aligned}\tau_I &= \frac{-\frac{1}{2}c\left(1+e^{\frac{c}{K}\Delta x}\right)-\frac{K}{\Delta x}\left(1-e^{\frac{c}{K}\Delta x}\right)}{\frac{c^2}{\Delta x}\left(1-e^{\frac{c}{K}\Delta x}\right)} \\ &= \frac{\Delta x}{2c}\frac{\left(e^{\frac{c}{K}\Delta x}+1\right)}{\left(e^{\frac{c}{K}\Delta x}-1\right)}-\frac{K}{c^2}.\end{aligned}$$

Due to $\coth(a)=\frac{e^{2a}+1}{e^{2a}-1}$ and $Pe=\frac{c\Delta x}{2K}$, it follows directly that

$$\tau_I=\frac{\Delta x}{2c}\left(\coth(Pe)-\frac{1}{Pe}\right).$$

Note that for this deduction of the coth-formula the solution of a characteristic polynomial of order $k-1$ was necessary, where k is the "bandwidth" of the equation. In this case, $k=3$ and the terms u_{i-1}, u_i and u_{i+1} are present. For higher-order FEM, k increases, e.g. for quadratic polynomials $k=5$, including the additional terms u_{i-2} and u_{i+2}. Then, a deduction of an optimal τ_I-formula can not be found conveniently taking this way, because the characteristic polynomial is of order 4.

Taylor series expansion This alternative obtains the coth-formula via a Taylor series expansion. Rearrange the above difference equation, such that

$$\begin{aligned}c\left(-\frac{1}{2}u_{i-1}+\frac{1}{2}u_{i+1}\right)+\frac{K}{\Delta x}\left(-u_{i-1}+2u_i-u_{i+1}\right)+ \\ \tau_I\frac{c^2}{\Delta x}\left(-u_{i-1}+2u_i-u_{i+1}\right) &= 0\end{aligned}$$

and insert the following Taylor series expansions for u_{i-1} and u_{i+1}

$$\begin{aligned}u_{i-1} &= u_i-\Delta x\frac{\partial u_i}{\partial x}+\frac{\Delta x^2}{2!}\frac{\partial^2 u_i}{\partial x^2}-\frac{\Delta x^2}{3!}\frac{\partial^3 u_i}{\partial x^3}+\dots \\ u_{i+1} &= u_i+\Delta x\frac{\partial u_i}{\partial x}+\frac{\Delta x^2}{2!}\frac{\partial^2 u_i}{\partial x^2}+\frac{\Delta x^2}{3!}\frac{\partial^3 u_i}{\partial x^3}+\dots.\end{aligned}$$

This leads to

$$\begin{aligned}c\left[-\frac{1}{2}\left(u_i-\Delta x\frac{\partial u_i}{\partial x}+\frac{\Delta x^2}{2!}\frac{\partial^2 u_i}{\partial x^2}-\dots\right)+\frac{1}{2}\left(u_i+\Delta x\frac{\partial u_i}{\partial x}+\frac{\Delta x^2}{2!}\frac{\partial^2 u_i}{\partial x^2}+\dots\right)\right] \\ \left(\frac{K}{\Delta x}+\tau_I\frac{c^2}{\Delta x}\right)\left[-\left(u_i-\Delta x\frac{\partial u_i}{\partial x}+\frac{\Delta x^2}{2!}\frac{\partial^2 u_i}{\partial x^2}-\dots\right)+2u_i\right. \\ \left.-\left(u_i+\Delta x\frac{\partial u_i}{\partial x}+\frac{\Delta x^2}{2!}\frac{\partial^2 u_i}{\partial x^2}+\dots\right)\right] &= 0.\end{aligned}$$

With knowledge of the exact solution $u_i=C_1e^{c/Kx_i}+C_2$ and $\frac{\partial^n u_i}{\partial x^n}=C_1\left(\frac{c}{K}\right)^n e^{c/Kx_i}$ this expression becomes after a number of manipulations

$$c\left[\sinh\left(\frac{c\Delta x}{K}\right)\right]-\frac{2}{\Delta x}\left(K+\tau_Ic^2\right)\left[\cosh\left(\frac{c\Delta x}{K}\right)-1\right]=0.$$

Rearranging for τ_I gives the desired result

$$\begin{aligned}\tau_I &= \frac{\Delta x}{2c}\frac{\sinh\left(\frac{c\Delta x}{K}\right)}{\cosh\left(\frac{c\Delta x}{K}\right)-1}-\frac{K}{c^2}\\ &= \vdots\\ &= \frac{\Delta x}{2c}\left(\coth\left(Pe\right)-\frac{1}{Pe}\right),\end{aligned}$$

where the last steps have already been shown in the previous subsection.

We summarize the different ways to obtain the well-known coth-formula:

1. Solve the resulting difference equation, defined by the weak form of the problem, exactly and obtain τ_I with knowledge of the exact solution of the differential equation at x_i. Therefore, the characteristic polynomial has to be solved which is only possible (convenient) for linear FEM, because then the polynomial is only of second order.

2. Make a Taylor series analysis of the difference equation. The result is obtained using the knowledge of the exact solution and its derivatives at x_i. This approach is applicable to any difference equation, i.e. it can also be used to obtain nodally exact solutions also for higher-order interpolations. However, the deduction is rather lengthy due to the Taylor series expansion.

3. Use the approach from the previous subsection, i.e. the knowledge of the exact solution at all nodes $(\ldots, x_{i-1}, x_i, x_{i+1}, \ldots)$.This approach is also applicable to higher-order interpolations, and the nodally exact stabilization parameters τ_I are obtained in a very simple way. See subsection 4.1.8 for the determination of the nodally exact τ_I for quadratic elements.

4.1.5 τ for Irregular Node Distributions

The well-known optimal coth-formula for the stabilization parameter only leads to nodally exact approximations in case of regular node distributions, i.e. Δx is constant. However, one can also deduce a formula for irregular node distributions. Consider the element situation of Figure 5, but with $\Delta x_l = x_i - x_{i-1}$ and $\Delta x_r = x_{i+1} - x_i$ and $\Delta x_l \neq \Delta x_r$. Then the result of the integration becomes

$$\begin{aligned}\int_\Omega w_i \frac{\partial \mathbf{N}^T}{\partial x} d\Omega &= \left[\begin{array}{ccccc}\cdots & -\frac{1}{2} & 0 & \frac{1}{2} & \cdots\end{array}\right],\\ \int_\Omega \frac{\partial w_i}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} d\Omega &= \left[\begin{array}{ccccc}\cdots & -\frac{1}{\Delta x_l} & \frac{1}{\Delta x_l}+\frac{1}{\Delta x_r} & -\frac{1}{\Delta x_r} & \cdots\end{array}\right].\end{aligned}$$

The same procedure from above can be applied then, finally leading to the following expression for τ_I:

$$\tau_I = \frac{1}{2c}\frac{e^{\frac{c}{K}\Delta x_r}-e^{-\frac{c}{K}\Delta x_l}}{\frac{1}{\Delta x_l}e^{-\frac{c}{K}\Delta x_l}-\left(\frac{1}{\Delta x_l}+\frac{1}{\Delta x_r}\right)+\frac{1}{\Delta x_r}e^{\frac{c}{K}\Delta x_r}}-\frac{K}{c^2}.$$

Setting $\Delta x = \Delta x_l = \Delta x_r$ will enable a number of simplifications leading to the coth-formula of subsection 4.1.3. It can be seen that also for irregular point distributions the stabilization formula

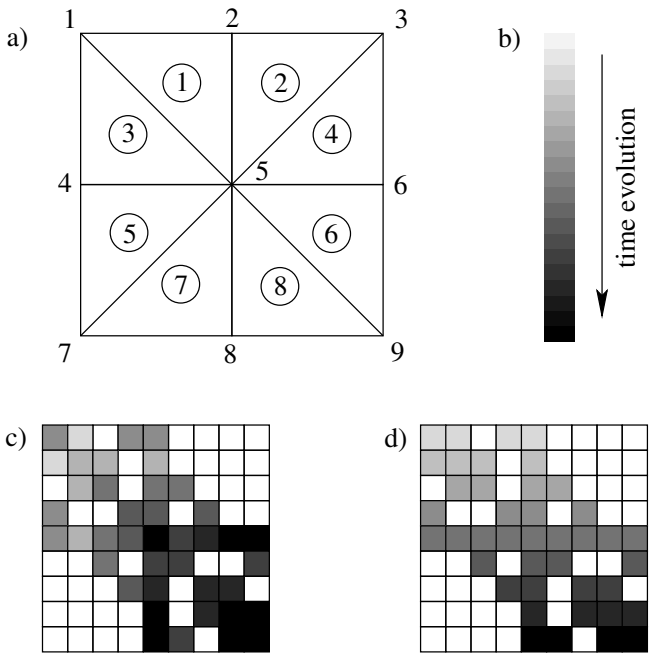


Figure 9: Difference between the typical element integration of the FEM and line-by-line integration. There is no assembly of the system of equations in the latter.

has a local character, as it is independent of boundary conditions and any node positions except the neighboring nodes.

4.1.6 Element vs. Nodal Stabilization

Up to this moment the stabilization parameter τ has been computed for each equation of the system of equations. Each equation refers to a certain node. Therefore, this stabilization is called *nodal* stabilization and the stabilization parameter is labeled τ_I. In practice, for the FEM, often an *element* stabilization is preferred, the stabilization parameter is then labeled τ_e. This is due to the fact that in the FEM one builds the system of equations more efficiently by assembling element matrices rather than integrating the whole system line by line. These different methodologies —both leading to the same systems of equations in the absence of stabilization terms— are symbolically depicted in Figure 9 to intuitively clarify this aspect.

Note that as a consequence, in FEM one prefers element stabilization whereas in MMs, where no elements are existent, nodal stabilization must be taken.

For the concrete problem of solving the 1D advection-diffusion equation with linear FEM one finds the following difference of the stabilization term between nodal and element stabilization

$$\sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau_I c^2 \frac{\partial w_I}{\partial x} \frac{\partial \mathbf{N}^T}{\partial x} \quad d\Omega = \quad \tau_I c^2 \left[\begin{array}{ccccc} \cdots & -\frac{1}{\Delta x_l} & \frac{1}{\Delta x_l} + \frac{1}{\Delta x_r} & -\frac{1}{\Delta x_r} & \cdots \end{array} \right],$$

$$\sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau_e c^2 \frac{\partial w_I}{\partial x} \frac{\partial \mathbf{N}^T}{\partial x} \quad d\Omega = \quad c^2 \left[\begin{array}{ccccc} \cdots & -\frac{\tau_{e,l}}{\Delta x_l} & \frac{\tau_{e,l}}{\Delta x_l} + \frac{\tau_{e,r}}{\Delta x_r} & -\frac{\tau_{e,r}}{\Delta x_r} & \cdots \end{array} \right],$$

where $\tau_{e,l}$ and $\tau_{e,r}$ are the element stabilization parameters of the left and the right element of node I. The following formula is taken for the computation of the element stabilization parameters

$$\begin{aligned}\tau_e &= \frac{\Delta x_e}{2c}\left(\coth\left(\frac{c\Delta x_e}{2K}\right) - \frac{2K}{c\Delta x_e}\right) \\ &= \frac{\Delta x_e}{2c}\left(\coth\left(Pe_e\right) - \frac{1}{Pe_e}\right),\end{aligned}$$

which is the one derived in subsection 4.1.3 for the nodal stabilization parameter τ_I. Clearly, in case of regular point distributions ($\Delta x = \text{const}$), element and nodal stabilization become identical, because then $\tau_I = \tau_{e,l} = \tau_{e,r}$. But for irregular node distributions it is *impossible* to obtain nodally exact results with element stabilization, because then the information of the relative up- and downstream positions of the neighboring nodes is needed and this cannot be obtained from only one element. Then, only nodal stabilization with τ_I from subsection 4.1.5 can give nodally exact approximations.

4.1.7 The Role of the Downstream Node

The resulting formulas for the stabilization parameters of the previous subsections are rather complicated. Simplifications, to extract the most important characteristics, are possible with help of estimates. In the following, it is assumed that the problem is convection-dominated, $c \gg K$. Also assume $c\Delta x \gg K$ (equivalently: $Pe \gg 1$), because otherwise, the mesh would be fine enough so that stabilization would not be needed.

With these assumptions the following estimate for the coth-formula of subsection 4.1.3 holds

$$\begin{aligned}\tau_I &= \frac{\Delta x}{2c}\left(\coth\left(Pe\right) - \frac{1}{Pe}\right) \\ &\approx \frac{\Delta x}{2c}\left(1 - 0\right) \\ &\approx \frac{\Delta x}{2c}.\end{aligned}$$

For the formula of subsection 4.1.5 for irregular node distributions

$$\tau_I = \frac{1}{2c}\frac{e^{\frac{c}{K}\Delta x_r} - e^{-\frac{c}{K}\Delta x_l}}{\frac{1}{\Delta x_l}e^{-\frac{c}{K}\Delta x_l} - \left(\frac{1}{\Delta x_l} + \frac{1}{\Delta x_r}\right) + \frac{1}{\Delta x_r}e^{\frac{c}{K}\Delta x_r}} - \frac{K}{c^2},$$

one can estimate $e^{-\frac{c}{K}\Delta x_l} \approx 0$, as well as $\frac{K}{c^2} \approx 0$ and $e^{\frac{c}{K}\Delta x_{l/r}} \gg \frac{1}{\Delta x_{l/r}}$, hence

$$\begin{aligned}\tau_I &\approx \frac{1}{2c}\frac{e^{\frac{c}{K}\Delta x_r}}{\frac{1}{\Delta x_r}e^{\frac{c}{K}\Delta x_r}} \\ &\approx \frac{\Delta x_r}{2c}.\end{aligned}$$

Thus, the stabilization relies most importantly on the *downstream node*. In cases, where stabilization is needed, the distance of the downstream node is of high importance, whereas the upstream node plays almost no role. Although the nodally exact solution cannot be obtained after these simplifications, the stabilization based only on the downstream node is still very successful.

The situation for element stabilization —in regular or irregular node distributions— can be estimated as follows

$$\begin{aligned}\tau_e &= \frac{\Delta x_e}{2c}\left(\coth\left(Pe_e\right) - \frac{1}{Pe_e}\right)\\ &= \frac{\Delta x_e}{2c}.\end{aligned}$$

The contribution of the stabilization terms to the overall system matrix with element and nodal stabilization are compared using the estimated results:

$$\begin{aligned}\sum^{n_{el}} \int_{\Omega_e} \tau_I c^2 \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} d\Omega &= \frac{1}{2}c\left[\begin{array}{ccccc}\cdots & -\frac{\Delta x_r}{\Delta x_l} & \frac{\Delta x_r}{\Delta x_l}+\frac{\Delta x_r}{\Delta x_r} & -\frac{\Delta x_r}{\Delta x_r} & \cdots\end{array}\right]\\ &= \frac{1}{2}c\left[\begin{array}{ccccc}\cdots & -\frac{\Delta x_r}{\Delta x_l} & \frac{\Delta x_r}{\Delta x_l}+1 & -1 & \cdots\end{array}\right],\\ \sum^{n_{el}} \int_{\Omega_e} \tau_e c^2 \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} d\Omega &= \frac{1}{2}c\left[\begin{array}{ccccc}\cdots & -\frac{\Delta x_l}{\Delta x_l} & \frac{\Delta x_l}{\Delta x_l}+\frac{\Delta x_r}{\Delta x_r} & -\frac{\Delta x_r}{\Delta x_r} & \cdots\end{array}\right]\\ &= \frac{1}{2}c\left[\begin{array}{ccccc}\cdots & -1 & 2 & -1 & \cdots\end{array}\right].\end{aligned}$$

Again, with $\Delta x_l = \Delta x_r$ the equality of nodal and element stabilization becomes obvious. However, with $\Delta x_l \neq \Delta x_r$ only the entry -1 of the stabilization term which belongs to the downstream node is equal for both stabilizations. They are different for the node itself and for the upstream node.

Remark 1 Although element stabilization cannot be deduced in a mathematically consistent way from nodal stabilization —especially not in multi-dimensional cases as shown in subsection 4.2—, results are also very satisfactory. We believe that this stems from the fact that the relevant downstream node is stabilized "correctly", i.e. this node is stabilized equally with nodal and element stabilization in case of the estimated results for the stabilization parameter. The fact that other entries of the stabilization term matrix do not agree, seems to be of less importance. This aspect is confirmed with later results.

4.1.8 τ for Quadratic Elements

In this section it is shown that it is also possible to obtain optimal stabilization parameters τ to obtain nodally exact solutions with quadratic elements (and any other). Only an outline of the deduction is given, the procedure is exactly the same as in subsection 4.1.3 for linear elements. Again, a regular node distribution is assumed. Starting point is

$$\tau_I = -\frac{\left[c\int_\Omega w_I \frac{\partial \mathbf{N}^T}{\partial x} + K\int_\Omega \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} - K\oint_\Gamma w_I \frac{\partial \mathbf{N}^T}{\partial x}\right]\mathbf{u}^{\mathrm{ex}}}{\left[\sum_{e=1}^{n_{el}} c^2 \int_{\Omega_e}\frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} - cK\int_{\Omega_e}\frac{\partial w_I}{\partial x}\frac{\partial^2 \mathbf{N}^T}{\partial x^2}\right]\mathbf{u}^{\mathrm{ex}}},$$

which was an intermediate result of subsection 4.1.3. Note that for quadratic elements the system of equations has *two* different difference equations instead of only one for linear elements, see Figure 10. This is because there are 3×3 element matrices and there is one difference equation I_a which has a 5 node stencil and another I_b with gives a three node stencil only. Clearly, each of

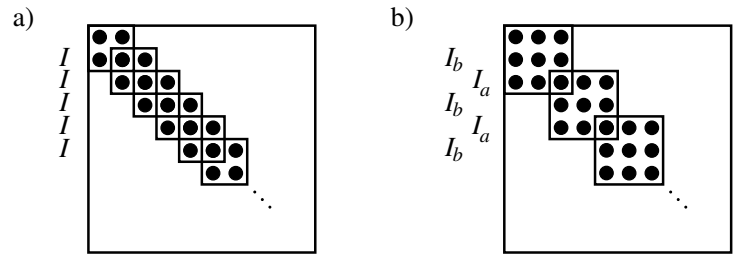


Figure 10: Comparison of different matrix structures for a) linear and b) quadratic elements. One can see that in case of quadratic elements two different difference stencils I_a and I_b arise (for a regular node distribution), whereas linear elements have only one difference stencil I, being the same for all nodes.

the two equations requires an individual τ.

