Mathematisches Forschungsinstitut Oberwolfach

Report No. 41/2015

DOI: 10.4171/OWR/2015/41

Recent Developments in the Numerics of Nonlinear Hyperbolic Conservation Laws

Organised by
Rémi Abgrall, Zürich
Willem Hundsdorfer, Amsterdam
Andreas Meister, Kassel
Thomas Sonar, Braunschweig

13 September – 19 September 2015

Abstract. The development of reliable numerical methods for the simulation of real life problems requires both a fundamental knowledge in the field of numerical analysis and a proper experience in practical applications as well as their mathematical modeling.

Thus, the purpose of the workshop was to bring together experts not only from the field of applied mathematics but also from civil and mechanical engineering working in the area of modern high order methods for the solution of partial differential equations or even approximation theory necessary to improve the accuracy as well as robustness of numerical algorithms.

Mathematics Subject Classification (2010): 34-99, 35F25, 35L65, 41A10, 76-99.

Introduction by the Organisers

The workshop *Recent Developments in the Numerics of Nonlinear Hyperbolic Conservation Laws*, organised by Rémi Abgrall (Zürich), Willem Hundsdorfer (Amsterdam), Andreas Meister (Kassel) and Thomas Sonar (Braunschweig) was held September 14th–September 19th, 2015. This meeting was well attended with over 50 participants with broad geographic representation from all continents.

Since modern numerical methods like Discontinuous Galerkin or Spectral Element Finite Difference methods are based on orthogonal polynomials on simplices and use modal filters as well as methods for edge detection and many more mathematical devices from different areas of research we decided to invite renowned

researchers from numerical methods for partial differential equations and approximation theory. Furthermore, to couple mathematical precision with a large range of applicability we also invited scientist from engineering departments working in the field of numerical schemes.

The talks ranged from new Runge-Kutta solvers, new filters and edge detection algorithms, Discontinuous Galerkin methods, Spectral Difference methods, Finite Difference operators, implicit solvers, and finite volume methods to the modeling of shallow water flow, viscous as well as inviscid fluid flow and solid mechanics. Discussions were lively and many different research areas met for the first time resulting in interesting talks and contacts.

The workshop was a tremendous success and we are looking forward to repeat this kind of conference in Oberwolfach again in a few years.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1049268, "US Junior Oberwolfach Fellows".

Recent Developments in the Numerics of Nonlinear Hyperbolic Conservation Laws

Table of Contents

Abstracts

Remove the mass matrix from finite element calculations of covection dominated transport equations

Rémi Abgrall
(joint work with P. Bacigaluppi, S. Tokareva)

We are interested in the numerical approximation of problems of the type: find the solution $u : \Omega \times \mathbb{R}^+ \to \mathbb{R}^p$, where Ω is a regular open set of $\mathbb{R}^d$, $d = 1, 2, 3$ and $p \in \mathbb{N}$ of:

$$\begin{aligned} &\frac{\partial \mathbf{u}}{\partial t} + \text{ div } \mathbf{f}(\mathbf{u}) = 0 \\ &\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}) \text{ in } \mathbf{x} \in \Omega \\ &\mathbf{u} = g \text{ on } \partial\Omega \end{aligned} \tag{1}$$

Let us give a more precise meaning to this. In order to impose a condition $u = g$ on the inflow boundary, we assume, for any real a and b, and any vector $\mathbf{n}$, the existence of $\overline{\nabla_u \mathbf{f}}(a, b)$ such that

- $\overline{\nabla_u \mathbf{f}}(a, a) = \nabla_u \mathbf{f}(a) \cdot \mathbf{n}$ and,
- $\mathbf{f}(b) \cdot \mathbf{n} - \mathbf{f}(b) \cdot \mathbf{n} = \overline{\nabla_u \mathbf{f}}(a, b)\ (b - a)$.

This is a reminscence of the Roe average. As soon as $\mathbf{f}$ is C^1, such average exists and is unique. Then the boundary conditions are set such that

$$\max\big(0, \overline{\nabla_u \mathbf{f}}(\mathbf{u}, g) \cdot \mathbf{n}\big)(\mathbf{u} - g) = 0 \text{ on } \Gamma = \partial\Omega \tag{2}$$

We introduce the flux $\mathcal{F}$ defined by:

$$\mathcal{F}(a, b) = \frac{1}{2}\left(\mathbf{f}(a) \cdot \mathbf{n} + \mathbf{f}(b) \cdot \mathbf{n} + \big|\overline{\nabla_u \mathbf{f}}\big|(a, b)\right).$$

Solutions of (1) should be understood in the weak sense. We introduce the space

$$C^1_{0,t}(\Omega \times \mathbb{R}^+) = \{\varphi \in C^1(\Omega \times \mathbb{R}^+) \text{ such that there exists } T_u \text{for which for any } \mathbf{x} \in \Omega, u(\mathbf{x}, t) = 0 \text{ when } t > T_u\}.$$

The weak form of (1) is: find $\mathbf{u} \in L^1(\Omega \times \mathbb{R}^+) \cap L^\infty_{loc}(\Omega \times \mathbb{R}^+)$ such that for any $\varphi \in C^1_{0,t}(\Omega \times \mathbb{R}^+)$

$$\int_{\Omega \times \mathbb{R}^+} \nabla\varphi \cdot \mathbf{f}(u) d\mathbf{x}dt + \int_\Omega \varphi(\mathbf{x}, 0) u_0(\mathbf{x}) d\mathbf{x} - \int_{\Gamma \times \mathbb{R}!^+} \varphi(\mathbf{x}, t)\big(\mathbf{f}(\mathbf{u}) \cdot \mathbf{n} - \mathcal{F}(\mathbf{u}, g, \mathbf{n})\big) d\omega = 0 \tag{3}$$

Typical examples of (1) are the Burgers equation where $p = 1$, $d = 2$ and $\mathbf{f}(u) = \big(\frac{u^2}{2}, u\big)$, the Kurganov-Petrov-Petrova system where $p = 1$, $d = 2$ and $\mathbf{f}(u) = \big(\cos u, \sin u\big)$, the Euler equations of fluid dynamics where $d = 1, 2, 3$, $p = 2 + d$ in the simplest case, $u = (\rho, \rho\vec{v}, E)^T$ where ρ is the density, $\vec{v}$ the

velocity field and E the total energy, i.e. the sum of the internal energy e and the kinetic energy. The flux is then

$$\mathbf{f}(u)\begin{pmatrix}\rho\vec{v}\\ \rho\vec{v}\otimes\vec{v}+p\ \mathrm{Id}\\ (E+p)\vec{v}\end{pmatrix}.$$

The system is closed by the equation of state $p = p(e, \rho)$ which has itself to satisfy convexity assumptions [7]. In what follows, for ease of exposition, we are only interested in scalar solutions, i.e $p = 1$, but the extension to the general case is straightforward.

When one is interested in solving the unsteady problem (1) using a finite element method, several choices need to be made. First, one has to define the approximation space V^h in which we are looking for the approximate solution. In order to define it, one first need to consider a tessellation $\mathcal{T}_h$ of Ω, i.e. a set of elements (typically simplices or more complicated geometrical objects) such that the intersection of two elements has empty interior when they are not identical, and such that they cover Ω [1].Once this choice is met, V^h is usually defined as:

$$V^h = \{ \text{ for any } K, u^h_{|K} \text{ is polynomial of degree } k\}.$$

Here again we assume that the polynomial degree is the same on any element, but this assumption might not be essential. The next step is to decide whether we are interested in globally continuous function, and in that case V^h must be replaced by

$$V^h = \{ \text{ for any } K, u^h_{|K} \text{ is polynomial of degree } k\} \cap C^0(\Omega),$$

or if the global continuity requirement is not required.

In order to approximate the weak form (2), discretization choices need to be made according to the structure of V^h. In the discontinuous case,a typical example is the Discontinuous galerkin method, see [6, 5]. In the continuous case, a typical example is the stabilized Galerkin methods, i.e. either the stream-line diffusion method [8, 9], or [4] where different stabilisation technique are used or the purely nonlinear method [3, 1, 2] which is is original design to compute steady discontinuous solutions.

In most cases, one arrives to solving problems of the type

$$M\frac{dU}{dt} + F = 0 \tag{4}$$

where U is the vector of unknowns, F is the approximation of the divergence term and M is a mass matrix. When the Discontinuous Galerkin method is employed, M is block diagonal, so computing M^{-1} is not a big deal. This is not the case in the continuous setting where M is only sparse. After that, one can use standard ODE solvers.

The purpose of the talk was to show how to avoid to invert the mass matrix, extending the work of [10]. More precisely, instead of considering (4), we introduce a slight modification of it which does not need any inversion of mass matrix, and

[1]Here we implicitely make the assumption that Ω is polygonal, but this is not essential.

this modification has the property that the original accuracy is kept, as well as the conservation properties of the method. With respect to a standard approximation, one has to slightly change the approximation representation: the basis functions φ_σ associated to any degree of freedom σ must be such that for any K contained in the support of φ_σ, we must have

$$\int_K \varphi_\sigma d\mathbf{x} > 0.$$

This condition is not met in general for $\mathbb{P}^k$ simplifies elements , except for $k = 1$, but in the case $k > 0$ one can replace them, without modifying V^h, by Bézier basis functions. In the case of $\mathbb{Q}^k$ elements, the degree of free

We have show several examples, in the linear (station case) and the non linear case (KPP problem), that the method is effective and lead to the expected accuracy (for linear problems).

References

[1] R. Abgrall and D. de Santis. High-order preserving residual distribution schemes for advection-diffusion scalar problems on arbitrary grids. *SIAM I. Sci. Comput.*, 36(3):A955–A983, 2014.

[2] R. Abgrall and D. de Santis. Linear and non-linear high order accurate residual distribution schemes for the discretization of the steady compressible navier-stokes equations. *J. Comput. Phys.*, 283:329–359, 2015.

[3] R. Abgrall, A. Larat, and M. Ricchiuto. Construction of very high order residual distribution schemes for steady inviscid flow problems on hybrid unstructured meshes. *J. Comput. Phys.*, 230(11):4103–4136, 2011.

[4] E. Burman and P. Hansbo. Edge stabilisation for Galerkin approximantion of convection-diffusion-reaction problems. *Computer methods in applied mechanics and engineering*, 193:1437–1453, 2004.

[5] B. Cockburn, S. Hou, and C.-W. Shu. TVB Runge-Kutta local projection discontinuous finite element method for conservation laws IV: the multidimensional case. *Math. Comp.*, 54:545–581, 1990.

[6] B. Cockburn and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: General framework. *Math. Comp.*, 52:411–435, 1989.

[7] E. Godlewski and P.-A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws.* New York, NY: Springer, 1996.

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

[9] C. Johnson, U. Nävert, and J. Pitkäranta. Finite element methods for linear hyperbolic problems. *Computer methods in applied mechanics and engineering*, 45:285–312, 1985.

[10] Mario Ricchiuto and Rémi Abgrall. Explicit Runge-Kutta residual-distribution schemes for time dependent problems. *Journal of Computational Physics*, 229(16):5653–5691, 2010.

Efficient Time Integration of IMEX Type using Exponential Integrators for Compressible, Viscous Flow Simulation

Veronika Diba
(joint work with Andreas Meister, Sigrun Ortleb, Philipp Birken)

The simulation of unsteady compressible, viscous flow in real applications such as the flow around a windturbine or gas quenching is still a challenging task due to the geometrical complexity and the size of the resulting systems [7, 8]. The spatial discretization of the compressible Navier-Stokes equations using the Discontinuous Galerkin method leads to a large system of ordinary differential equations (ODEs)

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{U}(t) = \mathrm{F}(\mathrm{U}(t)), \quad \mathrm{U} : \mathbb{R} \to \mathbb{R}^N,\ \mathrm{F} : \mathbb{R}^N \to \mathbb{R}^N,\ t \in [t_0, T]. \tag{1}$$

That system suffers from the so-called geometry-induced stiffness due to the drastically varying scales of the grid elements. Accordingly, explicit time integration methods become inefficient, since their time step size is determined by the smallest grid element to ensure stability [1]. Implicit time integration allows much greater time step sizes but at a crucial increase of the computational cost and memory requirement per step. As shown by Kanevsky et. al. [2] the key to an efficient time integration scheme in that scope promises to be the domain-based implicit-explicit (IMEX) idea, which suggests to split the computational domain into a stiff region associated with an implicit time integration scheme and a nonstiff region linked to explicit time stepping, see figure 1. The stiff region consists of the smallest grid elements and the nonstiff of the bigger ones.

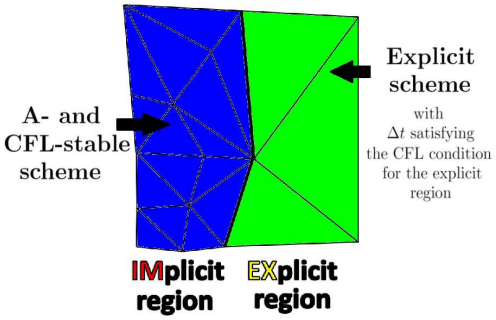


Figure 1. IMEX idea

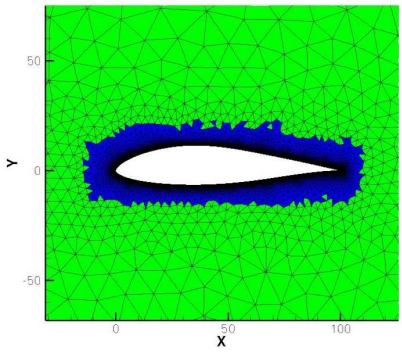


Figure 2. Wind turbine

Our goal is the construction of an IMEX type method, which is efficient even without preconditioning and thus allowing for good parallelization. The recently developed EPIRK schemes by Tokman [6] seem to be suited for the application on the implicit region. They belong to the class of exponential integrators while applying Krylov subspace projections and other strategies to gain efficiency. Their

advantages are that they are theoretically unconditionally A-stable, do not need a preconditioner, can be implemented in a Jacobian-free mode, and are shown to be more efficient than cassical schemes like explicit Runge-Kutta (RK), BDF or Rosenbrock methods for sufficiently stiff problems [3].

An EPIRK scheme of stage s with a set of coefficients $b_i, a_{ij}, g_{ij}, p_{sij}$ for $i, j \in \{1, 2, \ldots, s\}$ can be written as

$$\mathrm{U}_{n+1} = \mathrm{U}_n + b_1\psi_{s1}(g_{s1}\mathbf{A}_n h)h\mathrm{F}_n + h\sum_{k=2}^{s} b_k\psi_{sk}(g_{sk}\mathbf{A}_n h)\Delta^{(k-1)}\mathrm{R}_n(\mathrm{U}_n) \tag{2}$$

$$\mathrm{Y}_i = y_n + a_{i1}\psi_{i1}(g_{i1}\mathbf{A}_n h)h\mathrm{F}_n + h\sum_{j=2}^{i} a_{ij}\psi_{ij}(g_{ij}\mathbf{A}_n h)\Delta^{(j-1)}\mathrm{R}_n(\mathrm{U}_n), \tag{3}$$

$i = 1, \ldots, s-1$, where $\mathrm{R}_n(\mathrm{Y}) := \mathrm{F}(\mathrm{Y}) - \mathrm{F}_n - \mathbf{A}_n(\mathrm{Y} - \mathrm{U}_n)$, h denotes the time step size, $\mathbf{A}_n := \mathrm{F}'(\mathrm{U}_n)$ and $\Delta^j\mathrm{R}_n(\mathrm{U}_n)$ is the j-th forward difference through the nodes $U_n, \mathrm{Y}_1, \ldots, \mathrm{Y}_j$. At last, the functions ψ_{jk} for $j, k = 1, \ldots, s$ are defined by

$$\psi_{jk}(z) := \sum_{i=1}^{s} p_{jki}\varphi_i(z) \quad \text{and} \quad \varphi_i(z) := \sum_{k=0}^{\infty} \frac{z^k}{(k+i)!}, \quad i = 0, 1, \ldots, s.$$

Of great importance for applying those schemes in an IMEX context are the Courant-Friedrichs-Lewy (CFL-) stability, the conservativity, and the stability in practice, which have been examined by us.

As for the A-stability, we figured out that it is not given in practice any more, since we are approximating the φ-functions by a Krylov projection P_m into the Krylov subspace K_m

$$\varphi_k(\mathbf{A}_n h)\mathrm{F}_n \approx P_m(\varphi_k(\mathbf{A}h)\mathrm{F}_n) = \sum_{i=0}^{m-1} \frac{(h\mathbf{A}_n)^i}{(i+k)!}F_n,$$

$k = 0, 1, \ldots, s$. Consequently, the stability function consists of the Krylov projection P_m of $e^{\lambda h}$:

$$\mathrm{U}_{n+1} = P_m(e^{\lambda h})\mathrm{U}_n.$$

The stability region can be determined by computing the roots $z \in \mathbb{C}$ of

$$\sum_{j=0}^{m-1} \frac{z^j}{j!} - e^{i\theta}, \quad \theta \in [0, 2\pi).$$

Obviously, the stability region depends on the dimension m of the Krylov subspace and increases with increasing m.

The CFL-stability also depends on m, which was shown by us for the one-dimensional advection-diffusion equation discretized by Finite Differences. The CFL-condition for a linear advection with speed a discretized on an equidistant grid with mesh width Δx was shown to be given by $h \leq \frac{m\Delta x}{|a|}$. Unconditional CFL-stability is given for $m \geq N - 1$, if N is the size of the resulting ODE system.

Now let's have a look at the conservativity. We call a time stepping scheme *conservative*, if

$$\sum_{j=1}^{N}(\mathrm{U}_{n+1})_j = \sum_{j=1}^{N}(\mathrm{U}_n)_j \quad \forall\, n \in \mathbb{N}_0 \tag{4}$$

holds for a system (1) with a right-hand side satisfying

$$\sum_{j=1}^{N}\big(\mathrm{F}(\mathrm{U})\big)_j = 0 \quad \forall\, \mathrm{U} \in \mathbb{R}^N.$$

We showed that the EPIRK schemes fulfill that property (4).

Now we can apply those schemes in an IMEX-context, what means that we first split the right-hand side of the ODE system (1) in the following way:

$$\mathrm{F}(\mathrm{U}) = \mathrm{F}^E(\mathrm{U}) + \mathrm{F}^I(\mathrm{U}).$$

In F^E we put all the terms connected to the nonstiff region, i.e. all integrals over the larger grid elements and the other terms will be collected in the implicit part F^I.

Our first idea was to apply an explicit Runge-Kutta method to the vector F^E and an EPIRK scheme to F^I. That IMEX-EPIRK scheme turned out to be of at most first order of convergence and can be written as

$$\mathrm{U}_{n+1} = \mathrm{U}_n + h(\mathrm{F}^E + \varphi_1(h\mathbf{A}_n^I)\mathrm{F}_n^I). \tag{5}$$

It is conservative, if the so-called flux-splitting strategy [5] is applied to F. On the implicit region it is CFL-stable for sufficiently high Krylov subspace dimension m and is theoretically A-stable. In practice the stability region for the implicit part corresponds to that of the underlying EPIRK scheme resulting from $\mathrm{F}^E = 0$.

To construct methods of higher order than one in our IMEX context utilizing exponential integrators, consider the splitEPIRK schemes [4]: Given a split system of ODEs of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{U}(t) = \mathrm{F}(\mathrm{U}) = \mathbf{L}\mathrm{U}(t) + \mathrm{N}(\mathrm{U}(t)), \quad \mathbf{L} \in \mathbb{R}^{N\times N},$$

the *splitEPIRK schemes* are given as in equations (2),(3) with the difference that the matrix $\mathbf{L}$ is used instead of $\mathbf{A}$ and the function N replaces R. To apply those methods in the domain-based IMEX-manner, first split F into two parts $\mathrm{F}(\mathrm{U}) = \mathrm{F}^E(\mathrm{U}) + \mathrm{F}^I(\mathrm{U})$ as before using the flux splitting strategy to ensure conservativity. Now define

$$\mathbf{L}_n = (\mathrm{F}^I)'(\mathrm{U}_n), \quad \mathrm{N}(\mathrm{U}) = \mathrm{F}(\mathrm{U}) - \mathbf{L}_n\mathrm{U},$$

so that the application of the splitEPIRK method to $\mathrm{F}(\mathrm{U}) = \mathbf{L}_n\mathrm{U} + \mathrm{N}(\mathrm{U})$ is feasible. Those IMEX-splitEPIRK schemes have similar properties as the IMEX-EPIRK schemes on the implicit region and can be written as an RK scheme on the explicit region. Therefore, the CFL-condition for the underlying RK scheme determines the global time step size.

First numerical experiments for the flow around a wind turbine profile in two dimensions (see figure 2) were carried out showing the experimental order of convergence two for the IMEX-EPIRK scheme (5) in that test case. The reason for that is that the underlying EPIRK scheme is already of second order and is applied to the implicit region, where the main changes of the numerical solution seem to be located. Extensive numerical experiments and comparison with other schemes is subject of our current and future work as well as the task of finding criteria for deviding the domain into an explicit and an implicit region.

References

[1] P. Birken, G. Gassner, M. Haas, and C.-D. Munz, *Efficient Time Integration for Discontinuous Galerkin Methods for the Unsteady 3D Navier-Stokes Equations*, in European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2012) (2012), 1–20.

[2] A. Kanevsky, M. H. Carpenter, D. Gottlieb, and J. S. Hesthaven, *Application of implicit-explicit high order Runge-Kutta methods to discontinuous-Galerkin schemes*, J. Comp. Phys. **225** (2007), 1753–1781.

[3] J. Loffeld and M. Tokman, *Comparative performance of exponential, implicit, and explicit integrators for stiff systems of ODEs*, J. Comput. Appl. Math. **241** (2013), 45–67.

[4] G. Rainwater and M. Tokman, *A new class of split exponential propagation iterative methods of Runge-Kutta type (sEPIRK) for semilinear systems of ODEs*, J. Comput. Phys. **269** (2014), 40–60.

[5] M. Schlegel, O. Knoth, M. Arnold, and R. Wolke, *Numerical solution of multiscale problems in atmospheric modeling,* Appl. Numer. Math. **62** (2012), 1531–1543.

[6] M. Tokman, *A new class of exponential propagation iterative methods of Runge–Kutta type (EPIRK)*, J. Comput. Phys. **230** (2011), 8762–8778.

[7] P. E. Vincent and A. Jameson, *Facilitating the Adoption of Unstructured High-Order Methods Amongst a Wider Community of Fluid Dynamicists*, Math. Model. Nat. Phenom. **6** (2011), 97–140.

[8] Z. J. Wang, K. Fidkowski, R. Abgrall, F. Bassi, D. Caraeni, A. Cary, H. Deconinck, R. Hartmann, K. Hillewaert, H. T. Huynh, N. Kroll, and G. May, *High-order CFD methods: current status and perspective*, Int. J. Numer. Meth. Fluids **72** (2013), 811–845.

Space-time discontinuous Galerkin method for nonlinear hyperbolic conservation laws: residual error estimates

Vít Dolejší

(joint work with Filip Roskovec, Miloslav Vlasák)

We deal with the numerical solution of nonlinear hyperbolic problems with the aid of the fully implicit *space-time discontinuous Galerkin method* (STDGM). This technique is based on piecewise polynomial but discontinuous approximation which provides sufficient accurate and stable numerical solutions, see [3]. Moreover, it offers a great flexibility in mesh adaptation, namely it simply treats different grids on different time levels.

The numerical solution of nonlinear partial differential equations (PDEs) by the space-time discontinuous Galerkin method (STDG) is influenced by three types of errors:

- *space* (or spatial) *error* resulting from the space semi-discretization of the given PDEs by the DG method,
- *time* (or temporal) *error* resulting from the discretization of the arising ODEs system with the aid of the time DG scheme,
- *algebraic error* (including rounding errors) resulting from the inexact solution of the corresponding nonlinear algebraic systems at each time step.

In order to ensure the accuracy as well as the efficiency of the numerical method, these errors should be balanced.

In [4], we developed a residual based error estimation technique, which is able to identify the spatial, temporal and algebraic errors for convection-diffusion problems. Here, we present its extension to purely hyperbolic problems. These estimates are based on supremum

$$\eta := \sup_{\psi \in X_h \, \psi \neq 0} \frac{A_h(\tilde{w}_{h\tau}, \psi)}{\|\psi\|_X}, \tag{1}$$

where A_h is the form representing the STDG discretization of the given problem, $\tilde{w}_{h\tau}$ is the approximate solution and ψ is the test function. By a suitable choice of the (finite-dimensional) space X_h, we define

- space-time-algebraic error,
- space-algebraic error,
- time-algebraic error,
- algebraic error.

The delicate is the choice of the norm $\|\cdot\|_X$ in (1). Based on theoretical considerations and numerical experiments, we use the broken $H^1(0,T;H^1(\Omega))$-norm which leads to the estimate of the error in the $L^2(0,T;L^2(\Omega))$-norm. The presented numerical experiments verify that these estimates identify the corresponding ingredients of the total error.

Moreover, we developed an efficient algorithm which

- stops an iterative algorithm, which solves the arising nonlinear algebraic system at each time step, when the computational error is not essentially influenced by the algebraic error,
- chooses the size of the time step such that the computational error is not essentially influenced by the temporal error,
- adapts the given triangular grid such that the error estimate is under the given tolerance.

DGM can simply handle with different polynomial approximation degree and also with anisotropic grids. The use of the so-called *anisotropic hp*-adaptation can considerably decrease the number of degrees of freedom of many hyperbolic problems, where the exact solution is typically piecewise regular but discontinuous.

In [1], we developed the anisotropic *hp*-adaptation approach for the solution of boundary value problems. This technique is based on the optimization of a given *hp*-grid such that the interpolation error measured in the $L^q(\Omega)$-norm ($q \geq 1$) is under a given tolerance. This technique was extended in [2] to the numerical solution of non-stationary problems, where a sequence of *hp*-meshes is generated

such that the interpolation error in the discrete $L^\infty(0,T;L^q(\Omega))$-norm is limited. If an "unsuccessful time step" is detected, then a re-meshing has to be carry out. Here it is advantageous to use STDGM which couples two time slaps in a weak sense. The preliminary numerical example is presented.

References

[1] V. Dolejší. Anisotropic *hp*-adaptive method based on interpolation error estimates in the L^q-norm. *Appl. Numer. Math.*, 82:80–114, 2014.

[2] V. Dolejší. Anisotropic *hp*-adaptive discontinuous Galerkin method for the numerical solution of time dependent PDEs. *Applied Mathematics and Computation*, 2015. to appear, DOI: 10.1016/j.amc.2014.12.099.

[3] V. Dolejší and M. Feistauer. *Discontinuous Galerkin Method – Analysis and Applications to Compressible Flow*. Springer Series in Computational Mathematics 48. Springer International Publishing Switzerland, 2015.

[4] V. Dolejší, F. Roskovec, and M. Vlasák. Residual based error estimates for the space-time discontinuous Galerkin method applied to the compressible flows. *Comput. Fluids*, 117:304–324, 2015.

Skew-Symmetric Entropy-Stable Discontinuous Galerkin Methods for Hyperbolic Problems with Source Terms

Gregor Gassner

(joint work with Niklas Wintermeyer, Andrew Winters, David Kopriva)

Fluid flows in lakes, rivers, and near coastlines are of interest in oceanography and climate modelling. For such flows the vertical scales of motion are much smaller than the horizontal scales. From this and the assumption of hydrostatic balance, the Euler equations can be simplified to the shallow water equations. If the fluid flows over a non-constant bottom topography the shallow water equations may be written as a hyperbolic system of balance laws

$$\partial_t u + \partial_x f + \partial_y g = s. \tag{1}$$

For problems with non-constant bottom topographies $s \neq 0$ the preservation of non-trivial steady-state solutions is an important feature of a numerical approximation. Such preservation is particularly troublesome for discontinuous bottom topographies. One important steady-state for the shallow water equations is the so-called "lake at rest" solution, where a non-moving flat surface stays flat for all times. Discretisations which exactly preserve such steady-state solutions are said to be "well balanced". This is an important property when simulating problems, where the solution is a small perturbation of such a steady-state. If the method is not well balanced, this may result in numerically generated perturbations that are larger in magnitude the the actual perturbations of the solution, i.e. a so-called "numerical storm".

