Adaptive FE Methods for Conservation Equations

Ralf Hartmann

Abstract. We present an approach to solving conservation equations by the adaptive discontinuous Galerkin finite element method (DG method). Using a global duality argument and Galerkin orthogonality, we obtain a residual-based representation for the error with respect to an arbitrary functional of the solution. This results in local indicators that can be evaluated numerically and which are used for adaptive mesh refinement. In this way, very economical meshes can be generated which are tailored to the cost-efficient computation of the quantity of interest. We demonstrate the main ingredients of this approach to a posteriori error estimation and test the quality of the error estimator and the efficiency of the meshes by some numerical examples.

1. Introduction

We consider a general hyperbolic system of conservation equations,

$$\nabla \cdot F(u(x)) = 0, \qquad x \in \Omega, \tag{1}$$

on a bounded domain $\Omega \in \mathbb{R}^d$, where $u : \mathbb{R}^d \to \mathbb{R}^m$ and $F : \mathbb{R}^m \to \mathbb{R}^{m \times d}$. Let this problem be equipped with suitable boundary conditions, e.g. for the scalar case $u{=}g$ on the inflow boundary $\partial\Omega^- := \{x \in \partial\Omega | (F'(u) \cdot n)(x) < 0\}$.

The aim of this work is to apply the basic ideas of local residual-based a posteriori error estimation to discontinuous Galerkin (DG) methods for hyperbolic problems. Prior work on a posteriori error estimation and the derivation of rigorous error estimators for finite element methods with streamline diffusion (SDFEM) was done for example by Rannacher [6] and by Houston, Rannacher, Süli [4]. For an overview of the development of DG methods, we refer to Cockburn et al. [3] and the references therein.

2. Discretization

Let $\mathbb{T}_h$ be a regular triangulation of the domain Ω into cells K, where h denotes the maximal cell diameter. Multiplying problem (1) by a test function v and integration

The author acknowledges the support by the DFG Priority Research Program 'Analysis and Numerics for Conservation Laws' and the SFB 359 at the IWR, University of Heidelberg.

by parts on each cell K results in

$$\sum_{K\in\mathbb{T}_h}\Big[-(F(u),\nabla v)_K+(F(u)\cdot n,v)_{\partial K}\Big]=0\qquad\forall v\in V, \tag{2}$$

for $u\in V$, where V denotes the natural solution space and $n|_{\partial K}$ the outward unit normal to the cell boundary ∂K. This problem is approximated by a Galerkin method using a sequence of finite dimensional subspaces $V_h\subset V$. Hence for discretizing this problem, we replace the exact solution u and the test functions v by discrete functions $v_h\in V_h$ and $u_h\in V_h$, respectively. Here we choose $V_h=V_h^p$ to be the finite element space of discontinuous piecewise polynomial functions of a fixed degree p. Since the function u_h is discontinuous on faces between two neighboring cells, we also need to substitute the flux term $F(u_h)\cdot n$ by a consistent and conservative numerical flux $H(u_h,\hat{u}_h,n)$. This leads to the following discrete problem that seeks a $u_h\in V_h^p$, such that $a(u_h,v_h)=0\quad\forall v\in V_h^p$, with

$$a(u_h,v_h):=\sum_{K\in\mathbb{T}_h}\Big[-(F(u_h),\nabla v_h)_K+(H(u_h,\hat{u}_h,n),v_h)_{\partial K}\Big], \tag{3}$$

where $\hat{u}_h$ denotes the value of u_h on the neighboring cell if the face is an inner face and $\hat{u}_h=g$ a suitable boundary value on the boundary. The form $a(\cdot,\cdot)$ is linear in its second argument and in general nonlinear in its first.

Remark 2.1. *In the case of a nonlinear problem this discretization needs to be stabilized to avoid spurious oscillations and overshoots at discontinuities of the solution. This stabilization may for example be done by using shock-capturing, e.g. the method proposed by Jaffre, Johnson and Szepessy* [5].