Evaluating the above expression for equation I_a and inserting the exact solution evaluated at the nodes $\mathbf{u}^{\text{ex}}$ leads to

$$\begin{aligned}
\tau_{I_a} &= -\frac{\left(c\left[\ldots \;\; \frac{1}{6} \;\; -\frac{2}{3} \;\; 0 \;\; \frac{2}{3} \;\; -\frac{1}{6} \;\; \ldots\right] + \frac{K}{\Delta x}\left[\ldots \;\; \frac{1}{3} \;\; -\frac{8}{3} \;\; \frac{14}{3} \;\; -\frac{8}{3} \;\; \frac{1}{3} \;\; \ldots\right]\right)\mathbf{u}^{\text{ex}}}{\left(\frac{c^2}{\Delta x}\left[\ldots \;\; \frac{1}{3} \;\; -\frac{8}{3} \;\; \frac{14}{3} \;\; -\frac{8}{3} \;\; \frac{1}{3} \;\; \ldots\right] - \frac{cK}{\Delta x^2}\left[\ldots \;\; 4 \;\; -8 \;\; 0 \;\; 8 \;\; -4 \;\; \ldots\right]\right)\mathbf{u}^{\text{ex}}} \\
&= \vdots \\
&= \frac{\Delta x}{2c}\frac{\frac{2}{3}\sinh(2Pe) - \frac{8}{3}\sinh(Pe) - \frac{1}{Pe}\left[\frac{2}{3}\cosh(2Pe) - \frac{16}{3}\cosh(Pe) + \frac{14}{3}\right]}{\frac{2}{3}\cosh(2Pe) - \frac{16}{3}\cosh(Pe) + \frac{14}{3} - \frac{1}{Pe}\left[-4\sinh(2Pe) + 8\sinh(Pe)\right]}.
\end{aligned}$$

The same can be done for equation I_b

$$\begin{aligned}
\tau_{I_b} &= -\frac{\left(c\left[\ldots \;\; -\frac{2}{3} \;\; 0 \;\; \frac{2}{3} \;\; \ldots\right] + \frac{K}{\Delta x}\left[\ldots \;\; -\frac{8}{3} \;\; \frac{16}{3} \;\; -\frac{8}{3} \;\; \ldots\right]\right)\mathbf{u}^{\text{ex}}}{\left(\frac{c^2}{\Delta x}\left[\ldots \;\; -\frac{8}{3} \;\; \frac{16}{3} \;\; -\frac{8}{3} \;\; \ldots\right] - \frac{cK}{\Delta x^2}\left[\ldots \;\; 0 \;\; 0 \;\; 0 \;\; \ldots\right]\right)\mathbf{u}^{\text{ex}}} \\
&= \vdots \\
&= \frac{\Delta x}{2c}\frac{-\frac{8}{3}\sinh(Pe) - \frac{1}{Pe}\left[-\frac{16}{3}\cosh(Pe) + \frac{16}{3}\right]}{-\frac{16}{3}\cosh(Pe) + \frac{16}{3}}.
\end{aligned}$$

Applying these two τ_I definitions leads to nodally exact solutions as can be seen from the left part of Figure 11. In the right part the two definitions are compared with the coth-version for linear FEM. Most importantly, it is found that there are two different limits of τ_{I_a} and τ_{I_b}. Consequently, choosing only *one* τ for the stabilization seems inadequate.

Some conclusions for element stabilization with τ_e are possible. Having one τ_e for each element matrix does also not consider the two different limits of the two different types I_a and I_b of equations. However, in practice, this is still standard, see e.g. [50], where it is pointed out that for quadratic elements τ_e may be multiplied by one half. Looking at the two limits in Figure 11, it becomes clear, why this particular value may be chosen. However, a treatment of the element

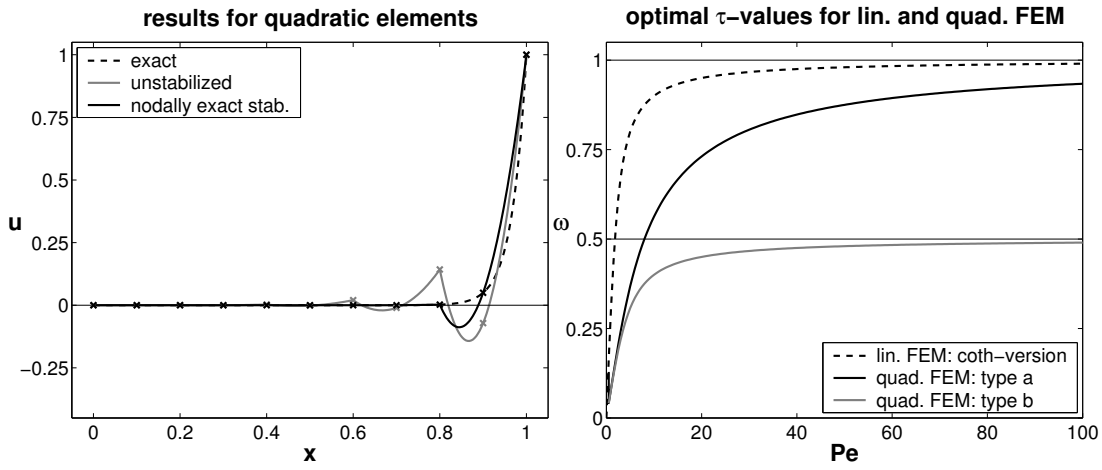


Figure 11: a) Nodally exact results with quadratic elements. b) Comparison of the different $\tau = \frac{\Delta x}{2c}\omega$ versions of linear and quadratic elements. The limits for $Pe \longrightarrow \infty$ are different for τ_{I_a} and τ_{I_b}.

equations as

$$\text{standard: } \tau_e \begin{pmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{pmatrix} \qquad \text{proposed: } \begin{pmatrix} \tau_{e_a} \begin{pmatrix} \times & \times & \times \end{pmatrix} \\ \tau_{e_b} \begin{pmatrix} \times & \times & \times \end{pmatrix} \\ \tau_{e_a} \begin{pmatrix} \times & \times & \times \end{pmatrix} \end{pmatrix}$$

seems more adequately, because then the different limits can be considered respectively. The advantage of this proposal can also be verified with numerical experiments. τ_{e_a} and τ_{e_b} are chosen equivalently to τ_{I_a} and τ_{I_b} by replacing Δx and Pe with the corresponding element numbers. This gives in case of a regular node distribution also for element stabilization nodally exact values. To the authors knowledge, this is the first nodally exact scheme (using standard element stabilization) which is directly applicable for higher order elements. Generalization to other higher-order elements in one dimension is straightforward. Clearly, only for nodal exactness a choice of τ_{e_a} and τ_{e_b} according to the above complicated formulas is necessary. But an obvious simplification would be to choose

$$\tau_{e_a} = \frac{\Delta x_e}{2c}\left(\coth\left(\frac{c\Delta x_e}{2K}\right) - \frac{2K}{c\Delta x_e}\right) \qquad \tau_{e_b} = \frac{1}{2}\tau_{e_a},$$

which approximates the exact τ_{e_a} and τ_{e_b} reasonably taking into account the different limits.

Thus, it can be seen that one can find some valuable proposals from nodal stabilization for element stabilization.

4.1.9 Minimization of Norms

It may seem arbitrary to call the stabilization parameter τ_I "optimal" when the approximation finds the nodally exact solution. The quality of an approximation is usually measured by norms. It depends on the choice of a particular norm whether one may call an approximation "optimal" or not.

Let us try to find the "optimal" stabilization parameter τ_I with respect to the approximation error in the L_2 norm

$$\|\tilde{u} - u\|_{L_2} \longrightarrow \min.$$

Inserting the ansatz $\tilde{u} = \mathbf{N}^T\mathbf{u}$ and differentiating for $\mathbf{u}$ leads to the following system of equations

$$\begin{aligned} \int_\Omega \mathbf{N}\mathbf{N}^T d\Omega \mathbf{u} &= \int_\Omega \mathbf{N} u^{\text{ex}} d\Omega \\ \mathbf{M}_{L_2}\mathbf{u} &= \mathbf{b}. \end{aligned}$$

Consequently, to compute the "optimal" stabilization parameter τ_I with respect to the L_2 norm, $\mathbf{u} = \mathbf{M}_{L_2}^{-1}\mathbf{b}$ has to be inserted in the expression for τ_I (instead of $\mathbf{u} = \mathbf{u}^{\text{ex}}$ for the nodally exact solution). Then, due to the matrix $\mathbf{M}_{L_2}$, the resulting expression depends on the position of all nodes, instead of only the neighboring nodes.

The conclusion is that determining "optimal" τ_I with respect to certain norms leads in general to a global stabilization criterion, whereas an "optimal" τ_I leading to a nodally exact solution leads only to a local stabilization criterion.

One may ask, which particular norm is minimized in case of a nodally exact solution? For general interpolations, this question may not be answered, but in case of linear FEM, the H_1 seminorm gives exactly the desired result. It follows that

$$\int_\Omega \frac{\partial \mathbf{N}}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} d\Omega \mathbf{u} = \int_\Omega \frac{\partial \mathbf{N}}{\partial x} u^{\text{ex}} d\Omega$$

and the nodally exact solution $\mathbf{u}^{\text{ex}}$ is the result. Then, the deduction of τ presented in subsection 4.1.3 does not change and leads for a regular node distribution to the coth-formula.

4.2 Two-dimensional Advection-Diffusion Equation

4.2.1 Modification of the Weak Form

The 2D advection-diffusion equation is considered in the following to extract important features for multi-dimensional problems. In strong form the equation is

$$c_i \frac{\partial u}{\partial x_i} - K_{ij}\frac{\partial^2 u}{\partial x_i \partial x_j} = 0,$$

where the summation convention has been used with $i, j \leq 2$. Boundary conditions are not considered for simplicity. Analogously to subsection 4.1.2 artificial diffusion $\widetilde{K}$ is introduced at this step. The important feature of the SUPG stabilization is that it introduces artificial diffusion only in streamline direction. Perpendicular to the flow direction no diffusion will be added. This property is realized with an artificial diffusion tensor $\widetilde{K}_{ij}$ defined as

$$\widetilde{K}_{ij} = \widetilde{K}\frac{c_i c_j}{c_k c_k}.$$

Remark 2 It is not trivial to realize that this structure gives the desired result. Therefore, a two-dimensional example is considered here, with an advection direction of $\tan\alpha = c_x/c_y$ as shown in Figure 12a). Artificial diffusion shall be introduced in streamline direction only, hence

$\widetilde{\mathbf{K}}_{\xi\eta} = \begin{pmatrix} \widetilde{K} & 0 \\ 0 & 0 \end{pmatrix}$. To obtain the artificial diffusion tensor in the xy-coordinate system a standard tensor transformation is applied, i.e.

$$\begin{aligned}
\widetilde{K}_{ij} = \widetilde{\mathbf{K}}^{xy} &= \mathbf{T}^T \widetilde{\mathbf{K}}^{\xi\eta} \mathbf{T} \quad \text{with } \mathbf{T} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \\
&= \widetilde{K} \begin{pmatrix} \cos^2\alpha & \sin\alpha\cos\alpha \\ \cos\alpha\sin\alpha & \sin^2\alpha \end{pmatrix} \\
&= \frac{\widetilde{K}}{c_x^2 + c_y^2} \begin{pmatrix} c_x^2 & c_x c_y \\ c_y c_x & c_y^2 \end{pmatrix} \\
&= \widetilde{K} \frac{c_i c_j}{c_k c_k} \quad \text{with } i,j,k \leq n_{sd} = 2.
\end{aligned}$$

The weak form of the problem becomes

$$\begin{aligned}
\int_\Omega \mathbf{w} \left(c_i \frac{\partial \widetilde{u}}{\partial x_i} - \left(K_{ij} + \widetilde{K}_{ij} \right) \frac{\partial^2 \widetilde{u}}{\partial x_i \partial x_j} \right) d\Omega &= \mathbf{0} \\
\left[\int_\Omega c_i \mathbf{w} \frac{\partial \mathbf{N}^T}{\partial x_i} - \widetilde{K}_{ij} \mathbf{w} \frac{\partial^2 \mathbf{N}^T}{\partial x_i \partial x_j} - K_{ij} \mathbf{w} \frac{\partial^2 \mathbf{N}^T}{\partial x_i \partial x_j} d\Omega \right] \mathbf{u} &= \mathbf{0} \\
\left[\int_\Omega c_i \mathbf{w} \frac{\partial \mathbf{N}^T}{\partial x_i} d\Omega + \int_\Omega \widetilde{K}_{ij} \frac{\partial \mathbf{w}}{\partial x_j} \frac{\partial \mathbf{N}^T}{\partial x_i} d\Omega + \int_\Omega K_{ij} \frac{\partial \mathbf{w}}{\partial x_j} \frac{\partial \mathbf{N}^T}{\partial x_i} d\Omega - \oint_\Gamma K_{ij} \mathbf{w} \frac{\partial \mathbf{N}^T}{\partial x_i} n_j d\Gamma \right] \mathbf{u} &= \mathbf{0}.
\end{aligned}$$

In the last step, the divergence theorem is applied on the second order terms. The same remarks from section 4.1 about the negligence of the artificial diffusion boundary term and the effect only in the element interiors apply here as well.

Factorizing the first two terms and inserting the definition of $\widetilde{K}_{ij}$ gives

$$\begin{aligned}
\left[\int_\Omega \left(c_i \mathbf{w} + \widetilde{K}_{ij} \frac{\partial \mathbf{w}}{\partial x_j} \right) \left(\frac{\partial \mathbf{N}^T}{\partial x_i} \right) d\Omega + \ldots \right] \mathbf{u} &= \mathbf{0} \\
\left[\int_\Omega \left(c_i \mathbf{w} + \widetilde{K} \frac{c_i c_j}{c_k c_k} \frac{\partial \mathbf{w}}{\partial x_j} \right) \left(\frac{\partial \mathbf{N}^T}{\partial x_i} \right) d\Omega + \ldots \right] \mathbf{u} &= \mathbf{0} \\
\left[\int_\Omega \left(\mathbf{w} + \frac{\widetilde{K}}{c_k c_k} c_j \frac{\partial \mathbf{w}}{\partial x_j} \right) \left(c_i \frac{\partial \mathbf{N}^T}{\partial x_i} \right) d\Omega + \ldots \right] \mathbf{u} &= \mathbf{0} \\
\left[\int_\Omega \left(\mathbf{w} + \tau c_j \frac{\partial \mathbf{w}}{\partial x_j} \right) \left(c_i \frac{\partial \mathbf{N}^T}{\partial x_i} \right) d\Omega + \ldots \right] \mathbf{u} &= \mathbf{0}.
\end{aligned}$$

Note that the summation convention is also applied to indices which occur more often than twice. The term $\frac{\widetilde{K}}{c_k c_k}$ is renamed with the stabilization parameter τ. Again, the artificial diffusion tensor can be interpreted as a modification of the test function of the advection term. Prescribing consistency of the weak form requires the modification of all other test function as well, hence

$$\begin{aligned}
\Bigg[\int_\Omega \left(\mathbf{w} + \tau c_j \frac{\partial \mathbf{w}}{\partial x_j} \right) c_i \frac{\partial \mathbf{N}^T}{\partial x_i} d\Omega + \int_\Omega K_{ij} \frac{\partial}{\partial x_j} \left(\mathbf{w} + \tau c_j \frac{\partial \mathbf{w}}{\partial x_j} \right) \frac{\partial \mathbf{N}^T}{\partial x_i} d\Omega \\
- \oint_\Gamma K_{ij} \left(\mathbf{w} + \tau c_j \frac{\partial \mathbf{w}}{\partial x_j} \right) \frac{\partial \mathbf{N}^T}{\partial x_i} n_j d\Gamma \Bigg] \mathbf{u} &= \mathbf{0}.
\end{aligned}$$

Re-application of the divergence theorem finally gives

$$\int_\Omega \left(\mathbf{w} + \tau c_j \frac{\partial \mathbf{w}}{\partial x_j}\right)\left(c_i \frac{\partial \widetilde{u}}{\partial x_i} - K_{ij}\frac{\partial^2 \widetilde{u}}{\partial x_i \partial x_j}\right) d\Omega = \mathbf{0}.$$

This is the SUPG stabilized weak form of the advection-diffusion equation (which also holds for the 3D case where $i, j \leq 3$). More precisely, considering the effect of the stabilizing terms only in the element interiors,

$$\int_\Omega \mathbf{w}\left(c_i \frac{\partial \widetilde{u}}{\partial x_i} - K_{ij}\frac{\partial^2 \widetilde{u}}{\partial x_i \partial x_j}\right) d\Omega + \sum_{e=1}^{n_{el}} \int_{\Omega_e} \left(\tau c_j \frac{\partial \mathbf{w}}{\partial x_j}\right)\left(c_i \frac{\partial \widetilde{u}}{\partial x_i} - K_{ij}\frac{\partial^2 \widetilde{u}}{\partial x_i \partial x_j}\right) d\Omega = \mathbf{0}.$$

4.2.2 Weighting the Modification: Stabilization Parameter τ

The deduction of the modification of the weak form from the previous subsection was done for multi-dimensions and arbitrary constant diffusion tensors K_{ij}. The determination of τ is done for the 2D case with $K_{ij} = K\delta_{ij}$ only. The strong and stabilized weak form become

$$\begin{aligned} c_x \frac{\partial u}{\partial x} + c_y \frac{\partial u}{\partial y} - K\frac{\partial^2 u}{\partial x^2} - K\frac{\partial^2 u}{\partial y^2} &= 0, \\ \int_\Omega \mathbf{w}\left(c_x \frac{\partial \widetilde{u}}{\partial x} + c_y \frac{\partial \widetilde{u}}{\partial y} - K\frac{\partial^2 \widetilde{u}}{\partial x^2} - K\frac{\partial^2 \widetilde{u}}{\partial y^2}\right) d\Omega + & \\ \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau\left(c_x \frac{\partial \mathbf{w}}{\partial x} + c_y \frac{\partial \mathbf{w}}{\partial y}\right)\left(c_x \frac{\partial \widetilde{u}}{\partial x} + c_y \frac{\partial \widetilde{u}}{\partial y} - K\frac{\partial^2 \widetilde{u}}{\partial x^2} - K\frac{\partial^2 \widetilde{u}}{\partial y^2}\right) d\Omega &= \mathbf{0}. \end{aligned}$$

With the same methodology of subsection 4.1.3 the stabilization parameter is computed in the following. Assume, the boundary conditions allow the exact solution to be of the kind

$$u^{\text{ex}}(x, y) = C_1 e^{\frac{c_x}{K}u + \frac{c_y}{K}v} + C_2.$$

Then, for nodally exact solutions, the formula for the nodal stabilization parameter τ_I becomes

$$\tau_I = -\frac{\left[\int_\Omega w_I \left(c_x \frac{\partial \mathbf{N}^T}{\partial x} + c_y \frac{\partial \mathbf{N}^T}{\partial y} - K\frac{\partial^2 \mathbf{N}^T}{\partial x^2} - K\frac{\partial^2 \mathbf{N}^T}{\partial y^2}\right) d\Omega\right] \mathbf{u}^{\text{ex}}}{\left[\sum_{e=1}^{n_{el}} \int_{\Omega_e} \left(c_x \frac{\partial w_I}{\partial x} + c_y \frac{\partial w_I}{\partial y}\right)\left(c_x \frac{\partial \mathbf{N}^T}{\partial x} + c_y \frac{\partial \mathbf{N}^T}{\partial y} - K\frac{\partial^2 \mathbf{N}^T}{\partial x^2} - K\frac{\partial^2 \mathbf{N}^T}{\partial y^2}\right) d\Omega\right] \mathbf{u}^{\text{ex}}}.$$

The diffusion terms in the denominator cancel out for linear FEM, whereas in the nominator the divergence theorem is applied to these terms:

$$\tau_I = -\frac{\left[\int_\Omega c_x w_I \frac{\partial \mathbf{N}^T}{\partial x} + c_y w_I \frac{\partial \mathbf{N}^T}{\partial y} + K\frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} + K\frac{\partial w_I}{\partial y}\frac{\partial \mathbf{N}^T}{\partial y} d\Omega\right] \mathbf{u}^{\text{ex}}}{\left[\sum_{e=1}^{n_{el}} \int_{\Omega_e} c_x^2 \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial x} + c_x c_y \frac{\partial w_I}{\partial x}\frac{\partial \mathbf{N}^T}{\partial y} + c_y c_x \frac{\partial w_I}{\partial y}\frac{\partial \mathbf{N}^T}{\partial x} + c_y^2 \frac{\partial w_I}{\partial y}\frac{\partial \mathbf{N}^T}{\partial y} d\Omega\right] \mathbf{u}^{\text{ex}}}.$$

Evaluating this expression for each node results in nodally exact approximations.