Besides the specific steady-state preservation, a general aspect of a numerical method is its robustness and hence the ability to produce accurate and stable results, even in the presence of non-linear discontinuities. While robustness is often relative straight forward for low order approximations such as e.g. first or

second order finite volume discretisations, it still remains an open challenge for many so-called high order methods, such as e.g. the nodal discontinuous Galerkin method. Due to variational crimes, Galerkin discretisations with high polynomial degree approximations suffer from aliasing errors. They suffer, because a positive benefit of high order polynomial approximations is a resulting low inherent numerical dissipation. Thus, while low numerical dissipation is certainly desired for accuracy reasons it is at the same time the reason that errors due to variational crimes are not masked and hence have an impact on the stability of the resulting discretisation. In a Galerkin type discretisation, aliasing errors typically occur when non-linear functions are approximated by polynomial interpolation and not by polynomial L_2-projection. If the number of interpolation (or quadrature) points are not sufficiently high to resolve the function, parts of higher modes of the function are transferred to lower modes, see [1]. This issue gets magnified in situations where the approximation is unresolved, e.g. shocks or vortical driven flows (turbulence). Without proper de-aliasing techniques, aliasing errors then accumulate in every time step and can even drive non-linear instabilities, i.e. cause simulations to crash, e.g. [1, 2].

A clean de-aliasing technique is the proper projection of non-linear functions onto the polynomial space by choosing the number of interpolation (quadrature) points sufficiently large to recover the analytical projection. This so-called polynomial de-aliasing is quite common in spectral methods and was first proposed in the context of discontinuous Galerkin methods by Kirby and Karniadakis in 2003. At first glance, this technique is quite simple (to implement) and the cleanest de-aliasing technique as it removes the source of aliasing, i.e. the variational crimes. It is consequently quite effective in stabilising computations e.g. shown in [2]. A practical issue for this type of de-aliasing is the actual choice of number of sample points to remove the interpolation errors. The right number depends on the type of function that is projected and on the polynomial approximation space. Due to computational efficiency, it is desired that the number of evaluation points is as small as possible. A strategy where the number of points are chosen arbitrarily "high enough" to be on the safe side is obviously not recommended. However, the task to determine the minimum amount of evaluation points to *guarantee* stability is in the general case very difficult, and in some cases even impossible. If we consider for instance the compressible Euler equations, the Galerkin Ansatz is a piecewise polynomial for all conservative quantities. The Galerkin discretisation needs to project the non-linear flux function onto the space of polynomials. However, the flux function of the compressible Euler equations are rational polynomials with respect to the conservative quantities. Thus, independent of the number of interpolation (quadrature) points, the projection is never analytically exact. As we are dealing with computers anyway, one can relax the exactness of the projection to achieving machine round-off errors. However, it is clear in this case that then the number of necessary evaluation points highly depend on the actual solution. For very smooth and well resolved parts of the solution it

is expected that the number is much smaller in comparison to parts of the solution containing e.g. shocks. Thus, ideally, an automatic and adaptive approach would be required to rigorously satisfy the "machine round-off" requirement. It is clear that such an approach would substantially complicate the discretisation (for instance load balancing in a parallel computation). Thus, up to now, the most common approach in the community is to choose the number of evaluation points in a pre-processor step of the simulation by an educated guess (or with trial and error, i.e. user experience) and hope for the best during the simulation. It can then be observed that in certain cases one can get away with a much lower number of evaluation points while retaining highly accurate results, e.g. [2]. While polynomial de-aliasing seems to be the natural and cleanest way of de-aliasing for discontinuous Galerkin methods, there are still many open questions left and further research is necessary.

Recently, another de-aliasing technique was introduced for (nodal) discontinuous Galerkin discretisations. The starting point is to accept that the projection is not a proper L_2-projection but rather an interpolation. We get highly effective nodal discontinuous Galerkin methods for smooth problems, if we collocate the polynomial Ansatz and the interpolation operator. Thus, the number of interpolation points are chosen assuming that the non-linear flux is of same polynomial degree as the polynomial Ansatz. As discussed above, this obviously introduces aliasing errors and for strongly non-linear problems the stability issues are severe. However, by restricting the discontinuous Galerkin discretisation to this specific case, it is possible to show the summation-by-parts property of the resulting operator [3]. This enables the construction of nodal collocated discontinuous Galerkin discretisation based on *skew-symmetric* formulations of the underlying partial differential equations, e.g. [3, 4]. These new skew-symmetric nodal discontinuous Galerkin discretisations are provably entropy/energy stable without the assumption of an exact evaluation of the projection.

In case of hyperbolic balance laws, i.e the shallow water equations, it is possible to derive specific skew-symmetric formulations that are not only provable entropy stable, but also compatible with the well balanced property [5]. It can be shown that for numerical robustness, both properties of the discretisation are desirable [6]. By augmenting the skew-symmetric formulation with suitable surface discretisations of the fluxes and the sources, it is possible to extend the novel skew-symmetric discontinuous Galerkin method to handle bottom topographies with (grid aligned) discontinuities on unstructured curvilinear quadrilateral grids [7].

References

[1] R.M. Kirby and G.E. Karniadakis, *De-aliasing on non-uniform grids: algorithms and applications*, Journal of Computational Physics **191** (2003), 249–264.

[2] Gregor J. Gassner and Andrea D. Beck, *On the accuracy of high-order discretizations for underresolved turbulence simulations*, Theoretical and Computational Fluid Dynamics **27** (2013), 221–237.

[3] G. Gassner, *A Skew-Symmetric Discontinuous Galerkin Spectral Element Discretization and Its Relation to SBP-SAT Finite Difference Methods*, SIAM Journal on Scientific Computing **35** (2013), A1233-A1253.
[4] D.A. Kopriva and G.J. Gassner, *An Energy Stable Discontinuous Galerkin Spectral Element Discretization for Variable Coefficient Advection Problems*, SIAM Journal on Scientific Computing **36** (2014), A2076-A2099.
[5] G.J. Gassner, A.R. Winters and David A. Kopriva, *A Well Balanced and Entropy Conservative Discontinuous Galerkin Spectral Element Method for the Shallow Water Equations*, accepted in Applied Mathematics and Computation.
[6] A.R. Winters and G.J. Gassner, *A Comparison of Two Entropy Stable Discontinuous Galerkin Spectral Element Approximations for the Shallow Water Equations with Non-Constant Topography*, accepted in Journal of Computational Physics.
[7] Niklas Wintermeyer, Andrew R. Winters, Gregor J. Gassner and David A. Kopriva, *An Entropy Stable Nodal Discontinuous Galerkin Method for the Two Dimensional Shallow Water Equations on Unstructured Curvilinear Meshes with Discontinuous Bathymetry*, submitted to Journal of Computational Physics.

Mimetic discretization of the Lie derivative

Marc Gerritsma
(joint work with Jeroen Kunnen)

The Lie derivative $\mathcal{L}_X\omega^{(k)}$ of a differential k-form $\omega^{(k)}$ on a manifold $\mathcal{M}$ represents convective transport along the flow generated by the vector field $X \in \mathfrak{X}(\mathcal{M})$. Using Cartan's magic formula we have

$$\mathcal{L}_X\omega^{(k)} = \iota_X \mathrm{d}\omega^{(k)} + \mathrm{d}\iota_X\omega^{(k)};,$$

where ι_X is the interior product – contraction of the vector field with the k-form – and d represents the exterior derivative.

A discrete representation of the exterior derivative is given by the coboundary operator, δ, see [1, 2]. This fully discrete representation is possible because the exterior derivative is an intrinsic operator, i.e. coordinate independent and metric free. One therefore hopes that a similar construction is possible for the Lie derivative, because the Lie derivative is also an intrinsic operator. By imposing that the discrete Lie derivative is linear in the differential forms, linear in the vector fields and satisfies a Leibniz rule

$$\mathcal{L}_X(\omega^{(k)} \wedge \eta^{(l)}) = \mathcal{L}_X(\omega^{(k)}) \wedge \eta^{(l)} + \omega^{(k)} \wedge \mathcal{L}_X(\eta^{(l)}) \,,$$

we present a discrete Lie derivative L_X which commutes with reduction operator $\mathcal{R}$

$$\begin{array}{ccc} \Lambda^k(\mathcal{M}) & \xrightarrow{\mathcal{L}_X} & \Lambda^k(\mathcal{M}) \\ \mathcal{R}\downarrow & & \downarrow\mathcal{R} \\ C^k & \xrightarrow{\mathsf{L}_X} & C^k \end{array}$$

This reuiqres that also the reduction operator $\mathcal{R}$ need to be redefined. For a 0-form $f^{(0)}$ the reduction consisted of sampling the function values at discrete points. In the new reduction operator we sample the function values of $f^{(0)}$ with

all its partial derivatives $g_i = \partial f(p)/\partial x^i$. Where the old reduction mapped onto real-valued 0-cochains, the new reduction produces vector-valued 0-cochains which are essentially jet spaces. By a suitable choice of multiplication of such vector-valued cochains the Leibniz rule is satisfied at the purely discrete level. In the 1 dimensional case, this construction can be extended to 1-forms.

Preliminary results of this approach were shown during the workshop.

References

[1] M. Gerritsma, *Edge functions for spectral element methods* , Lecture Notes in Computational Science and Engineering, **76** LNCSE (2011), 199–207.

[2] J. Kreeft, A. Palha and M. Gerritsma *Mimetic framework on curvilinear quadrilaterals of arbitrary order*, arxiv:1111.4304 (2011), 1–69.

Challenges and Opportunities in Long-Time Simulations of PDEs on Modern Parallel Computing Platforms

Matthias K. Gobbert

(joint work with Jonathan S. Graf, Samuel Khuvis, Bradford E. Peercy, Stefan Kopecz, Andreas Meister)

Advection-diffusion-reaction equations occur in a wide variety of applications, for instance fluid flow, heat transfer, spread of pollutants, and transport-chemistry problems. We consider a very general framework of this problem as a testbed to investigate the choices of numerical methods in the face of Dirac delta distributions among the source terms. We also demonstrate the ability of memory-efficient parallel implementations of these methods to solve the problem on extremely fine meshes efficiently using a cluster with state-of-the-art CPU nodes and cutting-edge hybrid CPU/GPU nodes. Our consideration of this problem is inspired by the need to simulate Calcium Induced Calcium Release (CICR) in a heart cell [4]. CICR describes a physiological process within a cell where calcium is able to activate calcium release from the sarcoplasmic reticulum into the cytosol, which is crucial for excitation-contraction coupling in the cardiac muscle.

As a testbed, we consider the system of coupled, non-linear, time-dependent advection-diffusion-reaction equations

$$u_t^{(i)} - \nabla \cdot \left(D^{(i)} \nabla u^{(i)}\right) + \beta^{(i)} \cdot \left(\nabla u^{(i)}\right) = q^{(i)}, \quad i = 1, \ldots, n_s, \tag{1}$$

with functions $u^{(i)} = u^{(i)}(\mathbf{x}, t)$, $i = 1, \ldots, n_s$, of space $\mathbf{x} \in \Omega \subset \mathbb{R}^3$ and time $0 \le t \le t_{\text{fin}}$ representing the concentrations of the n_s species. The diffusivity matrices $D^{(i)} = \text{diag}(D_{11}^{(i)}, D_{22}^{(i)}, D_{33}^{(i)}) \in \mathbb{R}^{3\times 3}$ consists of positive entries and are assumed to dominate the advection velocity vectors $\beta^{(i)} \in \mathbb{R}^3$, so that numerical methods for parabolic problems are always justified. The right-hand side $q^{(i)}$ in (1) is written in a way that distinguishes the different dependencies and effects as

$$q^{(i)}(u^{(i)}, \ldots, u^{(n_s)}, \mathbf{x}, t) = s^{(i)}(u^{(i)}, \mathbf{x}, t) + r^{(i)}(u^{(1)}, \ldots, u^{(n_s)}). \tag{2}$$

The reaction terms $r^{(i)}(u^{(1)}, \ldots, u^{(n_s)})$ are, in general, non-linear autonomous functions of all species and couple the reaction equations in the system (1). The terms $s^{(i)}(u^{(i)}, \mathbf{x}, t)$ in (2) contain many thousands of point sources modeled by Dirac delta distributions on a large lattice throughout the cell. This crucial feature of the model is responsible for many of the challenges, including the need to discretize the domain of the cell with a very fine mesh to accommodate the large number of point sources, since they are the crucial driver of the physiological effects.

In [4], a finite volume method for advection-diffusion-reaction systems with smooth and non-smooth sources was introduced. Efficient parallel computing is necessary to enable long-time simulations of complex nonlinear models on high-resolution meshes, such as for calcium induced calcium release in a heart cell modeled by a system of advection-diffusion-reaction equations in three space dimensions. A method of line approach with sophisticated time-stepping and matrix-free Newton-Krylov methods is well suited for parallel computing. With demands to push performance further, we need to use modern parallel hardware with hierarchical memory access, data transmission via InfiniBand, and more, whose optimal use is inherently challenging.

The UMBC High Performance Computing Facility has a cluster with over 300 nodes. The newest 72 nodes include hybrid nodes with high-end NVIDIA GPUs as well as with cutting-edge 60-core Intel Phi accelerators, which offer great research opportunities and potentially significant speedup of the computational kernels in simulation tools such as PDE solver for time-dependent problems.

Table 1 shows a historical comparison of observed wall clock run times for the application problem on clusters purchased in years ranging from 2009 to 2013. For context, the ODE solver NDFk, $1 \leq k \leq 5$, uses 73,123 time steps, and the linear solver is BiCGSTAB. The timings of the CPU runs show dramatic improvement over time for the serial run as well as by using up to 32 nodes with up to 16 computational cores per node. However, the application demands a longer final time than the 1,000 ms here as well as demands more than the 3 chemical species modeled here. Therefore, yet better, more efficient, parallel code is needed. Specifically, it is an interesting research opportunity to study sophisticated numerical methods and parallelization that take full advantage of available state-of-the-art architectures!

Table 2 shows the complete CPU performance study up to 32 nodes on the 2013 portion of the cluster, from which the summary data in the previous table are taken. We see here that the needed run times for a finer mesh become excessive, namely more than the maximum allowed run time of 5 days on the cluster. The power of parallel computing is visible when the run time can be reduced to about 27 hours by using all cores on 1 node or even to about 1:40 hour on 32 nodes (all 16 cores per node).

Table 3 in turn shows the power of using hybrid CPU/GPU nodes for the same meshes as in the previous table. Specifically, focusing on the finest mesh, the first row in the table starts with the run time of about 27 hours using all 16 CPU

TABLE 1. Historical comparison of observed wall clock times in HH:MM:SS for parabolic nonlinear three-species application problems: CPU only results. Mesh resolution $N_x \times N_y \times N_z = 64 \times 64 \times 256$, system dimension 3,257,475 [2]. Asterisk indicates use of 8 processes per node to enable run.

Cluster, method	serial 1 core time	(1 node) all cores time (speedup)	32 node 1 core per node time (speedup)	32 node all cores time (speedup)
tara (2009), FEM	67:04:28	09:17:32 (7.22)	02:15:03 (29.80)	00:29:06 (138.29)
tara (2009), FVM	47:46:46	07:31:51 (6.34)	01:41:46 (28.17)	00:25:54 (110.68)
maya (2009), FVM	32:02:14	05:49:56 (5.49)	01:05:40 (29.27)	00:17:57 (107.09)
maya (2010), FVM	30:51:54	05:46:48 (5.34)	01:04:22 (28.77)	00:20:37 (89.83)
maya (2013), FVM	25:02:01	02:25:55 (10.29)	00:53:17 (28.19)	00:18:33 (80.97)*

TABLE 2. Observed wall clock times in HH:MM:SS for parabolic nonlinear three-species application problems on maya 2013 on CPU nodes by number of nodes [2]. ET indicates "excessive time required" (more than 5 days), N/A indicates that the case is not feasible.

(a) Mesh resolution $N_x \times N_y \times N_z = 32 \times 32 \times 128$, DOF = 421,443

	1 node	2 nodes	4 nodes	8 nodes	16 nodes	32 nodes
1 proc. per node	02:22:19	01:10:46	00:36:09	00:19:17	00:10:54	00:06:37
2 proc. per node	01:11:43	00:35:48	00:19:13	00:10:45	00:06:47	00:04:48
4 proc. per node	00:37:52	00:19:27	00:11:07	00:06:59	00:05:18	00:04:41
8 proc. per node	00:21:35	00:11:28	00:07:24	00:05:41	00:05:36	N/A
16 proc. per node	00:12:58	00:07:24	00:06:30	00:07:26	N/A	N/A

(b) Mesh resolution $N_x \times N_y \times N_z = 64 \times 64 \times 256$, DOF = 3,257,475

	1 node	2 nodes	4 nodes	8 nodes	16 nodes	32 nodes
1 proc. per node	25:02:01	12:25:01	06:10:14	03:07:33	01:37:56	00:53:17
2 proc. per node	12:25:07	06:11:38	03:08:20	01:37:41	00:52:27	00:30:36
4 proc. per node	06:32:39	03:16:02	01:41:55	00:55:03	00:31:50	00:21:03
8 proc. per node	03:52:48	01:53:44	01:00:20	00:34:24	00:22:30	00:18:33
16 proc. per node	02:25:55	01:10:26	00:39:04	00:25:46	00:21:29	N/A

(c) Mesh resolution $N_x \times N_y \times N_z = 128 \times 128 \times 512$, DOF = 25,610,499

	1 node	2 nodes	4 nodes	8 nodes	16 nodes	32 nodes
1 proc. per node	ET	ET	69:15:37	34:51:02	17:31:44	08:59:06
2 proc. per node	ET	69:46:29	35:16:03	17:45:06	09:00:47	04:47:14
4 proc. per node	72:31:51	36:34:34	18:36:29	09:32:04	05:01:44	02:50:34
8 proc. per node	42:01:27	26:23:03	11:03:41	05:46:44	03:09:23	01:56:47
16 proc. per node	26:53:37	13:56:38	07:21:17	03:54:47	02:17:48	01:40:35

cores on 1 node. Using instead one GPU reduces this time to about 15:32 hours, showing the advantage possible with this state-of-the-art accelerator. The numbers in parentheses show speedup over the 1 node/16 processes run time, and it becomes apparent that using two GPUs in the hybrid node is even more advantageous. On the one hand, continuing to look at the speedups over the 1 node/16 processes

TABLE 3. Observed wall clock times in HH:MM:SS for parabolic nonlinear three-species application problem on maya 2013 with CPU and hybrid CPU/GPU nodes. Wall clock time (speedup) against one 16-core node using the finite volume method with BiCG-STAB [1]. CPU-only on $128 \times 128 \times 512$: 01:40:35 on 32 nodes (16 cores).

nodes (GPU/node)	$32 \times 32 \times 129$	$64 \times 64 \times 256$	$128 \times 128 \times 512$
1 node (16 cores)	00:12:58	02:25:55	26:53:37
1 node (1 GPU)	00:16:47 (0.77)	01:44:13 (1.40)	15:32:34 (1.73)
1 node (2 GPUs)	00:11:47 (1.10)	00:59:53 (2.44)	08:18:18 (3.24)
2 nodes (16 cores)	00:07:24	01:10:26	13:56:38
2 nodes (1 GPU)	00:12:20 (1.05)	00:58:12 (2.51)	08:14:26 (3.26)
2 nodes (2 GPUs)	00:09:27 (1.37)	00:35:31 (4.11)	04:25:55 (6.07)
4 nodes (16 cores)	00:06:30	00:39:04	07:21:17
4 nodes (1 GPU)	00:09:46 (1.33)	00:34:36 (4.22)	04:20:56 (6.18)
4 nodes (2 GPUs)	00:08:25 (1.54)	00:24:21 (5.99)	02:28:00 (10.90)
8 nodes (16 cores)	00:07:26	00:25:46	03:54:47
8 nodes (1 GPU)	00:08:32 (1.52)	00:24:58 (5.84)	02:24:46 (11.15)
8 nodes (2 GPUs)	00:08:10 (1.59)	00:20:27 (7.14)	01:31:22 (17.66)
16 nodes (16 cores)	N/A	00:21:29	02:17:48
16 nodes (1 GPU)	00:08:19 (1.56)	00:20:45 (7.03)	01:30:06 (17.91)
16 nodes (2 GPUs)	00:08:11 (1.58)	00:19:57 (7.31)	01:06:17 (24.34)

run shows that it pays to use all available 16 hybrid nodes with two GPUs each. On the other hand, also comparing run times with and without GPUs on the same number of nodes shows a clear advantage of combing CPUs and GPUs over just using CPUs. In fact, the final number of about 1:06 hour using 16 hybrid CPU/GPU nodes is faster than the time of about 1:40 hours using 32 CPU nodes!

REFERENCES

[1] Xuan Huang, *An MPI-CUDA Implementation of a Model for Calcium Induced Calcium Release in a Three-Dimensional Heart Cell on a Hybrid CPU/GPU Cluster*, Ph.D. Thesis, Department of Mathematics and Statistics, University of Maryland, Baltimore County, 2015.

[2] Xuan Huang and Matthias K. Gobbert, *Parallel Performance Studies for a Three-Species Application Problem on the Cluster maya*, Technical Report number HPCF–2015–8, UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2015, `hpcf.umbc.edu`.

[3] Xuan Huang, Matthias K. Gobbert, Bradford E. Peercy, Stefan Kopecz, Philipp Birken, and Andreas Meister, *Order Investigation of Scalable Memory-Efficient Finite Volume Methods for Parabolic Advection-Diffusion-Reaction Equations with Point Sources*, in preparation.

[4] Jonas Schäfer, Xuan Huang, Stefan Kopecz, Philipp Birken, Matthias K. Gobbert, and Andreas Meister, *A Memory-Efficient Finite Volume Method for Advection-Diffusion-Reaction Systems with Non-Smooth Sources*, Numerical Methods Partial Differential Equations, **31** (2015), 143–167.

Invariant domains and first-order continuous finite element approximation for hyperbolic systems

Jean-Luc Guermond
(joint work with Bojan Popov)

1. Introduction

We propose a numerical method to solve general hyperbolic systems in any space dimension using forward Euler time stepping and continuous finite elements on non-uniform grids. The properties of the method are based on the introduction of an artificial dissipation that is defined so that any convex invariant sets containing the initial data is an invariant domain for the method. Our technique extends to continuous finite elements the work of [Hoff(1979), Hoff(1985)], and [Frid(2001)]. The invariant domain property is proved for any hyperbolic system provided a CFL condition holds. The solution is also proved to satisfy a discrete entropy inequality for every admissible entropy of the system. The method is formally first-order accurate in space and can be made high-order in time by using any Strong Stability Preserving technique.

1.1. **The problem.** Consider the following hyperbolic system

$$\begin{cases} \partial_t \boldsymbol{u} + \nabla\cdot\boldsymbol{f}(\boldsymbol{u}) = 0, \quad \text{for } (\boldsymbol{x},t) \in \mathbb{R}^d\times\mathbb{R}_+. \\ \boldsymbol{u}(\boldsymbol{x},0) = \boldsymbol{u}_0(\boldsymbol{x}), \quad \text{for } \boldsymbol{x} \in \mathbb{R}^d. \end{cases} \tag{1}$$

where the dependent variable $\boldsymbol{u}$ takes values in $\mathbb{R}^m$ and the flux $\boldsymbol{f}$ takes values in $(\mathbb{R}^m)^d$. Here $\boldsymbol{u}$ is considered as a column vector $\boldsymbol{u} = (u_1,\dots,u_m)^{\mathsf{T}}$. The flux is a matrix with entries $f_{ij}(\boldsymbol{u})$, $1 \le i \le m$, $1 \le j \le d$ and $\nabla\cdot\boldsymbol{f}$ is a column vector with entries $(\nabla\cdot\boldsymbol{f})_i = \sum_{1\le j\le d} \partial_{x_j} f_{ij}$. For any $\boldsymbol{n} = (n_1\dots,n_d)^{\mathsf{T}} \in \mathbb{R}^d$, we denote $\boldsymbol{f}(\boldsymbol{u})\cdot\boldsymbol{n}$ the column vector with entries $\sum_{1\le l\le d} n_l f_{il}(\boldsymbol{u})$, where $i \in \{1{:}m\}$.

To simplify questions regarding boundary conditions, we assume that either periodic boundary conditions are enforced, or the initial data is compactly supported. In both cases we denote by D the spatial domain where the approximation is constructed. In the case of periodic boundary conditions, D is the d-torus. In the case of the Cauchy problem, D is a compact, polygonal portion of $\mathbb{R}^d$ large enough so that the domain of influence of $\boldsymbol{u}_0$ is always included in D over the entire duration of the simulation.

1.2. **Assumptions and definitions.** We assume that (1) is such that there is a clear notion for the solution of the Riemann problem. That is to say there exists an (nonempty) admissible set $\mathcal{A} \subset \mathbb{R}^m$ such that for any pair of states $(\boldsymbol{u}_L,\boldsymbol{u}_R) \in \mathcal{A}\times\mathcal{A}$ and any unit vector $\boldsymbol{n}$ in $\mathbb{R}^d$, the following one-dimensional Riemann problem

$$\partial_t \boldsymbol{u} + \partial_x(\boldsymbol{f}(\boldsymbol{u})\cdot\boldsymbol{n}) = 0, \quad (x,t) \in \mathbb{R}\times\mathbb{R}_+, \qquad \boldsymbol{u}(x,0) = \begin{cases} \boldsymbol{u}_L, & \text{if } x < 0 \\ \boldsymbol{u}_R, & \text{if } x > 0, \end{cases} \tag{2}$$

has a unique (physical) solution, which we henceforth denote $\boldsymbol{u}(\boldsymbol{n}, \boldsymbol{u}_L, \boldsymbol{u}_R)$. We make the following assumption:

(3) The unique solution of (2) has a finite speed of propagation for any $\boldsymbol{n}$, i.e., there is $\lambda_{\max}(\boldsymbol{n}, \boldsymbol{u}_L, \boldsymbol{u}_R)$ such that

$$\boldsymbol{u}(x,t) = \begin{cases} \boldsymbol{u}_L, & \text{if } x \le -t\lambda_{\max}(\boldsymbol{n}, \boldsymbol{u}_L, \boldsymbol{u}_R) \\ \boldsymbol{u}_R, & \text{if } x \ge t\lambda_{\max}(\boldsymbol{n}, \boldsymbol{u}_L, \boldsymbol{u}_R). \end{cases}$$

We now introduce the notions of invariant sets and invariant domains. Our definitions are slightly different from those in [Chueh et al.(1977)], [Hoff(1985)], [Smoller(1983)], [Frid(2001)]. We will associate invariant sets only with solutions of Riemann problems and define invariant domains only for an approximation process of (1).

Definition 1 (Invariant set)**.** *We say that a set $A \subset \mathcal{A} \subset \mathbb{R}^m$ is invariant for* (1) *if for any pair $(\boldsymbol{u}_L, \boldsymbol{u}_R) \in A \times A$, any unit vector $\boldsymbol{n} \in \mathbb{R}^d$, and any $t > 0$, the average of the entropy solution of the Riemann problem* (2) *over the Riemann fan, say, $\frac{1}{t(\lambda_m^+ - \lambda_1^-)} \int_{\lambda_1^- t}^{\lambda_m^+ t} \boldsymbol{u}(\boldsymbol{n}, \boldsymbol{u}_L, \boldsymbol{u}_R)(x,t)\,\mathrm{d}x$, remains in A.*

Note that, the above definition implies that given $t > 0$ and any interval I such that $(\lambda_1^- t, \lambda_m^+ t) \subset I$, we have that $\frac{1}{I} \int_I \boldsymbol{u}(\boldsymbol{n}, \boldsymbol{u}_L, \boldsymbol{u}_R)(x,t)\,\mathrm{d}x \in A$.

We now introduce the notion of invariant domain for an approximation process. Let $\boldsymbol{X}_h \subset L^1(\mathbb{R}^d; \mathbb{R}^m)$ be a finite-dimensional approximation space and let $S_h : \boldsymbol{X}_h \ni \boldsymbol{u}_h \longmapsto S_h(\boldsymbol{u}_h) \in \boldsymbol{X}_h$ be a discrete process over $\boldsymbol{X}_h$. Henceforth we abuse the language by saying that a member of $\boldsymbol{X}_h$, say $\boldsymbol{u}_h$, is in the set $A \subset \mathbb{R}^m$ when actually we mean that $\{\boldsymbol{u}_h(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathbb{R}\} \subset A$.