Inserting a continuous exact solution u into the discrete problem and using the consistency of the flux, $H(u,u,n)=F(u)\cdot n$, it reduces to equation (2), i.e., $a(u,v)=0\quad\forall v\in V$. Many different numerical fluxes are known. However, the error estimation approach that will be discussed in the following section, does not depend on the specific choice of the numerical flux as the consistency of a numerical flux is sufficient for the consistency of the numerical method which implies *Galerkin orthogonality* of the FE method:

$$a(u,v_h)-a(u_h,v_h)=0\qquad\forall v_h\in V_h^p. \tag{4}$$

For the error $e=u-u_h$, we now define a bilinear form $L(u,u_h;\cdot,\cdot)$ by

$$L(u,u_h;e,v)=a(u,v)-a(u_h,v)=\int_0^1 a'[su+(1-s)u_h](e,v)ds, \tag{5}$$

where $a'[w](\cdot,v)$ denotes the functional derivative of the form $a(\cdot,v)$. In terms of this bilinear form L the Galerkin orthogonality may be written as

$$L(u,u_h;e,v_h)=0\qquad\forall v_h\in V_h^p, \tag{6}$$

meaning that the error of the FE discretization is orthogonal (with respect to the bilinear form L) to the discrete function space V_h^p.

Remark 2.2. *The functional derivative $a'[w]$ is well defined for continuous functions w (since then the numerical flux $H(w, \hat{w}, n)$ reduces to $F(w) \cdot n$, which is differentiable with respect to w for many hyperbolic problems). But $a'[w]$ is not defined for discontinuous functions w, as in general the numerical flux $H(w, \hat{w}, n)$ is not differentiable with respect to w. Therefore $a'[w]$ needs to be replaced by an approximation, e.g. by a suitable difference quotient.*

3. A posteriori error estimation

We are interested in controlling the error of the numerical solution measured not only in a global norm but in terms of a given target functional $J(\cdot)$. $J(u)$ is the physical quantity of interest and might be, for example, the mean flow across a line, a point value of the solution, or the drag of a body emersed in a fluid. To highlight the main features of our approach, here we consider only the case of a linear target functional. Nonlinear functionals need linearization but can be treated in a similar way.

For representing the error $J(u) - J(u_h) = J(e)$, we define a $z \in V$ to be the solution of the dual problem

$$L(u, u_h; w, z) = J(w), \qquad \forall w \in V. \tag{7}$$

We assume that the dual problem is well posed, which is known at least for scalar problems in (x, t) plane with convex fluxes, see [7]. Then, we can choose the test function $w = e$, use Galerkin orthogonality (6), integrate by parts on each cell and arrive at the error representation

$$\begin{aligned} J(e) = L(u, u_h; e, z) &= L(u, u_h; e, z - z_h), \\ &= \sum_{K \in \mathbb{T}_h} \Big[(R(u_h), z - z_h)_K + (r(u_h), z - z_h)_{\partial K} \Big], \end{aligned} \tag{8}$$

for any discrete approximation $z_h \in V_h^p$ to the dual solution z. Here,

$$R(u_h) := -\nabla \cdot F(u_h) \qquad \text{and} \qquad r(u_h) := F(u_h) \cdot n - H(u_h, \hat{u}_h, n). \tag{9}$$

denote the cell and face residuals, respectively.

One possibility of a posteriori error estimation is to use now a priori estimates for the dual problem to get rid of the terms that contain the dual solutions. The problem is that a priori error estimates of the dual solution are only available for few target functionals, as for example the global L^1-norm, and furthermore require some regularity assumptions for the dual solution that are not satisfied in general. Another drawback is that these a priori estimates represent the worst case scenario and do not use the structure of the specific primal and dual solution under consideration. Therefore they generally lead to over-estimation of the error in the quantity of interest $J(\cdot)$.

An alternative is to proceed as follows: Use the triangle inequality, the Cauchy-Schwarz inequality, and some standard interpolation approximations

$$\|z - z_h\|_K + h^{1/2}\|z - z_h\|_{\partial K} \leq \min_{0<k<p+1} \left\{ c_I^k h_K^k \|\nabla^k z\|_K \right\}, \tag{10}$$

as we can choose $z_h = I_h z \in V_h$, an interpolation of z. This results in

$$|J(e)| \leq \sum_{K \in \mathbb{T}_h} \rho_K \omega_K \tag{11}$$

$$\rho_K = \|R(u_h)\|_K + h_K^{-\frac{1}{2}} \|r(u_h)\|_{\partial K}, \qquad \omega_K = \min_{0\leq k\leq p+1} \left\{ c_I^k h_K^k \|\nabla^k z\|_K \right\},$$

where the local residual terms ρ_K can be evaluated numerically, whereas the weights ω_K that contain the dual solution may be replaced by some numerical approximation.