A rotation of the coordinate system, where the main direction ξ is rotated in streamline

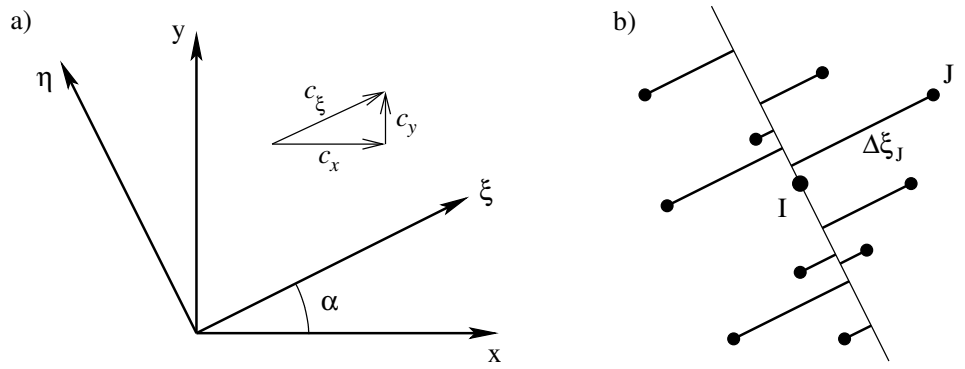


Figure 12: a) Shift of the coordinate system in streamline direction, b) distances in streamline direction $\Delta\xi_i$ between nodes i and node I.

direction as shown in Figure 12a) gives

$$\tau_I = -\frac{\left[\int_\Omega c_\xi w_I \frac{\partial \mathbf{N}^T}{\partial \xi} + K\frac{\partial w_I}{\partial \xi}\frac{\partial \mathbf{N}^T}{\partial \xi} + K\frac{\partial w_I}{\partial \eta}\frac{\partial \mathbf{N}^T}{\partial \eta} d\Omega\right]\mathbf{u}^{\mathrm{ex},\xi\eta}}{\left[\sum_{e=1}^{n_{el}}\int_{\Omega_e} c_\xi^2 \frac{\partial w_I}{\partial \xi}\frac{\partial \mathbf{N}^T}{\partial \xi} d\Omega\right]\mathbf{u}^{\mathrm{ex},\xi\eta}},$$

with $c_\xi = \sqrt{c_x^2 + c_y^2}$ and $c_\eta = 0$. Also $\mathbf{u}^{\mathrm{ex}}$ has been transformed into the new coordinate system, leading to $\mathbf{u}^{\mathrm{ex},\xi\eta}$, which is defined as

$$\mathbf{u}^{\mathrm{ex},\xi\eta} = C_1 e^{\frac{c_\xi}{K}\boldsymbol{\xi}} + C_2,$$

where $\boldsymbol{\xi}$ is the distance of the nodes in streamline direction. This becomes clear from Figure 12b).

Note that the obtained expression for τ_I is almost identical with the one-dimensional case. The only term, which does not rely on dimension ξ is the diffusion term $K\frac{\partial w_I}{\partial \eta}\frac{\partial \mathbf{N}^T}{\partial \eta}$, but with the assumption of advection-domination, it can be neglected. It can be seen that this equation is only dependent on the relative distances $\Delta\xi$ of the nodes around node I, i.e. nodes that are inside the same elements than node I. Only then, the evaluation of the scalar products in the nominator and denominator of the τ_I-formula lead to non-zero contributions. Note that in the 1D case, the neighboring nodes with a contribution to the scalar product are only the left and right node of node I. This can be seen from Figure 13. Therefore, it is clear that the reduction to the well-known coth-formula is not possible in this 2D case.

4.2.3 Element vs. Nodal Stabilization

In practice, for the stabilization of the 2D advection-diffusion equation —and other differential equations— with FEM, element stabilization is preferred over nodal stabilization. It is defined for this 2D case as

$$\tau_e = \frac{h_e}{2c_\xi}\left(\coth\left(\frac{c_\xi h_e}{2K}\right) - \frac{2K}{c_\xi h_e}\right),$$

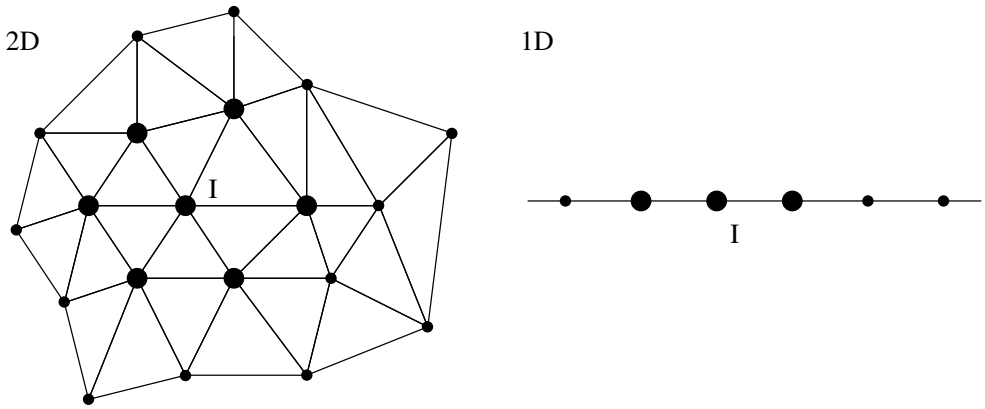


Figure 13: The large nodes contribute to the scalar products in the τ_I-formulas in one and two dimensions (for linear FEM).

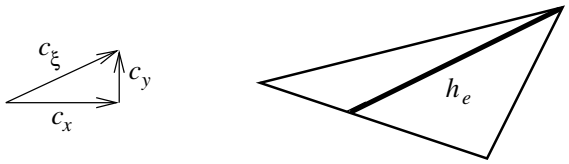


Figure 14: Element length h_e of a triangular element.

$$
\begin{aligned}
h_e &= 2\left(\sum_{\alpha}^{n_e}\left|\frac{c_i}{c_\xi}\frac{\partial N_\alpha}{\partial x_i}\right|\right)^{-1}, \\
c_\xi &= \|c_i\| = \sqrt{\sum_i c_i^2}.
\end{aligned}
$$

The parameter h_e can be interpreted as the element length in streamline direction as can be seen in Figure 14. Thus, in multi-dimensions, the stabilization of elements is still closely related to the one-dimensional case.

4.2.4 Relevant Downstream Node

It turns out that for the assumption of advection-domination and "coarse" meshes, i.e. for cases when stabilization is required, the scalar products of the τ_I-formula of subsection 4.2.2 are dominated by a certain node. The contribution to the scalar product of this node is much larger than all the other influences, hence τ_I relies most importantly on it. This node J can be identified by $\Delta\xi_J > \Delta\xi_i \ \forall i \in Q_I,\ i \neq J$, with Q_I being the set of all nodes that are inside the same elements than node I. That means node J is the neighboring node with the maximum distance in downstream direction from node I. The reason why this node dominates the scalar product is because

$$e^{\frac{c_\xi}{K}\Delta\xi_J} \gg e^{\frac{c_\xi}{K}\Delta\xi_i},\ \forall i \in Q_I,\ i \neq J$$

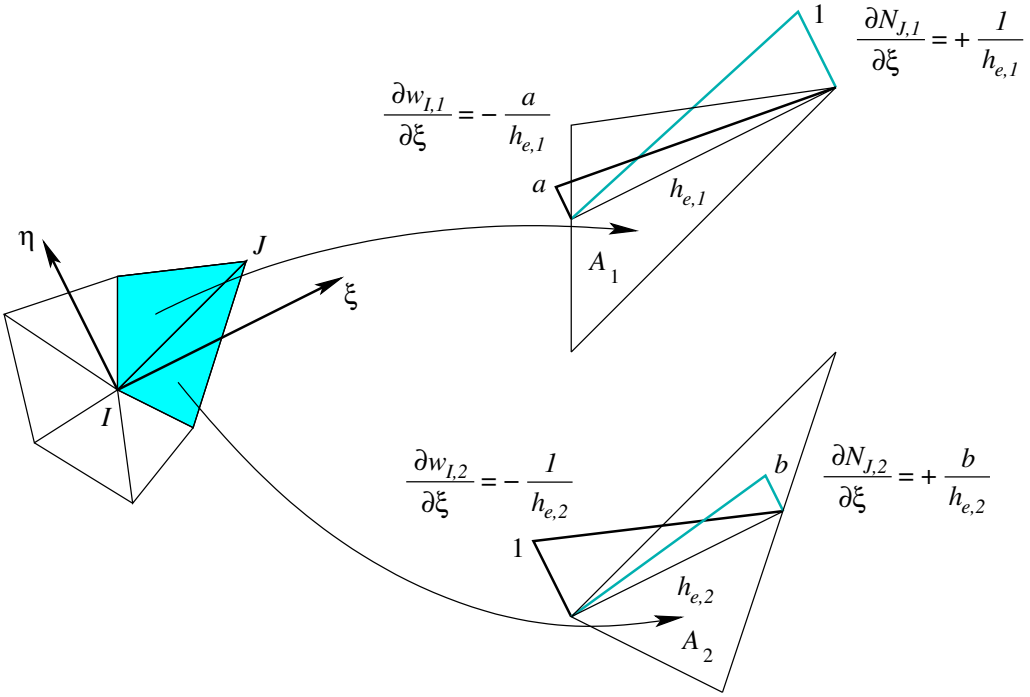


Figure 15: Element situation around node I. The integrals of the elements belonging to the relevant downstream node J can be evaluated to obtain an estimate for τ_I.

also for $\Delta\xi_J$ being only slightly larger than other $\Delta\xi_i$, because $\frac{c_\xi}{K}$ is large. It is also obvious that in the rare case where n nodes have the same maximum distance from node I, then n terms of the scalar product are relevant. Node J is called *relevant downstream node*. Assuming there exists such a node, τ_I can be estimated as

$$\begin{aligned}\tau_I &\approx -\frac{\int_\Omega c_\xi w_I \frac{\partial N_J}{\partial \xi} + K \frac{\partial w_I}{\partial \xi}\frac{\partial N_J}{\partial \xi} + K \frac{\partial w_I}{\partial \eta}\frac{\partial N_J}{\partial \eta} d\Omega}{\sum_{e=1}^{n_{el}} \int_{\Omega_e} c_\xi^2 \frac{\partial w_I}{\partial \xi}\frac{\partial N_J}{\partial \xi} d\Omega} \\ &\approx -\frac{\int_\Omega w_I \frac{\partial N_J}{\partial \xi} d\Omega}{c_\xi \int_\Omega \frac{\partial w_I}{\partial \xi}\frac{\partial N_J}{\partial \xi} d\Omega}.\end{aligned}$$

This integration can be evaluated by considering an element situation as shown in Figure 15:

$$\begin{aligned}\tau_I &= -\frac{1}{c_\xi}\frac{\int_{\Omega_{e,1}} w_I \frac{\partial N_J}{\partial \xi} d\Omega + \int_{\Omega_{e,2}} w_I \frac{\partial N_J}{\partial \xi} d\Omega}{\int_{\Omega_{e,1}} \frac{\partial w_I}{\partial \xi}\frac{\partial N_J}{\partial \xi} d\Omega + \int_{\Omega_{e,2}} \frac{\partial w_I}{\partial \xi}\frac{\partial N_J}{\partial \xi} d\Omega} \\ &= -\frac{1}{c_\xi}\frac{\frac{1}{3}A_1\frac{1}{h_{e,1}} + \frac{1}{3}A_2\frac{b}{h_{e,2}}}{-A_1\frac{a}{h_{e,1}}\frac{1}{h_{e,1}} - A_2\frac{1}{h_{e,2}}\frac{b}{h_{e,2}}}.\end{aligned}$$

When comparing nodal and element stabilization, there is —in contrast to the 1D situation— no way to obtain equality between the resulting final equations I. For nodal stabilization, there is only one τ_I for the whole equation I. But there is no element situation with the same element stabilization parameter τ_e for all elements surrounding node I (and which, in addition, equals

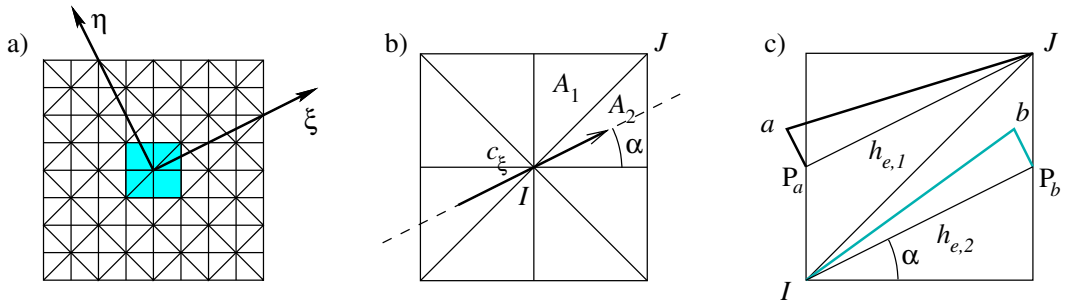


Figure 16: For a triangular mesh as shown in a) very good results can be expected for a flow direction of $\alpha = 26.565$ or $\alpha = 63.435$, because then $A_1 = A_2$, $h_{e,1} = h_{e,2}$, $a = w_I(\mathrm{P}_a) = 1/2$ and $b = N_J(\mathrm{P}_b) = 1/2$, hence $\tau_{e,1} = \tau_{e,2} = \tau_I$ for large advection-diffusion ratios.

τ_I). However, note that the stabilization parameter τ_I is related to the element stabilization parameter τ_e for the element containing the relevant downstream node J. The estimates for the element stabilization parameters containing the relevant downstream node J are

$$\tau_{e,1} = \frac{h_{e,1}}{2c_\xi} \qquad \tau_{e,2} = \frac{h_{e,2}}{2c_\xi}.$$

Remark 3 It is possible to construct a particular element and flow situation, where $\tau_{e,1} = \tau_{e,2} = \tau_I$, i.e. where the nodal stabilization parameter equals the element stabilization parameters belonging to the element containing the relevant downstream node. The matrix entry (I, J) of the stabilization matrix is then equal in both cases. For this element situation $h_{e,1} = h_{e,2}$, $A_1 = A_2$, $a = b = 1/2$ is required, then it is obvious that $\tau_{e,1} = \tau_{e,2} = \tau_I$. In Figure 16 such an element and flow situation is shown. $\tau_{e,1}$, $\tau_{e,2}$ and τ_I can be calculated as $\frac{\sqrt{5}}{16}/c_\xi$ for this situation.

Remark 4 Assuming that the stabilization of the relevant downstream node is of highest importance, then it may be predicted that using the element mesh as in Figure 16a) particularly good results can be obtained for a flow direction of $\alpha = \arctan(0.5) = 26.565$ and $\alpha = 90 - \arctan(0.5) = 63.435$. Figure 17 proves that this assumption is correct. In this figure, the nodal error norm $\varepsilon = \sqrt{\sum_{i=1}^{n} (\widetilde{u}_i - u(x_i))^2}$ is plotted in dependence of the flow direction. It can be seen that for growing advection-diffusion ratios a clear minimum can be found converging to the predicted flow directions.

Consequently, the stabilization of the relevant downstream node for convection-dominated cases is of high importance, whereas the other nodes are not so important. This is the reason why element stabilization works so successful although it can not be motivated mathematically from the nodal stabilization.

4.3 Non-linear Model Equations

In the previous subsections it was shown how to prevent oscillatory solutions of the linear advection-diffusion equation, with help of stabilized weak forms. The idea to add artificial diffusion to the problem and to maintain consistency leads for all differential equations to a modification of the

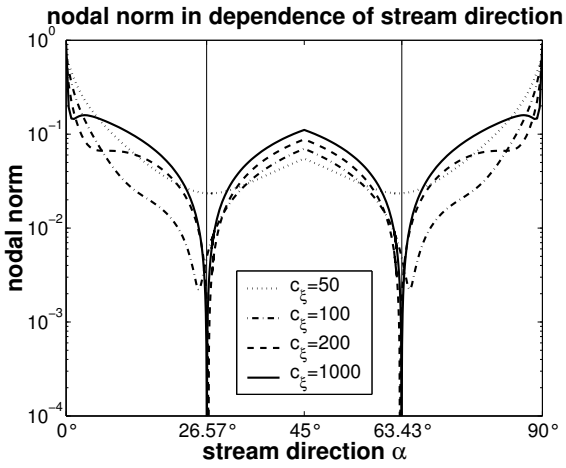


Figure 17: Dependence of the nodal error norm on the flow direction α. For growing advection-diffusion ratios and a flow direction of $\alpha = 26.565$ or $\alpha = 63.435$ the element stabilization of the elements belonging to the relevant downstream node comes closer to the optimal nodal stabilization, leading to super-convergence.

weak form of the kind

$$\int_\Omega \left(\mathbf{w} + \tau \mathcal{L}_{adv}\mathbf{w}\right)\left(\mathcal{L}\widetilde{u} - f\right) d\Omega = \mathbf{0}.$$

However, the deduction of the stabilization parameters τ has been performed with the knowledge of the exact solution of the advection-diffusion equation and is therefore problem-specific. For other differential equations the optimal stabilization parameter leading to nodally exact solutions is different depending on their exact solution.

In the following, the stabilization parameter shall be deduced for the one-dimensional nonlinear Burgers equation. Strong form and SUPG-stabilized weak form are

$$\begin{aligned} u\frac{\partial u}{\partial x} - K\frac{\partial^2 u}{\partial x^2} &= 0, \\ \int_\Omega \mathbf{w}\left(\widetilde{u}\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega + \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau\left(\widetilde{u}\frac{\partial \mathbf{w}}{\partial x}\right)\left(\widetilde{u}\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega &= \mathbf{0}. \end{aligned}$$

The exact solution is known as

$$u\left(x\right) = -2KC_1\frac{e^{2C_1 x} - C_2}{e^{2C_1 x} + C_2}.$$

Calculating the nodally exact stabilization parameter τ_I for this solution in case of a regular node distribution —analogously to subsection 4.1.3— and linear FEM gives after a large number

of modifications

$$\begin{aligned}\tau_I &= -\frac{\int_\Omega w_I \left(\widetilde{u}\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega}{\sum_{e=1}^{n_{el}} \int_{\Omega_e} \left(\widetilde{u}\frac{\partial w_I}{\partial x}\right)\left(\widetilde{u}\frac{\partial \widetilde{u}}{\partial x} - K\frac{\partial^2 \widetilde{u}}{\partial x^2}\right) d\Omega} \\ &= \vdots \\ &= -\frac{1}{6}\Delta x \frac{e^{2C_1 x} - C_2}{e^{2C_1 x} + C_2} \coth\left(C_1 \Delta x\right) - K.\end{aligned}$$

Note that this nodally exact stabilization parameter still depends on the constants C_1 and C_2. Thus, in contrast to the linear advection-diffusion case, there is a dependency of τ_I on the boundary conditions and the criterion is consequently *global*, instead of local. This can also be shown for the multi-dimensional Burgers equation.