Definition 2 (Invariant domain)**.** *A convex invariant set $A \subset \mathcal{A} \subset \mathbb{R}^m$ is said to be an invariant domain for the process S_h if and only if for any state $\boldsymbol{u}_h$ in A, the state $S_h(\boldsymbol{u}_h)$ is also in A.*

For scalar conservation equations the notions of invariant sets and invariant domains are closely related to the maximum principle. In the case of nonlinear systems, the notion of maximum principle does not apply and must be replaced by the notion of invariant domain. To the best of our knowledge, the definition of invariant sets for the Riemann problem was introduced in [Nishida(1968)], and the general theory of positively invariant regions was developed in [Chueh et al.(1977)]. Applications and extensions to numerical methods were developed in [Hoff(1979), Hoff(1985)] and [Frid(2001)].

2. The method

2.1. **The finite element space.** We want to approximate the solution of (1) with continuous finite elements. Let $(\mathcal{T}_h)_{h>0}$ be a shape-regular sequence of affine matching meshes. The elements in the mesh sequence are assumed to be generated from a finite number of reference elements denoted $\widehat{K}_1, \ldots, \widehat{K}_\varpi$. For example, the mesh $\mathcal{T}_h$ could be composed of a combination of triangles and parallelograms

in two space dimensions ($\varpi = 2$ in this case); it could also be composed of a combination of tetrahedra, parallelepipeds, and triangular prisms in three space dimensions ($\varpi = 3$ in this case). The affine diffeomorphism mapping $\widehat{K}_r$ to an arbitrary element $K \in \mathcal{T}_h$ is denoted $T_K : \widehat{K}_r \longrightarrow K$ and its Jacobian matrix is denoted $\mathbb{J}_K$, $1 \le r \le \varpi$. We now introduce a set of reference Lagrange finite elements $\{(\widehat{K}_r, \widehat{P}_r, \widehat{\Sigma}_r)\}_{1\le r\le \varpi}$ (the index $r \in \{1{:}\varpi\}$ will be omitted from now on to alleviate the notation). Then we define the scalar-valued and vector-valued Lagrange finite element spaces

$$P(\mathcal{T}_h) = \{v \in \mathcal{C}^0(D;\mathbb{R}) \mid v_{|K}\circ T_K \in \widehat{P},\ \forall K \in \mathcal{T}_h\}, \qquad \boldsymbol{P}(\mathcal{T}_h) = [P(\mathcal{T}_h)]^m. \tag{4}$$

where $\widehat{P}$ is the reference polynomial space defined on $\widehat{K}$ (note that the index r has been omitted). Denoting $n_{\rm sh} := \dim \widehat{P}$ and denoting by $\{\widehat{\boldsymbol{a}}_i\}_{i\in\{1:n_{\rm sh}\}}$ the Lagrange nodes of $\widehat{K}$, we assume that the space $\widehat{P}$ is such that

$$\min_{1\le\ell\le n_{\rm sh}} \widehat{v}(\widehat{\boldsymbol{a}}_\ell) \le \widehat{v}(\widehat{\boldsymbol{x}}) \le \max_{1\le\ell\le n_{\rm sh}} \widehat{v}(\widehat{\boldsymbol{a}}_\ell), \quad \forall \widehat{v} \in \widehat{P}, \forall \widehat{\boldsymbol{x}} \in \widehat{K}. \tag{5}$$

Denoting by $\mathbb{P}_1$ and $\mathbb{Q}_1$ the set of multivariate polynomials of total and partial degree at most 1, respectively; the above assumption holds for $\widehat{P} = \mathbb{P}_1$ when K is a simplex and $\widehat{P} = \mathbb{Q}_1$ when K is a parallelogram or a cuboid. This assumption holds also for first-order prismatic elements in three space dimensions.

Let $\{\boldsymbol{a}_i\}_{i\in\{1:I\}}$ be the collection of all the Lagrange nodes in the mesh $\mathcal{T}_h$, and let $\{\varphi_i\}_{i\in\{1:I\}}$ be the corresponding global shape functions. Recall that $\{\varphi_i\}_{i\in\{1:I\}}$ forms a basis of $P(\mathcal{T}_h)$ and $\varphi_i(\boldsymbol{a}_j) = \delta_{ij}$. The Lagrange interpolation operator in $\mathcal{P}(\mathcal{T}_h)$ is denoted $\Pi_h : \mathcal{C}^0(\overline{D}) \longrightarrow \mathcal{P}(\mathcal{T}_h)$. Recall that $\Pi_h(v) = \sum_{1\le i\le I} v(\boldsymbol{a}_i)\varphi_i$. We denote by S_i the support of φ_i and by $|S_i|$ the measure of S_i, $i \in \{1{:}I\}$. We also define $S_{ij} := S_i \cap S_j$ the intersection of the two supports S_i and S_j. Let E be a union of cells in $\mathcal{T}_h$; we define $\mathcal{I}(E) := \{j \in \{1{:}I\} \mid |S_j \cap E| \neq 0\}$ the set that contains the indices of all the shape functions whose support on E is of nonzero measure.

Let $\mathcal{M} \in \mathbb{R}^{I\times I}$ be the consistent mass matrix with entries $\int_{S_{ij}} \varphi_i(\boldsymbol{x})\varphi_j(\boldsymbol{x})\,\mathrm{d}x$, and let $\mathcal{M}^L$ be the diagonal lumped mass matrix with entries

$$m_i := \int_{S_i} \varphi_i(\boldsymbol{x})\,\mathrm{d}x. \tag{6}$$

The partition of unity property implies that $m_i = \sum_{j\in\mathcal{I}(S_i)} \int \varphi_j(\boldsymbol{x})\varphi_i(\boldsymbol{x})\,\mathrm{d}x$, i.e., the entries of $\mathcal{M}^L$ are obtained by summing the rows of $\mathcal{M}$.

2.2. **The scheme.** Let $\boldsymbol{u}_{h0} \in \mathcal{P}(\mathcal{T}_h)$ be a reasonable approximation of $\boldsymbol{u}_0$ (we shall be more precise in the following sections). Let $n \in \mathbb{N}$, τ be the time step, t^n be the current time, and let us set $t^{n+1} = t^n + \tau$. Let $\boldsymbol{u}_h^n = \sum_{i=1}^I \mathbf{U}^n\varphi_i \in \mathcal{P}(\mathcal{T}_h)$ be the space approximation of $\boldsymbol{u}$ at time t^n. We propose to compute $\boldsymbol{u}_h^{n+1} =$

$\sum_{i=1}^{I} \mathbf{U}^{n+1}\varphi_i$ by

$$m_i \frac{\mathbf{U}_i^{n+1} - \mathbf{U}_i^n}{\tau} + \int_D \nabla\cdot(\Pi_h \boldsymbol{f}(\boldsymbol{u}_h^n))\varphi_i \,\mathrm{d}x - \sum_{j\in\mathcal{I}(S_i)} d_{ij}\mathbf{U}_j^n = 0, \tag{7}$$

where the lumped mass matrix is used for the approximation of the time derivative. The coefficient d_{ij} is an artificial viscosity defined as follows for $i \neq j$:

$$d_{ij} := \max(\lambda_{\max}(\boldsymbol{n}_{ij}, \mathbf{U}_i, \mathbf{U}_j)\|\boldsymbol{c}_{ij}\|_{\ell^2}, \lambda_{\max}(\boldsymbol{n}_{ji}, \mathbf{U}_j, \mathbf{U}_i)\|\boldsymbol{c}_{ji}\|_{\ell^2}), \tag{8}$$

where recall that $\lambda_{\max}(\boldsymbol{n}_{ij}, \mathbf{U}_i, \mathbf{U}_j)$ is defined in the assumption (3), and $d_{ii} := \sum_{i\neq j\in\mathcal{I}(S_i)} -d_{ji}$. The vector-valued coefficient $\boldsymbol{c}_{ij} \in \mathbb{R}^d$ is defined by

$$\boldsymbol{c}_{ij} = \int_D \varphi_i \nabla\varphi_j \,\mathrm{d}x, \tag{9}$$

Remark 1 (Conservation)**.** *The definition* $d_{ii} := \sum_{i\neq j\in\mathcal{I}(S_i)} -d_{ji}$ *implies that* $\sum_{j\in\mathcal{I}(S_i)} d_{ji} = 0$*, which in turn implies conservation, i.e.,* $\int_D \boldsymbol{u}_h^{n+1}\,\mathrm{d}x = \int_D \boldsymbol{u}_h^n\,\mathrm{d}x + \int_D \nabla\cdot(\Pi_h \boldsymbol{f}(\boldsymbol{u}_h^n))\,\mathrm{d}x$ *since* $d_{ij} = d_{ji}$*.*

2.3. **Stability results.** Upon defining $h_K := \mathrm{diam}(K)$, the global maximum mesh size is denoted $h = \max_{K\in\mathcal{T}_h} h_K$. The local minimum mesh size, $\underline{h}_K$, for any $K \in \mathcal{T}_h$ is defined as follows:

$$\underline{h}_K := \frac{1}{\max_{i\neq j\in\mathcal{I}(K)} \|\nabla\varphi_i\|_{\boldsymbol{L}^\infty(S_{ij})}}, \tag{10}$$

and the global minimum mesh size is $\underline{h} := \min_{K\in\mathcal{T}_h} \underline{h}_K$. Due to the shape regularity assumption, the quantities $\underline{h}_K$ and h_K are uniformly equivalent, but it will turn out that using $\underline{h}_K$ instead of h_K gives a sharper estimate of the CFL number. Let $n_{\mathrm{sh}} := \mathrm{card}(\mathcal{I}(K))$ and let us define $\vartheta_K := \frac{1}{n_{\mathrm{sh}}-1}$. Note that

$$0 < \vartheta_{\min} := \min_{(\mathcal{T}_h)_{h>0}} \min_{K\in\mathcal{T}_h} \vartheta_K < +\infty, \tag{11}$$

since there are at most ϖ reference elements defining the mesh sequence. We also introduce the mesh-dependent quantities

$$\mu_{\min} := \min_{K\in\mathcal{T}_h} \min_{i\in\mathcal{I}(K)} \frac{1}{|K|}\int_K \varphi_i(\boldsymbol{x})\,\mathrm{d}x, \quad \mu_{\max} := \max_{K\in\mathcal{T}_h} \max_{i\in\mathcal{I}(K)} \frac{1}{|K|}\int_K \varphi_i(\boldsymbol{x})\,\mathrm{d}x. \tag{12}$$

Note that $\mu_{\min} = \mu_{\max} = \frac{1}{n_{\mathrm{sh}}} = \frac{1}{d+1}$ for meshes uniquely composed of simplices and $\mu_{\min} = \mu_{\max} = 2^{-d}$ for meshes uniquely composed of parallelograms and cuboids. The main result is as follows:

Theorem 1. *Let* $A \subset \mathcal{A}$ *be an invariant set for* (1) *in the sense of Definition 1. Assume that* $\lambda_{\max}(A) := \max_{\boldsymbol{n}\in S^{d-1}(0,1)} \max_{\boldsymbol{u}_L,\boldsymbol{u}_R\in A} \lambda_{\max}(\boldsymbol{n}, \boldsymbol{u}_L, \boldsymbol{u}_R) < \infty$ *and* A *is convex. Assume that* $\boldsymbol{u}_{h0} \in A$ *and* τ *is such that* $2\tau\frac{\lambda_{\max}(A)}{\underline{h}}\frac{\mu_{\max}}{\mu_{\min}\vartheta_{\min}} \leq 1$*, then*

(i) A *is an invariant domain for the solution process* $\boldsymbol{u}_h^n \longmapsto \boldsymbol{u}_h^{n+1}$ *for all* $n \geq 0$*.*
(ii) Given $n \geq 0$ *and* $i \in \{1{:}I\}$*, let* $B \subset A$ *be a convex invariant set such that* $\mathbf{U}_l^n \in B$ *for all* $l \in \mathcal{I}(S_i)$*, then* $\mathbf{U}_i^{n+1} \in B$*.*

(iii) The following discrete entropy inequality holds for any admissible entropy pair $(\eta, \boldsymbol{q})$*, any* $n \geq 0$ *and any* $i \in \{1{:}I\}$*:*

$$\frac{m_i}{\tau}(\eta(\mathbf{U}_i^{n+1}) - \eta(\mathbf{U}_i^n)) + \int_D \nabla{\cdot}(\Pi_h \boldsymbol{q}(\boldsymbol{u}_h^n))\varphi_i \,\mathrm{d}x + \sum_{i \neq j \in \mathcal{I}(S_i)} d_{ij}\eta(\mathbf{U}_j^n) \leq 0.$$

References

[Chueh et al.(1977)] K. N. Chueh, C. C. Conley, and J. A. Smoller. Positively invariant regions for systems of nonlinear diffusion equations. *Indiana Univ. Math. J.*, 26(2):373–392, 1977.

[Frid(2001)] H. Frid. Maps of convex sets and invariant regions for finite-difference systems of conservation laws. *Arch. Ration. Mech. Anal.*, 160(3):245–269, 2001.

[Hoff(1979)] D. Hoff. A finite difference scheme for a system of two conservation laws with artificial viscosity. *Math. Comp.*, 33(148):1171–1193, 1979.

[Hoff(1985)] D. Hoff. Invariant regions for systems of conservation laws. *Trans. Amer. Math. Soc.*, 289(2):591–610, 1985.

[Nishida(1968)] T. Nishida. Global solution for an initial boundary value problem of a quasilinear hyperbolic system. *Proc. Japan Acad.*, 44:642–646, 1968.

[Perthame and Shu(1996)] B. Perthame and C.-W. Shu. On positivity preserving finite volume schemes for Euler equations. *Numer. Math.*, 73(1):119–130, 1996.

[Smoller(1983)] J. Smoller. *Shock waves and reaction-diffusion equations*, volume 258 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science]*. Springer-Verlag, New York-Berlin, 1983.

Asymptotic theory of reduced MHD models for fusion plasmas

Hervé Guillard

Magnetohydrodynamics (MHD) is a macroscopic theory describing electrically conducting fluids. An important application of MHD is the study of the stability of fusion plasmas in tokamak machines. A tokamak is a toroidal device in which hydrogen isotopes in the form of a plasma reaching a temperature of the order of the hundred of millions of Kelvins is confined thanks to a very strong applied magnetic field.

In its simplest form (ideal MHD and barotropic fluid) the hyperbolic MHD system writes :

$$\rho(p)\frac{D}{Dt}\boldsymbol{u} + \nabla(p + \boldsymbol{B}^2/2) - (\boldsymbol{B}.\nabla)\boldsymbol{B} = 0 \quad (1.1)$$

$$\frac{D}{Dt}\boldsymbol{B} - (\boldsymbol{B}.\nabla)\boldsymbol{u} + \boldsymbol{B}\nabla.\boldsymbol{u} = 0 \quad (1.2)$$

$$\frac{1}{\gamma p}\frac{D}{Dt}p + \nabla.\boldsymbol{u} = 0 \quad (1.4)$$

In these equation, $\boldsymbol{u}$ the velocity, $\boldsymbol{B}$ the magnetic field and p is the pressure and ρ is the density that, thanks to the barotropic assumption is only a function of pressure. The notation $D./Dt$ stands for the material derivative that is defined by $D\cdot/Dt = \partial_t \cdot + (\boldsymbol{u}.\nabla)\cdot$.

Instead of the system (1) the plasma fusion community has introduced the so-called

reduced MHD systems. A prototypical example of these models is the following one introduced by Strauss in [1].

$$\frac{\partial U}{\partial t} + [\varphi, U] - [\psi, J] - \frac{\partial}{\partial z} J = 0 \quad (2.1)$$

$$\frac{\partial \psi}{\partial t} + [\varphi, \psi] - \frac{\partial \varphi}{\partial z} = 0 \quad (2.2)$$

Here, ϕ is the electric potential, ψ the magnetic flux and U and J are defined by :

$$U = \nabla_\perp^2 \varphi \text{ and } J = \nabla_\perp^2 \psi$$

The notation $[.,.]$ is a poisson bracket defined by $[f, g] = \boldsymbol{e}_z.\nabla_\perp f \times \nabla_\perp g$ where $\boldsymbol{e}_z$ denote the unit vector in the toroidal direction while $\nabla_\perp$ and $\nabla_\perp^2$ are respectively the gradient and Laplace operator in the poloidal direction (the term poloidal denotes the plane perpendicular to $\boldsymbol{e}_z$ i.e the plane generated by the unit vectors $\boldsymbol{e}_x$ and $\boldsymbol{e}_y$).

Using the theory of singular limit of hyperbolic PDE with a large parameter introduced in [2], I show that (2) is a rigorous approximation of the full system (1) when the magnetic field is high. More precisely, I prove the following result :
Theorem 2 : Let ε be the ratio of the poloidal magnetic field over the toroidal magnetic field and assume that the initial velocity, magnetic field and pressure are defined by :

$$\begin{cases} \boldsymbol{u}(0, \boldsymbol{x})/V_A = \varepsilon(\boldsymbol{v}^0(\boldsymbol{x}) + \varepsilon \boldsymbol{v}^1(\varepsilon, \boldsymbol{x})) \\ \boldsymbol{B}(0, \boldsymbol{x})/B_0 = \boldsymbol{e}_z + \varepsilon(\mathcal{B}^0(\boldsymbol{x}) + \varepsilon \mathcal{B}^1(\varepsilon, \boldsymbol{x})) \\ p(0, \boldsymbol{x})/p_0 = \bar{p} + \varepsilon(q^0(\boldsymbol{x}) + \varepsilon q^1(\varepsilon, \boldsymbol{x})) \end{cases}$$

where $\bar{p}$ is a constant, the functions $\boldsymbol{v}^0, \mathcal{B}^0, q^0$ and $\boldsymbol{v}^1, \mathcal{B}^1, q^1$ are bounded in H^s and where the 0-th order initial data verifies :

$$\begin{cases} \nabla_\perp.\boldsymbol{v}^0(\boldsymbol{x}) = 0 & (3.1) \\ \nabla_\perp.\mathcal{B}_\perp^0(\boldsymbol{x}) = 0 & (3.2) \\ \exists f(z) \text{ such that } \mathcal{B}_z^0(\boldsymbol{x}) = f(z) - q^0(\boldsymbol{x}) & (3.3) \end{cases}$$

then the solution of the full MHD system (1) exists for a time T independent of ε and this solution converges in H^{s-1} to the solution of the reduced system. Details can be found in [3]

References

[1] H. R. Strauss, *Nonlinear, three-dimensional magnetohydrodynamics of non circular tokamaks*, Phys. Fluids,**19**, (1976), 134,

[2] S. Klainerman and A. Majda, *Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids*, Communications on Pure and Applied Mathematics, **34**,(1981),481–524

[3] H. Guillard, *The mathematical theory of reduced MHD models for fusion plasmas*, Inria Research Report RR-8715, (2015), hal-01145009.

Thermo-mechanically coupled systems using high-order schemes

Stefan Hartmann

Constitutive models in solid mechanics frequently are formulated by means of ordinary differential equations or differential-algebraic equations, which determine the evolution of so-called internal variable. Some of these internal variables are embedded in the stress-strain relation, where the stress has to be inserted into the balance of linear momentum (partial differential equation determining the deformation of the material body). In quasi-static formulations the weak form of the balance of linear momentum is used for drawing on finite elements. This concept of applying the principle of virtual displacements and a time-integration step of Backward-Euler type is mainly attributed to [1]. In the theses of [2, 3] this concept is interpreted as the application of the method of vertical lines, where in the first step the spatial discretization is carried out, followed by the time-discretization drawn on the internal variable of the constitutive equations of evolutionary-type. The classification and the relation to models of small strains is proposed in [4]. There, the resulting system of differential-algebraic equations (DAE-system) is solved using stiffly accurate, diagonally-implicit Runge-Kutta methods, where the resulting coupled system of equations is solved using the Multilevel-Newton algorithm, see [5, 6].

In temperature-dependent problems, the constitutive equations additionally depend on the temperature, which is connected to the balance of energy resulting in the transient heat equation. Here, it is exemplarily referred to [7]. The extension to the interpretation of applying the method of vertical lines is, for example, applied in [8], where the weak form of the heat equation defines a system of differential equations of first order as well yielding

$$\boldsymbol{F}(t,\boldsymbol{y}(t),\dot{\boldsymbol{y}}(t)) := \left\{ \begin{array}{c} \boldsymbol{g}(t,\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{q}) \\ \boldsymbol{C}_{\mathrm{p}}(t,\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{q})\dot{\boldsymbol{\Theta}}(t) - \boldsymbol{r}_{\Theta}(t,\boldsymbol{u},\dot{\boldsymbol{u}},\boldsymbol{\Theta},\boldsymbol{q}) \\ \dot{\boldsymbol{q}}(t) - \boldsymbol{r}_{\mathrm{q}}(t,\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{q}) \end{array} \right\} = \boldsymbol{0}. \tag{1}$$

$\boldsymbol{g}$ defines the discretized equilibrium conditions, $\boldsymbol{u}$ the unknown nodal displacements, $\boldsymbol{\Theta}$ the nodal temperatures, $\boldsymbol{q}$ the internal variables at all spatial integration points, $\boldsymbol{C}_{\mathrm{p}}$ the heat capacity matrix and t denotes the time.

In the second step of the method of vertical lines, the time-discretization has to be exploited, i.e. an appropriate method to solve the DAE-system (1) has to be looked for. In [8] SDIRK-methods are applied combined with the p-version of finite elements, where high-order polynomial approach in the spatial domain is performed. Of course, further methods can be investigated. BDF-methods are discussed in [9], see the literature cited therein as well. In [10] a comparison between DIRK, Rosenbrock and half-explicit Runge-Kutta methods is carried out indicating that a DIRK-method of 2nd order combined with time-adaptivity is an appropriate scheme for a huge number of applications (in this respect see [11] as well). All these investigations have been carried out using a monolithic approach using one code, i.e. nodal displacements and temperatures are solved in the same step. The question, whether there is a connection to staggered solution schemes

(partitioned approach), is treated in [12]. There, accelerating techniques of Gauss-Seidel schemes solving the resulting coupled non-linear system is discussed.

In the case of considering inertia effects, Eq.$(1)_1$ must be written in the form

$$\boldsymbol{M}\ddot{\boldsymbol{u}}(t) - \boldsymbol{g}(t, \boldsymbol{u}, \boldsymbol{\Theta}, \boldsymbol{q}) = \boldsymbol{0}, \tag{2}$$

i.e. the problem under consideration changes to a coupled system of ODEs of 1st and 2nd order. In the presentation two approaches are discussed. First, the transformation of 2nd order ODEs into a system of 1st order ODE, to which a DIRK-method is applied, and, second, a modification of the generalized-α method proposed in [13] is considered, see [14]. Both methods work for specific problems well.

References

[1] J. C. Simo, R. L. Taylor, Consistent tangent operators for rate-independent elastoplasticity, *Computer Methods in Applied Mechanics and Engineering* **48** (1985), 101–118.

[2] J. Wittekindt, *Die numerische Lösung von Anfangs-Randwertproblemen zur Beschreibung inelastischen Werkstoffverhaltens*. Doctoral thesis, Dept. Math., Univ. of Darmstadt (1991).

[3] P. Fritzen, *Numerische Behandlung nichtlinearer Probleme der Elastizitäts- und Plastizitätstheorie*. Doctoral thesis, Dept. Math., Univ. of Darmstadt (1997).

[4] P. Ellsiepen, S. Hartmann, *Remarks on the Interpretation of Current Non-linear Finite-Element-Analyses as Differential-Algebraic Equations*, International Journal for Numerical Methods in Engineering **51** (2001) 679–707.

[5] N. B. G. Rabbat, A. L. Sangiovanni-Vincentelli, H. Y. Hsieh, A Multilevel Newton Algorithm with Macromodeling and Latency for the Analysis of Large-Scale Nonlinear Circuits in the Time Domain, *IEEE Transactions on Circuits and Systems* **26** (1979) 733–740.

[6] Hartmann, S., A remark on the application of the Newton-Raphson method in non-linear finite element analysis, *Computational Mechanics* **36** (2005) 100–116.

[7] J. C. Simo, C. Miehe, Associative coupled thermoplasticity at finite strains: Formulation, numerical analysis and implementation, *Computer Methods in Applied Mechanics and Engineering* **98** (1992) 41–104.

[8] T. Netz, S. Hartmann, A monolithic finite element approach using high-order schemes in time and space applied to finite strain thermo-viscoelasticity, *Computers & Mathematics with Applications* **70**(7) (2015) 1457–1480.

[9] S. Eckert, H. Baaser, D. Gross, O. Scherf, A BDF2 integration method with stepsize control for elastoplasticity, *Computational Mechanics* **34** (2004) 377–386.

[10] S. Rothe, A.-W. Hamkar, K. J. Quint, S. Hartmann, Comparison of diagonal-implicit, linear-implicit and half-explicit Runge-Kutta methods in non-linear finite element analyses, *Archive of Applied Mechanics* **82** (2012), 1057 – 1074.

[11] A.-W. Hamkar, S. Hartmann, J. Rang, A stiffly accurate Rosenbrock-type method of order 2 applied to FE-analyses in finite strain viscoelasticity, *Applied Numerical Mathematics* **62** (2012) 1837 – 1848.

[12] S. Rothe, P. Erbts, A. Düster, S. Hartmann, Monolithic and partitioned coupling schemes for thermo-viscoplasticity, *Computer Methods in Applied Mechanics and Engineering* **293** (2015) 375 – 410.

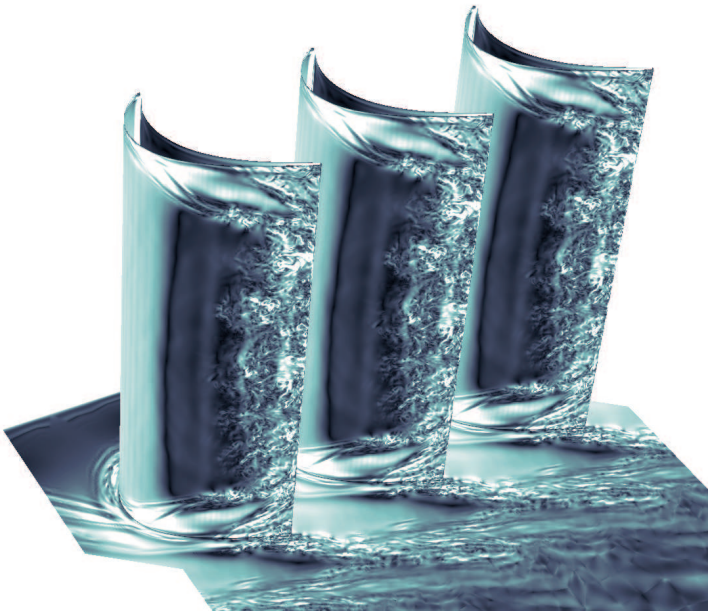
[13] J. Chung, G. M. Hulbert, A time integration algorithm for structural dynamics with improved numerical dissipation: The generalized-α method. *Journal of Applied Mechanics* **60**(2) (1993) 371 – 375.

[14] J. Rang, Coupling generalised-α methods: analysis, adaptivity, and numerics. in *V International Conference on Computational Methods for Coupled Problems in Science and Engineering, Coupled Problems 2013*, (Eds.) S. Idelsohn, M. Papadrakakis, B. Schrefler, (2013)

Development of an industrial simulation tool based on the discontinuous Galerkin Method in view of LES of turbomachinery.

Koen Hillewaert

(joint work with Corentin Carton de Wiart, Jean-Sbastien Cagnone, Michel Rasquin)



Skin friction on the blade of the T106C cascade at Re=110.000, computed by 4th order accurate DGM using Argo

The present contribution details key aspects in the development of a simulation platform, Argo, based on the *discontinuous Galerkin method (DGM)*, in view of *Large Eddy Simulation (LES)* of turbulent flows in turbomachinery. The aim is to exploit the capacity of DGM to guarantee high precision, required for reliable LES, on unstructured meshes, needed for representing the complex geometry of the machine. The algorithmic compacity and structured nature of the DGM can furthermore be exploited to provide excellent serial and parallel efficiency, a very important issue considering the computational resource requirements of LES. The industrial interest for the LES of turbomachinery flows stems from the need for predicting off-design regimes, transitional flow, instabilities and noise but also from the decreased reliability of the *Reynolds Averaged Navier Stokes (RANS)* turbulence models due to high increases in blade loading.

An important part of the research is focused on computational efficiency. An implicit time integration is used, in combination with Jacobian-free Newton-GMRES,

preconditioned with block Jacobi or block ILU. By freezing the preconditioner over several Newton iterations or even time steps, the computational cost is mainly determined by residual assembly and the application of the preconditioner.