Our experience is that already the first step of this estimation, using the triangle inequality, leads to over-estimation, as it suppresses any cancellation effect that occurs in the error representation (8) through the summation of the cell terms. For this reason, we do *not* use this error estimator in this work but try to evaluate the original error representation (8) as accurately as possible. However, this cannot be done directly, since the (unknown) dual solution z is involved, which must be approximated numerically. For that, we linearize the dual problem (7) by replacing the (unknown) exact solution u in $L(u, u_h; w, z)$ by the numerical solution u_h, yielding a perturbed dual problem

$$L(u_h, u_h; w, \tilde{z}) = a'[u_h](w, \tilde{z}) = J(w), \qquad \forall w \in V. \tag{12}$$

The discretization of this problem reads as follows: Find $\tilde{z}_h \in \tilde{V}_h$ such that

$$a'[u_h](w, \tilde{z}_h) = J(w), \qquad \forall w \in \tilde{V}_h, \tag{13}$$

with a discrete function space $\tilde{V}_h$ for the dual problem that will be specified in Remark 3.2 below. Replacing the exact dual solution z in (8) by this numerical approximation $\tilde{z}_h$ now leads to the following approximate error representation η:

$$J(e) \approx \sum_{K \in \mathbb{T}_h} \eta_K =: \eta, \tag{14}$$

with

$$\eta_K = (R(u_h), \tilde{z}_h - z_h)_K + (r(u_h), \tilde{z}_h - z_h)_{\partial K}. \tag{15}$$

Remark 3.1. *Note, that the approximate error representation* (14) *is not a strict upper bound to the true error but an approximation only.*

Remark 3.2. *As we calculate the dual solution on the same mesh as the primal solution, the functions in the discrete space $\tilde{V}_h$ of the dual problem must be of higher degree than the discrete functions used to approximate the primal problem; otherwise we would have $\eta = 0 \not\approx J(e)$.*

Remark 3.3. *The dual problem transports information along the characteristics of the primal problem in the opposite direction. As a simple example, the dual solution to the linear convection equation* $\nabla \cdot (\beta u) = 0$ *and the target functional*

$$J(u) = \int_{\partial\Omega^+} \beta \cdot n \, u\psi \, ds \tag{16}$$

is the solution to the problem $-\beta \cdot \nabla z = 0$ *with boundary condition* $z = \psi$ *on* $\partial\Omega^+$*. The dual solution traces back the domain of influence corresponding to the target functional, in the sense that the solution and the local residuals at all points within the support of the dual solution may affect the value and the error, respectively, of the target functional. The residuals in* (15) *are multiplied by weights including the dual solution. These weights describe* quantitatively *to what extent the residuals contribute to the error in the target functional.*

4. Numerical results

The local cell terms $|\eta_K|$, with η_K as in (15), can be evaluated and used as cell-wise indicators for adaptive mesh refinement. As these indicators include the local residuals multiplied by weights, they will be referred to as *weighted indicators* in the following. An adaptive algorithm that equilibrates these indicators by refining and coarsening the mesh leads to meshes that are tailored to the cost efficient computation of the quantity of interest $J(u)$.

In this section we show some numerical results, using the deal.II FE library, see [2] and [1], to illustrate the approach of a posteriori error estimation and to compare the efficiency of the adaptive mesh refinement using the weighted indicators with that using traditional error indicators, i.e. just the local residuals $\eta_K^{trad} = \|R(u_h)\|_K + \|r(u_h)\|_{\partial K}$. As the quality of the approximate error representation (14) depends on how good the dual solution is approximated numerically, we show examples including smooth as well as very 'rough' dual solutions. The test cases and target functionals are chosen sufficiently simple such that it is still possible to recognize the structure of the dual solution.