The situation is even worse for differential equations where no exact solution is in general known, i.e. for the Navier-Stokes equations and many others. Then, there is obviously no way to obtain formulas for τ that result in the nodally exact solution.

We conclude that the advection-diffusion equation was a particularly well-suited model equation for the stabilization parameter, because one obtains a local stabilization criterion. In general, this cannot be found for other differential equations. Therefore, in practice, one uses very successfully the same stabilization formulas, which were obtained for the advection-diffusion equations, for the stabilization of any other differential equations that may show oscillatory behavior. One can only expect that this works satisfactory if it is locally justified to approximate the non-linear situation with the linear advection-diffusion equation.

4.4 Instationary Model Equations

At the beginning of this section it became apparent that first order advective terms require stabilization. First order derivatives, however, do not only occur in convective terms, but also in time-dependent terms $\frac{\partial u}{\partial t}$ of instationary problems. Then, in general the same considerations as for advective terms have to be made for the time term. Looking at the 2D stationary advection equation with its stabilized weak form

$$c_x \frac{\partial u}{\partial x} + c_y \frac{\partial u}{\partial y} = 0, \qquad \int_\Omega \left[\mathbf{w} + \tau\left(c_x \frac{\partial \mathbf{w}}{\partial x} + c_y \frac{\partial \mathbf{w}}{\partial y}\right)\right]\left(c_x \frac{\partial \tilde{u}}{\partial x} + c_y \frac{\partial \tilde{u}}{\partial y}\right) = \mathbf{0},$$

it is obvious that the 1D instationary advection problem must have a similar stabilized weak form of

$$\frac{\partial u}{\partial t} + c_x \frac{\partial u}{\partial x} = 0, \qquad \int_\Omega \left[\mathbf{w} + \tau\left(\frac{\partial \mathbf{w}}{\partial t} + c_x \frac{\partial \mathbf{w}}{\partial x}\right)\right]\left(\frac{\partial \tilde{u}}{\partial t} + c_x \frac{\partial \tilde{u}}{\partial x}\right) = \mathbf{0}.$$

Thus it can be seen that also a stabilization in time direction is necessary if linear FEM is also used in time direction, or if "centered" finite difference time stepping schemes are chosen. Clearly, for "upwind" (=Euler forward) time stepping schemes, i.e. where

$$\frac{\partial \tilde{u}\left(\mathbf{x}, t\right)}{\partial t} \approx \frac{\tilde{u}\left(\mathbf{x}, t_{n+1}\right) - \tilde{u}\left(\mathbf{x}, t_n\right)}{\Delta t},$$

no stabilization in time-direction is necessary because this scheme is already diffusive enough. In

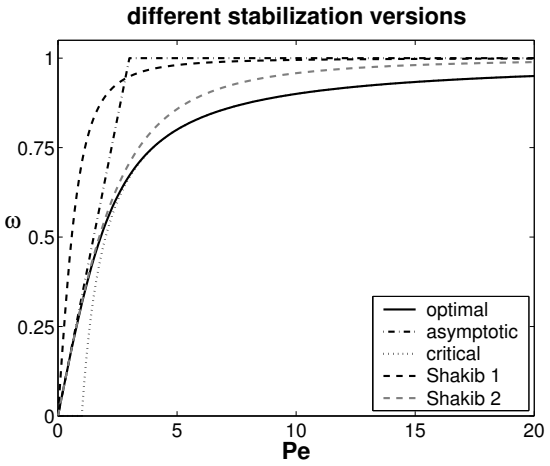


Figure 18: Alternative versions of the stabilization parameter $\tau = \frac{\Delta x}{2c}\omega$.

the following, the time derivative has to be taken into account for the operator $\mathcal{L}_{adv}$ whenever stabilization in time direction is necessary, i.e.

$$\mathcal{L}_{adv} = \frac{\partial}{\partial t} + c_i \frac{\partial}{\partial x_i}.$$

4.5 Alternative Versions of τ

In practice, alternative versions of the stabilization parameter are used instead of the "optimal" coth-version, which is due to the fact that they are less time-consuming to compute. They can be considered as approximations of the coth-formula and are shown in Figure 18. Instead of

$$\tau = \frac{\Delta x}{2c}\left(\coth\left(Pe\right) - \frac{1}{Pe}\right) = \frac{\Delta x}{2c}\omega,$$

only ω ("diffusion correction factor" [50]) is visualized as a function of the element Peclet number. Here, only the 1D stabilization case is considered.

- optimal version, first in [8]
$$\omega = \left(\coth\left(Pe\right) - \frac{1}{Pe}\right)$$

- doubly asymptotic approximation [24, 7]
$$\omega = \begin{cases} Pe/3, & -3 \leq Pe \leq 3 \\ \operatorname{sgn}\left(Pe\right), & Pe > 3 \end{cases}$$

- critical approximation [24, 8, 7]
$$\omega = \begin{cases} -1 - 1/Pe, & Pe < 1 \\ 0, & -1 \leq Pe \leq 1 \\ 1 - 1/Pe, & Pe > 1 \end{cases}$$

- versions of Shakib [50]

$$\omega_1 = \left(1 + \frac{1}{Pe^2}\right)^{-1/2} \quad \omega_2 = \left(1 + \frac{9}{Pe^2}\right)^{-1/2}$$

The first version of Shakib, ω_1, is maybe the one most often used in practice:

$$\tau = \frac{\Delta x}{2c}\omega_1 = \frac{\Delta x}{2c}\left(1 + \frac{1}{Pe^2}\right)^{-1/2} = \left[\left(\frac{2c}{\Delta x}\right)^2 + \left(\frac{4K}{\Delta x^2}\right)^2\right]^{-1/2}.$$

The two terms in the right expression can be interpreted as the advection-dominated and diffusion-dominated limit [54]. It can be seen that the dependency on the mesh size Δx in the advection-dominated case is $\tau \approx \frac{\Delta x}{2c}$, hence $O(\Delta x)$ (for equivalent estimates for the optimal version leading to the same result $\frac{\Delta x}{2c}$ see subsection 4.1.7), while it is in the diffusion-dominated case $\tau \approx \frac{\Delta x^2}{4K}$, hence $O\left(\Delta x^2\right)$. This aspect is re-considered in section 5.

The generalization of the Shakib version to element stabilization, multi-dimensions and instationary problems is straightforward:

$$\tau_e = \left[\left(\frac{2}{\Delta t}\right)^2 + \left(\frac{2\left\|\mathbf{c}\right\|}{h_e}\right)^2 + \left(\frac{4K}{{h_e}^2}\right)^2\right]^{-1/2}.$$

Note that for transient-dominated case, $\tau \approx \frac{\Delta t}{2}$ there is a relation of $O(\Delta t)$. This term can easily be explained as follows: Having an operator $\mathcal{L}_{adv} = c\frac{\partial}{\partial x}$, τ can be estimated by $\tau \approx \frac{\Delta x}{2c}$, hence having $\mathcal{L}_{adv} = \frac{\partial}{\partial t}$ gives $\tau \approx \frac{\Delta t}{2}$. In case of a mixed operator $\mathcal{L}_{adv} = \frac{\partial}{\partial t} + c_i\frac{\partial}{\partial x_i}$, then τ depends on the dominating term, i.e. in the transient-dominated case on $O(\Delta t)$. Thus, the formula of Shakib can be interpreted as a switch between estimates for advection, diffusion or transient dominated cases. Taking this viewpoint, other versions of the kind [54]

$$\tau_e = \left[\left(\frac{2}{\Delta t}\right)^r + \left(\frac{2\left\|c\right\|}{h_e}\right)^r + \left(\frac{4K}{{h_e}^2}\right)^r\right]^{-1/r}$$

can be motivated. Tezduyar *et al.* have a number of recent publications where further generalizations for the computation of stabilization parameters have been presented, see [54].

4.6 Summary

The results and conclusions of this section are summarized briefly:

- Finite difference context: Upwinding is first-order accurate, over-diffusive and monotone; central differences are second-order accurate, under-diffusive and oscillatory. Consequently, smart introduction of artificial diffusion is a possible way to obtain higher-order accurate and oscillation-free results.

- Finite element context: Introduce artificial diffusion to the weak form, interpret this as a modification of the test function of the convection term and enforce consistency by applying this modified test function to *all* terms in the weak form (this is what the SUPG stabilization does). Then, the term "artificial diffusion" is not applicable exactly any longer.

- The added stabilization terms are weighted with a stabilization parameter τ. The determination of τ is the critical aspect of stabilization schemes as it defines the "right" amount of the added artificial diffusion.

- For the one-dimensional advection-diffusion equation, approximated with linear FEM on a regular node distribution, the coth-formula for τ gives nodally exact results. There exist a number of alternative, similar versions of this formula.

- One has to separate nodal and element stabilization. The first requires one τ_I for the whole equation of a certain node, the second requires one τ_e for each element and then, several different τ-values are inside one equation. The choice, which of these stabilizations is chosen, is dependent on the numerical scheme: FEM works with element stabilization, for MMs, instead, nodal stabilization is required. However, also nodal stabilization can be applied to FEM, leading to indistinguishable results.

- The following dependency of τ on the node distribution holds: $\tau = O(\Delta x)$ for the convection-dominated limit and $\tau = O(\Delta x^2)$ for the diffusion-dominated limit. The latter has not been deduced above, but can be easily shown for the coth-formula by computing $\frac{\partial \omega(Pe)}{\partial Pe} \longrightarrow \frac{1}{3}$ for $Pe \longrightarrow 0$. Then, it is justified to approximate $\omega = \coth(Pe) - \frac{1}{Pe}$ at $Pe = 0$ by a linear function of $\omega = \frac{1}{3}Pe$, hence $\tau = \frac{\Delta x}{2c}\frac{1}{3}Pe = \frac{1}{12}\frac{\Delta x^2}{K} = O(\Delta x^2)$.

- The formulas obtained for τ in case of the simple, linear advection-diffusion models define local criteria, as they only rely on neighboring nodes and are independent of boundary conditions.

- Generalization to multi-dimensions and other problems is often straightforward, justified *a posteriori* by successful approximations.

5 Theoretical Approach: Error Analysis for the FEM

In this section, the aim is to outline some important aspects and results of mathematical analysis of stabilization methods in the FEM context. The illustrative understanding from the previous section is restricted to very simple situations and it was found that the artificial diffusion aspect can be used as a starting point, but does not fully describe the character of the resulting stabilizations. Mathematical analysis of convection-dominated and mixed problems gives a firm theoretical basis for stabilization methods and explains why Bubnov-Galerkin (and central differences) approaches lead to problems. However, it should be mentioned that mathematical analysis often fails in giving precise definitions for parameters —in case of stabilization most importantly for the stabilization parameters τ— and provides rather general conclusions. For example, one may show that a stabilization method converges for $\tau > 0$, however, it may not be shown which particular choice gives reliable results. Clearly, for "very small" τ results will be almost as oscillatory and useless as standard Bubnov-Galerkin methods.

In the literature, the stabilized schemes for convection-dominated problems are often summarized as *streamline-diffusion methods* [11, 17]. Then, there is at least a modification of the test function of $\mathbf{w} + \tau \mathcal{L}_{adv}\mathbf{w} = \mathbf{w} + \tau \mathbf{c} \cdot \nabla \mathbf{w}$, which is the case for the SUPG and GLS method.

5.1 General Remarks

Mathematical analysis of numerical methods applied to certain problems often has the aim to prove consistency, stability (or coercitivity) and convergence in certain norms. Local and global error estimates describe important features of the method, such as the convergence rate.

Frequently used tools in the mathematical analysis are inequalities from functional analysis, for example inverse estimates and Poincaré-Friedrichs inequalities. The first bounds higher derivatives by lower derivatives and hold only for finite-dimensional function spaces. The second bounds lower derivatives by higher ones, scaled by a domain-dependent coefficient.

Inequalities hold under certain conditions, often certain positive constants and mesh parameters, e.g. length of the longest element side, aspect ratios, limited distortion, uniform or quasi-uniform meshes etc. The constants, being dependent on the discrete operator, i.e. the problem under consideration, the shape functions, element geometries etc., are typically labeled C and the mesh dependent parameter h (above, Δx was frequently used analogously). For a number of sharp estimates for C and h in certain problems, see [18].

5.2 Outline of Standard Techniques in Mathematical Analysis of Stabilized Problems

In the following, our intention is to give a rough description of the standard procedure of mathematical analysis of FEMs. Let Ψ denote the set of elements resulting from the finite element discretization of the computational domain Ω into subdomains Ω_e, $e = 1, 2, \ldots n_{el}$, where n_{el} is the number of elements. With Ψ the finite dimensional space

$$H^{1h} = \left\{ \phi^h \left| \phi^h \in C^0\left(\overline{\Omega}\right), \phi^h|_{\Omega_e} \in P_k\left(\Omega_e\right), \forall \Omega_e \in \Psi \right. \right\}$$

is associated with k representing the order of polynomial interpolation. The trial and test function spaces are defined as

$$\begin{aligned} S_u^h &= \left\{ u^h \left| u^h \in \left(H^{1h}\right)^{n_{sd}},\ u^h = g^h \text{ on } \Gamma_g \right. \right\}, \\ V_u^h &= \left\{ w^h \left| w^h \in \left(H^{1h}\right)^{n_{sd}},\ w^h = 0 \text{ on } \Gamma_g \right. \right\}, \end{aligned}$$

where g^h is the discretized Dirichlet boundary condition along Γ_g and n_{sd} is the number of space dimension. Then, the finite element problem consists in finding $u^h \in S_u^h$ such that $\forall w^h \in V_u^h$

$$B\left(u^h, w^h\right) = L\left(w^h\right),$$

where $B\left(\cdot,\cdot\right)$ and $L\left(\cdot\right)$ describe the problem under consideration. Consistency (also called orthogonality of the error) requires

$$B\left(e, w^h\right) = 0 \quad \forall w^h \in V_u^h,$$

where e is the error in the finite element approximation, $e = \tilde{u} - u$. This is, in general, immediately fulfilled. Stability (or coercivity) can be proven with

$$B\left(w^h, w^h\right) \geq \left|\left|\left|w^h\right|\right|\right|^2 \quad \forall w^h \in V_u^h,$$

where $|||\cdot|||$ is a norm, in which later convergence will be shown. If the norm is chosen such that $\left|||w^h|||^2 := B\left(w^h, w^h\right)$, this is immediately fulfilled. In [29] and [50] it is pointed out that in the SUPG, stability is less straightforward than with GLS, because one needs specific properties on τ and an additional inverse estimate.

The global error estimate —i.e. the proof of convergence— is then of the form [36]

$$\|e\|_{L_2} \leq Ch^s \|u\|_{H^r}.$$

The FEM employing interpolations of order k has a rate of convergence which is called "optimal", if $r = s = k + 1$. This is the case for Bubnov-Galerkin FEMs applied to elliptic problems, such as those arising frequently in the context of structural analysis. However, for hyperbolic problems such as advection-diffusion problems, there is a "gap" of $r - s = 1$ (apart from stability problems) [36]. In contrast, with the streamline diffusion method, i.e. with stabilized methods, the "gap" is only $r - s = 1/2$. Note that this is still suboptimal. Approaches that simply add artificial diffusion to the convection-diffusion problem, can only be of first order, i.e. $s = 1$, see e.g. [36].

One has to separate local and global error estimates. We cite from [46] for a nice description in case of hyperbolic problems.

> Localization results state that effects are propagated in the discrete problem approximately as in the continuous problem, i.e. approximately along the characteristics. More precisely it is proved that the influence of a source in the discrete problem decays with the distance d to the source like $\exp(-Cd/h)$ in any direction with a positive component in the upwind direction and like $\exp\left(-Cd/\sqrt{h}\right)$ in directions orthogonal to the streamlines (crosswind directions). Alternatively, these results can be phrased as local error estimates.

The standard Bubnov-Galerkin method applied to hyperbolic problems does not allow such local estimates —in contrast to elliptical and parabolic problems— and effects may propagate through the whole domain with little damping. However, the stabilized FEMs allow local error estimates also for the hyperbolic case.

5.3 Review of Some Specific Problems

Scalar advective-diffusive problems One of the first mathematical analysis of the streamline diffusion FEM was done by Nävert in [46] and Johnson, Nävert and Pitkäranta in [36] for convection-diffusion problems. In [11], Franca *et al.* show global convergence proofs for the advective-diffusive model stabilized with SUPG or GLS, considering the results for C and h that were later published in [18]. The Peclet number is computed slightly different from the standard way to include the effect of the specific finite element polynomial employed. I.e., this convergence analysis is applicable to higher order elements, too.

The GLS stabilization for this class of problems has been analyzed in [29] by Hughes, Franca and Hulbert.

Stokes flow In subsection 2.2 it has been explained why problems arise for mixed variational formulations and because we restrict ourselves to Stokes flow and the incompressible Navier-Stokes equations in this paper, only "circumventing Babuška-Brezzi condition methods" [12] are considered. These method possess bilinear forms which are coercive on the primal variable and the Lagrange multiplier. Then, there is no need to fulfill the Babuška-Brezzi condition any longer.

In [28], Hughes *et al.* prove convergence of a stabilization method that has later been labeled as PSPG [53]. The resulting system of equations is unsymmetric. Stabilization with the GLS, leading to a symmetric system of equations, has been analyzed in [27]. Note that therefore more stabilization terms are added than with PSPG.

Incompressible Navier-Stokes equations In the case of the incompressible Navier-Stokes equations, stabilization due to the advection term is needed as well as due to circumvention of the Babuška-Brezzi condition. The study of these equations was shown in [37] by Johnson and Saranen for the stream function-vorticity form, which has several drawbacks such as the definition of the boundary conditions.

Hansbo and Szepessy make the analysis of the incompressible Navier-Stokes equations in [17] for the velocity-pressure formulation, which is most often used in practice. They restrict themselves to the convection-dominated case only, including the incompressible Euler equations (where no diffusion at all is present). Franca *et al.* analyze in [10] a linearized form of the incompressible Navier-Stokes equations with the same main characteristics as in [11] for the advective-diffusive model (see above), i.e. being applicable to higher-order FEM.

General advective-diffusive systems The compressible Euler and Navier-Stokes equations are part of this general class of advective-diffusive systems. Hughes, Franca and Mallet analyze in [31] the convergence of the SUPG generalized to these problems (as introduced in [32]). They base their study on symmetrized systems based on [30] —using entropy variables—, having important features for the mathematical study. Note that for advective-diffusive system the stabilization parameter becomes a matrix $\boldsymbol{\tau}$ [50, 32] to treat the modal components of each equation in an adequate and individual way. In this context, it has been referred to as "matrix of intrinsic time-scales" [32].

The GLS has been studied in [29] and [50] for advective-diffusive systems.

Summary Stabilized methods have been examined in a large variety of problems. For all analyses, there are some agreeing results.