Optimal data structures that exploit the underlying structured nature of DGM are then explained. During residual assembly, the storage of the solution in the form of matrices allows to use matrix-matrix products to provide highly efficient parametric operations. A simple reinterpretation of these matrices subsequently results in excellent data alignment for the element specific operations and the evaluation of the constitutive laws. A dedicated matrix format has been developed which is much more efficient for the block-sparse matrices with very large entries, characteristic for the residual Jacobian of DGM, than off-the-shelve packages. Finally by hiding ghost cell communication behind the volume operations, high parallel strong scaling is obtained, allowing to use the largest available machines.

An equally important part concerns turbulence modeling. Heuristic arguments are presented to indicate the need for high precision discretizations for LES in order to minimize interaction between the discretization and the *Subgrid Scale (SGS)* turbulence model. This interaction is unavoidable - even for non-dissipative schemes - since LES are by definition severely underresolved. Furthermore an explanation for the success of *Implicit LES (ILES)* modeling strategies - in which the discretisation replaces an explicit SGS model - for DGM is provided. The accuracy of the DGM/ILES approach is then illustrated on canonical test cases, both on high and low quality meshes, indicating similar accuracy as tailored academic codes. Finally some practical turbomachinery applications are shown.

To conclude important topics of future numerical research are outlined. These include the study of the interaction between shock capturing strategies and turbulence modeling, the development of wall-modeling strategies in combination with ILES and finally the elaboration of adaptation criteria for statistical data resulting from LES computations, where due to the underresolution, classical error estimates can not be used.

References

[1] C. Carton de Wiart, K. Hillewaert, L. Bricteux and G. Winckelmans, *Implicit LES of free and wall bounded turbulent flows based on the discontinuous Galerkin/symmetric interior penalty method*, International Journal of Numerical Methods in Fluids **78** (2015) 335–354

[2] K. Hillewaert and C. Carton de Wiart, *The discontinuous Galerkin method as an enabling technology for DNS and LES of industrial aeronautical applications (invited lecture)*, In "Advances in Simulation of Wing and Nacelle Stall - Results of the Closing Symposium of the DFG Research Unit FOR 1066, December 1-2, 2014, Braunschweig, Germany", Notes on Numerical Fluid Mechanics and Multidisciplinary Design, **131**, Springer, 2016.

Strong stability preserving linear multistep methods with variable step size

David I. Ketcheson
(joint work with Yiannis Hadjimichael, Adrian Németh, & Lajos Lóczi)

A linear multistep method (LMM) provides an approximate numerical solution to the initial value problem

$$u'(t) = f(u(t)) \qquad u(t_0) = u_0. \tag{1}$$

With a fixed step size, the method takes the form

$$u_n = \sum_{j=0}^{k-1} (\alpha_j u_{n-k+j} + h\beta_j f(u_{n-k+j})) \qquad n \geq k. \tag{2}$$

Strong Many authors have studied and developed strong stability preserving (SSP) LMMs. These LMMs are high order methods that guarantee the preservation of properties like positivity, contractivity, or monotonicity (relative to any convex functional), provided that the same property is satisfied under a forward Euler step. The SSP method provides this guarantee under the step size restriction

$$h \leq \mathcal{C} h_0, \tag{3}$$

where $\mathcal{C}$ is the SSP coefficient of the method and h_0 is the step size for strong stability when using the forward Euler method.

In practical numerical integration, it is often useful to adapt the step size based on local (in time) accuracy or stability considerations. This is true in particular for SSP methods, since the value h_0 typically depends on the solution u. A variable step size LMM takes the form

$$u_n = \sum_{j=0}^{k-1} \left(\alpha_{j,n} u_{n-k+j} + h_n \beta_{j,n} f(u_{n-k+j})\right). \tag{4}$$

In the present work we generalize some existing optimal SSP LMMs to allow for variation of the step size while maintaining the SSP property. This is challenging since the SSP coefficient depends on the method coefficients α, β, the method coefficients depend on the step size sequence, and the step size sequence depends on the SSP coefficient. Nevertheless, we are able to develop methods of second and third order that are provably convergent, optimal in a greedy sense (they maximize the size of allowed SSP step at each step), and for which the integration is guaranteed to terminate in a finite number of steps. The last result is non-trivial due to the nonlinear feedback between the previous step size sequence and the permissible size of the next step.

Details of this work can be found in the preprint [1].

References

[1] Y. Hadjimichael, D.I. Ketcheson, L. Lóczi, and A. Németh, *Strong stability preserving explicit linear multistep methods with variable step size*, arXiv:1504.04107.

Quantification of spurious mixing and dissipation and the effect of vertically adaptive meshes

Knut Klingbeil

(joint work with Mahdi Mohammadi-Aragh, Ulf Gräwe, Hans Burchard)

Truncation errors in numerical models can cause the violation of basic conservation properties of the continuous solution. The discretisation in terms of Finite-Volumes (FVs) offers the straightforward implementation of advection schemes conserving global first moments (momentum and tracer). But, usually for monotonicity reasons, these schemes do not conserve the associated global second moments (kinetic energy and tracer variance). The local quantification of spurious variance decay is essential for detecting hotspots of spurious dissipation and mixing as well as the comparison with their physically induced counterparts. For the first-order upstream (FOU) scheme Morales Maqueda and Holloway (2006) quantified the local variance decay as the variance destroyed by the final recombination of advected subvolumes inside one FV-cell. Burchard and Rennau (2008) generalised an alternative derivation for the variance decay of the 1D FOU scheme and proposed a diagnostic method (hereafter BR08) for arbitrary 3D advection schemes. The application of this method to quantify spurious dissipation in 1D was demonstrated by Burchard (2012).

Klingbeil et. al. (2014) argued for a general analysis method (hereafter K14) of discrete variance decay (DVD) that can be used for the quantification of both, spurious and physically induced variance decay. It was shown that the BR08 approach does not recover a physically sound definition of discrete variance within a FV-cell. In contrast, the K14 method extends the idea of Morales Maqueda and Holloway to arbitrary transport schemes by considering also the possible variance growth during the decomposition of a FV-cell into subvolumes (see Fig. 1).

Considering the discrete second moments associated with the subvolumes, it is possible to calculate the DVD rate $\chi_i^{(n+1)}(\varphi)$ within a single FV-cell,

$$\frac{V_i^{(n+1)}\left(\varphi_i^{(n+1)}\right)^2 - V_i^{(n)}\left(\varphi_i^{(n)}\right)^2}{\triangle t} + \left[A_{i'}\left(J_{\varphi^2}\right)_{i'}\right]_{i'=i-1/2}^{i'=i+1/2} = -V_i^{(n+1)}\chi_i^{(n+1)}(\varphi)\,, \tag{1}$$

with the discrete (for brevity here only advective) fluxes of second moments $\left(J_{\varphi^2}\right)_{i+1/2}$ defined directly in terms of the associated fluxes of first moments $J_{i+1/2}(\varphi)$:

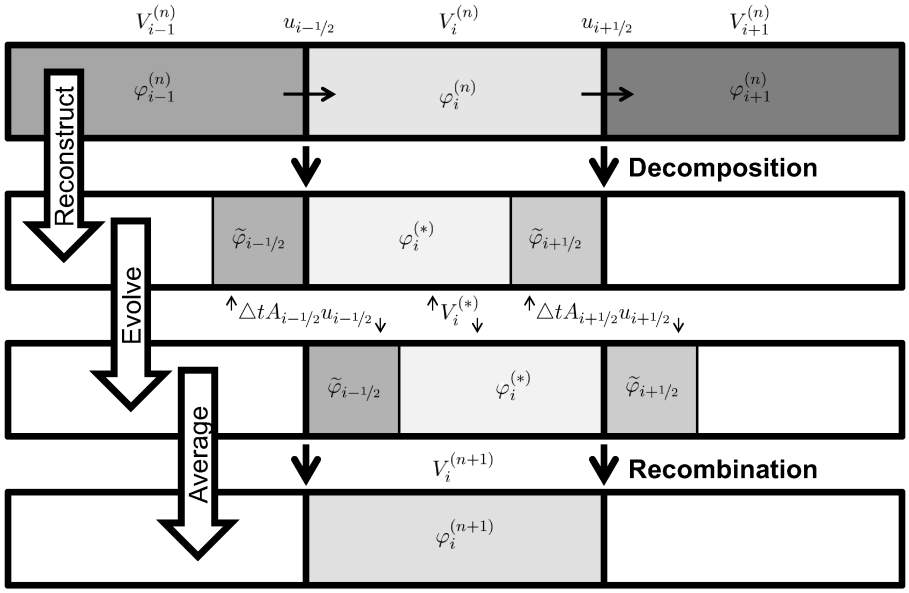


FIGURE 1. Advection in a FV-framework (taken from Klingbeil et. al. (2014)): The schema shows the exemplary decomposition of FV-cells into subvolumes, to be advected, and their recombination on a fixed grid for uniform 1D flow. V, A, u, φ and $\tilde{\varphi}_{i\pm1/2}$ denote the volumes and interfacial areas of the FV-cells, the velocity, the prognostic quantity and its approximated interfacial values.

$$\left(J_{\varphi^2}\right)_{i+1/2} = \tilde{\varphi}_{i+1/2} J_{i+1/2}\left(\varphi\right) = u_{i+1/2}\left(\tilde{\varphi}_{i+1/2}\right)^2. \tag{2}$$

Based on this new analysis method Klingbeil et. al. (2014) also presented, how the spurious dissipation of kinetic energy on 3D staggered grids can be diagnosed from the individual DVD rates of the single velocity components.

The analysis methods have been implemented into the coastal ocean model GETM (Burchard and Bolding, 2002; Hofmeister et. al. 2010; Klingbeil and Burchard, 2013). By providing valuable insight into the local strength and origin of spurious dissipation and mixing in realistic ocean modelling applications (Rennau and Burchard, 2009; Hofmeister et. al. 2011; Gräwe et. al. 2015, Mohammadi-Aragh et al. 2015), the analysis methods promote the assessment of advection

schemes and subgrid-scale closures as well as the development of new numerical techniques reducing spurious variance decay.

For example, Hofmeister et. al. (2010) developed vertically adaptive meshes that increase the resolution in the vicinity of vertical density gradients (thereby reducing truncation errors during the vertical advection) and also consider the lateral evolution of the density field. The latter offers a tendency to isopycnal coordinates (model layers aligned with levels of constant density), which decrease the spurious mixing during the advection along the layers (ideally no density gradients) and during the vertical advection (ideally no grid-related vertical transport).

Direct feedback of the diagnosed DVD rates to the transport schemes and mesh adaptation is planned to be investigated in the future.

References

[1] Burchard, H., 2012. Quantification of numerically induced mixing and dissipation in discretisations of shallow water equations. International Journal on Geomathematics, 3, 51–65.

[2] Burchard, H., Bolding, K., 2002. GETM – a General Estuarine Transport Model. Scientific Documentation. Technical Report EUR 20253 EN. European Commission.

[3] Burchard, H., Rennau, H., 2008. Comparative quantification of physically and numerically induced mixing in ocean models. Ocean Modelling, 20, 293–311.

[4] Gräwe, U., Holtermann, P., Klingbeil, K., Burchard, H., 2015. Advantages of vertically adaptive coordinates in numerical models of stratified shelf seas. Ocean Modelling, 92, 56–68.

[5] Hofmeister, R., Beckers, J.M., Burchard, H., 2011. Realistic modelling of the exceptional inflows into the central Baltic Sea in 2003 using terrain-following coordinates. Ocean Modelling, 39, 233–247.

[6] Hofmeister, R., Burchard, H., Beckers, J.M., 2010. Non-uniform adaptive vertical grids for 3D numerical ocean models. Ocean Modelling, 33, 70–86.

[7] Klingbeil, K., Burchard, H., 2013. Implementation of a direct nonhydrostatic pressure gradient discretisation into a layered ocean model. Ocean Modelling, 65, 64–77.

[8] Klingbeil, K., Mohammadi-Aragh, M., Gräwe, U., Burchard, H., 2014. Quantification of spurious dissipation and mixing – Discrete Variance Decay in a Finite-Volume framework. Ocean Modelling, 81, 49–64.

[9] Mohammadi-Aragh, M., Klingbeil, K., Brüggemann, N., Eden, C., Burchard, H., 2015. The impact of advection schemes on restratifiction due to lateral shear and baroclinic instabilities. Ocean Modelling, 94, 112–127.

[10] Morales Maqueda, M.A., Holloway, G., 2006. *Second-order moment advection scheme applied to Arctic Ocean simulation.* Ocean Modelling, 14, 197–221.

[11] Rennau, H., Burchard, H., 2009. Quantitative analysis of numerically induced mixing in a coastal model application. Ocean Dynamics, 59, 671–687.

Provably Stable Nodal Discontinuous Galerkin Methods on Curved Elements

David A. Kopriva
(joint work with Andrew R. Winters, Gregor J. Gassner)

To get accurate solutions to fluid flow problems one can use *high order approximations* (e.g. $4 \leq \text{Order} \lesssim 20$). High order methods have great promise, among the most important being the potential for greatly improved computational efficiency when compared to low order methods. Unfortunately, the positive features of high order discretizations, like low dissipation, mean they often lack the robustness needed for industrial level computations. For this reason, there is a need for provably stable variants.

We are developing arbitrary order provably stable discontinuous Galerkin spectral element methods for curved elements on both static and moving domains. The starting point is the system of conservation laws

$$\mathbf{q}_t + \nabla \cdot \vec{\mathbf{f}} = \mathbf{q}_t + \sum_{i=1}^{3} \frac{\partial \mathbf{f}_i}{\partial x_i} = 0 \tag{1}$$

on a three dimensional domain with moving boundaries, $\Omega(\vec{x}, t)$, where $\vec{x} = (x_1, x_2, x_3) = (x, y, z) = x\hat{x} + y\hat{y} + z\hat{z}$. Here we denote the solution and flux vector components by bold face and spatial vectors by overbars. In this work, we consider linear fluxes of the form $\mathbf{f}_i = A_i \mathbf{q} \quad i = 1, 2, 3$. The domain, which may have moving boundaries is subdivided into hexahedral elements each of which is mapped to the reference element $E = [-1, 1]^3$. On the reference element,

$$\frac{\partial \mathcal{J}\mathbf{q}}{\partial \tau} + \sum_{i=1}^{3} \frac{\partial}{\partial \xi^i} \left(\mathcal{A}^i \mathbf{q} \right) = 0 \tag{2}$$

where

$$\mathcal{A}^i \left(\vec{\xi}, \tau \right) = \tilde{A}^i - J\vec{a}^i \cdot \vec{x}_\tau I, \tag{3}$$

$$\tilde{A}^i = \mathcal{J}\vec{a}^i \cdot \sum_{j=1}^{3} A_j \hat{x}_j, \tag{4}$$

$\mathcal{J}$ is the Jacobian of the transformation, $\vec{a}^i$ is the contravariant basis vector and $\vec{x}_\tau$ is the change due to the motion of the mesh.

The key is to rewrite the system of equations in skew symmetric form, which on the reference element is the average of the conservative and nonconservative form

$$\frac{1}{2} \left\{ \frac{\partial \mathcal{J}\mathbf{q}}{\partial \tau} + \frac{\partial \mathcal{J}}{\partial \tau}\mathbf{q} + \mathcal{J}\frac{\partial \mathbf{q}}{\partial \tau} \right\} + \frac{1}{2} \sum_{i=1}^{3} \frac{\partial}{\partial \xi^i} \left(\mathcal{A}^i \mathbf{q} \right) + \frac{1}{2} \left\{ \sum_{i=1}^{3} \frac{\partial}{\partial \xi^i} \left(\mathcal{A}^i \right) \mathbf{q} + \sum_{i=1}^{3} \mathcal{A}^i \frac{\partial \mathbf{q}}{\partial \xi^i} \right\} = 0. \tag{5}$$

A weak form is then created by multiplying by a test function $\vec{\phi} \in \mathbb{L}^2$ and integrating over the reference element. Finally, a polynomial ansatz is made where

the solution, flux, metric terms and Jacobian are approximated by polynomials of degree N. All terms with derivatives of the solution are integrated by parts and the boundary fluxes that appear are replaced by a numerical flux. The resulting weak formulation for the solution is

$$\begin{aligned}\frac{1}{2}\left(\frac{\partial \mathbb{I}^N\left(J\mathbf{Q}\right)}{\partial\tau}+J\frac{\partial\mathbf{Q}}{\partial\tau},\vec{\phi}\right)_N &+\left(\vec{\mathbf{F}}^*\right)^T\vec{\phi}\Big|_{\partial E,N}-\frac{1}{2}\sum_{i=1}^{3}\left(\vec{\mathbf{F}}^i,\frac{\partial}{\partial\xi^i}\vec{\phi}\right)_N\\ &-\frac{1}{2}\sum_{i=1}^{3}\left(\mathbf{Q},\frac{\partial\mathbb{I}^N\left(\mathbb{I}^{\mathbb{N}}\left(\mathcal{A}^i\right)\vec{\phi}\right)}{\partial\xi^i}\right)_N=0,\end{aligned} \tag{6}$$

where $\vec{\mathbf{F}}^*$ is the numerical flux, and the discrete inner products are Gauss-Lobatto quadratures.

The main result is that the approximation is stable, and satisfies the discrete $\mathbb{L}^2$ energy bound

$$\begin{aligned}\|\mathbf{Q}(T)\|^2_{J,N}+\int_0^T\sum_{\substack{boundary\\faces}}\left\{\sum_{r,s=0}^{N}W_rW_s\left\|\sqrt{|A|_{\vec{m}(r,s)}}\mathbf{Q}_{\vec{m}(r,s)}-\left(\sqrt{|A|_{\vec{m}(r,s)}}\right)^{-1}\left|A^-\right|_{\vec{m}(r,s)}\mathbf{Q}_\infty\right\|_2^2\right\}dt\\ \leqslant\|\mathbf{Q}(0)\|^2_{J,N}+\int_0^T\sum_{\substack{boundary\\faces}}\left\{\sum_{r,s=0}^{N}W_rW_s\mathbf{Q}_\infty^T\left|A^-\right|^2_{\vec{m}(r,s)}\mathbf{Q}_\infty\right\}dt\end{aligned} \tag{7}$$

so that when $\mathbf{Q}_\infty=0$, the approximation is *stable*,

$$\|\mathbf{Q}(T)\|_{J,N}\leqslant\|\mathbf{Q}(0)\|_{J,N} \tag{8}$$

Additional properties are that the skew symmetric approximation remains conservative and preserves constant states for constant coefficient problems.

Enhancements of numerical schemes with tetrahedral-based mesh adaptation

Adrien Loseille

Metric-based mesh adaptation has been used to capture accurately anisotropic physical phenomena while optimizing the ratio between accuracy and CPU time. A lot of 3D successful examples on real-life problems have already proved its efficiency [1, 2, 6]. However, one question remains: are the adaptive computations really anisotropic or optimal ? Apart from its simplicity, this question raises, as we will see, many other capital issues: assessment of the numerical solution, convergence of the computation at the theoretical order, automatic detection and capturing of all the scales of the solution, ... Consequently, answering positively to these questions is not straightforward as we face both theoretical and practical difficulties. In the sequel, we review the main problematics and how we address them with the continuous mesh framework.

The report is organized as follows. In the first section, we motivate the use of a continuous mesh model and state the main result to derive the optimal mesh minimizing the interpolation error. In the second section, this framework is applied to the case of numerical solutions, and the enhancements of mesh adaptation are illustrated on two numerical examples.

1. Controlling the interpolation error: an ill-posed problem

In its more general form, the problem of mesh adaptation consists in finding the mesh $\mathcal{H}$ of Ω that minimizes a given error for a given function u. We consider here the linear interpolation error $u - \Pi_h u$ controlled in $\mathbf{L}^p$ norm so that the problem reads:

$$\text{Find } \mathcal{H}_{opt} \text{ having } N \text{ nodes such that } E(\mathcal{H}_{opt}) = \min_{\mathcal{H}} \|u - \Pi_h u\|_{\mathbf{L}^p(\Omega)} \tag{1}$$

As it, (1) is a global combinatorial problem which turns out to be intractable practically. Indeed, both topology and vertices location need to be optimized. Consequently, simpler problems are considered to approximate the solution. All these strategies have in common the resolution of a local problem. Consequently, such error minimizations are equivalent to a steepest descent algorithm that converges only to a local minimum. This drawback arises because of considering directly the minimization on a discrete mesh. To overcome this issue, we define a continuous mesh model and continuous interpolation. From a theoretical point of view, (1) is then recast as a continuous problem solvable thanks to a calculus of variations. From a practical point of view, the optimal solution is a metric-tensor that is directly used by any adaptive mesh generator to generate a computational adapted mesh.

A continuous mesh [3, 4] of a domain Ω is a Riemannian metric space $\mathbf{M} = (\mathcal{M}(\mathbf{x}))_{\mathbf{x}\in\Omega}$, where at each point $\mathcal{M}(\mathbf{x})$ is a definite positive matrix. To emphasize the analogy with a discrete mesh, we rewrite $\mathbf{M}$ as :

$$\mathcal{M}(\mathbf{x}) = d^{\frac{2}{3}}(\mathbf{x})\, \mathcal{R}(\mathbf{x}) \begin{pmatrix} r_1^{-\frac{2}{3}}(\mathbf{x}) & & \\ & r_2^{-\frac{2}{3}}(\mathbf{x}) & \\ & & r_3^{-\frac{2}{3}}(\mathbf{x}) \end{pmatrix} {}^t\mathcal{R}(\mathbf{x}),$$

where $\mathcal{R}(\mathbf{x})$ is the eigenvectors matrix and $(\lambda_i)_{i=1,3}$ the eigenvalues of $\mathcal{M}(\mathbf{x})$, the density d is equal to: $d = (h_1 h_2 h_3)^{-1} = (\lambda_1\lambda_2\lambda_3)^{\frac{1}{2}}$, the anisotropic quotients r_i are equal to: $r_i = h_i^3 (h_1 h_2 h_3)^{-1}$. The following result is then used to recast (1) in a continuous setting:

Theorem 1.1. *For any tetrahedron K such all its edges have a length one in $\mathcal{M}$, i.e., ${}^t\mathbf{e}_i\, \mathcal{M}\, \mathbf{e}_i = 1$, the interpolation error of a twice differentiable function u on K is bounded by:*

$$\|u - \Pi_h u\|_{\mathbf{L}^1(K)} \leq \frac{\sqrt{2}}{240} \det(\mathcal{M}^{-\frac{1}{2}})\, \mathrm{trace}(\mathcal{M}^{-\frac{1}{2}}\, H_u\, \mathcal{M}^{-\frac{1}{2}}), \tag{2}$$

where H_u is the Hessian of u.

We note that the right hand side of (2) depends only on the local metric and Hessian of the solution. It is then possible to use this bound to set a well-posed global optimization problem where the unknown becomes the continuous mesh:

$$\text{Find } \mathbf{M}_{\mathbf{L}^p} = \min_{\mathbf{M}} \int_\Omega \left(trace(\mathcal{M}^{-\frac{1}{2}}(\mathbf{x})\, H_u(\mathbf{x})\, \mathcal{M}^{-\frac{1}{2}}(\mathbf{x})) \right)^p \mathrm{d}\mathbf{x} \tag{3}$$

under the constraint:

$$\mathcal{C}(\mathbf{M}) = \int_\Omega d = N.$$

The constraint on the complexity is added to avoid the trivial solution where all h_i are zero which provides a null error. The constant and term $\det(\mathcal{M}^{-\frac{1}{2}})$ of (2) are dropped off the functional as they only account for the density and are taken into account in the constraint directly. Contrary to the discrete analysis, this problem can be solved globally by using a calculus of variations. In [4], it is proved that (3) admits a unique solution. Deriving practically a discrete mesh from the optimal continuous one consists in generating a unit-mesh by using an adaptive mesh generator [5]. In addition, the following properties hold:

Theorem 1.2. *Let u be a twice continuously differentiable function defined on $\Omega \subset \mathbb{R}^3$, the optimal continuous mesh $\mathbf{M}_{\mathbf{L}^p}(u)$ minimizing (3) reads locally:*

$$\mathcal{M}_{\mathbf{L}^p} = D_{\mathbf{L}^p}\ \det(|H_u|)^{\frac{-1}{2p+3}}\, |H_u|,\ \textit{with}\ D_{\mathbf{L}^p} = N^{\frac{2}{3}} \left(\int_\Omega \det(|H_u|)^{\frac{2p}{2p+3}} \right)^{-\frac{2}{3}}. \tag{4}$$

It verifies the following properties:

- $\mathbf{M}_{\mathbf{L}^p}(u)$ *is unique*
- $\mathbf{M}_{\mathbf{L}^p}(u)$ *is locally aligned with the eigenvectors basis of H_u and has the same anisotropic quotients as H_u*
- $\mathbf{M}_{\mathbf{L}^p}(u)$ *provides an optimal explicit bound of the interpolation error in $\mathbf{L}^p$ norm:*

$$\|u - \Pi_h u\|_{\mathbf{L}^p(\Omega)} = 3\, N^{-\frac{2}{3}} \left(\int_\Omega \det\left(|H_u|\right)^{\frac{p}{2p+3}} \right)^{\frac{2p+3}{3p}}. \tag{5}$$

- *For a sequence of continuous meshes having an increasing complexity with the same orientation and anisotropic quotients $(\mathbf{M}^N_{\mathbf{L}^p}(u))_{N=1\ldots\infty}$, the asymptotic order of convergence verifies:*

$$\|u - \Pi_h u\|_{\mathbf{L}^p(\Omega)} \leq \frac{Cst}{N^{2/3}}. \tag{6}$$

2. Enhancements for the numerical schemes

When a numerical solver is used to solve a partial differential equation, the continuous solution u is not known and only the numerical approximation u_h is provided. It is then impossible to apply directly the previous error estimate on u_h which is only piecewise linear. Instead, we recover a solution $R_h(u_h)$ which is a better approximation of u_h in a given norm $\|.\|$:

$$\|u - R_h u_h\| \leq \alpha \|u - u_h\| \quad \text{where } \ 0 \leq \alpha < 1\,.$$

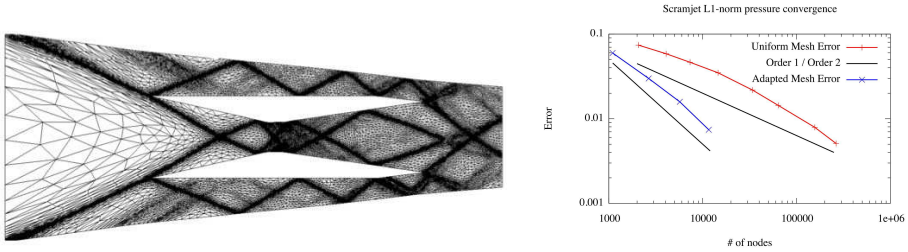


FIGURE 1. Illustration of the second order of convergence and early capturing on a supersonic flow inside a scramjet.

Various recovery operators exist, see [7]. Then, from the triangle inequality we deduce:

$$\|u - u_h\| \ \leq \ \frac{1}{1-\alpha}\|R_h u_h - u_h\| \, .$$

If the reconstruction operator R_h has the property: $\Pi_h R_h \phi_h = \phi_h$ for all piecewise linear function ϕ_h, we can then bound the approximation error of the solution by the interpolation error of the reconstructed function $R_h u_h$:

$$\|u - u_h\| \leq \frac{1}{1-\alpha}\|R_h u_h - \Pi_h R_h u_h\| = \frac{6\mathcal{N}^{-\frac{2}{3}}}{1-\alpha}\left(\int_\Omega \det\left(|H_{R_h u_h}|\right)^{\frac{p}{2p+3}}\right)^{\frac{2p+3}{3p}} .$$

In Figures 1 and 2, we verify that the properties of Theorem (1.2) hold for numerical solutions issued from Computational Fluid Dynamics. In Figure 1, a supersonic flow is considered inside a scramjet geometry while in Figure 2, a supersonic flow is computed around a F15 geometry equipped with a lowboom quiet-spike. We observe that a second order of convergence is obtained on a sequence of adapted meshes whereas only a convergence of order one is obtained on a sequence of uniform meshes due the discontinuities (shocks) in the flow field. In addition, using adaptive meshes allows us to early capture the smallest scales of the solution, by reducing the solver dissipation, see Figure 2.

REFERENCES

[1] J. Bruchon, H. Digonnet, T. Coupez, *Using a signed distance function for the simulation of metal forming process: formulation of the contact condition and mesh adaptation. From Lagrangian approach to an Eulerian approach*, Int. J. Numer. Meth. Engng 78 (8) (2009) 980–1008.

[2] C. Dobrzynski, P. Frey, *Anisotropic Delaunay Mesh Adaptation for Unsteady Simulations*, In Proceedings of the17th International Meshing Roundtable, Springer, 177–194, 2008.