4.1. Linear advection equation

We consider the linear advection equation $\nabla \cdot (\beta u) = 0$ on $[0,2] \times [0,1] \subset \mathbb{R}^2$, with $\beta = \tilde{\beta}/|\tilde{\beta}|$, $\tilde{\beta}(x,y) = (y, 1-x)$ for $x < 1$, $\tilde{\beta}(x,y) = (2-y, x-1)$ elsewhere, and with prescribed boundary values on the inflow boundary ($u(x,0) = 1$ for $\frac{1}{8} < x < \frac{3}{4}$ and zero boundary values elsewhere). The vector field and the primal solution is displayed in Fig. 1. The two jumps of the discontinuous boundary function are transported along the characteristic directions given by the vector field. Suppose that we are interested in the values of the solution on the part $\frac{1}{4} < y < 1$ of the right outflow boundary. Let us therefore take $J(u) = \int_{\partial\Omega^+} u\psi \, ds$ as target functional, where ψ is chosen to be very smooth, $\psi(2,y) = \exp\left((\frac{3}{8})^{-2} - ((y - \frac{5}{8})^2 - \frac{3}{8})^{-2}\right)$ for $\frac{1}{4} < y < 1$, and 0 elsewhere. The exact value of this target functional is $J(u) = 0.1928\ldots$. As the weighting function ψ is smooth also the corresponding

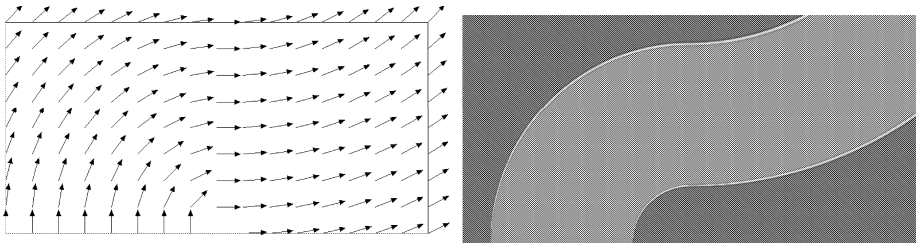


FIGURE 1. Vector field β and primal solution

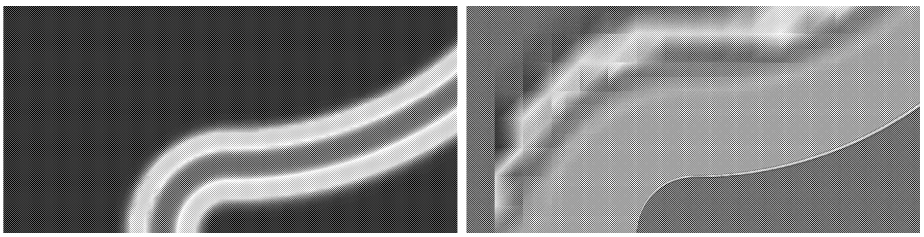


FIGURE 2. Dual solution (left) and numerical primal solution (right).

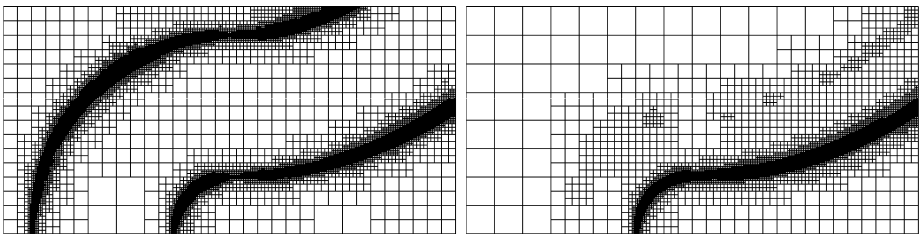


FIGURE 3. Mesh refined by the traditional and weighted indicators.

dual solution (see Fig. 2, left) is smooth. Fig. 2, right, shows the numerical solution (DG(1)) on the adaptively refined mesh using weighted indicators (see right mesh of Fig. 3). The refinement takes place at the position of only one of the jumps. The second jump is not resolved at all as it does not belong to the domain of influence of the target functional and the residuals in the neighborhood of this jump do *not* contribute to the error in the target functional. Comparing the two meshes in Fig. 3, it is obvious that the mesh on the right is more efficient for evaluating the value of the target functional than the mesh refined with traditional indicators. Table 1 shows the results of the adaptive refinement using the weighted indicators (15) with $\tilde{z}_h$ being a DG(2) numerical dual solution. Note, that the exact error $J(e)$ in the target functional and the approximate error representation η almost coincide and the efficiency index is remarkably close to 1.