For best rate of convergence, it follows for the design of the stabilization parameter τ that $\tau = O\left(h^2\right)$ for the diffusion-dominated case (including Stokes flow) and $\tau = O\left(h\right)$ for the advection-dominated case. It can be found from the references given above that the same functional dependency of the stabilization parameter τ for circumventing the Babuška-Brezzi condition can be found as for the stabilization of advection-dominated problems. It is thus not surprising that SUPG and PSPG stabilization parameters are often chosen identically. In GLS this comes automatically, because there only *one* τ arises to stabilize oscillations from advection-dominated regimes as well as from violating the Babuška-Brezzi condition.

Optimal convergence can be proven for the diffusion-dominated case. Only suboptimal convergence with a "gap" of $\frac{1}{2}$ in the convergence rate, can be shown for the stabilized FEMs applied to convection-dominated problems (i.e. linear interpolations have convergence of $O\left(h^{3/2}\right)$). However, it is pointed out in [34] that numerical studies suggest that also in the advection-dominated case, the L_2 error may actually be optimal in cases where the exact solution is sufficiently smooth.

It is pointed out in the recent literature, e.g. [56] and references therein, under which conditions one can expect higher-order convergence rates. The influence of the mesh is of high importance in these analyses, i.e. whether the mesh is uniform, quasi-uniform, streamline-orientated etc.

5.4 Conclusions for the Stabilization Parameter τ

In all of the stabilization methods described in section 3 a weighting factor for the stabilization terms arises, the stabilization parameter τ. In 5.3 it was found that mathematical analysis provides us with some important design criteria:

- $\tau > 0$ (this is underlined with the dimension of time of τ in certain problems, see e.g. [32])
- $\tau = O(h)$ for the advection-dominated case,
- $\tau = O\left(h^2\right)$ for the diffusion-dominated case.

This fully agrees with the choices of τ presented in 4.5.

However, *the optimal choice of this parameter remains to be an open question.* It depends on the problem under consideration and on the numerical method chosen to approximate this problem. We found the choice of τ being based on the following approaches:

- Model equation: This is shown in section 4, leading to the coth-formula that can be generalized to alternative similar functions (subsection 4.5) and multi-dimensions. This seems to be rather the engineering viewpoint and is justified *a posteriori* by reasonable results in a large number of applications.
- Error analysis: In this case, often the doubly asymptotic version is chosen for τ, satisfying the design criteria quoted above. This has the main purpose of enabling mathematical analysis in a comfortable way.
- Eigenvalue problems: In case of advective-diffusive systems, it has been shown that one way to obtain the stabilization matrix $\boldsymbol{\tau}$ comes from the solution of an eigenvalue problem [32].
- Green's function: In [25] Hughes explains the evaluation of τ from the element's Green's function. In practice, there is not much experience with this approach.
- Matrix and vector norms: Tezduyar and Osawa show in [54] new ways to compute the stabilization parameters. They use matrix and vector norms to calculate τ in a generalized way, also applicable to higher-order elements. They point out that it is possible to obtain stabilization parameters of element nodes, degrees of freedom or element equations rather than of elements only. However, this approach is very general and does not sharply define formulas for τ.

5.5 Relations of Stabilization Schemes to Other Areas

In the following, a short description of areas that are related to stabilization is given. The interested reader is referred to the references given below.

Least-squares FEM The least-squares FEM (LS-FEM) is a residual method, where the test functions are chosen, such that the squared residual of a differential equation is minimized with respect to a certain norm. The least-squares formulation is itself a stable formulation and does not require any additional stabilization in its basic form. Therefore, it can be considered as a universal approach for the numerical solution of all types of partial differential equations. However, there are major drawbacks of this formulation: One is that C^1 continuous shape functions must be employed, or, the more popular way that the differential equations have to be transformed to a system of first order differential equations, thereby allowing C^0 standard shape functions. Note that the GLS stabilization, although having many analogies to the LS-FEM, requires only C^0 shape functions from the beginning, due to the contribution in element interiors only.

Another drawback is that the results of the LS-FEM are often more diffusive than standard stabilized methods. Comparing stabilized methods and LS-FEM, it may be shown that the parameter τ, which is only indirectly existent in the LS-FEM, is larger for the LS-FEM, leading to more diffusive results [23].

Although optimal rates of convergence can be proven for many problems (compared with only sub-optimal rates of convergence for the stabilized methods for convection-dominated problems!), the computational effort seems currently to be higher for the LS-FEM than for the stabilized methods.

Bubble functions The relationship between the Galerkin method with bubble functions and stabilized finite element methods has been pointed out in [6] by Brezzi *et al.* and in [9] by Franca and Farhat. This aspect is also reconsidered in [25] from Hughes.

Wavelet RKPM Not surprisingly, as well as bubble functions in meshbased methods have a stabilizing influence, it has been shown by Li and Liu in [40, 41] that in the meshfree context wavelets can be used for stabilization. They introduce and analyze the so-called wavelet Reproducing Kernel Particle Method (Wavelet-RKPM).

Mini element In [10], Franca and Frey find a relationship under certain conditions of the so-called mini element and the SUPG stabilization method. Again, this approach can be reduced to the influence of bubbles.

Subgrid modeling and Green's function In [25], Hughes proves the derivation of stabilized methods from subgrid modeling concepts and explains the evaluation of τ from the elements Green's functions.

6 Extension to Meshfree Methods

The need for stabilization, described in section 2, is the same for all weighted residual methods, regardless whether they employ meshbased or meshfree shape functions. However, they raised in the finite element, i.e. meshbased, context and most of the experiences have been done for the FEM and its theoretical foundation is well studied.

The first steps in the area of MMs have been made rather three decades ago. A large number of different methods have been developed. For a classification and overview of MMs see [13]. In here, only MMs are considered whose shape functions are constructed from the Moving Least Squares (MLS) methodology and Reproducing Kernel Particle Method (RKPM). These shape functions form a partition of unity of n-th order. The resulting MM can further be characterized due to the following two criteria:

Choice of the test function In a weighted residual method, the test functions may be chosen rather arbitrary. In practice, only a few choices are standard: Taking a Dirac-δ function as a test function leads to collocation MMs, such as Smoothed Particle Hydrodynamics (SPH) [15, 43, 45] or the Finite Point Method (FPM) [47, 48]. Then, the weak form reduces to the strong form of the problem and no integration is required. This approach is fast in terms of calculation time but the accuracy is rather low and stabilization problems may arise in certain particle arrangements. Choosing the test function equal to the shape function leads to Bubnov-Galerkin MMs, such as the Element Free Galerkin (EFG) [4, 42] method. The integration of the weak form is quite time-consuming —especially due to the rational character of meshfree shape functions—, but the accuracy is high, in general remarkably higher than the corresponding FEM with the same order of consistency. Finally, Petrov-Galerkin MMs with test functions resulting from stabilized weak forms are mentioned as shown in the previous sections for the FEM. Then, the test function is constructed by modifying the shape function with certain perturbation terms.

Lagrangian or Eulerian viewpoint The standard way to work with MMs is still using problems in a Lagrangian formulation, i.e. the particles move with certain velocities through the domain. Then, the resulting methods show a very different behavior from Eulerian methods. Advection terms are not present in the differential equation any longer and stabilization is *not* needed. However, the Lagrangian viewpoint has a number a disadvantages: Boundary conditions are difficult to apply. At outflow boundaries, particles leave the domain and require special attention and even at standard boundary conditions care has to be taken with the formulation of reflection conditions. The Lagrangian viewpoint is those only in problems fully successful, where boundaries are of less importance. This is for example the case in impact calculations, being one of the most important applications of Lagrangian MMs.

Another disadvantage of them is the problem of particle clustering. One finds that after a while certain areas of the domain contain less and less particles while others accumulate them. This can result in severe problems with the MM applied. Prescribing a certain order of consistency e.g. in the MLS requires certain overlap of the particle supports. If the initial particle distribution becomes more and more distorted due to the clustering problem, the necessary overlap cannot be guaranteed and stability problems may arise.

Note that with Eulerian MMs neither problems with boundary conditions arise nor with particle clustering, because the particles are placed *a priori* in the domain and stay there during the calculation. If one desires a higher particle density in a certain area due to accuracy reasons, this can be done by controlled adaptive refinement.

Another aspect is the coupling of meshbased and meshfree methods. In problems, where the distortions are rather small such as in elastoplastic deformation problems, the Lagrangian viewpoint is applicable also to meshbased methods, such as the FEM. However, in fluid mechanics, there is no way to use the Lagrangian approach and move the whole mesh through the domain, maintaining its conformity. Therefore, for the FEM in fluid mechanics, the Eulerian or Arbitrary-Lagrangian-Eulerian (ALE) viewpoint is standard. Both require stabilization. Coupling Eulerian (or ALE) finite element formulations with MMs seems most natural when also Eulerian MMs are used.

We conclude that Eulerian-Galerkin MMs (including ALE-Galerkin MMs) seem to have promising features. It is our believe that a Galerkin formulation is desirable due to its high accuracy. The problem of time-consuming integration of the weak form can be handled by coupling meshfree and meshbased methods. Then, MMs are only applied where they are needed, i.e. in regions where a mesh is difficult to maintain. Standard meshbased methods should then be used in the other parts of the domain. They require an Eulerian (or ALE) viewpoint, which should therefore also be taken for the MM. The Eulerian viewpoint also enables to avoid severe problems of Lagrangian MMs, such as boundary treatment and particle clustering.

To successfully use Eulerian-Galerkin MMs the problem of stabilization has to be overcome. This is the main aspect of this paper. Stabilization schemes include two main ingredients: The modification of the weak form and the weighting of this modification in terms of the stabilization parameter τ.

The modification of the weak form, motivated in section 4 from introducing artificial diffusion in streamline direction only and confirmed in its usefulness in the theoretical studies, see section 5, applies absolutely equivalent in the context of MMs as well. Thus, one can use the same structures of the different stabilization schemes also for MMs.

However, it was found in section 4 that the stabilization parameter τ depends on the problem under consideration and the chosen numerical method to approximate this problem. It turned out that in fact, the advection-diffusion model leads to local stabilization parameters that can be generalized to other problems straightforward. However, the chosen numerical method is indeed important for the choice of τ. For example, using the coth-formula —or some alternative versions— without any adjustment or justification to any higher-order FEM will not give adequate results [10, 11, 50].

There have also been some approaches to stabilized MMs in the recent literature, see e.g. [16, 23]. There, standard stabilization schemes as those introduced in 3 are applied together with standard formulas to obtain τ. It is analyzed in the following under which circumstances the usage of standard formulas for τ —being based in the analysis of the linear FEM!— can be justified.

6.1 One-dimensional Advection-Diffusion Equation

As a starting point again the 1D advection-diffusion equation is used. The SUPG-stabilized weak form is

$$\int_\Omega \left(\mathbf{w} + \tau c \frac{\partial \mathbf{w}}{\partial x} \right) \left(c \frac{\partial \widetilde{u}}{\partial x} - K \frac{\partial^2 \widetilde{u}}{\partial x^2} \right) d\Omega = \mathbf{0}.$$

Note that the typical notation of stabilized weak forms with a sum over element interiors ($\sum_{e=1}^{n_{el}} \int_{\Omega_e}$) only applies to the FEM. In MMs, the shape functions are easily constructed to have C^1 continuity and the contribution of the stabilization can be considered in the whole domain rather than in element interiors only.

We follow exactly the procedure described in 4.1. The ansatz is $\tilde{u}(x) = \mathbf{N}^T(x)\,\mathbf{u}^{\text{fict}}$, where $\mathbf{N}$ are the meshfree shape functions and $\mathbf{u}^{\text{fict}}$ are the unknowns. They may be called *fictitious* unknowns due to the fact that meshfree shape functions in general lack Kronecker-δ property that is $N_i(x_j) \neq \delta_{ij}$. Consequently, the resulting vector of unknowns $\mathbf{u}^{\text{fict}}$ can not be interpreted as the final *real* result $\mathbf{u}^{\text{real}}$ of the approximation at the particle positions. Instead $\mathbf{u}^{\text{real}} = \mathbf{D}\mathbf{u}^{\text{fict}}$, where $\mathbf{D} = D_{ij} = N_i(x_j)$. Note that $\mathbf{D}$ is a sparse matrix, as the meshfree shape functions are local, i.e. they are nonzero only in their supports defined by the dilatation parameter ρ. However, $\mathbf{D}^{-1}$ is the inverse of a sparse matrix and is *full*.

Again, the aim is to find an "optimal" stabilization parameter τ_I. It was shown already in subsection 4.1.9 that "optimal" in the sense of minimizing with respect to a certain norm, leads to a global stabilization parameter *a priori*. Consequently, our aim is again to obtain the nodally exact solution in the following (which gave for the linear FEM the coth-formula, a local stabilization criterion). Therefore, one equation I of the system of equations is extracted and the ansatz is inserted:

$$\left[\int_\Omega \left(w_I + \tau_I c \frac{\partial w_I}{\partial x} \right) \left(c \frac{\partial \mathbf{N}^T}{\partial x} - K \frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) d\Omega \right] \mathbf{u}^{\text{fict}} = 0.$$

Rearranging for τ_I and introducing the knowledge of the exact solution of the problem results in

$$\tau_I = - \frac{\left[\int_\Omega (w_I) \left(c \frac{\partial \mathbf{N}^T}{\partial x} - K \frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) d\Omega \right] \mathbf{u}^{\text{ex,fict}}}{\left[\int_\Omega \left(c \frac{\partial w_I}{\partial x} \right) \left(c \frac{\partial \mathbf{N}^T}{\partial x} - K \frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) d\Omega \right] \mathbf{u}^{\text{ex,fict}}}.$$

The exact solution of the problem is known as $u^{\text{ex}}(x) = C_1 e^{\frac{c}{K}x} + C_2$, consequently $\mathbf{u}^{\text{ex,real}} = C_1 e^{\frac{c}{K}\mathbf{x}} + C_2$, where $\mathbf{x}$ are the positions of the particles. For $\mathbf{u}^{\text{ex,fict}}$ follows $\mathbf{u}^{\text{ex,fict}} = \mathbf{D}^{-1}\mathbf{u}^{\text{ex,real}}$, consequently,

$$\begin{aligned} \tau_I &= - \frac{\left[\int_\Omega (w_I) \left(c \frac{\partial \mathbf{N}^T}{\partial x} - K \frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) d\Omega \right] \mathbf{D}^{-1}\mathbf{u}^{\text{ex,real}}}{\left[\int_\Omega \left(c \frac{\partial w_I}{\partial x} \right) \left(c \frac{\partial \mathbf{N}^T}{\partial x} - K \frac{\partial^2 \mathbf{N}^T}{\partial x^2} \right) d\Omega \right] \mathbf{D}^{-1}\mathbf{u}^{\text{ex,real}}} \\ &= - \frac{\left[\int_\Omega (w_I) \left(c \frac{\partial \mathbf{N}^T\mathbf{D}^{-1}}{\partial x} - K \frac{\partial^2 \mathbf{N}^T\mathbf{D}^{-1}}{\partial x^2} \right) d\Omega \right] \mathbf{u}^{\text{ex,real}}}{\left[\int_\Omega \left(c \frac{\partial w_I}{\partial x} \right) \left(c \frac{\partial \mathbf{N}^T\mathbf{D}^{-1}}{\partial x} - K \frac{\partial^2 \mathbf{N}^T\mathbf{D}^{-1}}{\partial x^2} \right) d\Omega \right] \mathbf{u}^{\text{ex,real}}}. \end{aligned}$$

Remark 5 This expression for τ_I holds for arbitrary shape and test functions and arbitrary point distributions. It reduces for FEM interpolations due to their Kronecker-δ property, i.e. $\mathbf{D}^{-1}$ cancels out due to $\mathbf{D} = N_i(x_j) = \delta_{ij} = \mathbf{I}$. For linear FEM it was shwon in subsection 4.1 that

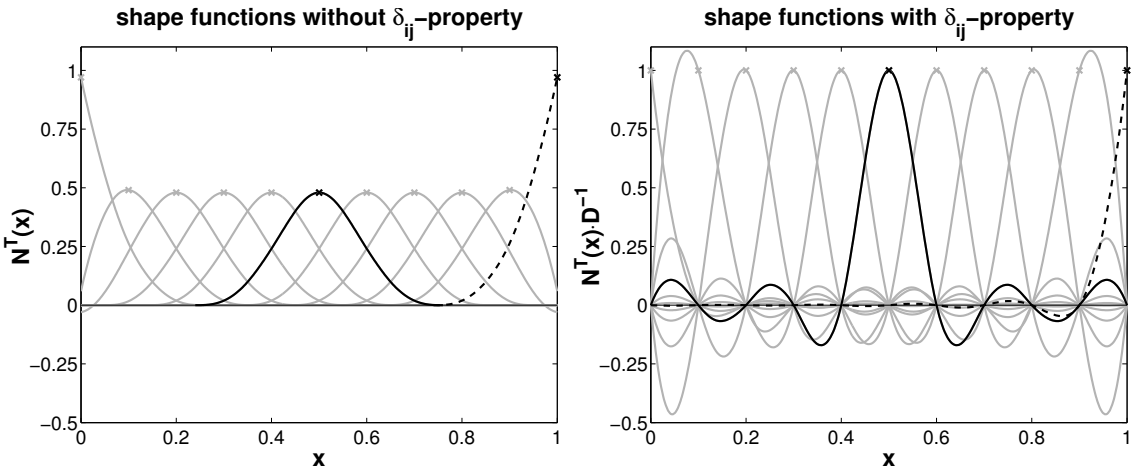


Figure 19: Local shape functions $\mathbf{N}^T$ without Kronecker-δ property and transformed global shape functions $\mathbf{N}^T\mathbf{D}^{-1}$ with Kronecker-δ property.

this reduces in case of a regular node distributions to the coth-formula.

This result is interpreted in the following. The expression for τ_I is rewritten as

$$\tau_I = -\frac{\left[\int_\Omega f_1\left(w_I\right) g\left(\mathbf{N}^T\mathbf{D}^{-1}\right) d\Omega\right] \mathbf{u}^{\mathrm{ex,real}}}{\left[\int_\Omega f_2\left(w_I\right) g\left(\mathbf{N}^T\mathbf{D}^{-1}\right) d\Omega\right] \mathbf{u}^{\mathrm{ex,real}}},$$

where f_1, f_2 and g are linear functions of the test and shape functions respectively. The meshfree test and shape functions, $\mathbf{w}$ and $\mathbf{N}$, have local supports. Consequently, there is a large number of zero-entries for every position x, i.e. the vectors $\mathbf{w}$ and $\mathbf{N}$ are sparse. However, $\mathbf{D}^{-1}$ is a full matrix, hence the term $\mathbf{N}^T\mathbf{D}^{-1}$ is a full vector. The term $\mathbf{N}^T\mathbf{D}^{-1}$ can be interpreted as the "globalized" meshfree shape functions having Kronecker-δ property. This can be depicted from Figure 19.

Remark 6 It is well known that locally defined meshfree shape functions with Kronecker-δ property can only be obtained with the standard procedure (i.e. with MLS or RKPM) by using singular weighting functions [39]. This has a number of severe numerical disadvantages and is only rarely used in practice. Here, the term "weighting functions" has the standard meaning in the MM context and may not be mixed up with the test functions of the weak form, which are frequently called weighting functions in the meshbased context.