[3] A. Loseille, F. Alauzet,*Continuous mesh framework. PartI: well-posed continuous interpolation error*, SIAM J. Numer. Anal. 49(1)(2011) 38-60.

[4] A. Loseille, F. Alauzet,*C*ontinuous mesh framework. Part II: validations and applications, SIAM J. Numer. Anal. 49(1)(2011)61–86.

[5] A. Loseille, R. Löhner, *Adaptive anisotropic simulations in aerodynamics*, In 48th AIAA Aerospace Sciences Meeting, 2010–169, Orlando, FL, USA, 2010.

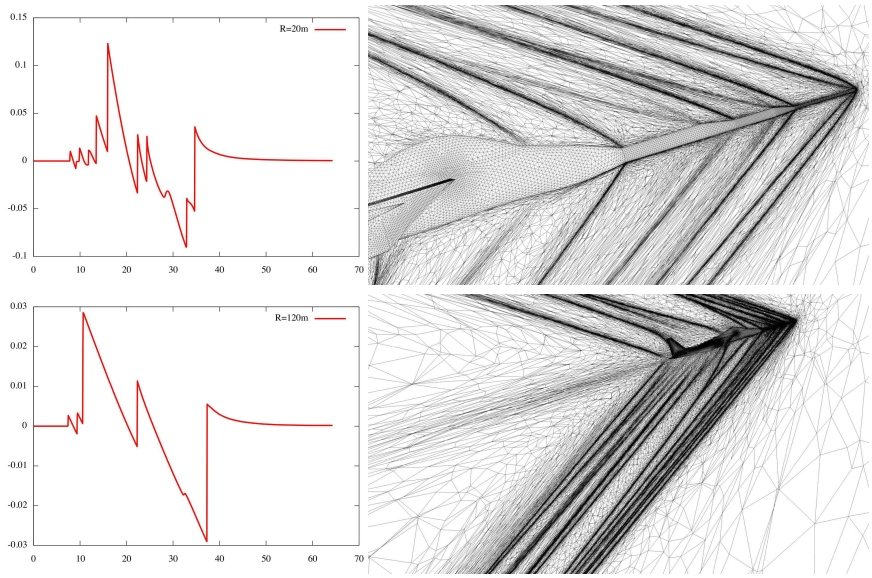


FIGURE 2. Illustration of the reduction of the numerical solver dissipation during the propagation of small scales shock waves.

[6] J.-F. Remacle, X. Li, M. Shephard, J. Flaherty, *Anisotropic adaptive simulation of transient flows using discontinuous Galerkin methods*, Int. J. Numer. Meth. Engng. 62 (2005) 899–923
[7] O. Zienkiewicz, J. Zhu, *The super convergent patch recovery and a posteriori error estimates. Part 1 : The recovery technique*, Int. J. Numer. Meth. Engng 33 (7) (1992) 1331–1364

Measure valued and Statistical solutions of systems of conservation laws.

SIDDHARTHA MISHRA

We consider nonlinear systems of conservation laws, such as the compressible Euler equations of gas dynamics, in several space dimensions. It is well known that discontinuities, such as shock waves, can form even when the initial data is smooth. Hence, the solutions are sought in a weak (distributional) sense. However, weak solutions are not unique and need to be supplemented by additional admissibility criteria, the so-called entropy conditions. Entropy solutions are shown to be well posed for systems of conservation laws in one space dimension, at least for initial data with sufficiently small total variation. However, there are no global existence results in several space dimensions. Furthermore, recent work has shown that infinitely many entropy solutions may exist for multi-dimensional systems of conservation laws.

A variety of numerical methods have been developed to discretize systems of conservation laws efficiently. These include finite volume (difference) schemes,

based on Riemann solvers, non-oscillatory reconstruction procedures such as TVD, ENO, WENO and strong stability preserving Runge-Kutta time stepping as well as Discontinuous Galerkin finite element methods. However, very few rigorous convergence results for these numerical methods are available.

In fact, in a recent paper [2], the authors showed that several state of the art numerical methods may not even converge to any function as the mesh is refined. Oscillations on finer and finer scales are observed and they impede convergence under mesh refinement. Based on these observations, it was postulated in [2] that *entropy measure valued solutions* are an appropriate solution framework for systems of conservation laws, particularly in several space dimensions. Entropy measure valued solutions are young measures i.e, parametrized probability measures, and were introduced by DiPerna in the mid 80s as a solution concept for conservation laws.

The main contribution of [2] lay in the design of an efficient algorithm to compute entropy measure valued solutions. This algorithm was based on exploiting the equivalence between Random fields and Young measures and on Monte Carlo type ensemble averaging. A set of criteria were proposed and it was shown that if the underlying numerical discretizations satisfied these criteria, the resulting approximate young measures generated by the algorithm, converged to an entropy measure valued solution as the mesh was refined. The arbitrary high-order ENO based schemes of [1] and entropy stable shock capturing schemes of [4] were examples of schemes that satisfied these criteria. Hence, the results of [2] provide a convergence framework for generic systems of conservation laws in several space dimensions.

A large number of numerical experiments were presented in [2]. These experiments validated the convergence framework by showing that statistical quantifies of interest, such as moments and one-point pdfs, converged under mesh refinement even when individual realizations of the approximation did not converge. Furthermore, non-atomic measure valued solutions were observed for atomic initial data. Hence, this spreading out of the measure by the solution operator implies that one might not be able to close the well-posedness of solutions of systems of conservation laws within integrable functions.

However, entropy measure valued solutions are not unique. It is argued in a recent paper [3] that the lack of information about spatial correlations in measure valued solutions allows non-uniqueness. Instead, the authors in [3] propose a much narrower concept, that of *Statistical solutions*, as solution frameworks for systems of conservation laws. Statistical solutions are time-parametrized probability measures on L^p. It was shown in [3] that this measure on infinite dimensional function spaces is equivalent to an infinite hierarchical family of correlation measures. Thus, the statistical solution embeds all possible correlation structures in the flow. The well-posedness of statistical solutions for scalar conservation laws was shown in [3]. Furthermore, it is shown that the ensemble based algorithm of [2] actually converges to a statistical solution, with the underlying measure valued solution being the first correlation marginal of the statistical solutions. As

moments, 1-point pdfs, two-structure functions, power spectra, increment pdfs, multi-point correlators are all admissible observables in the sense of [3], statistical solutions promises to be an appropriate framework for describing turbulent flows. Further properties of statistical solutions, including their well-posedness and long time behavior, are the subject of future work.

References

[1] U.S. Fjordholm, S. Mishra and E. Tadmor, Arbitrarily high order accurate entropy stable essentially non-oscillatory schemes for systems of conservation laws, *SIAM Jl. Num. Anal,* 50(2), 2012, 544–573.

[2] U. S. Fjordholm, R. Käppeli, S. Mishra and E. Tadmor, Construction of approximate entropy measure valued solutions for hyperbolic systems of conservation laws. *Preprint*, available as arXiv:1402.0909 [math.NA].

[3] U. S. Fjordholm, S. Lanthaler and S. Mishra, Statistical solutions of systems of conservation laws I. Theory, *Preprint*, 2015.

[4] A. Hiltebrand and S. Mishra, Entropy stable shock capturing streamline diffusion space-time discontinuous Galerkin (DG) methods for systems of conservation laws, *Numer. Math,* 126 (1), 2014, 103–151

Well Posed Problems and Boundary Conditions in Computational Fluid Dynamics

Jan Nordström

All numerical calculations will fail to provide a reliable answer unless the continuous problem under consideration is well posed. Well-posedness depends in most cases only on the choice of boundary conditions. In this paper we will highlight this fact by discussing well-posedness of the most important equations in fluid dynamics: the time-dependent compressible Navier-Stokes equations.

In particular, we will discuss *i*) how many boundary conditions are required, *ii*) where to impose them and *iii*) which form they should have. The procedure is based on the energy method and generalizes the characteristic boundary procedure for the Euler equations to the compressible Navier-Stokes equations.

Once the boundary conditions in terms of *i-iii*) are known, one issue remains; they can be imposed weakly or strongly. The weak and strong imposition is discussed for the continuous case. It will be shown that the weak and strong boundary procedures produce identical solutions and that the boundary conditions are satisfied exactly also in the weak procedure.

We conclude by relating the well-posedness results to energy-stability of the numerical approximation. It is shown that the results obtained for weak boundary conditions in the well-posedness analysis lead directly to corresponding stability results for the discrete problem.

Orthogonal Polynomials and their Application in Spectral Difference Methods

Philipp Öffner
(joint work with Thomas Sonar)

We are interested in numerical methods to solve hyperbolic conservation laws in two space dimensions

$$u_t(\boldsymbol{x},t) + \nabla \cdot F\left(u(\boldsymbol{x},t)\right) = 0, \qquad \boldsymbol{x} \in \Omega \subseteq \mathbb{R}^2, t \in \mathbb{R}^+.$$

We consider here the Spectral Difference Method described in [1] by the general classical orthogonal polynomials on triangles and their specific modal filters. The SD method is a very powerful and nowadays often used method to solve hyperbolic conservation laws. It combines the basic ideas of finite difference and spectral methods and the simple formulation and the possibility to reach spectral accuracy in the approximation renders it so attractive. The classical approach in the SD method uses Lagrange polynomials, but in [1] the authors employ Proriol-Koornwinder-Dubiner polynomials. We extend the method by using Appell-Proriol-Koornwinder polynomials, which are the general classical orthogonal polynomials on triangles.
These polynomials are defined on the unit triangle $\mathbb{T} := \{(x,y) \in \mathbb{R}^2 | x \geq 0, y \geq 0, x+y \leq 1\}$ and $h(x,y) := x^{\alpha-1}y^{\beta-1}(1-x-y)^{\gamma-\alpha-\beta}$ ($\alpha,\beta,\gamma \in \mathbb{N}, \gamma > \alpha+\beta-1$ and $\mathbb{N} = \{1,2,\cdots\}$) is the weight function, given in this domain. Note that we only consider $\alpha,\beta,\gamma \in \mathbb{N}$ for simplicity. In principle, $\alpha,\beta,\gamma \in \mathbb{R}_0^+$ is possible. The $P_m^{\alpha,\beta}$ are the classic Jacobi polynomials.

Definition 1 The polynomials $A_{m,l}(x,y)$, $m,l \in \mathbb{N}_0$, defined as

$$A_{m,l}(x,y) := P_m^{\alpha-1,2l+\gamma-\alpha}(1-2x)P_l^{p,\beta-1}\left(\frac{2y}{1-x}-1\right)(1-x)^l$$

on $\mathbb{T}$ are called **Appell-Proriol-Koornwinder polynomials** (APK polynomials).

In the first step we present two theoretical results. We prove spectral convergence for the APK series coefficients $\tilde{u}_k$ of a C^∞-function u. Spectral convergence or spectral accuracy means that the $\tilde{u}_k$ decay faster than any power k^{-j} for all $j \in \mathbb{N}$. Furthermore we also show spectral accuracy for the truncation error in a weighted L^2-norm and in the pointwise sense. This behaviour is the justification to utilize the APK polynomials in a spectral method as the SD method. The proofs can be found in [2].

Theorem 1 *Let u be a function in the space $H^{2k}(\mathbb{T},h), k \in \mathbb{N}$ where $h(x,y) = x^{\alpha-1}y^{\beta-1}(1-x-y)^{\gamma-\alpha-\beta}$ is the weight with $\alpha,\beta,\gamma \in \mathbb{N}$.*
The APK expansion of u is given by

$$P_N u(x,y) = \sum_{\substack{l+m \leq N \\ l,m \in \mathbb{N}_0}} \tilde{u}_{m,l} A_{m,l}(x,y); \quad \tilde{u}_{m,l} = \frac{(u;A_{m,l})_{\mathrm{L}^2(\mathbb{T},h)}}{(A_{m,l};A_{m,l})_{\mathrm{L}^2(\mathbb{T},h)}}.$$

The following estimates hold:

$$(A_{m,l}; A_{m,l})^{\frac{1}{2}}_{\mathrm{L}^2(\mathbb{T},h)} \cdot |\tilde{u}_{m,l}| = \mathcal{O}(\lambda_{m,l}^{-k}), \qquad (m+l) \to \infty,$$
$$||u - P_N u||_{\mathrm{L}^2(\mathcal{T},h)} = \mathcal{O}(N^{-2k}), \qquad N \to \infty.$$

Theorem 2 *Let u be a function in the space $H^{2k}(\mathbb{T},h) \cap C(\mathbb{T}), k \in \mathbb{N}$ where $h(x,y) = x^{\alpha-1}y^{\beta-1}(1-x-y)^{\gamma-\alpha-\beta}$ the weight with $\alpha,\beta,\gamma \in \mathbb{N}$. $P_N u(x,y)$ is the APK expansion of u.*
We have following pointwise estimate of the truncation error:

(a) *for $2\beta - 1 < \gamma - \alpha$ and $1 + \frac{\gamma-\alpha}{2} \leq \frac{3}{4} + \frac{3}{4}\alpha + \frac{\gamma-\alpha-\beta}{2} < k$:*

$$|u(x,y) - P_N u(x,y)| = \mathcal{O}(N^{-2k+\frac{3}{2}\alpha+\gamma-\alpha-\beta+\frac{1}{2}}), \quad \forall (x,y) \in \mathbb{T},$$

(b) *for $2\beta - 1 < \gamma - \alpha$ and $\frac{3}{4} + \frac{3}{4}\alpha + \frac{\gamma-\alpha-\beta}{2} \leq 1 + \frac{\gamma-\alpha}{2} < k$:*

$$|u(x,y) - P_N u(x,y)| = \mathcal{O}(N^{-2k+\gamma-\alpha+\frac{3}{2}}), \quad \forall (x,y) \in \mathbb{T},$$

(c) *for $\gamma - \alpha \leq 2\beta - 1$ and $1 + \frac{\gamma-\alpha}{2} \leq \frac{1}{4} + \frac{3}{4}\alpha + \frac{\beta}{2} < k$:*

$$|u(x,y) - P_N u(x,y)| = \mathcal{O}(N^{-2k+\frac{3}{2}\alpha+\beta-\frac{1}{2}}), \quad \forall (x,y) \in \mathbb{T},$$

(d) *for $\gamma - \alpha \leq 2\beta - 1$ and $\frac{1}{4} + \frac{3}{4}\alpha + \frac{\beta}{2} \leq 1 + \frac{\gamma-\alpha}{2} < k$:*

$$|u(x,y) - P_N u(x,y)| = \mathcal{O}(N^{-2k+\gamma-\alpha+\frac{3}{2}}), \quad \forall (x,y) \in \mathbb{T}.$$

We extend the SD method by the APK polynomials but have to consider another fact in the numerical calculation. In hyperbolic conservation laws discontinuities may arise in the solution even for smooth initial data. In spectral methods this means that the high-frequency coefficients decay slowly and lead to spurious oscillations (Gibbs phenomenon). These oscillations lead to stability issues in the SD method.
To stabilize the SD Methods we add a small viscosity term to the equation, which depends on the differential operator of the APK polynomials. Using a splitting approach to solve the viscosity formulation of the conservation law in every time step shows the equivalence to applying the extended SD Method, but at each time step the numerical solution is filtered by the exponential filter

$$\sigma\left(\frac{l+m}{N}\right) = e^{-\varepsilon_N N^{2p}\Delta t\left(\frac{l+m}{N}\right)^p\left(\frac{l+m+\gamma}{N}\right)^p}. \tag{1}$$

The reader finds the details in [3]. Finally we mention that this filter (1) depends on the variable γ . Hence every APK polynomial family has its specific modal filter (1).
In two well-known test cases, one for the Burgers equation and one for the Euler equation, we demonstrate that the selection of the APK polynomials and their specific filter (1) may have some positive influence on the accuracy and stability of the method. These numerical tests can be seen in [4].
In the last section we motivate why discrete orthogonal polynomials may be a good approach to improve spectral methods. With the discrete scalar product we don't make numerical errors in the calculation of the coefficients. We start to investigate

the Hahn polynomials as an example of classic discrete orthogonal polynomial and show spectral accuracy of the coefficients.

Theorem 3 *Let $-1 < \alpha, \beta \in \mathbb{R}$. $\tilde{Q}_n(x, \alpha, \beta, N)$ are the normalized Hahn polynomials in $[0, N]$. The Hahn expansion of the function u until degree $m \leq N$ is given by*

$$P_m u(x) = \sum_{n=0}^{m} \hat{u}_n \tilde{Q}_n(x).$$

Let $u \in C^\infty([-1, 1+N])$. We get

$$|\hat{u}_n| \leq \frac{1}{n^{2k}} C_{u,k}.$$

for all $k \in \mathbb{N}_0$. $C_{u,k}$ is a constant depending on u and k.

The idea of proof is partial summation and using the difference operator, for details see [3].

References

[1] A. Meister, S. Ortleb, Th. Sonar, M. Wirz *A comparison of the Discontinuous-Galerkin and Spectral-Difference Method on triangulations using PKD polynomials*, Journal of Computational Physics **231** (2012), 7722–7729.

[2] P. Öffner, Th. Sonar , *Spectral convergence for orthogonal polynomials on triangles*, Numerische Mathematik **124(4)** (2013), 701–721.

[3] P. Öffner, *Zweidimensionale klassische und diskrete orthogonale Polynome und ihre Anwendung auf spektrale Methoden zur Lsung von hyperbolischen Erhaltungsgleichungen* , TU Braunschweig (2015).

[4] P. Öffner, Th. Sonar, *Orthogonal polynomials on triangles in a Spectral Difference Method*, submitted (2015).

The Triangular Grid DG Scheme in SBP Framework and the Construction of Special Purpose RK Time Integrators

Sigrun Ortleb

In this talk, we deal with the SBP framework of numerical methods for hyperbolic conservation laws on the one hand and the construction of special time integration methods for certain applications on the other hand.

SBP Schemes. The framework of SBP operators yields several advantages for the numerical approximation of nonlinear hyperbolic equations. First of all, SBP schemes are linearly stable when combined with suitable boundary and interface treatments, second, they lead to conservative discretizations of split form conservation laws and in addition they can be related to quadrature rules satisfying the discrete divergence theorem. In this talk, popular high order schemes for nonlinear hyperbolic conservation laws are revisited in the summation-by-parts(SBP) framework. More precisely, it is shown that both the classical discontinuous Galerkin(DG) method based on Gauss nodes and the energy stable flux reconstruction(FR) schemes by Vincent et al. [7] can be regarded as generalized SBP schemes using the SBP definition of Del Rey Fernández et al. in [2]. In the 1D

case of a scalar conservation law $\frac{\partial}{\partial t}u + \frac{\partial}{\partial x}f(u) = 0$, both schemes can be put in the form

$$\mathbf{u}_t + \mathbf{D}\mathbf{f} = \mathbf{M}^{-1}[(f_h - f^*)\mathbf{\Phi}]_{x_i}^{x_{i+1}}$$

on each cell $[x_i, x_{i+1}]$, with different mass matrices $\mathbf{M}$ for DG or FR scheme. Herein, a set of Lagrange basis functions Φ_k is used and collected in the vector $\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_p)^T$, whereas $\mathbf{u}$ and $\mathbf{f}$ represent the corresponding nodal values of the approximate solution and the flux function, respectively. Furthermore, f_h is a suitable interpolation of f and f^* denotes the values of a numerical flux function. The matrices $\mathbf{M}, \mathbf{D}$ in this formulation then fulfill the requirements of generalized SBP schemes, i.e.

- $\mathbf{M}$ is symmetric positive definite,
- $\mathbf{D}$ approximates $\frac{\partial}{\partial x}$,
- $\mathbf{S} := \mathbf{M}\mathbf{D}$ is almost skew-symmetric, i.e. $\mathbf{S} + \mathbf{S}^T = \mathbf{B}$ with $(\mathbf{x}^\mu)^T \mathbf{B} \mathbf{x}^\nu = (x_{i+1})^{\mu+\nu} - (x_i)^{\mu+\nu}$.

In addition, a generalized SBP formulation for a discontinuous Galerkin method on triangular grids with non-diagonal norm matrix is given based on the definition of generalized SBP schemes on triangular grids in [4]. An SBP formulation can also be devised for the energy stable FR schemes on triangular grids constructed in [1]. Furthermore, as discussed with P. Vincent during the Oberwolfach Workshop, the the extended range of energy stable FR schemes given in [8] can be put in SBP framework as well. It may hence be conjectured that the SBP property is in fact a prerequisite in order to obtain high order energy stable schemes.

Time Integrators for Specific Applications.

Shallow water equations. In the majority of previous work on DG schemes for shallow water flows, explicit time stepping is implemented to deal with wetting and drying. Linear stability then usually requires rather small time steps. Especially for locally refined grids, the stiff system resulting from space discretization makes implicit or partially implicit time stepping absolutely necessary. As implicit schemes require a lot of computational time solving large systems of nonlinear equations, a much larger time step is necessary to beat explicit time stepping in terms of CPU time. However, for stability purposes, the DG scheme applied to the shallow water equations also has to guarantee non-negativity of the water height, accordant with the true physical solution. At wet/dry interfaces, this requirement also leads to restrictive time step constraints in case of implicit time integration. In this talk, a novel approach to positivity preservation is hence based on the so-called Patankar trick and guarantees non-negativity of the water height for any time step size while still preserving conservativity. In the DG context, consistency of the discretization as well as a truncation error of third order away from the wet-dry transition can be proven, see [5].

Mechanical fluid structure interaction. The simulation of fluid structure interaction(FSI) often necessitates the choice of different time step sizes for the time integration of the fluid equations on the one hand and structure equations on the other hand. For mechanical fluid structure interaction, Piperno devised a special

volume-discontinuous staggered scheme to enhance the stability of subcycling procedures with large implicit structural time step and several smaller fluid steps, see [6]. This scheme, as well as the more classical volume-continuous scheme can be put in a multirate Runge-Kutta framework. Thereby, different splittings of the right hand sides as well as different ordering of fluid and structure solves reflect the differences of these schemes. In this talk, we apply the third order multirate method of Günther et al. [3] to the classical piston test case of mechanical FSI. Here, a slight increase in combined fluid and structural energy occurs due to the non-symplectic implicit part of the time integrator. While for monolithic FSI schemes, energy stability is possible, it is an open question if higher order energy stable partitioned schemes can be found within the class of multirate partitioned Runge-Kutta methods. One approach could be the construction of a predictor-corrector scheme with a corrector specifically designed for stability. In addition, adapted boundary procedures at the fluid structure interface could be necessary for this purpose.

References

[1] P. Castonguay, P. E. Vincent, A. Jameson, *A new class of high-order energy stable flux reconstruction schemes for triangular elements*, J. Sci. Comput. **51** (2012), 224–256.

[2] D. C. Del Rey Fernández, P. D. Boom, D. W. Zingg, *A generalized framework for nodal first derivative summation-by-parts operators*, Journal of Computational Physics **266** (2014), 214–239.

[3] M. Günther, A. Kvaernø, P. Rentrop, *Multirate partitioned Runge-Kutta methods*, BIT Numerical Mathematics **41** (2001), 504–514.

[4] J. E. Hicken, D. C. Del Rey Fernández, D. W. Zingg, *Multidimensional summation-by-parts operators: general theory and application to simplex elements*, arXiv:1505.03125 [math.NA] (2015)

[5] A. Meister, S. Ortleb, *On unconditionally positive implicit time integration for the DG scheme applied to shallow water flows*, Int. J. Numer. Meth. Fluids **76** (2014), 69–94.

[6] S. Piperno, *Explicit/implicit fluid/structure staggered procedures with a structural predictor and fluid subcycling for 2D inviscid aeroelastic simulations*, Int. J. Numer. Meth. Fluids **25** (1997), 1207–1226.

[7] P. E. Vincent, P. Castonguay, A. Jameson, *A new class of high-order energy stable flux reconstruction schemes*, J. Sci. Comput. **47** (2011), 50–72.

[8] P. E. Vincent, A. M. Farrington, F. D. Witherden FD, A. Jameson, *An extended range of stable-symmetric-conservative flux reconstruction correction functions*, Comput. Methods in Appl. Mech. Eng. **296** (2015), 248–272.

High-Order Methods for Turbulent Flow Simulations on Deforming Domains

Per-Olof Persson
(joint work with Luming Wang)

Introduction. A challenging problem in the numerical solution of PDEs is how to obtain high-order accuracy in the presence of moving domains with large deformations. The popular Arbitrary Lagrangian-Eulerian (ALE) method [1] can be viewed as a mapping-based approach which, together with appropriate treatment of the Geometric Conservation Law (GCL), allows for arbitrarily high orders of accuracy in both space and time. However, the method requires smooth mappings between the initial (or reference) frame and the actual physical configuration. These can be generated for domains undergoing moderate deformations, such as pitching and heaving airfoils, or structures with small deformations. But many other applications require topological changes to maintain a well-shaped mesh/transformation, e.g. rotating machinery or configurations involving multiple moving objects.

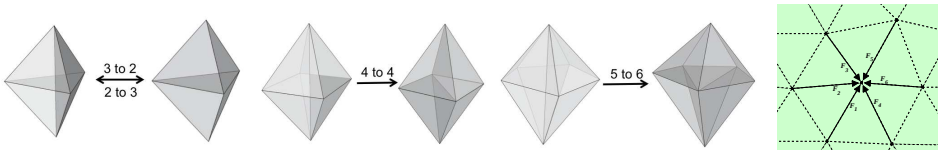
In our work [2, 3], we proposed a framework for solving systems of conservation laws on moving meshes to an arbitrary degree of accuracy in both space and time. The method is based on the assumption that the unstructured moving meshes can be produced by a sequence of entirely local operations [4]. This produces high-quality meshes throughout the simulation, and provides a simple description of the mesh changes between each timestep. Using this information we can construct efficient numerical schemes based on high-order discontinuous Galerkin formulations, and we consider both space-time and ALE/projection-based methods in 2D and 3D.

The DistMesh mesh generator. The so-called DistMesh method [4] is based on iterative updates of the positions of the N interior mesh nodes $\boldsymbol{p}^{(n)}$, $n = 1, \ldots, N$, driven by spring-based forces $\boldsymbol{F}_i$ for each edge i adjacent to node n, of the form:

$$\boldsymbol{p}^{(n+1)} = \boldsymbol{p}^{(n)} + \delta \sum_i \boldsymbol{F_i} \tag{1}$$

$$|\boldsymbol{F}_i(l)| = \begin{cases} k(l - l_0) & \text{if } l \geq l_0, \\ 0 & \text{if } l < l_0. \end{cases} \tag{2}$$

Here, δ is a pseudo-timestep, k is the spring constant, l_0 is the desired edge length, and l is the actual edge length. In addition, the mesh generator employs local element topology modifications, such as the simple edge flipping in 2D or more complex operations in 3D. Here, we assume that each element can be flipped once during each sweep, but multiple rounds during each time step.



Space-Time mesh generation. Noting that our moving meshes are defined by entirely local operations (node movements and element flips), we are able to define provably good space-time meshes. These are based on fully unstructured simplex elements, that is, tetrahedra in 2D+time and 4D simplices in 3D+time. We restrict our method to slab-based space-time meshing, that is, all elements have uniform size in the temporal direction. However, extending this to adaptivity in time using local refinement should be straight-forward, even for the 4D elements.

The method for choosing appropriate space-time elements is mostly combinatorial, and uses node data only to allow for choosing the highest quality configurations. For details, see [3].

Space-Time discontinuous Galerkin discretization. Using our space-time meshes, we define a high-order accurate discontinuous Galerkin (DG) method as follows. The scheme is based on a fully consistent discretization in both space and time, and allows for arbitrary mesh deformations and topology changes.