# cells	$J(e)$	η	$\eta/J(e)$
128	8.70e-03	8.29e-03	0.953
194	1.55e-03	1.48e-03	0.957
506	2.54e-04	2.66e-04	1.048
1286	3.12e-05	3.25e-05	1.040
3065	1.48e-06	1.37e-06	0.928
7451	5.56e-07	5.57e-07	1.003
16922	2.82e-07	2.84e-07	1.007

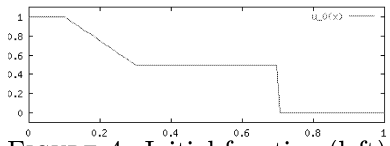
TABLE 1. DG(1) solution on adaptively refined grids using the weighted indicators including a DG(2) numerical dual solution.

Remark 4.1. *Solving the dual problem with elements of higher order than the primal problem is of course too much effort, since solving, e.g., the dual problem with DG(2) needs about 3 times the time of solving the primal problem with DG(1). The situation is different for nonlinear problems where the dual problem is still linear. Then the dual problem can be solved in relatively short time compared to the nonlinear iteration solution process of the primal problem.*

4.2. Scalar nonlinear problem

Let us consider the inviscid 1D Burgers equation in the (x,t) plane, that is represented in the form (1) by choosing $F(u) = (u, \frac{1}{2}u^2)$ and $\nabla = (\partial_t, \partial_x)^T$, such that $\nabla \cdot F(u) = u_t + uu_x = 0$. With prescribed initial function as in Fig. 4, the solution is as shown in Fig. 5, left. The slope of the linear part of the solution increases until a second shock develops and both shocks move to the right as time increases.

$$u_0(x) = \begin{cases} 1 & \text{for } x < 0.1, \\ -2.5x + 1.25 & \text{for } 0.1 < x < 0.3, \\ 0.5 & \text{for } 0.3 < x < 0.7, \\ 0 & \text{for } x > 0.7. \end{cases}$$



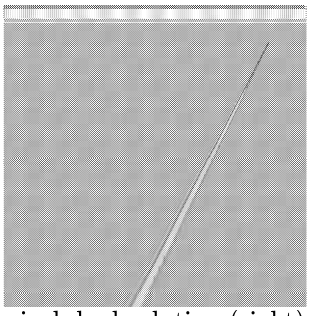


FIGURE 4. Initial function (left) and numerical dual solution (right).

Consider the point $x = 0.875$ at time $t = 0.875$ which is placed between the two shocks just a little time before they merge. Now the task might be to decide numerically whether one, two or none of the shocks already crossed this point at this specific time and calculate the value of the solution at this point to best accuracy, i.e. we take as target functional the point evaluation at the point of interest, $J(u) = u(x = 0.875, t = 0.875)$, for which the exact value is $J(u) = 0.5$.

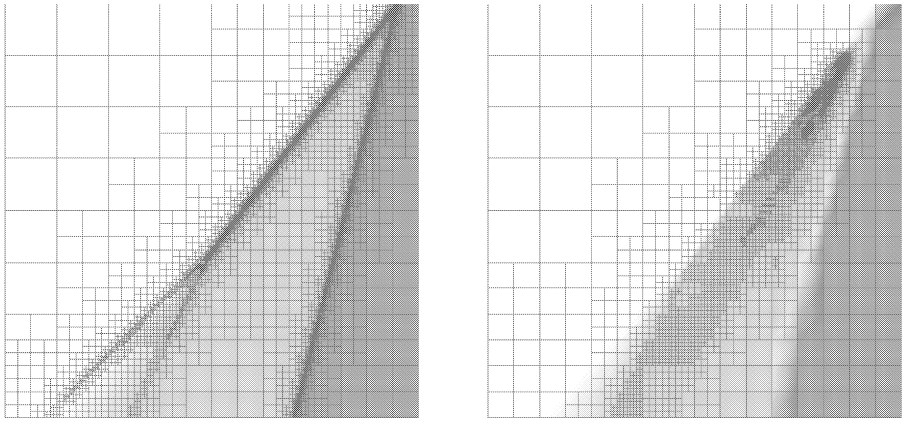


FIGURE 5. Primal solution on meshes refined by traditional indicator (left) and weighted indicator (right) with horizontal x-axis and vertical t-axis.