One finds an important influence of the $\mathbf{D}^{-1}$ matrix in the formula for the stabilization parameter τ_I. In the nominator and denominator are scalar expressions, i.e.

$$\int_\Omega \underbrace{f_i\left(w_I\right)}_{1\times 1} \underbrace{g\left(\mathbf{N}^T\mathbf{D}^{-1}\right)}_{1\times n} d\Omega \underbrace{\mathbf{u}^{\mathrm{ex,real}}}_{n\times 1}.$$

The vector $g\left(\mathbf{N}^T\mathbf{D}^{-1}\right)$ is a full vector, which is in contrast to shape functions having Kronecker-δ property. In the latter case, $g\left(\mathbf{N}^T\mathbf{D}^{-1}\right) = g\left(\mathbf{N}^T\right)$ and $\mathbf{N}^T$ is sparse. Evaluating the scalar product with $\mathbf{u}^{\mathrm{ex,real}}$ makes the difference obvious. Shape functions without Kronecker-δ property

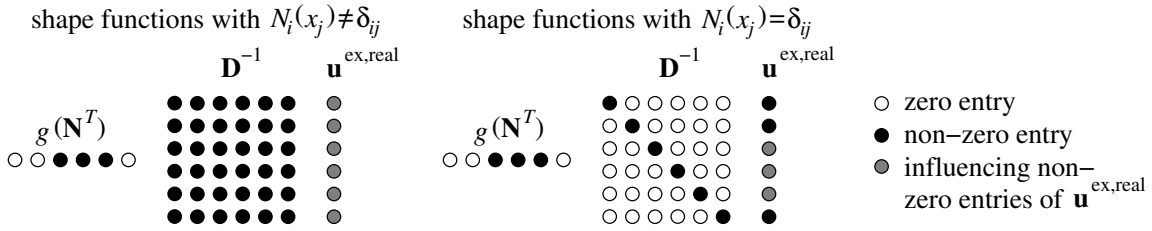


Figure 20: Evaluating the above shown symbolical scalar products makes clear that the nodally exact τ_I for shape functions without Kronecker-δ property can only be obtained with a global criterion, because all entries of $\mathbf{u}^{\text{ex,real}}$ have influence in the result.

have non-zero entries in the scalar-product for *all* components of the vector $\mathbf{u}^{\text{ex,real}}$, whereas, in contrast, shape functions with Kronecker-δ property only have non-zero entries for the *neighboring* nodes. This can be depicted symbolically from Figure 20.

Keeping in mind that $\mathbf{u}^{\text{ex,real}}$ is an exponential function, then the scalar product will depend more and more on the last entry of this vector as the convection-diffusion ratio $\frac{c}{K}$ grows, because then

$$u_n^{\text{ex,real}} \gg u_i^{\text{ex,real}} \qquad \forall n > i.$$

The entry of the last component of $\mathbf{u}^{\text{ex,real}}$ belongs to the node with the largest x-value, i.e. the global downstream node. We conclude that the stabilization parameter τ_I, leading to nodally exact solutions has a global character, i.e. it depends on all particle positions and for convection-dominated cases most importantly on the global downstream node. This is contrasted by shape functions with Kronecker-δ property, whose stabilization relies only on the local downstream node.

Consequently, it can not be expected in general that the simple coth-formula —or other simple alternative versions—, derived as a local stabilization criterion for linear FEM, works successfully also for MMs!

6.2 Small Dilatation Parameters

The shape functions of MMs depend highly on their pre-defined supports. Meshbased methods require nodes and the definition of a mesh which then indirectly defines the support size of the shape functions. In contrast, MMs require nodes and a definition of supports which is accomplished with help of the dilatation parameter ρ.

It is a well known fact that MLS shape functions in one dimension with first order consistency become more and more equal to the standard linear shape functions of the FEM, when the dilatation parameter ρ goes to $1 \cdot \Delta x$. This can be seen from Figure 21. Consequently, knowing that the coth-formula is the optimal choice for τ_I, leading to nodally exact solutions for the linear FEM, it can be seen that when $\rho \longrightarrow \Delta x$, the coth-formula becomes more and more suited also for MMs, i.e.

$$\begin{aligned} \rho \longrightarrow 1 \cdot \Delta x \quad &: \quad \mathbf{N}^{\text{MM}} \longrightarrow \mathbf{N}^{\text{lin FEM}} \\ &\Rightarrow \quad \tau_I^{\text{MM}} \longrightarrow \tau_I^{\text{lin. FEM}} = \frac{\Delta x}{2c} \left(\coth(Pe) - \frac{1}{Pe} \right). \end{aligned}$$

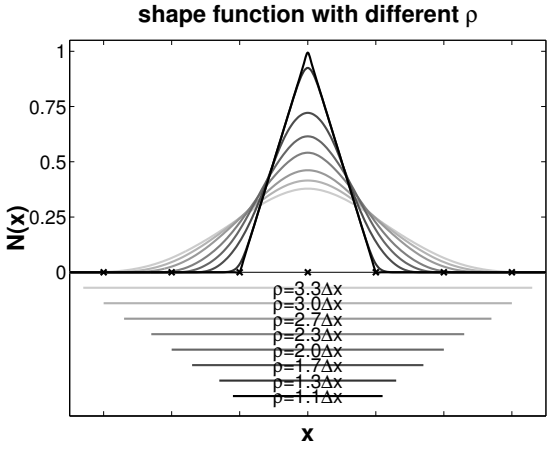


Figure 21: Meshfree shape function in a regular particle distribution with varying dilatation parameter ρ.

Note that the invertibility of the so-called "mass-matrix" of the MLS, which has to be solved at every point $\mathbf{x}$ of a domain to ensure consistency of the resulting shape functions, requires $\rho > 1 \cdot \Delta x$. Thus, the limit $\rho = 1 \cdot \Delta x$, where the coth-formula gives exactly the nodally exact solution, can never be reached. However, it may be concluded that for reasonable advection-diffusion ratios and "small" dilatation parameters a successful stabilization with standard formulas —derived for meshbased methods— can be obtained. The numerical results in subsection 6.4 confirm this assumption.

6.3 Stabilization in Multi-dimensions

The stabilization in more than one dimension has to take into account the particular support size in x- and y-direction and the shape of the support. As well as an element length in streamline direction h_e is introduced for meshbased methods, one may introduce here for meshfree methods a *support length* in streamline direction h_ρ. Figure 22 shows several possibilities to interpret h_ρ in case of rectangular supports.

- min-version:
$$h_\rho = \min\left(\rho_x, \rho_y\right)$$
- max-version:
$$h_\rho = \max\left(\rho_x, \rho_y\right)$$
- inner-ellipsoid-version:
$$h_\rho = \sqrt{\frac{\left(1 + \frac{c_y^2}{c_x^2}\right) \cdot \rho_y^2}{\left(\frac{c_y}{c_x}\right)^2 + \left(\frac{\rho_y}{\rho_x}\right)^2}}$$
- real-length-version:
$$h_\rho = \min\left(\frac{\rho_x}{|c_x|}, \frac{\rho_y}{|c_y|}\right) \cdot \sqrt{c_x^2 + c_y^2}$$

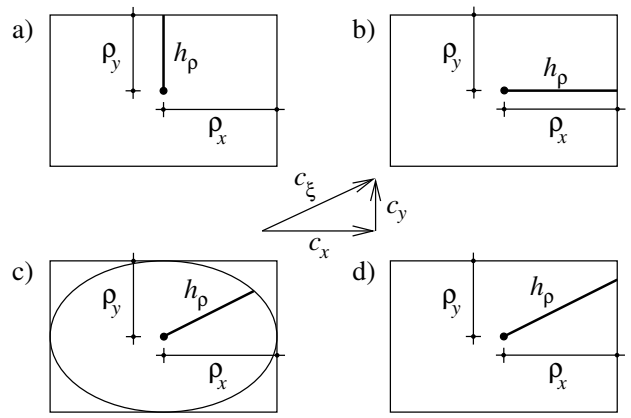


Figure 22: Different versions to compute the support length in streamline direction; a) min-version, b) max-version, c) inner-ellipsoid-version, d) real-length-version.

The support lengths for circular and ellipsoid supports can be directly red off from these formulas, too. Note that c_x and c_y are replaced in many non-linear problems by the velocities u and v. In our numerical experiments, it is found that particularly the min-version works very successful also for large aspect ratios ($\rho_x/\rho_y \gg 1$, or $\rho_y/\rho_x \gg 1$). See [51], for an interesting parallel for high aspect *elements*: Mittal also finds that the minimal edge length works better than other versions for h_e.

Especially the inner-ellipsoid-version and the real-length-version, applied to non-linear problems, degrade the convergence of the method considerably. This can be explained by noting that the supports in MMs are much larger than elements. Finding a certain u and v, being representative for the whole velocity field inside the support of a node is much more difficult, than determining an average element velocity. Also note that $\tau = f(u, v, h_\rho)$, i.e. τ introduces an additional non-linearity. This non-linearity is even higher for $h_\rho = f(u, v)$, as in the inner-ellipsoid version and the real-length version. These non-linearities are not considered in the Jacobian (or whatever method is applied to solve the non-linear problem under consideration). Instead, in marching from one iteration level n to $n+1$, τ depends on n only [54].

With knowledge of the support length, τ is easily computed as follows

$$\frac{h_\rho}{2\,|\mathbf{c}|}\left(\coth\left(Pe_\rho\right) - \frac{1}{Pe_\rho}\right), \text{ with } Pe_\rho = \frac{ch_\rho}{2K}.$$

Also any other alternative version from subsection 4.5 for the functional dependency of τ on Pe may be applied.

6.4 Numerical Results

Numerical results are shown for two different problems. The first is the one-dimensional advection-diffusion equation, the model equation of subsection 6.1. It is shown that we are able to obtain nodally exact solutions with meshfree shape functions, but therefore, a global stabilization criterion is needed. Then, it is shown that standard formulas for τ are successful only for small dilatation parameters which confirms our assumption from subsection 6.2.

The second problem are the incompressible Navier-Stokes equations in two dimensions. This non-linear problem requires stabilization of high convection-diffusion ratios and due to the circumvention of the Babuška-Brezzi condition. Small dilatation parameters are a crucial ingredient to obtain successful stabilized results. Using supports with too large dilatation parameters results in degradation of convergence and solutions that are still either too oscillatory or too diffusive. Our intention is to show that stabilization with small dilatation parameters

- smoothes out oscillations successfully
- does not degrade accuracy in cases where stabilization is not necessary
- works also for anisotropic supports.

6.4.1 1D Advection-Diffusion Equation

The one-dimensional advection diffusion equation is solved with 21 particles. The advection-diffusion ratio is $\frac{c}{K} = \frac{100}{1} = 100$. Figure 23a) shows the unstabilized results for two different dilatation parameters $\rho = 1.3\Delta x$ ("small") and $\rho = 3.3\Delta x$ ("large"). It can be seen that higher dilatation parameters lead to more oscillations, simply due to their higher Peclet number, $Pe_\rho = \frac{c\rho}{2K}$. Clearly, for both cases, stabilization is required.

Figure 23b) shows the nodally exact result, which can be obtained with the global stabilization criterion for τ_I derived in subsection 6.1:

$$\tau_I = -\frac{\left[\int_\Omega (w_I)\left(c\frac{\partial \mathbf{N}^T\mathbf{D}^{-1}}{\partial x} - K\frac{\partial^2 \mathbf{N}^T\mathbf{D}^{-1}}{\partial x^2}\right)d\Omega\right]\mathbf{u}^{\text{ex,real}}}{\left[\int_\Omega \left(c\frac{\partial w_I}{\partial x}\right)\left(c\frac{\partial \mathbf{N}^T\mathbf{D}^{-1}}{\partial x} - K\frac{\partial^2 \mathbf{N}^T\mathbf{D}^{-1}}{\partial x^2}\right)d\Omega\right]\mathbf{u}^{\text{ex,real}}}.$$

In Figure 23c) it can be seen that standard formulas for τ_I, like the coth-formula,

$$\begin{aligned}\tau_I &= \frac{\rho}{2c}\left(\coth\left(\frac{c\rho}{2K}\right) - \frac{2K}{c\rho}\right)\\ &= \frac{\rho}{2c}\left(\coth\left(Pe_\rho\right) - \frac{1}{Pe_\rho}\right),\end{aligned}$$

only lead to successful stabilization when the dilatation parameter is small. Compare Figure 23b) and c) to realize that for small dilatation parameters, the result of the complicated global criterion and the coth-criterion gives almost the same result. This, however, is not the case for the large dilatation parameter of $\rho = 3.3\Delta x$, where pronounced oscillations remain in the solution. These oscillations are clearly not a problem of the high gradient itself that could not be captured by shape functions with such a large dilatation parameter (then the result of Figure 23b) must have been oscillatory, too, and this is not the case), but results from the usage of unsuited stabilization parameters.

We conclude that our assumption of subsection 6.2 is confirmed: Only small dilatation parameters can be reliably used with the standard formulas of meshbased methods.

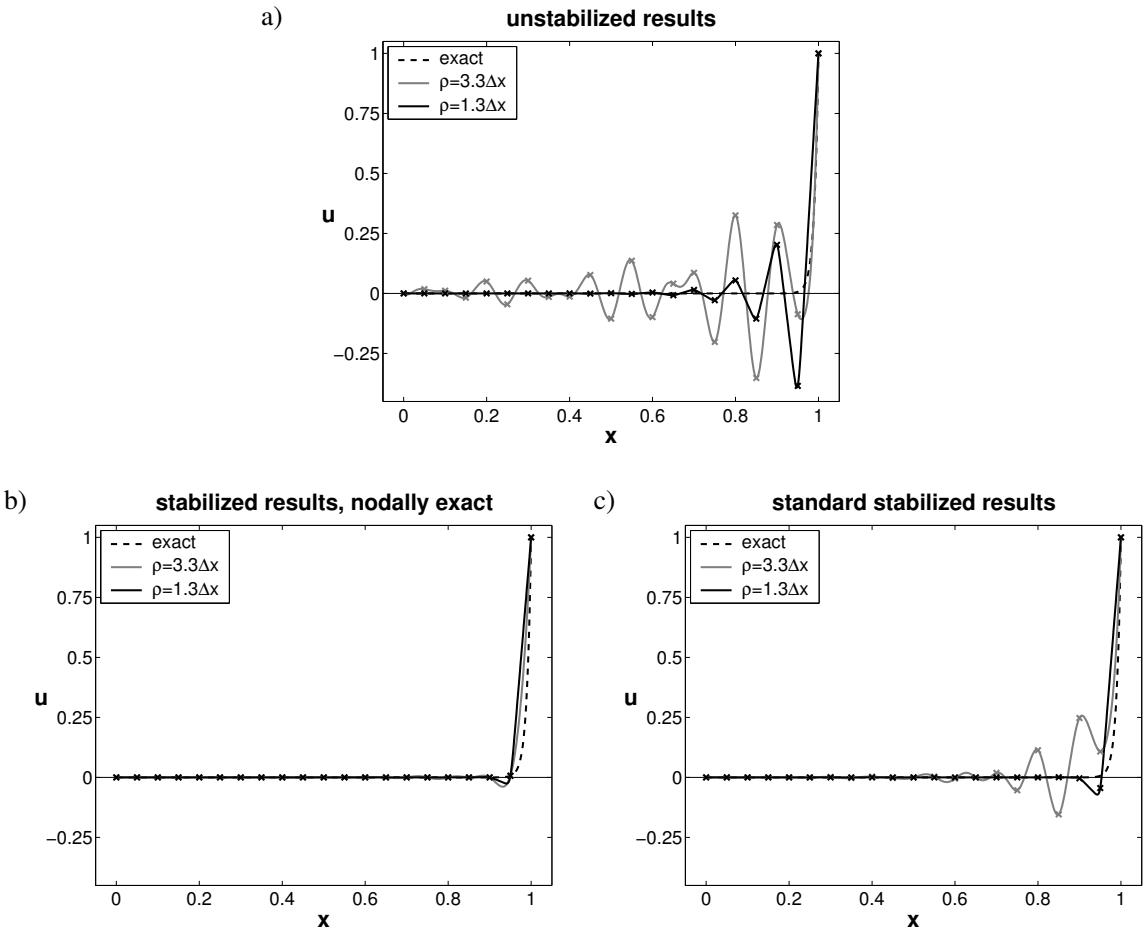


Figure 23: Results for the 1D advection-diffusion equation; a) without any stabilization, b) with the global stabilization criterion from subsection 6.1, c) with the local coth-formula from subsection 4.1.3.

6.4.2 Incompressible Navier-Stokes Equations in 2D

The incompressible Navier-Stokes equations are in strong form

$$\begin{aligned} \varrho\left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u} - \mathbf{f}\right) - \nabla\cdot\sigma &= 0 \\ \nabla\cdot\mathbf{u} &= 0, \end{aligned}$$

$$\sigma = -p\mathbf{I} + \mu\left(\nabla\mathbf{u} + (\nabla\mathbf{u})^T\right),$$

where $\mathbf{u} = (u, v)$ are the velocities, p is the pressure, σ the stress tensor and μ the dynamic viscosity. Stabilization is needed due to the advection term $\mathbf{u}\cdot\nabla\mathbf{u}$ and due to the incompressibility constraint $\nabla\cdot\mathbf{u} = 0$. This is realized by either using GLS stabilization or SUPG/PSPG stabilization. Then, the stabilized weak forms are:

- SUPG/PSPG stabilization

$$\begin{aligned} &\int_\Omega \left(\mathbf{w} + \tau_{\mathrm{SUPG}}\mathbf{u}\cdot\nabla\mathbf{w} + \tau_{\mathrm{PSPG}}\frac{1}{\varrho}\nabla q\right)\cdot \\ &\left[\varrho\left(\mathbf{u}\cdot\nabla\mathbf{u} - \mathbf{f}\right) - \nabla\cdot\sigma\right] d\Omega + \int_\Omega q\nabla\cdot\mathbf{u} d\Omega \;=\; 0 \end{aligned}$$

- GLS stabilization

$$\begin{aligned} &\int_\Omega \left(\mathbf{w} + \tau_{\mathrm{GLS}}\left(\mathbf{u}\cdot\nabla\mathbf{w} - \frac{1}{\varrho}\nabla\cdot\sigma\right)\right)\cdot \\ &\left[\varrho\left(\mathbf{u}\cdot\nabla\mathbf{u} - \mathbf{f}\right) - \nabla\cdot\sigma\right] d\Omega + \int_\Omega q\nabla\cdot\mathbf{u} d\Omega \;=\; 0 \end{aligned}$$

In general, we have $\tau_{\mathrm{SUPG}} = \tau_{\mathrm{PSPG}} = \tau_{\mathrm{GLS}}$. Then, the only difference between the SUPG/PSPG and GLS stabilization is in the following modification of the test functions

$$\begin{aligned} \text{SUPG/PSPG} \;&:\; \ldots\frac{1}{\varrho}\nabla q\ldots \\ \text{GLS} \;&:\; \ldots - \frac{1}{\varrho}\nabla\cdot\sigma\ldots = \ldots\frac{1}{\varrho}\nabla q - \frac{1}{\varrho}\nabla\cdot\mu\left(\nabla\mathbf{w} + (\nabla\mathbf{w})^T\right)\ldots, \end{aligned}$$

i.e. there are additional terms in the modification of the test function in the GLS stabilization, which result from the diffusion part.