Consider a system of conservation laws of the form

$$\frac{\partial \boldsymbol{u}}{\partial t} + \nabla_{\boldsymbol{X}} \cdot \boldsymbol{F}^{\text{inv}}(\boldsymbol{u}) = \nabla_{\boldsymbol{X}} \cdot \boldsymbol{F}^{\text{vis}}(\boldsymbol{u}, \nabla_{\boldsymbol{X}} \boldsymbol{u}) \tag{3}$$

with appropriate initial and boundary conditions. We write this in a space-time form and use the standard technique of splitting into a first-order system:

$$\nabla_{\boldsymbol{XT}} \cdot \tilde{\boldsymbol{F}}^{\text{inv}}(\boldsymbol{u}) = \nabla_{\boldsymbol{X}} \cdot \boldsymbol{F}^{\text{vis}}(\boldsymbol{u}, \boldsymbol{q}), \tag{4}$$

$$\nabla_{\boldsymbol{X}} \boldsymbol{u} = \boldsymbol{q}. \tag{5}$$

Next, we define the broken DG spaces $\mathcal{V}_T^h$ and Σ_T^h associated with a triangulation $\mathcal{T}_{[0,T]}^h = \{K\}$ of the space-time domain $\Omega[0,T]$ as:

$$\mathcal{V}_T^h = \{\boldsymbol{v} \in [L^2(\Omega[0,T])]^5 \mid \boldsymbol{v}|_K \in [\mathcal{P}_p(K)]^5 \ \ \forall K \in \mathcal{T}_{[0,T]}^h\}, \tag{6}$$

$$\Sigma_T^h = \{\boldsymbol{\sigma} \in [L^2(\Omega[0,T])]^{5\times 3} \mid \boldsymbol{\sigma}|_K \in [\mathcal{P}_p(K)]^{5\times 3} \ \ \forall K \in \mathcal{T}_{[0.T]}^h\}, \tag{7}$$

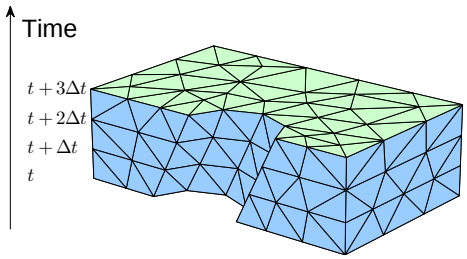
Now, discretize the first-order system using a standard DG formulation on the space-time domain $\Omega[t_1, t_2]$:

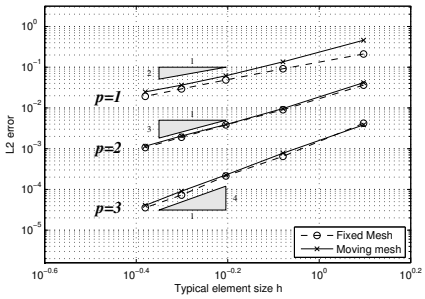
$$\begin{aligned} &-\int_K \tilde{\boldsymbol{F}}^{\text{inv}}(\boldsymbol{u}^h) : \nabla_{\boldsymbol{XT}} \boldsymbol{v}^h \, dx + \oint_{\partial K} (\widehat{\tilde{\boldsymbol{F}}^{\text{inv}} \cdot \boldsymbol{n}}) \cdot \boldsymbol{v}^h \, ds \\ &\qquad = -\int_K \boldsymbol{F}^{\text{vis}}(\boldsymbol{u}^h, \boldsymbol{q}^h) : \nabla_{\boldsymbol{X}} \boldsymbol{v}^h \, dx + \oint_{\partial K} (\widehat{\boldsymbol{F}^{\text{vis}} \cdot \boldsymbol{n}_s}) \cdot \boldsymbol{v}^h \, ds, \qquad \forall \boldsymbol{v}^h \in \mathcal{V}_T^h \end{aligned} \tag{8}$$

$$\int_K \boldsymbol{q}^h : \boldsymbol{\sigma}^h \, dx = -\int_K \boldsymbol{u}^h \cdot (\nabla_{\boldsymbol{X}} \cdot \boldsymbol{\sigma}^h) \, dx + \oint_{\partial K} (\widehat{\boldsymbol{u}^h} \otimes \boldsymbol{n}_s) : \boldsymbol{\sigma}^h \, ds, \qquad \forall \boldsymbol{\sigma}^h \in \Sigma_T^h. \tag{9}$$

For more details, including the details on how to assemble and solve the resulting nonlinear systems of equations, see [2,3].

Results. To demonstrate the accuracy of our proposed methods, we consider the standard test problem of an isentropic Euler vortex. We use polynomial degrees p of 1, 2, and 3, and obtain numerically the expected optimal orders of accuracy $p+1$. See [2,3] for demonstrations on more complex problems, including flow problems in both 2D and 3D which show the ability of our method to deal with complex domain motions.





References

[1] P.-O. Persson, J. Bonet, and J. Peraire, *Discontinuous Galerkin solution of the Navier-Stokes equations on deformable domains*, Comput. Methods Appl. Mech. Engrg., 198(17-20):1585–1595, 2009.

[2] L. Wang, P.-O. Persson, *A high-order discontinuous Galerkin method with unstructured space-time meshes for two-dimensional compressible flows on domains with large deformations*, Comput. & Fluids, Vol. 118, pp. 53-68, September 2015.

[3] L. Wang, *Discontinuous Galerkin Methods on Moving Domains with Large Deformations*, Ph.D. thesis, University of California, Berkeley, May 2015.

[4] P.-O. Persson and G. Strang, *A simple mesh generator in Matlab*, SIAM Rev., 46(2):329–345, 2004.

A robust high-order Lagrange-projection like scheme with large time steps for the isentropic Euler equations

Florent Renac

(joint work with Frederic Coquel, Christophe Chalons)

Introduction. We present an extension to high-order of a first-order Lagrange-projection like method for the approximation of the Euler equations introduced in Coquel *et al.* [2]. The method is based on a decomposition between acoustic and transport operators associated to an implicit-explicit time integration, thus relaxing the constraint of acoustic waves on the time step. Besides, the work in [2] circumvents the difficulties in the treatment of nonlinearities associated to the equation of state by using relaxation approximation [1]. The latter method approximates the nonlinear system with a linear or a quasi-linear enlarged system with stiff relaxation source terms.

The method in [2] is based on a first-order finite volume approximation. We propose here to use a discontinuous Galerkin (DG) method for the space discretization. These are particular aspects that make the DG method well suited. First, it is possible to use the numerical fluxes derived in [2] and therefore the present method may be viewed as a natural extension to high-order of the first-order method. Then, the effect of the numerical flux on the quality of the approximation is known to decrease as the polynomial degree p in the DG method increases [3, 4]. This avoids the use of local numerical parameters tuned at each interface of the mesh in order to lower the numerical diffusion induced by the first-order approximation [2]. This aspect is essential in our analysis to restore the PDE properties at the discrete level and to derive *a priori* conditions to preserve invariant domains and satisfy an entropy inequality by the present Lagrange-projection DG (LPDG) scheme.

Isentropic Euler equations. The discussion here focuses on the Euler equations for an isentropic gas in one space dimension. Let $\Omega = \mathbb{R}$ be the space domain and consider the following problem

$$\begin{cases} \partial_t \mathbf{u} + \partial_x \mathbf{f}(\mathbf{u}) = 0, & \text{in } \Omega \times (0,\infty), \\ \mathbf{u}(\cdot,0) = \mathbf{u}_0(\cdot), & \text{in } \Omega, \end{cases} \tag{1}$$

where $\mathbf{u} = (\rho, \rho u)^\top$ and $\mathbf{f}(\mathbf{u}) = (\rho u, \rho u^2 + \mathrm{p})^\top$represent the conservative variables and the nonlinear fluxes, with ρ the density and u the velocity. Equations (1) are supplemented with an equation of state for the pressure of the form $\mathrm{p} = \mathrm{p}(\tau)$, with $\tau = 1/\rho$ the specific volume. Assuming that $\mathrm{p}'(\tau) < 0$ and $\mathrm{p}''(\tau) > 0$ for all $\tau > 0$, the system in (1) is strictly hyperbolic with eigenvalues associated to nonlinear fields. Note that the total energy $\rho E(\mathbf{u}) = \rho e(1/\rho) + (\rho u)^2/2\rho$, with $e'(\tau) = -\mathrm{p}(\tau)$ the specific internal energy, is a strictly convex function. Physically relevant solutions to (1) must hence satisfy an inequality of the form

$$\partial_t \rho E + \partial_x(\rho E u + \mathrm{p}u) \le 0. \tag{2}$$

Space-time discretization. The DG method consists in defining a discrete weak formulation of problem (1). The domain is discretized with a uniform grid $\Omega_h = \cup_{j\in\mathbb{Z}} \kappa_j$ with cells $\kappa_j = [x_{j-1/2}, x_{j+1/2}]$, $x_{j+1/2} = (j+\frac{1}{2})h$ and $h > 0$ the space step. The approximate solution to (1) is sought under the form

$$\mathbf{u}_h(x,t) = \sum_{l=0}^{p} \phi_j^l(x)\mathbf{U}_j^l(t), \quad \forall x \in \kappa_j,\ \kappa_j \in \Omega_h,\ t \ge 0, \tag{3}$$

where $\mathbf{U}_j^l = (\rho_j^l, \rho U_j^l)^\top$ constitute the degrees of freedom in the element κ_j. The function space is spanned by using Lagrange interpolation polynomials $\phi_j^{0\le l\le p}$ in each element κ_j of the mesh.

Using a one-step first-order implicit-explicit time discretization for the acoustic and transport operators in (1), we introduce the following conservative approximation consistent in time and space with the Euler equations:

$$\mathbf{U}_j^{k,n+1} = \mathbf{U}_j^{k,n} - \lambda_k \Big[-\big\langle \mathbf{f}(\mathbf{u}_h^{n+1^-}), d_x\phi_j^k \big\rangle_j^p + \delta_{k,p}\mathbf{h}_{j+\frac{1}{2}}^{n+1^-} - \delta_{k,0}\mathbf{h}_{j-\frac{1}{2}}^{n+1^-} \Big], \tag{4}$$

with $\lambda_k = 2\lambda/\omega_k$, $\lambda = \Delta t/h$, $\mathbf{U}_j^{k,n} = \mathbf{U}_j^k(n\Delta t)$, and $\Delta t > 0$ the time step. The numerical fluxes are defined by

$$\mathbf{h}_{j+\frac{1}{2}} = \begin{pmatrix} u^\star_{j+\frac{1}{2}}\hat{\rho}_{j+\frac{1}{2}} \\ u^\star_{j+\frac{1}{2}}\hat{\rho u}_{j+\frac{1}{2}} + \Pi^\star_{j+\frac{1}{2}} \end{pmatrix}, \tag{5}$$

where a star symbol indicates the solution at interfaces of the Riemann problem associated to the relaxation approximation of the acoustic operator, while a hat symbol indicates an upwind flux. The time $t^{(n+1^-)}$ corresponds to a fictious time where the splitting between acoustic and transport operators is applied within the interval $(t^{(n)}, t^{(n+1)})$. The volume integrals in the discrete variational formulation are approximated by using a collocated Gaus-Lobatto quadrature rule of the form

$$\int_{\kappa_j} f(x)g(x)dx \simeq \langle f, g\rangle_j^p := \frac{h}{2}\sum_{l=0}^{p} \omega_l f(x_l)g(x_l),$$

with $\omega_{0\le l\le p} > 0$ and $x_{0\le l\le p}$ the weights and nodes of the quadrature.

The main results of the present work are as follows. Assume that $\rho_{j\in\mathbb{Z}}^{0\le k\le p,n} > 0$, then under some CFL condition on the time step, evaluated from the transport terms only, and one condition on the numerical parameter of the scheme, the LPDG scheme guaranties positivity of the mean value in cell of density of the numerical solution:

$$\overline{\rho}_j^{n+1} > 0, \tag{6}$$

and satisfies the following discrete entropy inequality

$$\rho E(\overline{\mathbf{u}}_j^{n+1}) - \overline{\rho E}_j^n + \lambda\Big(u^\star_{j+\frac{1}{2}}(\widehat{\rho E}_{j+\frac{1}{2}}^{n+1^-} + \Pi^\star_{j+\frac{1}{2}}) - u^\star_{j-\frac{1}{2}}(\widehat{\rho E}_{j-\frac{1}{2}}^{n+1^-} + \Pi^\star_{j-\frac{1}{2}})\Big) \le 0. \tag{7}$$

The above results allow to design linear limiting procedures to restore these properties at nodal values within elements [6]. Numerical experiments (see figure), obtained by using explicit SSP Runge-Kutta integration in time, support the conclusions of the analysis and highlight stability and robustness of the present LPDG scheme, though it allows the use of large time steps [5].

References

[1] F. Coquel, E. Godlewski, A. In, B. Perthame and P. Rascle, *Some new Godunov and relaxation methods for two-phase flow problems. Godunov methods*, Kluwer/Plenum, New York, 2001, 179–188.

[2] F. Coquel, Q. Long-Nguyen, M. Postel and Q. H. Tran, *Entropy-satisfying relaxation method with large time-steps for Euler IBVPs*, Math. Comput., **79** (2010), 1493–1533.

[3] J. Qiu, B.C. Khoo and C.-W. Shu. *A numerical study for the performance of the Runge-Kutta discontinous Galerkin method based on different numerical fluxes*, J. Comput. Phys., **26** (2006), 540–565.

[4] F. Renac, *Stationary discrete shock profiles for scalar conservation laws with a discontinuous Galerkin method*, SIAM J. Numer. Anal., **53** (2015), 1690–1715.

[5] F. Renac, *A robust high-order Lagrange-projection like scheme with large time steps for the isentropic Euler equations*, in preparation.

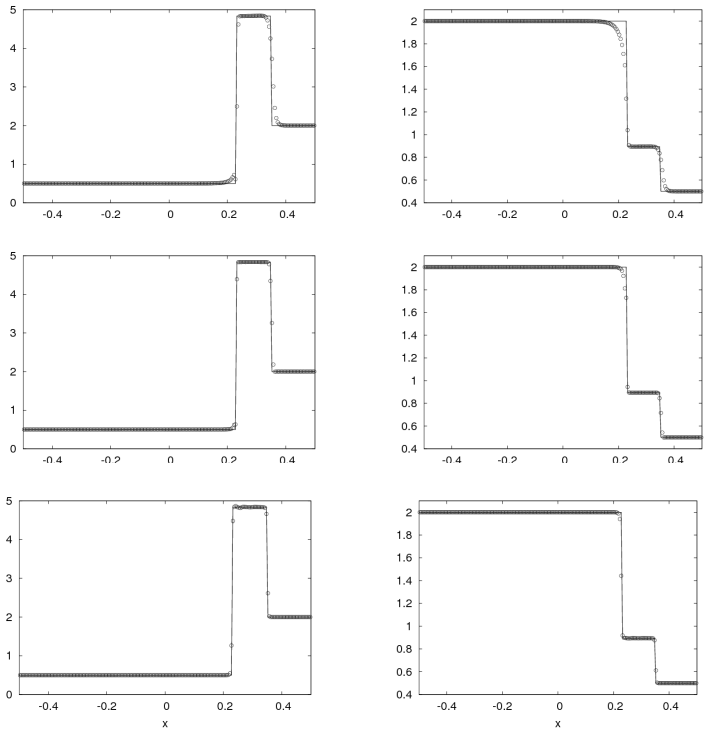


FIGURE 1. Riemann problem with two shock waves: numerical solution for density (left) and velocity (right) at time $t = 0.3$ for polynomial degrees $1 \leq p \leq 3$ from top to bottom, $h = \frac{1}{200}$.

[6] X. Zhang and C.-W. Shu, *On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes*, J. Comput. Phys., **229** (2010), 8918–8934.

Coupling Parallel Lattice Boltzmann Methods and Granular Dynamics for Complex Flows

ULRICH RÜDE

(joint work with Regina Ammer, Dominik Bartuschat, Martin Bauer, Simon Bogner, Ehsan Fattahi, Christian Godenschwager, Harald Köstler, Kristina Pickl, Tobias Preclik, Christoph Rettinger, Florian Schornbaum)

We present new parallel algorithms for coupling the dynamics of multi-body systems and hydrodynamics with the goal of developing computational models for two phase flows in 3D that have a fluid carrier phase and a suspended solid phase. Differerent from other work, we study here a fully resolved approach where each

particle is represented as an individual geometric entity. Moving particles are represented with a Lagrangian approach and are treated as rigid objects that can interact through frictional collisions. The fluid is simulated by the Lattice Boltzmann method (LBM) and interacts with the particulate phase through specific fluid-structure interaction techniques. Such fully resolved simulations can become extremely compute intensive and thus all research relies critically on suitable parallel algorithms and their efficient implementation on advanced supercomputer architectures.

For the massively parallel granular dynamics simulations on distributed memory architectures [31, 25] a domain partitioning approach is proposed [39, 40]. Special algorithms for massively parallel granular dynamics simulations [3, 22, 23] with fully resolved rigid particles on distributed memory architectures are developed, where collisions are geometrically detected [8, 12] and the physical response is modelled with hard contacts [30] using the paradigm of measure differential inclusions.

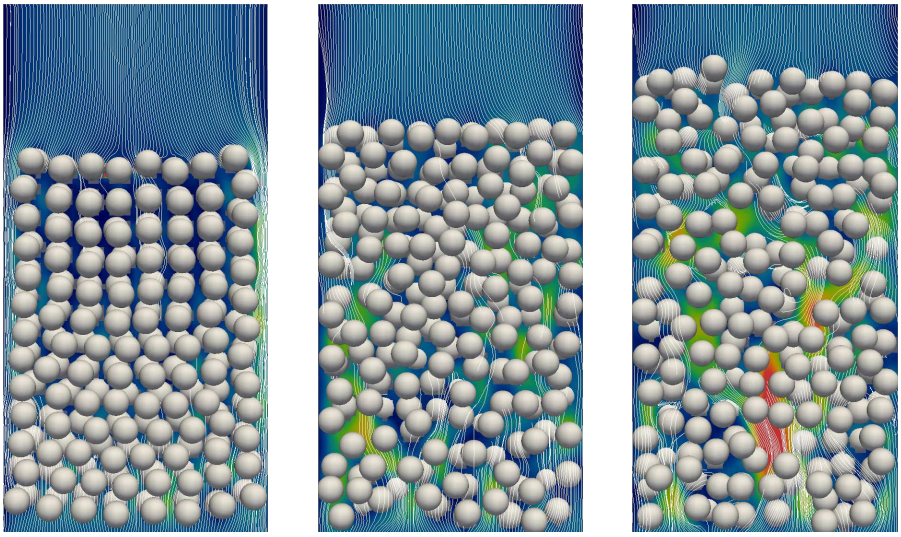
The global frictional multi-contact problem leads to a complementarity problem [3, 35] that is solved using a parallel non-linear block Gauss-Seidel method [37]. The algorithms employ a sophisticated communication protocol between processors that delegates algorithmic tasks such as contact treatment and position integration uniquely and robustly to the processors [36]. Communication overhead is minimized through aggressive message aggregation, leading to excellent strong and weak scaling [23]. The robustness and scalability is assessed on peta-scale supercomputers with up to 458 752 processor cores. Such granular media simulations can reach an unprecedented resolution of up to ten billion (10^{10}) non-spherical particles and contacts [37].

For flow simulation, the LBM has gained popularity as an alternative to classical Navier-Stokes solvers for computational fluid dynamics (CFD) [20, 7, 28, 13, 1]. With the LBM, the simulation domain is discretized with a uniform Cartesian grid. If the resolution of a three-dimensional simulation must be increased in space and time, then the total number of cells and the computational cost increase quickly. Many implementations of the LBM [19, 32, 21, 27, 17] are therefore designed for parallel computers. Going beyond scalability alone, a carefully crafted architecture-aware implementation of the LBM, as realized in the waLBerla framework [14, 10], can achieve excellent absolute performance and thus reduce the time to solution to reach a given computational objective. The waLBerla framework can discretize a complex flow geometry with in excess of a trillion (10^{12}) lattice cells on current petascale supercomputers [18].

Coupled simulations with particles embedded in the fluid become even more compute intensive. To limit the computational effort, many previous simulations were performed either with only few particles or in only two dimensions, or with simplified models for the fluid-particle interaction. Our new simulation model and its implementation, as presented here, extends these approaches and allows fluid simulations in 3D with millions of interacting particles [15, 16] suspended in the flow. Furthermore, our simulation framework can model particles of arbitrary

shape [5], includes the so-called lubrication correction [26, 9] and can be extended to include electrostatic effects [4]. Simulations of such complexity inevitably require the effective use of supercomputers. A thorough analysis demonstrates that our techniques achieve excellent computational performance [18, 4, 11].

An increasing number of real-life applications illustrate the generality and the power of the approach. These include the development of closure relations for macroscopic multiphase models [6], the study of self-propelled swimmers [33, 34], processing metal foams [24], patient specific blood flow [18], food technology [2], and additive manufacturing [29].



Simulated model fluidized bed with thin rectangular cross section.
208 spherical particles driven by inflow from below. LBM resolution $100 \times 30 \times 200$
computed with 25 000 time steps using 64 MPI threads.

References

[1] Aidun C, Clausen J (2010) Lattice-Boltzmann method for complex flows, Annual Review of Fluid Mechanics, 42, pp. 439–472.

[2] Anderl, D., Bauer, M., Rauh, C., Rüde, U., Delgado, A. (2014). Numerical simulation of adsorption and bubble interaction in protein foams using a lattice Boltzmann method. Food & function, 5(4), pp. 755–763.

[3] Anitescu M, Potra F (1997) Formulating dynamic multi-rigid-body contact problems with friction as solvable linear complementarity problems. Nonlinear Dynamics 14(3), pp. 231–247.

[4] Bartuschat, D and Rüde U (2015) Parallel multiphysics simulations of charged particles in microfluidic flows, Journal of Computational Science 8, pp. 1–19.

[5] Bartuschat D, Fischermeier, E., Gustavsson, K., Rüde, U. (2015) Two Computational Models for Simulating the Tumbling Motion of Elongated Particles in Fluids. arXiv preprint arXiv:1503.06869.

[6] Bogner S, Mohanty S, Rüde U (2015) Drag correlation for dilute and moderately dense fluid-particle systems using the lattice boltzmann method. International Journal of Multiphase Flow 68(0), pp. 71–79.
[7] Shiyi Chen and Doolen, G (1998) Lattice Boltzmann method for fluid flows, Annual Review of Fluid Mechanics 30, pp. 329–364.
[8] Cohen J, Lin M, Manocha D, Ponamgi M (1995) I-COLLIDE: An interactive and exact collision detection system for large-scale environments. In: Proceedings of the 1995 symposium on Interactive 3D graphics, ACM, pp. 189–ff.
[9] E. J. Ding, C. K. Aidun (2003) Extension of the lattice-Boltzmann method for direct simulation of suspended particles near contact, J. Stat. Phys. 112 (3), pp. 685–708.
[10] Feichtinger, C., Götz, J., Donath, S., Iglberger, K., Rüde, U. (2009). WaLBerla: Exploiting massively parallel systems for lattice Boltzmann simulations. In Parallel Computing, pp. 241–260. Springer London.
[11] Feichtinger, C., Habich, J., Köstler, H., Rüde, U., Aoki, T. (2015). Performance modeling and analysis of heterogeneous lattice Boltzmann simulations on CPU–GPU clusters. Parallel Computing, 46, pp. 1–13.
[12] Gilbert E, Johnson D, Keerthi S (1988) A fast procedure for computing the distance between complex objects in three-dimensional space. IEEE Journal of Robotics and Automation 4(2), pp. 193–203
[13] I. Ginzburg, F. Verhaeghe, D. d'Humières (2008) Study of simple hydrodynamic solutions with the two-relaxation-times lattice Boltzmann scheme, Commun. Comput. Phys. 3 (3), pp. 519–581.
[14] Götz, J., Feichtinger, C., Iglberger, K., Donath, S., Rüde, U. (2008). Large scale simulation of fluid structure interaction using Lattice Boltzmann methods and the physics engine. ANZIAM Journal, 50, pp. 166-188.
[15] Götz J., Iglberger K., Feichtinger C., Donath S., Rüde U (2010) Coupling multibody dynamics and computational fluid dynamics on 8192 processor cores, Parallel Comput. 36 (2) pp. 142–141.
[16] Götz J., Iglberger K., Stürmer M., Rüde U (2010) Direct Numerical Simulation of Particulate Flows on 294912 Processor Cores, in: Proc. 2010 ACM/IEEE Proc. Int. Conf. for High Performance Computing, Networking, Storage and Analysis, SC '10, IEEE, 2010, pp. 1–11.
[17] Groen D, Hetherington J, Carver H, Nash R, Bernabeu M, Coveney P (2013) Analysing and modelling the performance of the HemeLB lattice-Boltzmann simulation environment, Journal of Computational Science, 4 pp. 412–422.
[18] Godenschwager C, Schornbaum F, Bauer M, Köstler H, Rüde U (2013), A framework for hybrid parallel flow simulations with a trillion cells in complex geometries, in: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, SC '13, IEEE, pp. 35:1–35:12.
[19] Hasert M, Masilamani K, Zimny S, Klimach H, Qi J, Bernsdorf J, Roller S (2014) Complex fluid simulations with the parallel tree-based lattice boltzmann solver musubi, Journal of Computational Science, pp. 784 – 794.
[20] X. He, L.-S. Luo (1997) Lattice Boltzmann Model for the Incompressible Navier-Stokes Equation, J. Stat. Phys. 88 (3), pp. 927–944.
[21] Heuveline V, Latt J (2007) The OpenLB Project: An Open Source and Object Oriented Implementation of Lattice Boltzmann Methods, International Journal of Modern Physics C, 18, pp. 627–634.
[22] Iglberger K, Rüde U (2009) Massively parallel rigid body dynamics simulations. Computer Science-Research and Development 23(3-4), pp. 159–167
[23] Iglberger K, Rüde U (2011) Large-scale rigid body simulations. Multibody System Dynamics 25(1), pp. 81–95
[24] Körner, C., Pohl, T., Rüde, U., Thürey, N., Hofmann, T. (2005). FreeWIHR: Lattice Boltzmann methods with free surfaces and their application in material technology. In High

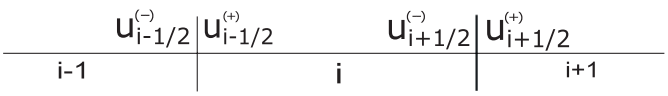
Performance Computing in Science and Engineering, Garching 2004, Springer Berlin Heidelberg, pp. 225–236.
[25] Koziara T, Bićanić N (2011) A distributed memory parallel multibody contact dynamics code. International Journal for Numerical Methods in Engineering 87(1–5), pp. 437–456
[26] A. J. C. Ladd, R. Verberg (2001 Lattice-Boltzmann Simulations of Particle-Fluid Suspensions, J. Stat. Phys. 104 (5), pp. 1191–1251.
[27] LB3D. http://ccs.chem.ucl.ac.uk/lb3d. Accessed: 2015-10-16.
[28] L.-S. Luo (1998) Unified Theory of Lattice Boltzmann Models for Nonideal Gases, Phys. Rev. Lett. 81 (8), pp. 1618–1621.
[29] M. Markl, R. Ammer, U. Ljungblad, U. Rüde, C. Körner (2013) Electron Beam Absorption Algorithms for Electron Beam Melting Processes Simulated by a Three–Dimensional Thermal Free Surface Lattice Boltzmann Method in a Distributed and Parallel Environment, Procedia Computer Science 18, pp. 2127–2136.
[30] Moreau J, Panagiotopoulos P (1988) Nonsmooth Mechanics and Applications, vol 302. Springer.
[31] Negrut D, Tasora A, Mazhar H, Heyn T, Hahn P (2012) Leveraging parallel computing in multibody dynamics. Multibody System Dynamics 27(1), pp. 95–117.
[32] Palabos. http://www.palabos.org/. Accessed: 2015-10-16
[33] Pickl K, J. Götz J, Iglberger K, Pande J, Mecke K, Smith A-S, Rüde U (2012) All good things come in threes–Three beads learn to swim with lattice Boltzmann and a rigid body solver , J. Comput. Sci. 3 (5) pp. 374–387.
[34] Pickl K, Hofmann M, Preclik T , Köstler H, Smith A.-S, Rüde U (2014) Parallel Simulations of Self-propelled Microorganisms, in Parallel Computing: Accelerating Computational Science and Engineering (CSE), M. Bader, A. Bode, H.-J. Bungartz, M. Gerndt, G. R. Joubert, and F. Peters, eds., vol. 25 of Advances in Parallel Computing, IOS Press, pp. 395–404.
[35] Popa C, Preclik T, Rüde U (2014) Regularized solution of LCP problems with application to rigid body dynamics. Numerical Algorithms, pp. 1–12.
[36] Preclik, T. (2014) Models and Algorithms for Ultrascale Simulations of Non-smooth Granular Dynamics. Dissertation, FAU Erlangen-Nürnberg.
[37] Preclik, T., Rüde, U (2015) Ultrascale simulations of non-smooth granular dynamics. Computational Particle Mechanics, pp. 1–24.
[38] Tasora A, Anitescu M (2011) A matrix-free cone complementarity approach for solving large-scale, nonsmooth, rigid body dynamics. Computer Methods in Applied Mechanics and Engineering 200(5), pp. 439–453.
[39] Visseq V, Martin A, Dureisseix D, Dubois F, Alart P (2012) Distributed nonsmooth contact domain decomposition (NSCDD): Algorithmic structure and scalability. In: Proceedings of the International Conference on Domain Decomposition Methods.
[40] Visseq V, Alart P, Dureisseix D (2013) High performance computing of discrete nonsmooth contact dynamics with domain decomposition. International Journal for Numerical Methods in Engineering 96(9), pp. 584–598.