# cells	$J(e)$
1024	9.62e-02
1012	6.10e-02
2086	-9.27e-03
4390	-5.64e-04
8527	-1.58e-05
16447	1.35e-06

# cells	$J(e)$	η	$\eta/J(e)$
1024	9.62e-02	8.46e-03	0.09
535	4.57e-02	4.33e-02	0.95
709	-1.03e-02	-9.59e-03	0.93
1312	-5.58e-04	-6.14e-04	1.10
2638	5.23e-06	5.30e-06	1.01
5755	3.96e-08	3.05e-08	0.77

FIGURE 6. $J(e)$ for adaptation using traditional and weighted indicator.

Fig. 5 shows the solutions on meshes refined by the traditional and the weighted indicators, respectively. One recognizes that in the mesh on the right hand side of Fig. 5 there is almost no refinement at the position of the two shocks. Both shocks and also the margin of the linear part of the solution are not well resolved. Most of the refinement takes place upstream of the point of interest, i.e. in the neighborhood of the support of the dual solution, see Fig. 4. The shocks in the neighborhood of that point are just sufficiently resolved such that the error at these shocks does not affect the solution at the point of interest. The tables in Fig. 6 show the number of cells and the exact errors $J(e)$ of the target functional for the traditional refinement and the refinement using the weighted indicators, respectively. We observe that the refinement using weighted indicators is much more cost-efficient for computing the quantity of interest. Furthermore notice that the quality of error estimation is quite good again, indicating that the error due to the linearization of the dual problem and the approximation of its rough solution is comparably small.

5. Conclusion

We have presented a new approach to solving conservation equations by an adaptive finite element method including a posteriori estimation of the error with respect to an arbitrary target functional. The mesh refinement is based on local weighted indicators derived from the error representation. These weighted indicators consist of the local residuals of the numerical solution multiplied by local weights derived from the solution of a dual problem which is adjusted to the quantity of interest. The weights depending on the dual solution describe quantitatively to what extent the residuals contribute to the error in the target functional.

The approximate error representation is found to be reliable for model problems and target functionals including point evaluations. Also the sign of the error is predicted correctly.

References

[1] W. Bangerth, R. Hartmann, and G. Kanschat. `deal.II` *Differential Equations Analysis Library, Technical Reference,* IWR, University of Heidelberg, Apr. 2001. Available from the URL: `http://gaia.iwr.uni-heidelberg.de/~deal/`.

[2] W. Bangerth and G. Kanschat. *Concepts for object-oriented finite element software – the `deal.II` library,* Preprint 99-43, SFB 359, University of Heidelberg, Oct. 1999.

[3] B. Cockburn, G. Karniadakis, and C.-W. Shu. *The development of discontinuous Galerkin methods,* In: B. Cockburn, G. Karniadakis, and C.-W. Shu, eds, *Discontinuous Galerkin Methods,* Lecture Notes in Computational Science and Engineering (Springer), **11** (1999), 3–50.

[4] P. Houston, R. Rannacher, and E. Süli. *A posteriori error analysis for stabilised finite element approximations of transport problems,* Comput. Meth. Appl. Mech. Engrg., **190(11-12)** (2000), 1483–1508.

[5] J. Jaffre, C. Johnson, and A. Szepessy. *Convergence of the discontinuous Galerkin finite element method for hyperbolic conservation laws,* Math. Models and Methods in Appl. Sciences, **5** (1995), 367–386.

[6] R. Rannacher. *A posteriori error estimation in least-squares stabilized finite element schemes,* Comput. Meth. Appl. Mech. Engrg., **166** (1998), 99–114.

[7] E. Süli. *A posteriori error analysis and adaptivity for finite element approximations of hyperbolic problems,* In D. Kröner et al, eds, *An introduction to recent developments in theory and numerics for conservation laws*, Lecture Notes in Comput. Sciences and Engrg (Springer), **5** (1999), 123–194.

Institute of Applied Mathematics,
University of Heidelberg,
Im Neuenheimer Feld 294,
D-69120 Heidelberg
E-mail address: `Ralf.Hartmann@iwr.uni-heidelberg.de`