Test case: driven cavity flow The driven cavity test case is a standard test case with benchmark solutions given in [14] for a variety of Reynolds numbers. In here, this problem is solved with $Re = 1000$. For a problem statement see Figure 24, showing also streamlines and pressure distribution for $Re = 1000$. In the sequel, only *velocity profiles* are studied at certain cuts through the two-dimensional domain. The standard MLS shape functions with first order consistency are applied to solve the SUPG/PSPG and GLS stabilized weak forms from above respectively.

The first results are produced with 21×21 regularly distributed particles. Figure 25 shows velocity profiles for u and v at $y = 0.95$, i.e. near the tangential flow boundary, where most of the oscillations occur. Two different dilatation parameters are shown, $\rho = 1.3\Delta x$ and $\rho = 2.3\Delta x$.

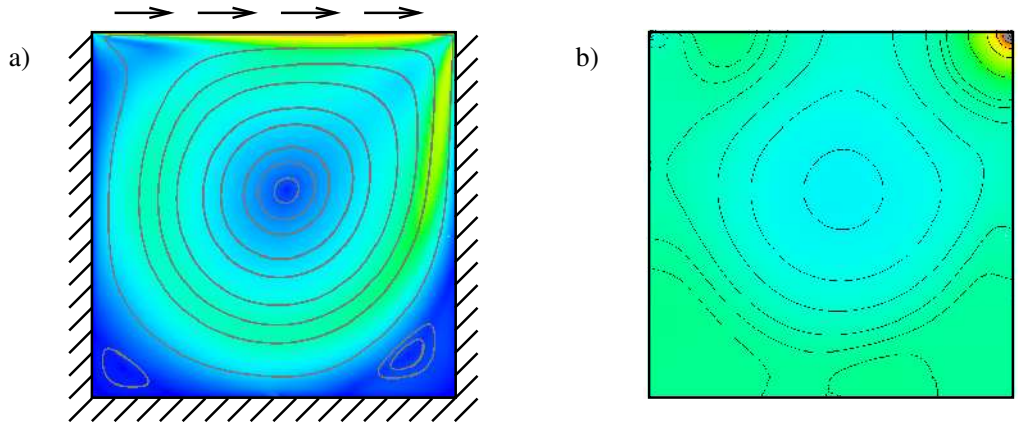


Figure 24: Problem statement of the driven cavity test case. As an example the velocity and pressure field is shown at $Re = 1000$.

Dilatation parameters $\rho > 2.7\Delta x$ converged either not at all or only very badly, underlining the need for small dilatation parameters, when standard formulas for τ_I are used.

One can clearly see that the oscillations apparent in the unstabilized result are smoothed out successfully, especially for the case where $\rho = 1.3\Delta x$. For $\rho = 2.3\Delta x$ one may see from the velocity profile for v that very slight oscillations remain in this case. Again, our assumption that shape functions with small dilatation parameters can be stabilized very successfully is confirmed.

Figure 26 shows the center velocity profiles for the case where $\rho = 1.3\Delta x$. It can be seen that although along these cuts no oscillations are apparent in the unstabilized case, the stabilized profiles give better results. This is because the oscillations in the unstabilized case near the tangential flow boundary degrade the overall solution.

We conclude that stabilization for shape functions with small dilatation parameters smoothes out oscillations successfully and leads to superior overall solutions than unstabilized calculations.

The next results are computed with 101×101 particles and $\rho = 1.3\Delta x$. With this large number of particles, stabilization is not needed at all, i.e. the unstabilized solution is already free of oscillations. The results show that stabilization does not degrade the accuracy when it is not needed. Figure 27 shows the center velocity profiles. It is interesting that unstabilized and SUPG stabilized results are indistinguishable, whereas GLS stabilized results are slightly more diffusive. This was confirmed in a number of additional computations.

Figure 28 shows a comparison of the benchmark solution with the meshfree solution (with $\rho = 1.3\Delta x$) and the solution from the P1/P1 triangular element with the same number of unknowns. For both numerical methods, SUPG/PSPG stabilization and an irregular node distribution as shown in Figure 29 has been used. The supports of the nodes are anisotropic with respect to the distance to the neighboring nodes,

$$\begin{aligned} \rho_{x,i} &= c \cdot \min\left(|x_j - x_i|\right), \ \forall i \neq j, \\ \rho_{y,i} &= c \cdot \min\left(|y_j - y_i|\right), \ \forall i \neq j, \end{aligned}$$

with $c = 1.6$. The min-version for the support length h_ρ performs best compared to the other

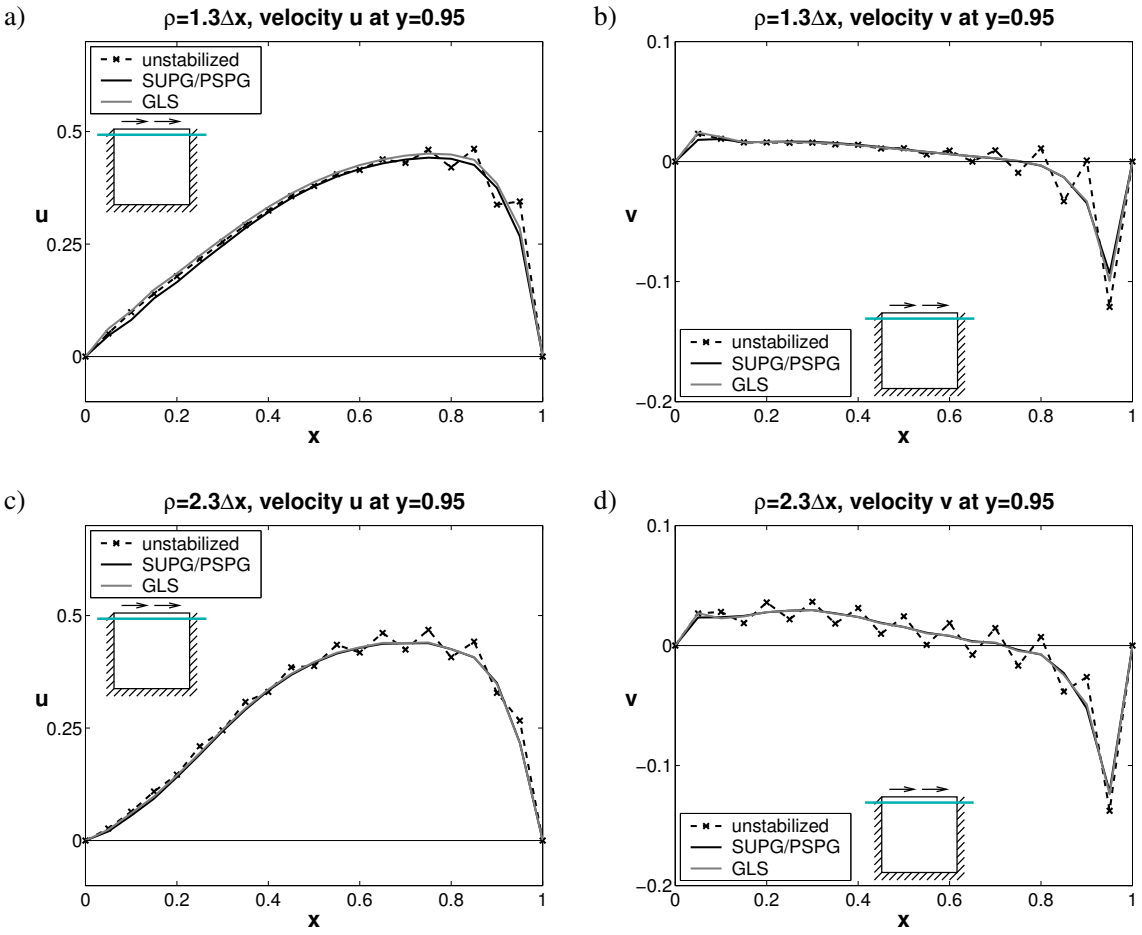


Figure 25: Velocity profiles for u and v near the tangential flow boundary at $y = 0.95$ for different dilatation parameters of $\rho = 1.3\Delta x$ and $\rho = 2.3\Delta x$.

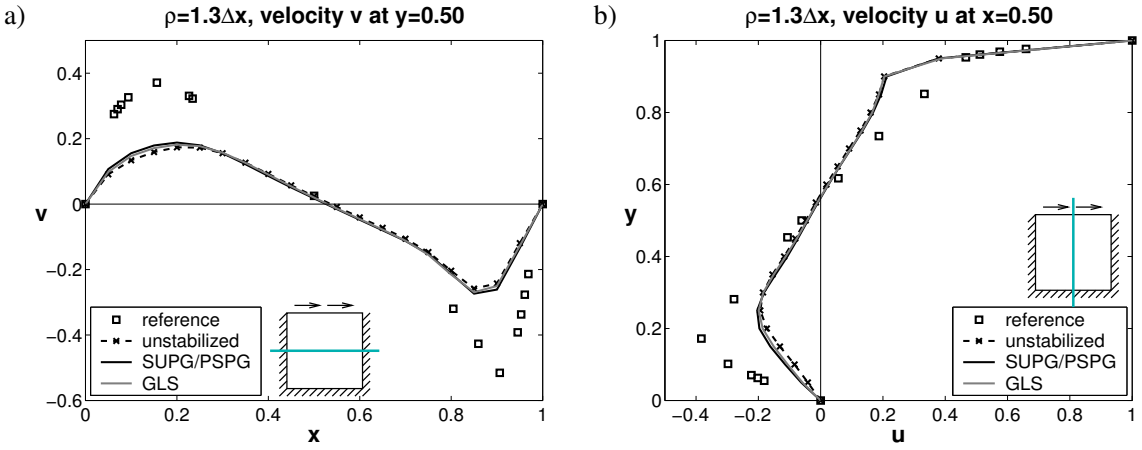


Figure 26: Velocity profiles for u and v along $y = 0.5$ and $x = 0.5$ respectively (for $\rho = 1.3\Delta x$ and 21×21 nodes).

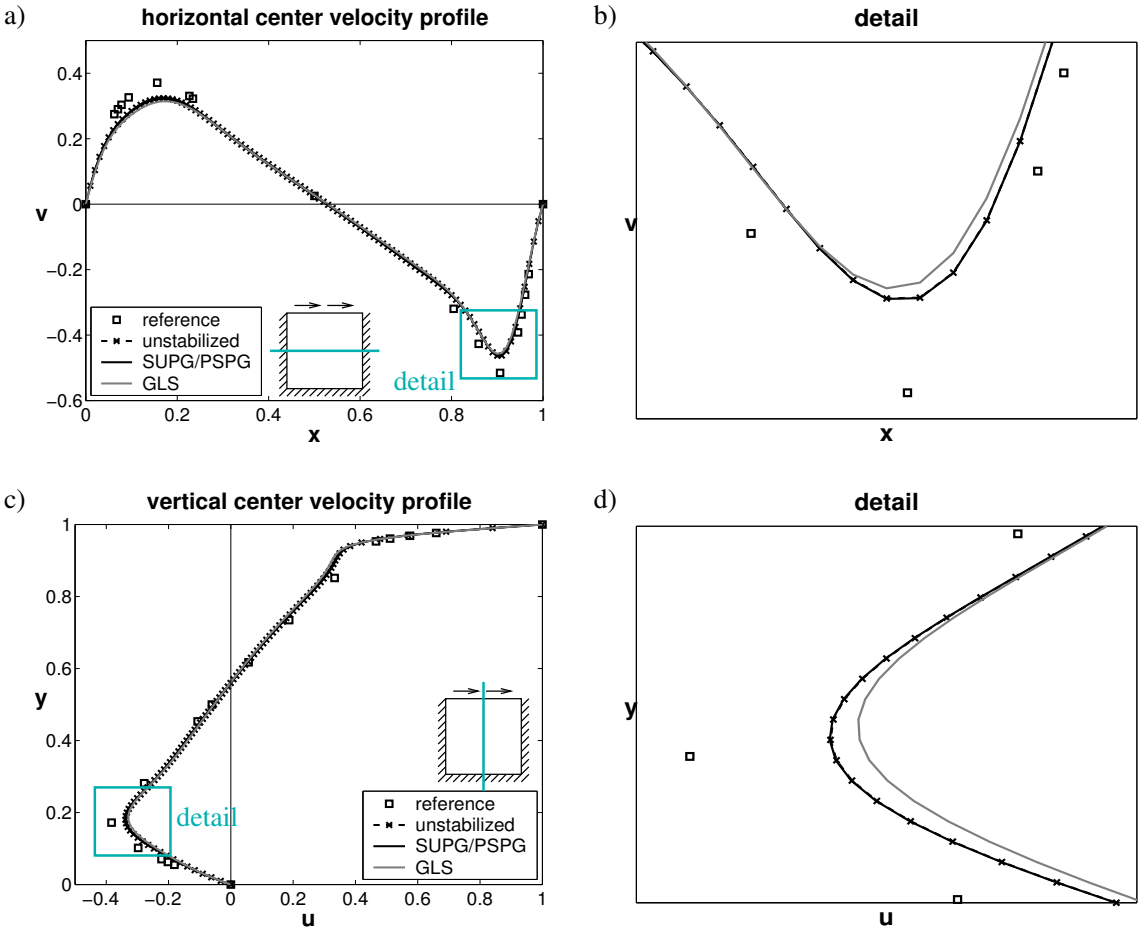


Figure 27: Velocity profiles for u and v along $y = 0.5$ and $x = 0.5$ respectively (for $\rho = 1.3\Delta x$ and 101×101 nodes). The details show that unstabilized and SUPG/PSPG-results are indistinguishable, whereas the GLS result is slightly more diffusive.

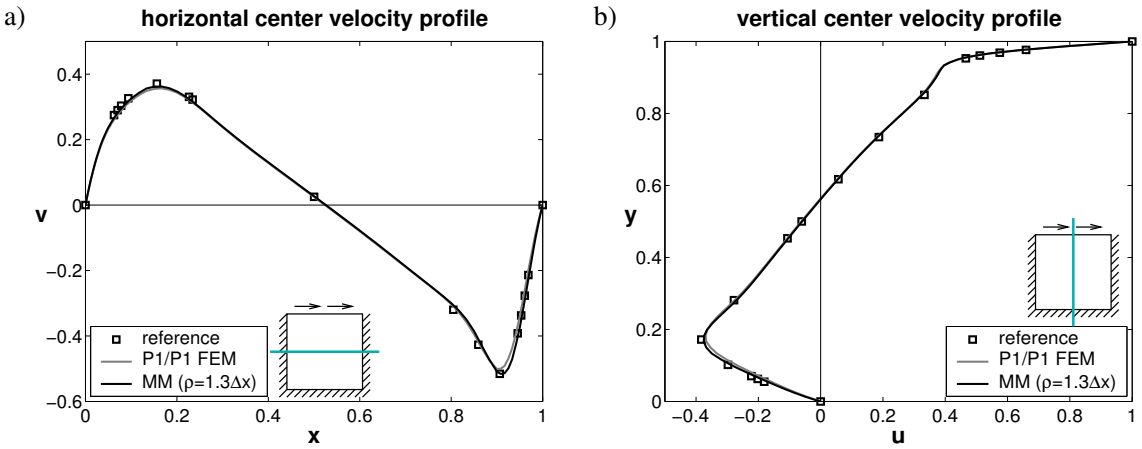


Figure 28: Velocity profiles for u and v along $y = 0.5$ and $x = 0.5$ respectively (for $\rho = 1.3\Delta x$ and 96×96 irregular nodes).

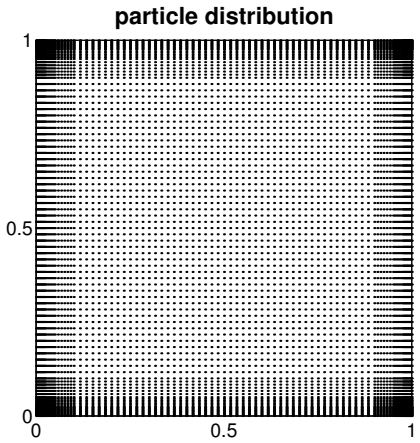


Figure 29: Irregular node distribution for the driven cavity test case (96×96 nodes).

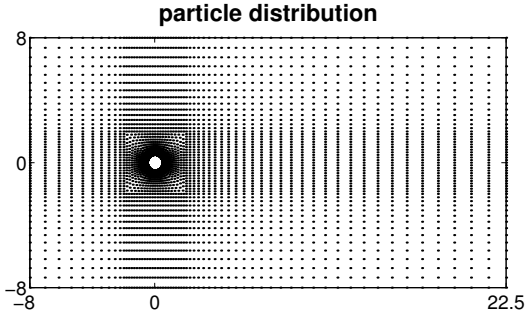


Figure 30: Irregular node distribution for the flow past a cylinder test case (6268 nodes).

h_ρ-versions. A clear convergence towards the benchmark solution can be found and it may be seen that the meshfree solution is more accurate than the P1/P1 element. Comparing the results for the regular 101×101 mesh with the irregular 96×96 mesh, one can clearly see the improvement in the solution for the anisotropic supports. Hence, as well as using high-aspect ratio elements in meshbased methods in order to resolve boundary layers successfully, high-aspect anisotropic supports should be used in the meshfree context analogously.

Test case: flow past a cylinder The "steady-state" solution for flow past a cylinder at $Re = 100$ is computed, as presented in [52]. Instationary computations at this Reynolds-number lead to periodic flow patterns known as the Kármánn vortex street, but this is not considered here. Inflow and outflow boundary conditions are applied on the left and right side of the rectangular domain respectively. Slip boundary conditions are applied at the upper and lower boundary, no-slip boundary conditions are applied at the cylinder surface. The cylinder is placed as shown in Figure 30, where also the irregular particle distribution for this test case is shown. Figure 31 shows a typical result for the velocity and pressure distribution around the cylinder. Further details of this test case are not described, because our interest is rather in the stabilization characteristics than in obtaining benchmark solutions.

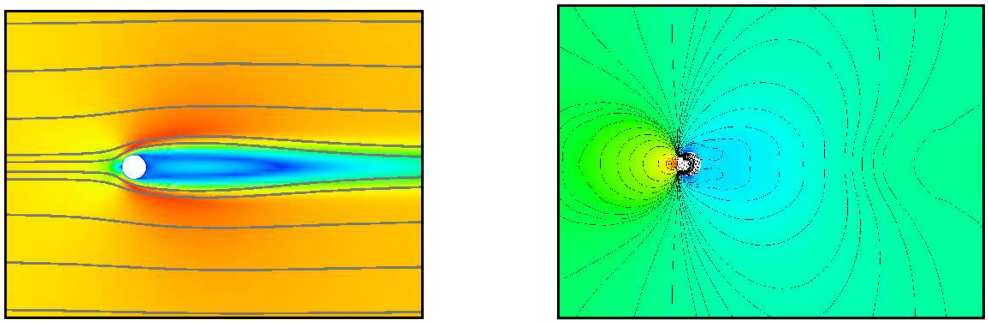


Figure 31: Stationary example solution for the velocity and pressure field around the cylinder at $Re = 100$.