Third-Order Limiter Functions in Finite Volume Methods

Birte Schmidtmann
(joint work with Manuel Torrilhon, Rémi Abgrall, Benjamin Seibold)

We are interested in the numerical solution of hyperbolic conservation laws on the most local compact stencil consisting of only nearest neighbors. In the Finite Volume setting, the main challenge is the reconstruction of the interface values $u_{i\pm1/2}$ which are crucial for the definition of the numerical flux functions and thus, for the order of accuracy of the resulting scheme.

Often, the functions of interest contain smooth parts as well as large gradients, discontinuities, or shocks. Treating such functions with high-order schemes may lead to undesired effects such as oscillations. However, what is required is a solution with sharp discontinuities while maintaining high-order accuracy in smooth regions. One possible way of achieving this is the use of limiter functions, which switch the reconstruction to lower order when necessary. In this work, we restrict ourselves to reconstructions based on three cells, which is the most compact stencil possible for higher-order reconstructions. The three input values are the cell averages of cell i, $\bar{u}_i(t^n) = \frac{1}{\Delta x}\int_{x_i-\Delta x/2}^{x_i+\Delta x/2} u(x,t^n)dx$ and its direct neighbors $\bar{u}_{i\pm1}$.



The left and right limits of the interface values $u_{i\pm1/2}$ are computed by

$$u^{(-)}_{i+1/2} = \bar{u}_i + \frac{1}{2}\,\phi(\theta_i)\delta_{i+\frac{1}{2}}, \tag{1a}$$

$$u^{(+)}_{i-1/2} = \bar{u}_i - \frac{1}{2}\,\phi(\theta_i^{-1})\delta_{i-\frac{1}{2}}, \tag{1b}$$

where $\delta_{i+\frac{1}{2}} = \bar{u}_{i+1} - \bar{u}_i$, $\delta_{i-\frac{1}{2}} = \bar{u}_i - \bar{u}_{i-1}$, $\theta_i = \delta_{i-\frac{1}{2}}/\delta_{i+\frac{1}{2}}$ and ϕ is a univariate limiter function depending on the local smoothness measure θ_i. For example, the Minmod limiter reads $\phi(\theta_i) = \max(0, \min(1, \theta_i))$ and the full-third-order reconstruction is given by

$$\phi_3(\theta_i) = (2+\theta_i)/3. \tag{2}$$

In this work, we develop a new limiter function based on the function introduced by Čada and Torrilhon [2]. This function was inspired by the local double logarithmic reconstruction function of Artebrant and Schroll [1].
The new limiter function, denoted by $\phi_{3\mathrm{L}}$, is third-order accurate in smooth parts of the solution. It reads

$$\phi_{3\mathrm{L}}(\theta_i) = \max\left(0, \min\left(\frac{2+\theta_i}{3}, \max\left(-\theta_i, \min\left(2\,\theta_i, \frac{2+\theta_i}{3}, 1.5\right)\right)\right)\right). \tag{3}$$

It is easy to see that it recovers the unlimited third-order reconstruction $\phi_3(\theta_i)$ for certain values of θ_i, namely for $\theta_i \approx 1$ and $\theta_i \approx -1$. Note that the function is non-zero in part of the $\theta_i < 0$ region, which means that it is not total variation diminishing (TVD). This non-zero part is kept intentionally, to avoid the so-called extrema clipping. Consider an extrema, this means $\theta \approx -1$, which corresponds to t_1, the green curve in Fig. 1. TVD limiters on the one hand, return 0 in this case, which means that the reconstructions Eq. (1) are first order accurate and the extremum is not captured correctly. The limiter $\phi_{3\mathrm{L}}(\theta_i)$, on the other hand, returns $(2+\theta_i)/3$, which means that it resolves extrema with third-order accuracy. The only issue concerning extrema-clipping, which $\phi_{3\mathrm{L}}(\theta_i)$ does not fully resolve, is that $\phi_{3\mathrm{L}}(\theta_i) = const.$ for $\theta_i = 0$ or $\pm\infty$, cf. t_2, the blue curve in Fig. 1. In

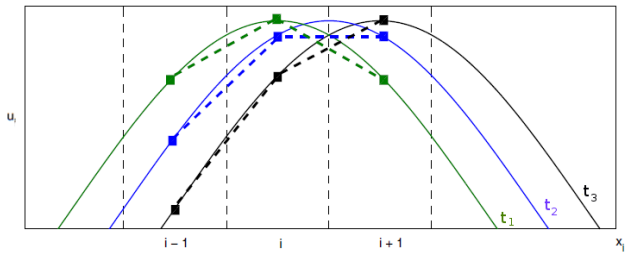


FIGURE 1. Two phenomena of extrema clipping of a smooth profile.

this case, the reconstruction is first order accurate. These values of θ_i occur when one of the slopes $\delta_{i\pm1/2}$ is zero and the other one is non-zero. This is the case for discontinuities or large gradients but can also appear when extrema are not well-resolved. The only difference is the value of the non-zero slope. It is large, when discontinuities are present, however, it is small close to extrema. Thus, instead of relying on their ratio θ_i, we should explicitly consider both normalized slopes and especially the size of the vector $(\delta_{i-1/2}, \delta_{i+1/2})$.
In a first step, we introduce the two-parameter framework, which makes the use of $\theta_i = \delta_{i-1/2}/\delta_{i+1/2}$ unnecessary. This has the advantage, that is does not include divisions by $\delta_{i\pm\frac{1}{2}}$ anymore. In this new framework, the limiter functions are now defined as

$$H(\delta_{i-\frac{1}{2}}, \delta_{i+\frac{1}{2}}) = \phi(\theta_i)\delta_{i+\frac{1}{2}} \tag{4a}$$

$$H(\delta_{i+\frac{1}{2}}, \delta_{i-\frac{1}{2}}) = \phi(\theta_i^{-1})\delta_{i-\frac{1}{2}}. \tag{4b}$$

E.g. the third-order reconstruction, Eq. (2) now reads $H_3(\delta_{i-1/2}, \delta_{i+1/2}) = (2\delta_{i+1/2} + \delta_{i-1/2})/3$ in the two-parameter framework.

Having written the limiter functions and thus, the reconstructions Eq. (1), in the two-parameter setting now allows us to define an indicator function η which measures the magnitude of the vector $(\delta_{i-1/2}, \delta_{i+1/2})$. In this manner, it is possible to distinguish between the case of a large gradient with one zero slope and the case of an extrema with one zero slope. The indicator function is defined as

$$\eta = \frac{\sqrt{\delta_{i-1/2}^2 + \delta_{i+1/2}^2}}{\alpha \Delta x^2} \quad \text{with } \alpha = \sqrt{\frac{5}{2} + \mathcal{O}(\Delta x)} \max_{x \in \Omega \setminus \Omega_d} |u_0''(x)|, \tag{5}$$

where Ω is the computational domain and Ω_d is a set of points where the initial condition $u_0(x)$ is discontinuous. The function η measures if the vector of undivided slopes is sufficiently small. In this case, the reconstruction can be switched to the full-3$^{\text{rd}}$-order reconstruction. This is summarized in the combined limiter function

$$H_{3\text{L}}^{(\text{c})}(\delta_{i-1/2}, \delta_{i+1/2}) := \begin{cases} H_3(\delta_{i-1/2}, \delta_{i+1/2}) & \text{if } \eta < 1 \\ H_{3\text{L}}(\delta_{i-1/2}, \delta_{i+1/2}) & \text{if } \eta \geq 1. \end{cases} \tag{6}$$

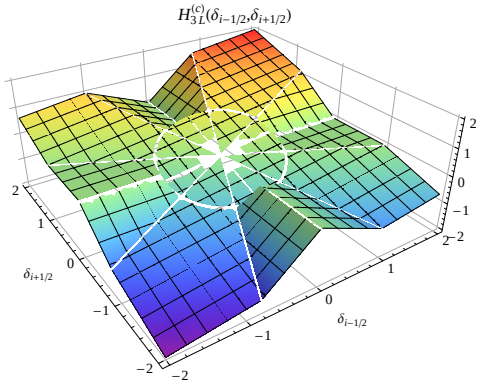


FIGURE 2. 3rd-order limiter function incl. the asymptotic region.

Figure 2 shows the combined limiter function $H_{3L}^{(c)}$.

Finally, we consider the Sod Shock Tube Problem [3] to show the excellent performance of the new limiter function. We compare the full-third-order reconstruction H_3, the combined limiter $H_{3L}^{(c)}$ and third-order WENO, WENO3, as introduced by Jiang and Shu [4]. Note that $\alpha = 0$ in the formulation of η, Eq. (5), which means that $H_{3L}^{(c)}$ reduces to H_{3L} in this case. Fig. 3 shows the solution at time $T_{\text{end}} = 0.8$ on a grid with $N = 100$ cells, CFL 0.95 and $\gamma = 1.4$. As expected, the full-third-order reconstruction yields overshoots at the discontinuities, which are well-removed by the limiter function $H_{3L}^{(c)}$. WENO3 does not yield overshoots, however, does not approach the true solution as well as $H_{3L}^{(c)}$.

REFERENCES

[1] R. Artebrant, H.J. Schroll *Conservative logarithmic reconstructions and finite volume methods*, SISC, **27(1)** (2005), 294-314.
[2] M. Čada, M. Torrilhon, *Compact third-order limiter functions for finite volume methods*, JCP **228(11)** (2009), 4118-4145.
[3] G.A. Sod, *A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws*, JCP **27(1)** (1978), 1-31.
[4] G.S. Jiang, C.W. Shu, *Efficient implementation of weighted ENO schemes*. JCP **126(1)** (1996), 202-228.

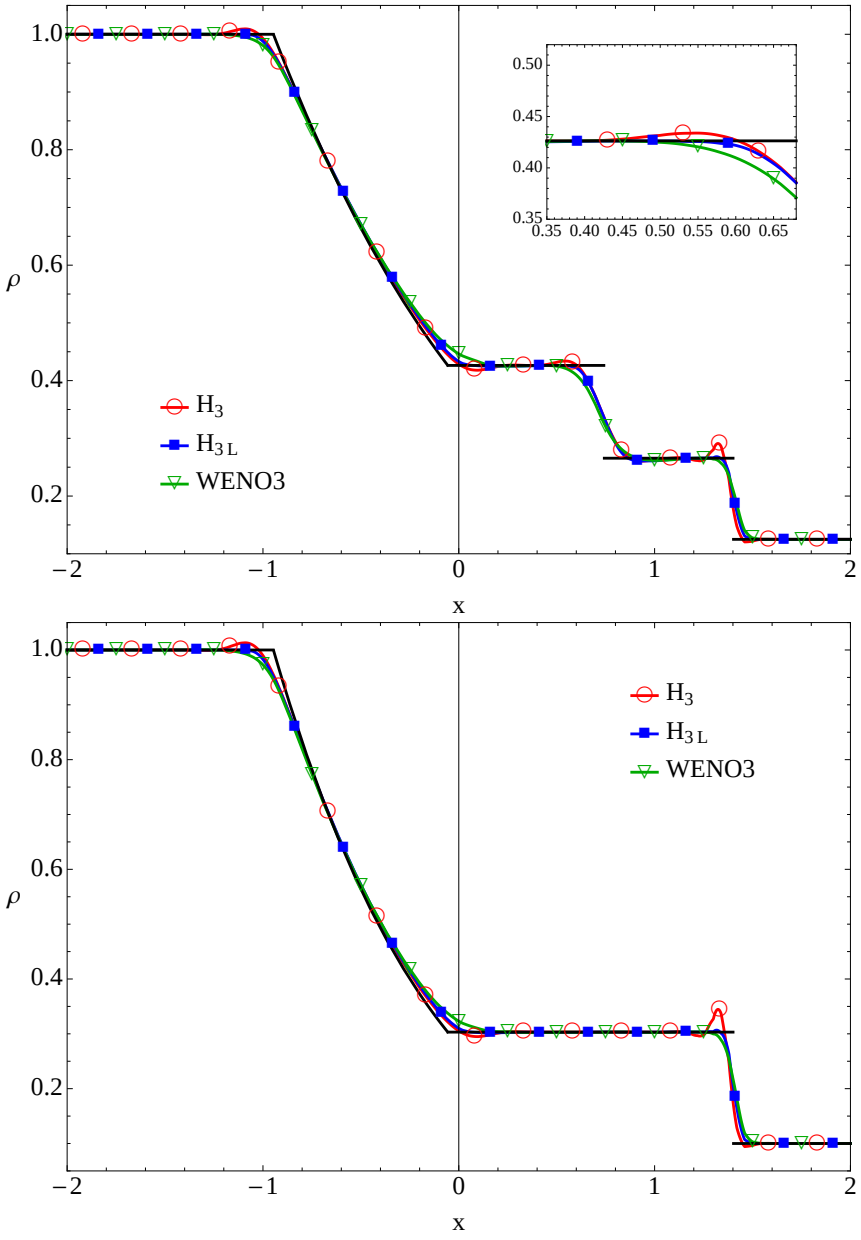


FIGURE 3. Sod's Problem with $N = 200$ grid cells, CFL 0.95, $T_{\text{end}} = 0.8$.

On maximum principle and positivity-preserving high order schemes

Chi-Wang Shu

We give a survey of our recent work, jointly with Xiangxiong Zhang and other collaborators, on the construction of uniformly high order accurate discontinuous Galerkin (DG) and weighted essentially non-oscillatory (WENO) finite volume (FV) schemes which satisfy strict maximum principle for nonlinear scalar conservation laws, passive convection in incompressible flows, and nonlinear scalar convection- diffusion equations, and preserve positivity for density and pressure for compressible Euler systems in computational fluid dynamics. These schemes are referred to as bound-preserving high order schemes.

In [5, 6], A general framework (for arbitrary order of accuracy) is established, which consists of the following 3 ingredients: (1) A first order accurate scheme which has the bound-preserving property under a suitable CFL condition; (2) a simple scaling limiter applied to the high order DG or FV method, involving only evaluations of the polynomial solution at certain quadrature points, which does not affect high order accuracy and guarantees the same bound-preserving property for the first order Euler forward time discretization under a modified CFL condition; (3) a strong stability preserving (SSP) high order Runge-Kutta or multistep time discretization which increases time accuracy and maintains the same bound-preserving property.

One remarkable property of this approach is that the second ingredient above is straightforward to extend two and higher dimensions on arbitrary triangulations [9], and the scaling limiter is local in the cell, thus it keeps the parallel efficiency of the original algorithm. The schemes constructed in this framework are extremely robust, especially for problems involving strong shocks or even δ-function singularities [3]. We will list applicability of the method for problems including arbitrary equations of state [1], source terms [7], shallow water equations [2], Lagrangian schemes [1], and positivity-preserving high order finite volume scheme and piecewise linear DG scheme for convection-diffusion equations [4, 10]. A survey can be found in [8].

References

[1] J. Cheng and C.-W. Shu, *Positivity-preserving Lagrangian scheme for multi-material compressible flow*, Journal of Computational Physics, **257** (2014), 143–168.

[2] Y. Xing, X. Zhang and C.-W. Shu, *Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations*, Advances in Water Resources, **33** (2010), 1476–1493.

[3] Y. Yang, D. Wei and C.-W. Shu, *Discontinuous Galerkin method for Krause's consensus models and pressureless Euler equations*, Journal of Computational Physics, **252** (2013), 109–127.

[4] X. Zhang, Y. Liu and C.-W. Shu, *Maximum-principle-satisfying high order finite volume weighted essentially nonoscillatory schemes for convection-diffusion equations*, SIAM Journal on Scientific Computing, **34** (2012), A627–A658.

[5] X. Zhang and C.-W. Shu, *On maximum-principle-satisfying high order schemes for scalar conservation laws*, Journal of Computational Physics, **229** (2010), 3091–3120.

[6] X. Zhang and C.-W. Shu, *On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes*, Journal of Computational Physics, **229** (2010), 8918–8934.
[7] X. Zhang and C.-W. Shu, *Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms*, Journal of Computational Physics, **230** (2011), 1238–1248.
[8] X. Zhang and C.-W. Shu, *Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments*, Proceedings of the Royal Society A, **467** (2011), 2752–2776.
[9] X. Zhang, Y. Xia and C.-W. Shu, *Maximum-principle-satisfying and positivity-preserving high order discontinuous Galerkin schemes for conservation laws on triangular meshes*, Journal of Scientific Computing, **50** (2012), 29–62.
[10] Y. Zhang, X. Zhang and C.-W. Shu, *Maximum-principle-satisfying second order discontinuous Galerkin schemes for convection-diffusion equations on triangular meshes*, Journal of Computational Physics, **234** (2013), 295–316.

Numerical bifurcation analysis of the macroscopic behavior in multiscale systems

Jens Starke
(joint work with Christian Marschler and Jan Sieber)

An implicit multiscale approach is suggested which enables numerical investigations of the macroscopic behavior of microscopically defined systems including continuation and bifurcation analysis on the coarse or macroscopic level where no explicit equations are available. This approach fills a gap in the analysis of many complex real-world applications where the microscopic system is too large for direct investigations and macroscopic equations cannot be obtained analytically.

This multiscale approach belongs to the so-called equation-free methods which were suggested by Kevrekidis (see [3] and references therein). Equation-free methods take advantage of a typical property of complex systems which show pattern formation namely that there is a fast convergence of many degrees of freedom to a low-dimensional slow manifold. This key property can be used for analytical as well as numerical dimension reduction, known in physics as adiabatic elimination or slaving principle. The suggested implicit equation-free variant reduces the numerical error significantly [4], [5]. It can be shown in the framework of slow-fast dynamical systems, using Fenichel theory, that the implicitly defined coarse-level time-stepper converges to the true dynamics on the slow manifold.

The information about the macroscopic behavior necessary for the numerical analysis on that level (where equations of motion are not explicitly available) is obtained by short simulations of the given microscopic model with a time dependent microscopic state $u(t_0) \in \mathbb{R}^N$. We assume now that the system behavior converges to an attractive low-dimensional slow manifold and the goal is to construct a macroscopic time-stepper Φ for the macroscopic behavior on the low-dimensional slow manifold by an appropriately initialized microscopic time-stepper M. In details, the macroscopic time-stepper Φ with $x(t_0 + t) = \Phi(t, x(t_0))$, where $x(t_0) \in \mathbb{R}^n$ describes the macroscopic state on the slow manifold and $n \ll N$, is

computed by using the appropriately initialized microscopic time-stepper M with $u(t_0+t) = M(t, u(t_0))$. This requires to change back and forth between the microscopic and macroscopic level where the map $\mathcal{R}$ from the microscopic description to the macroscopic one is called restriction while the reversed map $\mathcal{L}$ is called lifting:

$$\begin{array}{lccc} \text{macro-level} & x(t_0) & \xrightarrow{\Phi} & x(t_0+t) \\ & \downarrow \mathcal{L} & & \uparrow \mathcal{R} \\ \text{micro-level} & u(t_0) & \xrightarrow{M} & u(t_0+t) \end{array}$$

The definition of the restriction operator $\mathcal{R}$ is typically motivated by a problem specific scientific question but is usually straight forward while the definition of the lifting operator $\mathcal{L}$ often requires advanced problem based knowledge and intuition. As it is in general impossible to construct a lifting that maps on the slow manifold but only close-by, the direct application of the sketched procedure for Φ with

$$\Phi(t, x(t_0)) = \mathcal{R}(M(t_0, \mathcal{L}(x(t_0)))) \tag{1}$$

describes the behavior of the slow manifold instead of on it. A so-called healing step which brings the system state closer to the attractive slow manifold by forward integration with time t_{skip} improves this but changes also the position in the macroscopic phase space such that the behavior is investigated at another position than wanted. To avoid this and to be able to refer in the numerical evaluation to the appropriate states in phase space the following implicit method is suggested:

$$\mathcal{R}(M(t_{\text{skip}}, \mathcal{L}(y))) = \mathcal{R}(M(t_{\text{skip}}+t, \mathcal{L}(x))). \tag{2}$$

This equation defines the map $\Phi : x \mapsto y$ via y implicitly (cf. Fig. 1). The implicit

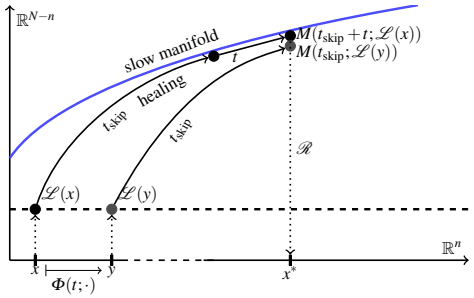


FIGURE 1. The yet unknown macroscopic state y defines the macroscopic time-stepper $\Phi : x \mapsto y$. Initializing the microscopic time-stepper M with state $\mathcal{L}(x)$, the macroscopic dynamics is observed for time t after applying a healing step for time t_{skip}. The restricted endpoint of this, i.e. the corresponding macroscopic position $x^\star$, is compared with the likewise healed and restricted position of $\mathcal{L}(y)$. With kind permission from Springer Science+Business Media: [5].

equation (2) is solved with Newton iterations. Depending on t_{skip} an error estimate can be given for this implicit equation-free method [4].

The approach allows one to formulate the macroscopic dynamics in an implicit way, to perform a coarse projective integration with an implicit Euler method (including time backward integration) as well as continuation with predictor-corrector methods and numerical bifurcation analysis to investigate the parameter dependent change of the qualitative behavior of the macroscopic dynamics. For the special case of the computation of macroscopic equilibria see also [6].

The practical benefit of the implicit equation-free method is demonstrated by performing a coarse bifurcation analysis of a microscopic particle model, the optimal velocity model [1], describing car traffic on single lane highways. The results include an equation-free continuation of traveling waves (cf. Fig. 2), identification of bifurcations as well as two-parameter continuations of bifurcation points. Fur-

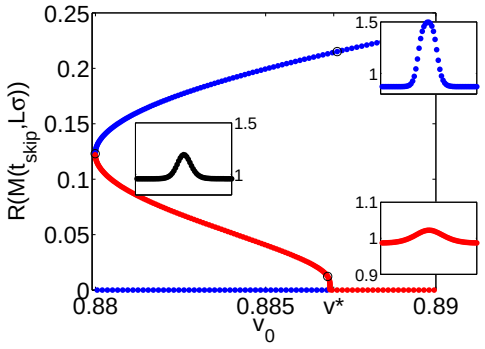


FIGURE 2. Coarse bifurcation diagram of a microscopic traffic model with healed quantities of standard deviation of car distances over parameter v_0 (amplitude of optimal velocity function for drivers behavior). The stable uniform flow (zero standard deviation) becomes unstable at $v^\star$ where a branch of unstable traveling waves is born which changes stability at a fold point. Headway profiles are depicted at positions of small circles. With kind permission from Springer Science+Business Media: [5].

ther results include an equation-free detection and two-parameter continuation of a Hopf bifurcation point in a particle model for pedestrian flow [2].

As the approach does not depend on the specific microscopic model, it can be used to different types and also very realistic models which makes it very powerful for real-world applications. In addition to the presented numerical approach, the implicit formulation might be also advantageous for analytical studies.

REFERENCES

[1] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, Y. Sugiyama, *Dynamical model of traffic congestion and numerical simulation*, Phys. Rev. E, **51**(2) (1995), 1035–1042.

[2] O. Corradi, P. Hjorth, J. Starke, *Equation-free detection and continuation of a Hopf bifurcation point in a particle model of pedestrian flow*, SIAM Journal on Applied Dynamical Systems **11**(3) (2012), 1007–1032.
[3] I. Kevrekidis, G. Samaey, *Equation-Free Multiscale Computation: Algorithms and Applications*, Annual Review of Physical Chemistry, **60**(1) (2009), 321–344.
[4] C. Marschler, J. Sieber, R. Berkemer, A. Kawamoto, J. Starke, *Implicit methods for equation-free analysis: Convergence results and analysis of emergent waves in microscopic traffic models*, SIAM Journal on Applied Dynamical Systems **13**(3) (2014), 1202–1238.
[5] C. Marschler, J. Sieber, P. Hjorth, J. Starke, *Equation-Free Analysis of Macroscopic Behavior in Traffic and Pedestrian Flow*, Pages 423–439 in M. Chraibi, M. Boltes, A. Schadschneider, A. Seyfried (Eds.): *Traffic and Granular Flow '13*, Springer International Publishing, Heidelberg, New York 2015.
[6] C. Vandekerckhove, B. Sonday, A. Makeev, D. Roose, I.G. Kevrekidis, *A common approach to the computation of coarse-scale steady states and to consistent initialization on a slow manifold*, Computers & Chemical Engineering **35**(10) (2011), 1949–1958.

(Adaptive) Tsunami Modelling with Discontinuous Galerkin Schemes

Stefan Vater
(joint work with Nicole Beisiegel, Jörn Behrens)

An important part in the numerical modeling of tsunami or storm surge events is the accurate and robust treatment of flooding and drying at the coast. Within the framework of the shallow water equations, such an algorithm should preserve the steady state of a fluid at rest, be mass conservative and should preserve the positivity of the fluid depth among other features of the exact solution.

While the inundation problem has been well investigated in the context of finite volume discretizations, it is still an active research area for discontinuous Galerkin (DG) methods. Here we present a novel treatment for explicit second-order Runge-Kutta-DG schemes which is based on a limiting approach [4]. The core of the method is a velocity based "limiting" of the momentum, which provides stable and accurate solutions in the computation of wetting and drying events. Additional limiting of the fluid depth ensures its positivity while preserving local mass conservation. A special flux modification in cells located at the wet/dry interface leads to a well-balanced method, which maintains the steady state at rest.

The problem of stably computing the velocity is addressed in this study in the context of a limiter based treatment of inundation events. The basic idea is that the momentum variable is modified on the basis of the resulting velocity distribution given a fixed (but already limited) distribution of the fluid depth. This results in a stable flux computation, which usually involves the computation of the velocity at some point. The general idea is borrowed from finite volume methods, where limiting in other than the primary flow variables often enhances the solution (see [2] and references therein). In case of compressible gas dynamics this approach can easily ensure that the pressure always remains non-negative and velocity and pressure stay constant across contact discontinuities. For shallow water flows limiting in surface elevation generally makes sense, since this quantity is constant for the steady state at rest [1]. Furthermore, limiting in the velocity

instead of momentum has often been employed in these methods (see e.g. [3]). However, in finite volume methods, these limited values are only used for the flux computation at the cell interfaces. In discontinuous Galerkin methods, on the other hand, the solution itself is limited and further used throughout the computations. Therefore, the non-trivial in-cell functional behavior of the velocity, which is the quotient of two polynomials, cannot be ignored in the limiting process. To the authors' knowledge this concept of limiting in other than the primary variables has not been thoroughly transferred to discontinuous Galerkin methods, yet.

When it comes to wetting and drying, i.e., parts of the domain have water depth $h = 0$, several problems arise which must be handled by the numerical algorithm. First, the wet-dry transition might be within a cell and cannot be exactly represented by a piecewise (smooth) polynomial DG discretization. The result is the occurrence of artificial gradients in the surface elevation that can influence the tendencies of the momentum equation and render the scheme unbalanced. For this problem, we propose a special flux modification in these cells. Furthermore, one must ensure that the fluid depth remains non-negative. Otherwise the shallow water equations are undefined at these points. To obtain such a behavior, we follow [6] who showed that under a suitable CFL condition one gets positivity in the mean. Positivity in the whole cell can be then obtained by the application of a positivity preserving (PP) limiter. The third and in the author's opinion least investigated problem is that near the wet/dry interface, both, fluid depth and momentum go to zero, which yields an ill-conditioned computation of the velocity $u = (hu)/h$ in these regions. This problem is solved by our velocity based "limiting" of the momentum.

The velocity based limiting is illustrated in Fig. 1, where we have compiled a possible configuration in surface elevation and momentum at the wet/dry interface in the top row. The resulting velocity distribution is displayed in the rightmost figure. As one can see, the velocity in the center cell has an unphysical extreme value at $x = 2$, where both, fluid depth and momentum become small. For the limiting, the cell mean values of h and (hu) are computed (magenta diamonds) and upper and lower limits for the velocity are derived (dashed magenta line). This results in two limited velocity and associated momentum distributions, which are marked as red with triangles and green with squares at the end points. The final distribution is the velocity distribution with the smallest in-cell variation – in this case it is the red one with triangles. As one can see the associated momentum distribution has a slightly bigger in-cell variation compared to the original one.

By extending the algorithm to two-dimensional triangular grids, one finds that in each cell there are three possible limited velocity distributions to consider instead of two in the one-dimensional case [5]. These arise by limiting the velocities in two vertexes and computing the resulting velocity in the third one by fixing the fluid depth and the mean value of the momentum. The other parts of the algorithm are equivalent to the one-dimensional case.

The performance of the method is verified by several test cases, and results are compared to classical second order hydrostatic reconstruction in the context of

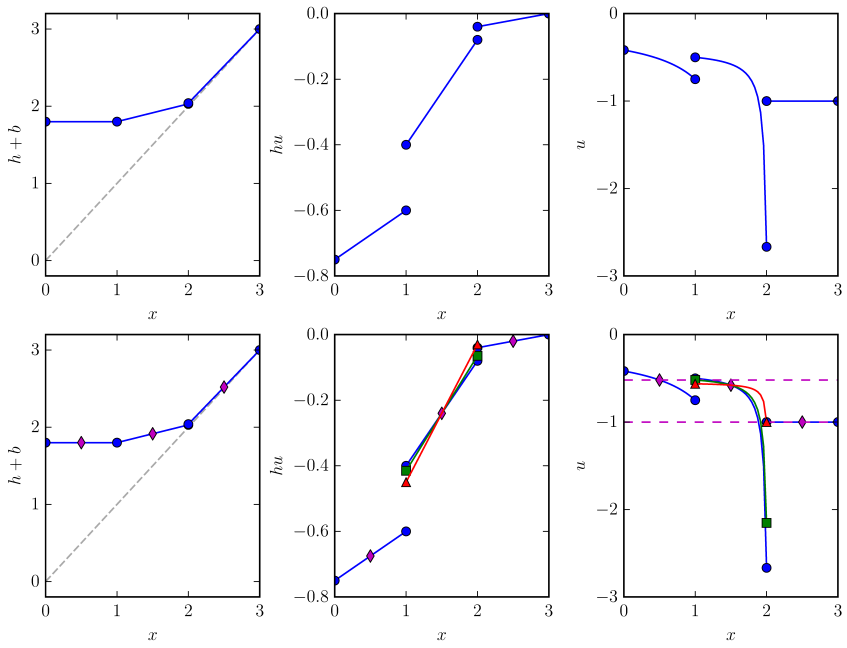


FIGURE 1. Visualization of the velocity based "limiting" of the momentum. Initial distribution in surface elevation, momentum and velocity (top row) and limiting procedure (bottom row). See the text for further explanation.

finite volume methods. These test cases demonstrate the well-balancing property of the method and its stability in case of rapid transition of the wet/dry interface. We also verify the conservation of mass and investigate the convergence characteristics of the scheme. Finally, the wetting and drying treatment is applied to quasi-realistic tsunami test problems.

The development of limiters for discontinuous Galerkin methods is still under heavy development. The velocity based "limiting" of the momentum resembles the limiting procedure in finite volume methods in other than the primary flow variables. We are optimistic that the general concept proposed might be transferable to other problems such as DG methods for the solution of inviscid compressible flow applications. An open question is the possibility to extend the proposed concept to higher than linear discontinuous Galerkin finite elements, which will be addressed in future research.

REFERENCES

[1] E. Audusse, F. Bouchut, M.O. Bristeau, R. Klein, and B. Perthame, *A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows* (2004), SIAM Journal on Scientific Computing **25**, 2050–2065.

[2] B. van Leer, *Upwind and high-resolution methods for compressible flow: From donor cell to residual-distribution schemes*, Communications in Computational Physics **1** (2006), 192–206.
[3] S. Noelle, N. Pankratz, G. Puppo, and J.R. Natvig, *Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows*, Journal of Computational Physics **213** (2006), 474–499.
[4] S. Vater, N. Beisiegel, and J. Behrens, *A Limiter-Based Well-Balanced Discontinuous Galerkin Method for Shallow-Water Flows with Wetting and Drying: One-Dimensional Case*, Advances in Water Resources **85** (2015), 1–13.
[5] S. Vater, N. Beisiegel, and J. Behrens, *A Limiter-Based Well-Balanced Discontinuous Galerkin Method for Shallow-Water Flows with Wetting and Drying: Two-Dimensional Case* (2015), in preparation.
[6] Y. Xing, X. Zhang, and C.-W. Shu, *Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations*, Advances in Water Resources **33** (2010), 1476–1493.

Automated parameters for troubled-cell indicators using outlier detection

Mathea J. Vuik
(joint work with Jennifer K. Ryan)

In general, solutions of nonlinear hyperbolic PDEs contain discontinuities or develop shocks at a certain time. One option for improving the numerical treatment of the spurious oscillations that occur near these artifacts is through the application of a limiter. The cells where such treatment is necessary are referred to as troubled cells, which are selected using a troubled-cell indicator. Three important methods for troubled-cell indication are the KXRCF shock detector [3], the minmod-based TVB indicator [2] and the modified multiwavelet troubled-cell indicator [5, 6]. For the multiwavelet approach, the global DG approximation on 2^n elements is decomposed into a sum of a global average and finer details on different levels:

$$u_h(x) = \sum_{j=0}^{2^n-1} \sum_{\ell=0}^{k} u_j^{(\ell)} \phi_\ell(\xi_j) = \sum_{\ell=0}^{k} s_{\ell 0}^0 \phi_\ell(x) + \sum_{m=0}^{n-1} \sum_{j=0}^{2^m-1} \sum_{\ell=0}^{k} d_{\ell j}^{n-1} \psi_{\ell j}^{n-1}(x).$$

Using scaled Legendre polynomials for the DG approximation, the multiwavelet basis $\{\psi_\ell\}_{\ell=0}^k$ follows from Alpert [1]. The scaling-function and multiwavelet coefficients can be efficiently computed from the DG coefficients using decomposition.

The multiwavelet coefficients of level $n-1$ are related to the jumps in (derivatives of) the DG approximation. We have shown that these coefficients equal

$$d_{\ell j}^{n-1} = 2^{-\frac{n-1}{2}} \sum_{m=0}^{k} c_{m\ell}^n \cdot \left(u_h^{(m)}(x_{2j+1/2}^+) - u_h^{(m)}(x_{2j+1/2}^-) \right), \tag{1a}$$

with

$$c_{m\ell}^n = \frac{2^{(-n+1)m}}{m!} \cdot \int_0^1 x^m \psi_\ell(x)\, dx, \tag{1b}$$

where $\ell = 0, \ldots, k$, $j = 0, \ldots, 2^{n-1} - 1$, and $u_h^{(m)}$ is the mth derivative of u_h [6]. Using a renumbering technique, we find in total 2^n coefficients for each mode ℓ in level $n-1$, denoted by $\tilde{d}_{\ell j}^{n-1}$, $\ell = 0, \ldots, k$, $j = 0, \ldots, 2^n - 1$.

The value of coefficient d_{kj}^{n-1} suddenly increases with respect to its neighboring values in the neighborhood of a discontinuity. This makes the coefficients useful for troubled-cell indication. We defined the modified multiwavelet troubled-cell indicator, which detects elements I_j and I_{j+1} if

$$|\tilde{d}_{kj}^{n-1}| > C \cdot \max\{|\tilde{d}_{kj}^{n-1}|, j = 0, \ldots, 2^n - 1\},\ C \in [0, 1],$$

where C prescribes the strictness of the indicator.

In general, a troubled-cell indicator only performs well as long as a suitable, problem-dependent parameter is chosen. Until now, these parameters could not be chosen automatically such that the indicator works well in a variety of situations.

The choice of the parameter has impact on the approximation: it determines the strictness of the troubled-cell indicator. An inappropriate choice of the parameter will result in detection (and limiting) of too few or too many elements. Detection of too few elements leads to spurious oscillations. If too many elements are detected, then the limiter is applied too often, and therefore the method is more costly and the approximation smooths out after a long time. The optimal parameter is chosen such that the minimal number of troubled cells is detected and the resulting approximation is free of non-physical spurious oscillations. In general, many tests are required to obtain this optimal parameter for each problem.

We determined that for each troubled-cell indicator the sudden increase or decrease of the indicator value with respect to the neighboring values is important for detection. Indication basically reduces to detecting the outliers of a vector (one dimension) or matrix (two dimensions). We do this by using Tukey's boxplot approach to detect which coefficients in a vector are straying far beyond others [4]. In order to take the local structure of the data into account, we split our global indication vector into local vectors (typically containing 16 coefficients). Our outlier-detection algorithm executes the following steps:

Algorithm 1 Outlier-detection algorithm using local vectors.

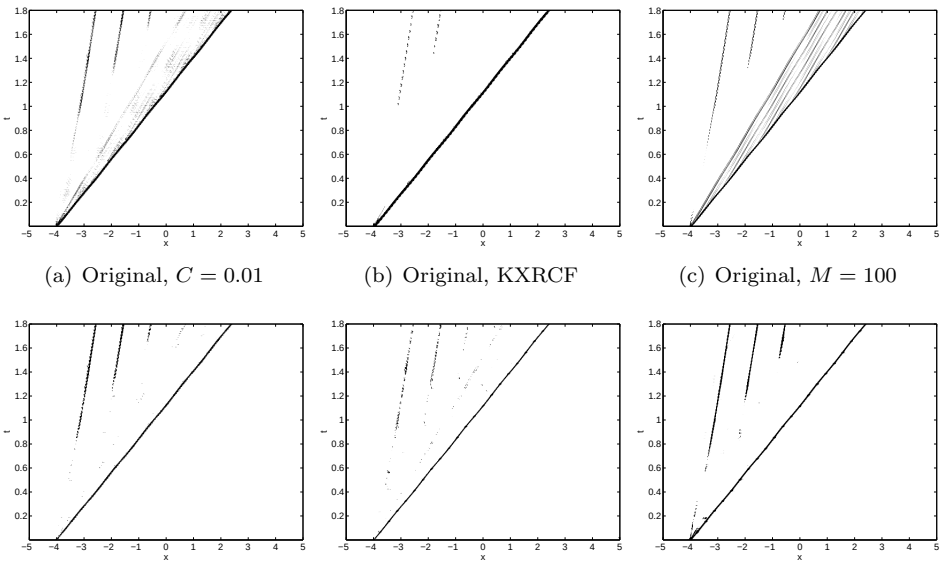
Send in a suitable troubled-cell indication vector $\mathbf{D}$.
Split this vector into local vectors, $\mathbf{d}$.
for all local vectors **do**
  Sort $\mathbf{d}$ to obtain $\mathbf{d}^s$.
  Compute the 25th and 75th percentile of the data (Q_1 and Q_3, respectively).
  Detect d_j^s in the smallest 25% of $\mathbf{d}^s$ if $d_j^s < Q_1 - 3(Q_3 - Q_1)$, and d_j^s in the biggest 25% of $\mathbf{d}^s$ if $d_j^s > Q_3 + 3(Q_3 - Q_1)$.
end for
Ignore detected outliers in the left half of the local region if not detected in the left-neighboring vector, similarly test the right half of the local region.

The whisker length 3 is chosen such that only a few 'false positives' are found if the data are well behaved. Using this technique the problem-dependent parameter that troubled-cell indicators require is no longer necessary [7].

In Figure 1 the performance of three original troubled-cell indicators can be compared to the new outlier-detection approach for the Shu-Osher problem (Euler equations). For this problem an initial discontinuity is moving to the right, thereby evolving (highly oscillatory) continuous regions and developing new shocks in the left side of the domain.

The first row of the figure consists of time-history plots of detected troubled cells using the original indicators with optimal parameter. Note that both the multiwavelet indicator with $C = 0.01$ and the minmod-based TVB indicator with $M = 100$ detect the highly-oscillatory region as being discontinuous. In this case, the KXRCF indicator gives more accurate results.

In the second row of the figure, the time-history plots are shown when the indication vectors are used in the outlier-detection algorithm. All three indication techniques detect the correct regions, and the approximations are as expected. This means that the outlier-detection algorithm is indeed able to replace the problem-dependent parameters in the original indicators.



(a) Original, $C = 0.01$ (b) Original, KXRCF (c) Original, $M = 100$

(d) Outlier, multiwavelets (e) Outlier, KXRCF value (f) Outlier, minmod-TVB

FIGURE 1. Detected troubled cells, Shu-Osher, $k = 2$, $\Delta x = 10/512$.

References

[1] B.K. Alpert, *A Class of Bases in L^2 for the Sparse Representation of Integral Operators*, SIAM Journal on Mathematical Analysis, **24** (1993), 246 – 262.

[2] B. Cockburn and C.-W. Shu, *TVB Runge-Kutta Local Projection Discontinuous Galerkin Finite Element Method for Conservation Laws II: General Framework*, Mathematics of Computation, **52** (1989), 411 – 435.

[3] L. Krivodonova, J. Xin, J.-F. Remacle, N. Chevaugeon, and J.E. Flaherty, *Shock detection and limiting with discontinuous Galerkin methods for hyperbolic conservation laws*, Applied Numerical Mathematics, **48** (2004), 323 – 338.

[4] J.W. Tukey, *Exploratory Data Analysis*, Addison-Wesley Publishing Company, 1977.

[5] M.J. Vuik and J.K. Ryan, *Multiwavelet troubled-cell indicator for discontinuity detection of discontinuous Galerkin schemes*, Journal of Computational Physics, **270** (2014), 138 – 160.

[6] M.J. Vuik and J.K. Ryan, *Multiwavelets and jumps in DG approximations*, in Spectral and High Order Methods for Partial Differential Equations ICOSAHOM '14, R.M. Kirby, M. Berzins and J.S. Hesthaven, eds., Springer, 2015.

[7] M.J. Vuik and J.K. Ryan, *Automated parameters for troubled-cell indicators using outlier detection*, ArXiv e-print `http://arxiv.org/abs/1504.05783`.

Subgrid Scale Models for Large Eddy Simulations Using Discontinuous High-Order Methods

Zhi J. Wang

(joint work with Yanan Li and Z. J. Wang)

In large eddy simulations (LES) of turbulent flows, large scale motions are resolved by the numerical simulation while the effect of the small scale motions is represented as sub-grid scale (SGS) stresses computed with SGS models. In order to assess the performance of SGS models in the context of high-order discontinuous methods, we perform a priori and a posteriori evaluations of five SGS models for the one dimensional Burgers' equation with the high-order flux reconstruction (FR) or correction procedure via reconstruction (CPR) method. It is shown that all models, except the scale similarity model (SSM) and the mixed model (MM), demonstrate very little correlation with the DNS results. The stability of the SSM is also investigated. The effects of numerical dissipation and the models' influence on LES are discussed. Based on the present study, we advocate the use of implicit LES (ILES) in LES in the context of discontinuous high-order methods, or any numerical methods with built-in numerical dissipation.

References

[1] J. Smagorinsky, *General circulation experiments with the primitive equations*, I. The basic experiment, Weather Rev **9**1–99.

[2] Lilly. DK, *A propose modification of the Germano subgrid-scale closure method*, Phys. Fluids **A 4** (1990), 633–635

[3] M. Germano, U. Piomelli, P. Moin, WH. Cabot, *A dynamic subgrid-scale eddy viscosity model*, Phys. Fluids **A 3** (1991), 1760–1765

[4] J. Bardina, JH. Ferziger, WC. Reynolds, *Improved subgrid scale models for large eddy simulation*, Am. Inst. Aeronaut. Astronaut **8**0–1357

[5] C. Meneveau, J. Katz, *Scale-invariance and turbulence models for large-eddy simulation*, Annu. Rev. Fluid Mech. (2000), 1–32
[6] H.T.. Huynh, *A flux reconstruction approach to high-order schemes including discontinuous Galerkin methods*, AIAA paper **2**007-4079
[7] SB. Pope, , Turbulent Flows **C**ambridge, UK: Cambridge Univ. Press
[8] S. Liu, C. Meneveau, J. Katz, *Experimental study of similarity subgrid-scale models of turbulence in the far-field of a jet*, Appl. Sci. Res. **54** 177–190

Edge detection approaches in numerical methods for conservation laws

Martina Wirz

The detection of discontinuities in the numerical solution of a given conservation law is an important tool to handle shocks or discontinuous initial data appropriately. In many applications the numerical data is given at non-equidistantly distributed points, such that classical edge detectors cannot be directly used. Furthermore, the correct choice of a sensible threshold value is difficult for almost all known shock indicators.
We will thus focus on an edge detector based on the conjugated Fourier partial sum

$$\widetilde{S}_n f(x) = -\mathrm{i} \sum_{|k|\leq n} \operatorname{sgn}(k) \hat{\mathbf{f}}_k \mathrm{e}^{\mathrm{i}kx} = f * \widetilde{D}_n(x) = \int_\Omega f(t) \widetilde{D}_n(x-t)\,\mathrm{d}t$$

which can be written as the convolution of a function f and the conjugated Dirichlet kernel $\widetilde{D}$ and has the property to converge pointwise to the jump height of f. Considering generalized kernels instead of the Dirichlet kernel as proposed in [3] yields to an accelaration of the convergence rate away from the discontinuity, which has been validated in the context of spectral methods in one dimension and extended to the quasi-two-dimensional approach in [4].
In order to extend this result to the fully two-dimensional case, we follow the idea of [7] in the classical case, where the conjugated Fourier partial sum in two variables was considered, and apply this to our edge detection technique for the generalized conjugated Fourier partial sum. This leads to improved convergence rates of the generalized conjugated Fourier partial sum in two dimensions $\widetilde{S}^\sigma_{nm} f(x,y)$ which can be seen in various testcases.
Furthermore, we apply this edge detector efficiently in the context of general high order methods where nodal or modal coefficients are given, e.g. for the Spectral-Difference- or Discontinuous-Galerkin-method on triangles [5]. Here, especially Proriol-Koornwinder-Dubiner-polynomials [2] are used due to an efficient modal filtering procedure, for which an exact convertion to Fourier coefficients is possible. Hence, we simply can take $\left|\widetilde{S}^\sigma_{nm} f(x,y)\right| < \varepsilon$ as an edge detector where ε is half the height of the smallest discontinuity. Several testcases show superior results of the Fourier-based detector compared to a common coefficient-based shock indicator. Especially the choice of the threshold value ε is much easier as in other approaches.

In case of *modal* coefficients given, the Fourier based edge detector needs additional computation steps to convert these to Fourier coefficients. We thus will focus on an extension of our results to generalized partial sums: How can we recognize an underlying discontinuity from arbitrary coefficients? To his end, we follow the idea of [6] and extend it for Jacobi polynomials, which already cover a broad class of classical polynomials and are also used in the context of the given Spectral-Difference- and DG-methods from [5]. Here, we especially construct a concentration factor such that $\sum_{I_k} \sigma_k \hat{f}_k \varphi_k(x)$ will converge to the jump height of the function f.

Since many numerical methods are based on *nodal* coefficients, we will also investigate on edge detection techniques based on nodal values in further research, as for example given in the polynomial annihilation [1].

Additionally, we start investigating the connection between modal filters and concentration factors, which are based on the same expansion $\sum_{I_k} \sigma_k \hat{f}_k \varphi_k(x)$ but using different indices I_k and filters resp. concentration factors σ_k.

References

[1] R. Archibald, A. Gelb, R. Saxena, D. Xiu, *Discontinuity detection in multivariate space for stochastic simulations.* Journal of Computational Physics **228** (2009), 2676–2689.

[2] M. Dubiner. *Spectral Methods on Triangles and Other Domains.* Journal of Scientific Computing **6** (1991), 345-390.

[3] A. Gelb, E. Tadmor, *Detection of Edges in Spectral Data.* Applied and Computational Harmonic Analysis **7** (1999), 101-135.

[4] A. Gelb, E. Tadmor, *Spectral reconstruction of one- and two-dimensional piecewise smooth functions from their discrete Data.* Mathematical Modeling and Numerical Analysis **36** (2002), 155–175.

[5] A. Meister, S. Ortleb, Th. Sonar, M. Wirz, *An extended Discontinuous Galerkin and Spectral Difference Method with modal filtering.* ZAMM **93** (6–7) (2013), 459–464.

[6] A. Viswanathan, A. Gelb, D. Cochran, *Iterative Design of Concentration Factors for Jump Detection.* JCP **51** (3) (2012), 631–649.

[7] F. Móricz, *Extension of a Theorem of Ferenc Lukács from Single to Double Conjugate Series.* Journal of Mathematical Analysis and Application **259** (2001), 582–592.

Reporter: Andreas Meister

Participants

Prof. Dr. Remi Abgrall
Institut für Mathematik
Universität Zürich
Winterthurerstrasse 190
8057 Zürich
SWITZERLAND

Prof. Dr. Martin Arnold
Institut für Mathematik
Martin-Luther-Universität
Halle-Wittenberg
06099 Halle
GERMANY

Paola Bacigaluppi
Institut für Mathematik
Universität Zürich
Winterthurerstrasse 190
8057 Zürich
SWITZERLAND

Roxana G. Crisovan
Institut für Mathematik
Universität Zürich
Winterthurerstrasse 190
8057 Zürich
SWITZERLAND

Veronika Diba
FB 10 -
Mathematik u. Naturwissenschaften
Universität Kassel
Heinrich-Plett-Strasse 40
34132 Kassel
GERMANY

Prof. Dr. Vit Dolejsi
Department of Numerical Mathematics
Charles University
Sokolovska 83
18675 Praha 8
CZECH REPUBLIC

Prof. Dr. Gregor Gassner
Mathematisches Institut
Universität zu Köln
Weyertal 86 - 90
50931 Köln
GERMANY

Dr. Marc I. Gerritsma
Department of Aerospace Engineering
Delft University of Technology
Kluyverweg 1
2629 HS Delft
NETHERLANDS

Prof. Dr. Matthias K. Gobbert
Department of Mathematics & Statistics
University of Maryland
Baltimore County
1000 Hilltop Circle
Baltimore, MD 21250
UNITED STATES

Rene Goertz
Institut für Mathematik
Universität Hildesheim
Samelsonplatz 1
31141 Hildesheim
GERMANY

Prof. Dr. Jean-Luc Guermond
Department of Mathematics
Texas A & M University
College Station, TX 77843-3368
UNITED STATES

Prof. Dr. Hervé Guillard
INRIA Méditerranée
2004 Route des Lucioles
06560 Sophia Antipolis Cedex
FRANCE

Dr. Maria Han Veiga
Departement Mathematik
ETH-Zentrum
Rämistrasse 101
8092 Zürich
SWITZERLAND

Prof. Dr. Stefan Hartmann
Institut für Technische Mechanik
Technische Universität Clausthal
Adolph-Roemer-Strasse 2A
38678 Clausthal-Zellerfeld
GERMANY

Dr. Koen Hillewaert
CENAERO
Rue des Frères Wright 29
6041 Gosselies
BELGIUM

Prof. Dr. Willem Hundsdorfer
Centrum Wiskunde & Informatica
(CWI)
Postbus 94079
1090 GB Amsterdam
NETHERLANDS

Dr. Tulin Kaman
Institut für Mathematik
Universität Zürich
Winterthurerstrasse 190
8057 Zürich
SWITZERLAND

Prof. Dr. David Ketcheson
CEMSE
4700 King Abdullah University of
Science and Technology (KAUST)
Thuwal 23955-6900
SAUDI ARABIA

Dr. Knut Klingbeil
Leibniz-Institut für Ostseeforschung
Warnemünde
Seestrasse 15
18119 Rostock
GERMANY

Dr. Stefan Kopecz
FB 10 -
Mathematik u. Naturwissenschaften
Universität Kassel
Heinrich-Plett-Strasse 40
34132 Kassel
GERMANY

Prof. Dr. David A. Kopriva
Department of Mathematics
Florida State University
Tallahassee, FL 32306-4510
UNITED STATES

Prof. Dr. Detlef Kuhl
Institut für Baustatik und Baudynamik
Universität Kassel
Mönchebergstrasse 7
34109 Kassel
GERMANY

Prof. Dr. Göktürk Kuru
ONERA - CFD and Aeroacoustics
Department
University of Bordeaux
29, avenue de la Division Leclerc
P.O. Box 72
92322 Chatillon Hauts-de-Seine
FRANCE

Dr. Adrien Loseille
INRIA Rocquencourt
Domaine de Voluceau
B. P. 105
78153 Le Chesnay Cedex
FRANCE

Kjetil Olsen Lye
Departement Mathematik
ETH-Zentrum
Rämistrasse 101
8092 Zürich
SWITZERLAND

Prof. Dr. Andreas Meister
FB 10 -
Mathematik u. Naturwissenschaften
Universität Kassel
Heinrich-Plett-Strasse 40
34132 Kassel
GERMANY

Prof. Dr. Siddhartha Mishra
Departement Mathematik
ETH-Zentrum
Rämistrasse 101
8092 Zürich
SWITZERLAND

Prof. Dr. Claus-Dieter Munz
Institut für Aero- und Gasdynamik
Universität Stuttgart
Pfaffenwaldring 21
70569 Stuttgart
GERMANY

Prof. Dr. Jan Nordström
Department of Mathematics
Linköping University
581 83 Linköping
SWEDEN

Dr. Philipp Öffner
Institut for Computational Mathematics
Technische Universität Braunschweig
Pockelsstrasse 14
38106 Braunschweig
GERMANY

Dr. Sigrun Ortleb
FB 10 -
Mathematik u. Naturwissenschaften
Universität Kassel
Heinrich-Plett-Strasse 40
34132 Kassel
GERMANY

Prof. Dr. Per-Olof Persson
Department of Mathematics
University of California
Berkeley CA 94720-3840
UNITED STATES

Hendrik Ranocha
Institut for Computational Mathematics
Technische Universität Braunschweig
Pockelsstrasse 14
38106 Braunschweig
GERMANY

Prof. Dr. Stefano Rebay
Dipartimento di Ingegneria
Meccanica e Industriale
Università di Brescia
Via Branze, 38
25123 Brescia
ITALY

Dr. Florent Renac
DMFN/NFLU
ONERA
29, Avenue de la Division Leclerc
92320 Chatillon Hauts-de-Seine
FRANCE

Dr. Mario Ricchiuto
INRIA Bordeaux
200, Ave. de la Vieille Tour
33405 Talence Cedex
FRANCE

Prof. Dr. Ulrich Rüde
Lehrstuhl für Informatik 10
(Systemsimulation)
Universität Erlangen-Nürnberg
Cauerstraße 11
91058 Erlangen
GERMANY

Prof. Dr. Steven Ruuth
Department of Mathematics & Statistics
Simon Fraser University
Burnaby, B.C. V5A 1S6
CANADA

Birte Schmidtmann
Computational Mathematics, Group
CCES
RWTH Aachen University
Schinkelstrasse 2
52062 Aachen
GERMANY

Prof. Dr. Chi-Wang Shu
Division of Applied Mathematics
Brown University
Box F
182 George Street
Providence, RI 02912
UNITED STATES

Prof. Dr. Thomas Sonar
Institut for Computational Mathematics
Technische Universität Braunschweig
Pockelsstrasse 14
38106 Braunschweig
GERMANY

Prof. Dr. Raymond J. Spiteri
Department of Computer Science
University of Saskatchewan
110 Science Place
Saskatoon Sask. S7N 5E6
CANADA

Jens Starke
School of Mathematical Sciences
Queen Mary
University of London
Mile End Road
London E1 4NS
UNITED KINGDOM

Dr. Svetlana Tokareva
Institut für Mathematik
Universität Zürich
Winterthurerstrasse 190
8057 Zürich
SWITZERLAND

Dr. Stefan Vater
CEN - Zentrum für Erdsystemforschung
und Nachhaltigkeit
Universität Hamburg
Raum 409.1
Grindelberg 5
20144 Hamburg
GERMANY

Prof. Dr. Peter Vincent
Department of Aeronautics
Imperial College London
South Kensington Campus
London SW7 2AZ
UNITED KINGDOM

Mathea J. Vuik
Faculty of EEMCS, DIAM
Delft University of Technology
Mekelweg 4
2628 CD Delft
NETHERLANDS

Prof. Dr. Zhi J. Wang
Department of Aerospace Engineering
University of Kansas
2120 Learned Hall
1530 W. 15th Street
Lawrence, KS 66045
UNITED STATES

Dr. Martina Wirz
Institut for Computational Mathematics
Technische Universität Braunschweig
Pockelsstrasse 14
38106 Braunschweig
GERMANY