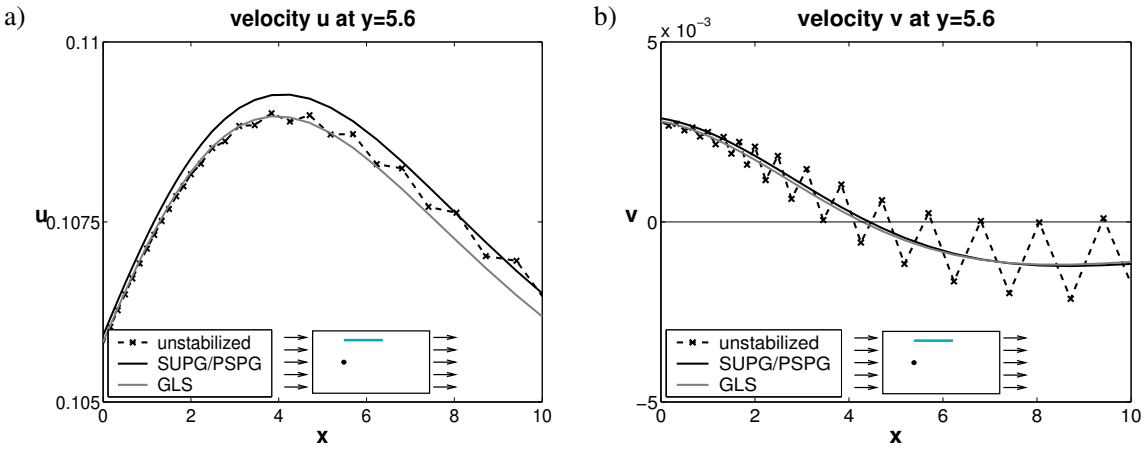


Figure 32: Velocity profiles for u and v at $y = 5.6$.

The supports of the meshfree shape functions are anisotropic as defined above for the irregular driven cavity test case. Figure 32 depicts oscillatory unstabilized velocity profiles for u and v at $y = 5.6$. Both, SUPG/PSPG and GLS stabilization smooth out the oscillations successfully.

We conclude that the stabilization with small dilatation parameters works successfully also for anisotropic supports.

7 Conclusion

The main purpose of this paper is to establish Eulerian meshfree methods. To apply these methods successfully, the problem of stabilization has to be overcome. In this paper we review the development of stabilized methods in the meshbased context and try to apply the same steps to meshfree methods.

The need for stabilization is described and the standard techniques that are used. Most importantly, the SUPG, PSPG and GLS stabilization methods are considered. It is developed in detail, how stabilization may be deduced in an illustrative way. Thereby, the local coth-formula for τ is obtained, being the basis of many alternative versions of τ-formulas. Element and nodal

stabilization are compared. The first is the standard procedure in finite element methods, the latter must be used for meshfree methods. Important steps and results of mathematical analysis of stabilized formulations are reviewed.

Then, with the knowledge of the most important aspects of stabilization in meshbased methods, we turn our attention to meshfree methods. It is found that standard stabilization schemes are well-suited for the stabilization of both meshfree and meshbased methods. However, the aspect of the stabilization parameter τ, being a crucial ingredient in stabilized methods, has to be considered with care. It can easily be shown that the same procedure to obtain the important coth-formula, being a local stabilization criterion, applied to meshfree methods results in a global criterion. That is, there is a dependence of τ on all nodes, most importantly the global downstream node. In contrast, local stabilization criteria only rely on the relative positions of neighboring nodes. Therefore, it can not be expected in general that standard formulas for τ, derived in the meshbased context, work also successfully with meshfree methods.

However, it is shown that meshfree shape functions with small dilatation parameters, i.e. with small supports, justify the usage of standard meshfree stabilization parameters τ also in the meshfree context. The numerical results agree very well with our theoretical approach.

We believe that the door is now a little more open, to use Eulerian meshfree methods to successfully approximate a large variety of differential equations. This is also one more step towards successful coupled meshfree-meshbased methods, where the advantages of both methods ban be unified. The current approaches are often restricted to fully Lagrangian formulations, not being applicable to fluid mechanics, where large distortions occur, or to mixed Eulerian-Lagrangian formulations which does not seem to be the most straightforward way. Coupled Eulerian meshfree-meshbased methods can be expected to generate promising results in the future.

8 References

References

[1] Aliabadi, S.K.; Ray, S.E.; Tezduyar, T.E.: SUPG finite element computation of viscous compressible flows based on the conservation and entropy variables formulations. *Comput. Mech.*, **11**, 300 – 312, 1993.

[2] Babuška, I.: Error-bounds for finite element method. *Numer. Math.*, **16**, 322 – 333, 1971.

[3] Belytschko, T.; Krongauz, Y.; Organ, D.; Fleming, M.; Krysl, P.: Meshless Methods: An Overview and Recent Developments. *Comp. Methods Appl. Mech. Engrg.*, **139**, 3 – 47, 1996.

[4] Belytschko, T.; Lu, Y.Y.; Gu, L.: Element-free Galerkin Methods. *Internat. J. Numer. Methods Engrg.*, **37**, 229 – 256, 1994.

[5] Brezzi, F.: On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers. *RAIRO Anal. Numér.*, **R-2**, 129 – 151, 1974.

[6] Brezzi, F.; Bristeau, M.O.; Franca, L.P.; Mallet, M.; Rogé, G.: A relationship between stabilized finite element methods and the Galerkin method with bubble functions. *Comp. Methods Appl. Mech. Engrg.*, **96**, 117 – 129, 1992.

[7] Brooks, A.N.; Hughes, T.J.R.: Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comp. Methods Appl. Mech. Engrg.*, **32**, 199 – 259, 1982.

[8] Christie, I.; Griffiths, D.F.; Mitchell, A.R.; Zienkiewicz, O.C.: Finite element methods for second order differential equations with significant first derivatives. *Internat. J. Numer. Methods Engrg.*, **10**, 1389 – 1396, 1976.

[9] Franca, L.P.; Farhat, C.: Bubble functions prompt unusual stabilized finite element methods. *Comp. Methods Appl. Mech. Engrg.*, **123**, 299 – 308, 1995.

[10] Franca, L.P.; Frey, S.L.: Stabilized finite element methods: II. The incompressible Navier-Stokes equations. *Comp. Methods Appl. Mech. Engrg.*, **99**, 209 – 233, 1992.

[11] Franca, L.P.; Frey, S.L.; Hughes, T.J.R.: Stabilized finite element methods: I. Application to the advective-diffusive model. *Comp. Methods Appl. Mech. Engrg.*, **95**, 253 – 276, 1992.

[12] Franca, L.P.; Hughes, T.J.R.: Two classes of mixed finite element methods. *Comp. Methods Appl. Mech. Engrg.*, **69**, 89 – 129, 1988.

[13] Fries, T.P.; Matthies, H.G.: Classification and Overview of Meshfree Methods. Informatikbericht-Nr. 2003-03, Technical University Braunschweig, (http://opus.tu-bs.de/opus/volltexte/2003/418/), Brunswick, 2003.

[14] Ghia, U.; Ghia, K.N.; Shin, C.T.: High-Re solutions for incompressible flow using the Navier-Stokes equations and a multi-grid method. *J. Comput. Phys.*, **48**, 387 – 411, 1982.

[15] Gingold, R.A.; Monaghan, J.J.: Kernel Estimates as a Basis for General Particle Methods in Hydrodynamics. *J. Comput. Phys.*, **46**, 429 – 453, 1982.

[16] Günther, F.C.: *A Meshfree Formulation for the Numerical Solution of the Viscous Compressible Navier-Stokes Equations*. Dissertation, Northwestern University, Evanston, IL, 1998.

[17] Hansbo, P.; Szepessy, A.: A velocity-pressure streamline diffusion finite element method for the incompressible Navier-Stokes equations. *Comp. Methods Appl. Mech. Engrg.*, **84**, 175 – 192, 1990.

[18] Harari, I.; Hughes, T.J.R.: What are C and h?: Inequalities for the analysis and design of finite element methods. *Comp. Methods Appl. Mech. Engrg.*, **97**, 157 – 192, 1992.

[19] Harten, A.: On the symmetric form of systems of conservation laws with entropy. *J. Comput. Phys.*, **49**, 151 – 164, 1983.

[20] Heinrich, J.C.; Huyakorn, P.S.; Zienkiewicz, O.C.; Mitchell, A.R.: An 'upwind' finite element scheme for two-dimensional convective transport equation. *Internat. J. Numer. Methods Engrg.*, **11**, 131 – 143, 1977.

[21] Hirsch, C.: *Numerical Computation of Internal and External Flows: Computational Methods for Inviscid and Viscous Flows*, Vol. 2. John Wiley & Sons, Chichester, 1988.

[22] Hirsch, C.: *Numerical Computation of Internal and External Flows: Fundamentals of Numerical Discretization*, Vol. 1. John Wiley & Sons, Chichester, 1988.

[23] Huerta, A.; Fernández-Méndez, S.M.: Time accurate consistently stabilized mesh-free methods for convection-dominated problems. *Internat. J. Numer. Methods Engrg.*, **50**, 1 – 18, 2001.

[24] Hughes, T.J.R.: A simple scheme for developing 'upwind' finite elements. *Internat. J. Numer. Methods Engrg.*, **12**, 1359 – 1365, 1978.

[25] Hughes, T.J.R.: Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Comp. Methods Appl. Mech. Engrg.*, **127**, 387 – 401, 1995.

[26] Hughes, T.J.R.; Brooks, A.N.: A multidimensional upwind scheme with no crosswind diffusion. In *ASME Monograph AMD-34*. (Hughes, T.J.R., Ed.), Vol. 34, ASME, New York, NY, 1979.

[27] Hughes, T.J.R.; Franca, L.P.: 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. Methods Appl. Mech. Engrg.*, **65**, 85 – 96, 1987.

[28] Hughes, T.J.R.; Franca, L.P.; Balestra, M.: A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuška-Brezzi condition: a stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations. *Comp. Methods Appl. Mech. Engrg.*, **59**, 85 – 99, 1986.

[29] Hughes, T.J.R.; Franca, L.P.; Hulbert, G.M.: A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/Least-squares method for advective-diffusive equations. *Comp. Methods Appl. Mech. Engrg.*, **73**, 173 – 189, 1989.

[30] Hughes, T.J.R.; Franca, L.P.; Mallet, M.: A new finite element formulation for computational fluid dynamics: I. Symmetric forms of the compressible Euler and Navier-Stokes equations and the second law of thermodynamics. *Comp. Methods Appl. Mech. Engrg.*, **54**, 223 – 234, 1986.

[31] Hughes, T.J.R.; Franca, L.P.; Mallet, M.: A new finite element formulation for computational fluid dynamics: VI. Convergence analysis of the generalized SUPG formulation for linear time-dependent multidimensional advective-diffusive systems. *Comp. Methods Appl. Mech. Engrg.*, **63**, 97 – 112, 1987.

[32] Hughes, T.J.R.; Mallet, M.: A new finite element formulation for computational fluid dynamics: III. The generalized streamline operator for multidimensional advective-diffusive systems. *Comp. Methods Appl. Mech. Engrg.*, **58**, 305 – 328, 1986.

[33] Hughes, T.J.R.; Mallet, M.: A new finite element formulation for computational fluid dynamics: IV. A discontinuity-capturing operator for multidimensional advective-diffusive systems. *Comp. Methods Appl. Mech. Engrg.*, **58**, 329 – 336, 1986.

[34] Hughes, T.J.R.; Mallet, M.; Mizukami, A.: A new finite element formulation for computational fluid dynamics: II. Beyond SUPG. *Comp. Methods Appl. Mech. Engrg.*, **54**, 341 – 355, 1986.

[35] Hughes, T.J.R.; Tezduyar, T.E.: Finite element methods for first-order hyperbolic systems with particular emphasis on the compressible Euler equations. *Comp. Methods Appl. Mech. Engrg.*, **45**, 217 – 284, 1984.

[36] Johnson, C.; Nävert, U.; Pitkäranta, J.: Finite element methods for linear hyperbolic problems. *Comp. Methods Appl. Mech. Engrg.*, **45**, 285 – 312, 1984.

[37] Johnson, C.; Saranen, J.: Streamline diffusion methods for the incompressible Euler and Navier-Stokes equations. *Math. Comput.*, **47**(175), 1 – 18, 1986.

[38] Kelly, D.W.; Nakazawa, S.; Zienkiewicz, O.C.: A note on upwinding and anisotriopic balancing dissipation in finite element approximations to convective diffusion problems. *Internat. J. Numer. Methods Engrg.*, **15**, 1705 – 1711, 1980.

[39] Lancaster, P.; Salkauskas, K.: Surfaces Generated by Moving Least Squares Methods. *Math. Comput.*, **37**, 141 – 158, 1981.

[40] Li, S.; Liu, W.K.: Reproducing Kernel Hierarchical Partition of Unity, Part I – Formulation and Theory. *Internat. J. Numer. Methods Engrg.*, **45**, 251 – 288, 1999.

[41] Li, S.; Liu, W.K.: Reproducing Kernel Hierarchical Partition of Unity, Part II – Applications. *Internat. J. Numer. Methods Engrg.*, **45**, 289 – 317, 1999.

[42] Lu, Y.Y.; Belytschko, T.; Gu, L.: A New Implementation of the Element Free Galerkin Method. *Comp. Methods Appl. Mech. Engrg.*, **113**, 397 – 414, 1994.

[43] Lucy, L.B.: A numerical approach to the testing of the fission thesis. *Astronom. J.*, **82**(12), 1013 – 1024, 1977.

[44] Mizukami, A.; Hughes, T.J.R.: A Petrov-Galerkin finite element method for convection-dominated flows: an accurate upwinding technique for satisfying the maximum principle. *Comp. Methods Appl. Mech. Engrg.*, **50**, 181 – 193, 1985.

[45] Monaghan, J.J.: An introduction to SPH. *Comput. Phys. Comm.*, **48**, 89 – 96, 1988.

[46] Nävert, U.: *Finite element methods for convection-diffusion problems.* Dissertation, Dep. of Computer Science, Chalmers University of Technology, Göteborg, Sweden, 1982.

[47] Oñate, E.; Idelsohn, S.; Zienkiewicz, O.C.; Taylor, R.L.: A Finite Point Method in Computational Mechanics. Applications to Convective Transport and Fluid Flow. *Internat. J. Numer. Methods Engrg.*, **39**, 3839 – 3866, 1996.

[48] Oñate, E.; Idelsohn, S.; Zienkiewicz, O.C.; Taylor, R.L.; Sacco, C.: A Stabilized Finite Point Method for Analysis of Fluid Mechanics Problems. *Comp. Methods Appl. Mech. Engrg.*, **139**, 315 – 346, 1996.

[49] Rice, J.G.; Schnipke, R.J.: A monotone streamline upwind finite element method for convection-dominated flows. *Comp. Methods Appl. Mech. Engrg.*, **48**, 313 – 327, 1985.

[50] Shakib, F.; Hughes, T.J.R.; Johan, Z.: A new finite element formulation for computational fluid dynamics: X. The compressible Euler and Navier-Stokes equations. *Comp. Methods Appl. Mech. Engrg.*, **89**, 141 – 219, 1991.

[51] S.Mittal: On the performance of high aspect ratio elements for incompressible flows. *Comp. Methods Appl. Mech. Engrg.*, **188**, 269 – 287, 2000.

[52] Tezduyar, T.E.: Stabilized Finite Element Formulations for Incompressible Flow Computations. In *Advances in Applied Mechanics*. (Hutchinson, J.W.; Wu, T.Y., Eds.), Vol. 28, Academic Press, New York, NY, 1992.

[53] Tezduyar, T.E.; Mittal, S.; Ray, S.E.; Shih, R.: Incompressible flow computations with stabilized bilinear and linear equal-order-interpolation velocity-pressure elements. *Comp. Methods Appl. Mech. Engrg.*, **95**, 221 – 242, 1992.

[54] Tezduyar, T.E.; Osawa, Y.: Finite element stabilization parameters computed from element matrices and vectors. *Comp. Methods Appl. Mech. Engrg.*, **190**, 411 – 430, 2000.

[55] Tezduyar, T.E.; Park, Y.J.: Discontinuity-capturing finite element formulations for nonlinear convection-diffusion-reaction equations. *Comp. Methods Appl. Mech. Engrg.*, **59**, 307 – 325, 1986.

[56] Zhou, G.: How accurate is the streamline diffusion finite element method? *Math. Comput.*, **66**(217), 31 – 44, 1997.

Technische Universität Braunschweig
Informatik-Berichte ab Nr. 2000-02

2000-02	J. Saperia, J. Schönwälder	Policy-Based Enhancements to the SNMP Framework
2000-03	A. Casties	Finite-Element-Interpolation der räumlichen Dichten eines Vielteilchensystems auf ungeordneten Gittern
2000-04	J. Koslowski	A 2-dimensional view of the Chu-construction
2000-05	S. Eckstein, P. Ahlbrecht, K. Neumann	Von parametrisierten Spezifikationen zu generierten Informationssystemen: ein Anwendungsbeispiel
2000-06	F. Strauß, J. Schönwälder, M. Mertens	JAX - A Java AgentX Sub-Agent Toolkit
2000-07	F. Strauß	Advantages and Disadvantages of the Script MIB Infrastructure
2000-08	T. Gehrke, U. Goltz	High-Level Sequence Charts with Data Manipulation
2000-09	T. Firley	Regular languages as states for an abstract automaton
2001-01	K. Diethers	Tool-Based Analysis of Timed Sequence Diagrams
2002-01	R. van Glabbeek, U. Goltz	Well-behaved Flow Event Structures for Parallel Composition and Action Refinement
2002-02	J. Weimar	Translations of Cellular Automata for Efficient Simulation
2002-03	H. G. Matthies, M. Meyer	Nonlinear Galerkin Methods for the Model Reduction of Nonlinear Dynamical Systems
2002-04	H. G. Matthies, J. Steindorf	Partitioned Strong Coupling Algorithms for Fluid-Structure-Interaction
2002-05	H. G. Matthies, J. Steindorf	Partitioned but Strongly Coupled Iteration Schemes for Nonlinear Fluid-Structure Interaction
2002-06	H. G. Matthies, J. Steindorf	Strong Coupling Methods
2002-07	H. Firley, U. Goltz	Property Preserving Abstraction for Software Verification
2003-01	M. Meyer, H. G. Matthies	Efficient Model Reduction in Non-linear Dynamics Using the Karhunen-Loève Expansion and Dual-Weighted-Residual Methods
2003-02	C. Täubner	Modellierung des Ethylen-Pathways mit UML-Statecharts
2003-03	T.-P. Fries, H. G. Matthies	Classification and Overview of Meshfree Methods
2003-04	A. Keese, H. G. Matthies	Fragen der numerischen Integration bei stochastischen finiten Elementen für nichtlineare Probleme
2003-05	A. Keese, H. G. Matthies	Numerical Methods and Smolyak Quadrature for Nonlinear Stochastic Partial Differential Equations
2003-06	A. Keese	A Review of Recent Developments in the Numerical Solution of Stochastic Partial Differential Equations (Stochastic Finite Elements)
2003-07	M. Meyer, H. G. Matthies	State-Space Representation of Instationary Two-Dimensional Airfoil Aerodynamics
2003-08	H. G. Matthies, A. Keese	Galerkin Methods for Linear and Nonlinear Elliptic Stochastic Partial Differential Equations
2003-09	A. Keese, H. G. Matthies	Parallel Computation of Stochastic Groundwater Flow
2003-10	M. Mutz, M. Huhn	Automated Statechart Analysis for User-defined Design Rules
2004-01	T.-P. Fries, H. G. Matthies	A Review of Petrov-Galerkin Stabilization Approaches and an Extension to Meshfree Methods