

Decomposition of the multidimensional Euler equations into advection equations

M. Fey

Research Report No. 95-14
December 1995

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Decomposition of the multidimensional Euler equations into advection equations

M. Fey

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 95-14

December 1995

Abstract

Based on a genuine multi-dimensional numerical scheme, called Method of Transport, we derive a form of the compressible Euler-equations, capable of a linearization for any space dimension. This form allows a rigorous error analysis of the linearization error without the knowledge of the numerical method. The generated error can be eliminated by special correction terms in the linear equations. Hence, existing scalar high order methods can be used to solve the linear equations and obtain high order accuracy in space and time for the non-linear conservation law.

In our approach, the scalar version of the method of transport is used to solve the linear equations. This method is multi-dimensional and reduces the solution of the partial differential equation to an integration process. Convergence histories presented at the end of the paper show that the numerical results agree with the theoretical predictions.

Keywords: conservation laws, flux vector splitting, characteristic surfaces

Subject Classification: 65M06, 76M25, 76N15

1 Introduction

For the class of scalar conservation laws, the theory of convergence and stability is well established for a large number of numerical methods. This is also true in view of error estimates and convergence properties, i.e. the order of convergence. Even for multi-dimensional calculations, there are more and more attempts to design high order schemes [3, 10, 19].

For the class of systems of conservation laws, the situation is quite different. For most of the existing schemes, there are only heuristic arguments, that these methods are of the same order as their scalar counterparts. In the case of numerical methods for multi-dimensional systems, it is even worse. Here, Strang [17] showed that most of the methods are limited to second order accuracy.

To obtain high order method for systems, it is necessary to take into account all sources of errors. In several space dimensions, there is mainly the dimension splitting error that restricts the order to at most two. But even in one space dimension, the error due to the linearization of a non-linear system plays an important role. This kind of error is introduced, if the method relies on approximate Riemann solver, e.g. flux-difference or flux-vector splitting.

Based on the derivation of a multi-dimensional method, called Method of Transport (MoT), we introduce a new form of the compressible Euler equations that is capable of linearization, i.e. decoupling into a finite number of advection equations, independent of the space dimension.

Most of the attempts to derive a 'truly' multi-dimensional method in the sense of a mesh independent discretization, start with the scalar advection equation, e.g. [3, 11, 15]. The adaption to the non-linear system case than becomes more complicated, since it requires some kind of linearized version of the equations. This is in most cases not obvious without reintroducing the coordinate axis.

The MoT on the other hand, is originally derived for the non-linear Euler-equations with the intention to include the physical propagation directions of the linearized equations, e.g. the characteristic hyper-surface or Monge cone, into the numerical method. Although the first version (c.f. [5]) was not quite competitive to the splitting approach because of the complexity and the large amount of computational work per grid point, it provided some insight in the discretization of multi-dimensional conservation laws. In fact, it leads to the linearization of the multi-dimensional system that is missing in the other attempts.

We are able to decompose the non-linear system of conservation laws

into a finite number of scalar advection equations with variable coefficients. This allows us to compare the Taylor expansion of the exact solution at time $t + \Delta t$ with the expansion of the exact solution of the linearized equations, to obtain the splitting and truncation errors. It simplifies a rigorous error analysis and makes it visible.

A direct comparison of the linearized equations and the non-linear Euler equations shows that the local approximation error is $O(\Delta t^2)$, i.e. we obtain a first order approximation of the non-linear system independent of the spatial discretization. Because of the special structure of the error terms they can be included into the linear equations to eliminate the resulting approximation error. These correction terms, added to the scalar advection equation, do not change the character of the equations, nor is their influence limited to a second order correction. It has been verified that these correction terms exist for the Euler equations at least up to third order accuracy.

In this paper we first recall the process of linearization in one space dimension including the flux-difference and flux-vector splitting methods [9, 16]. Then, using the ideas of the MoT we derive a simple decomposition of the multi-dimensional Euler equations into a set of linear advection equations. To obtain the linearization error in the smooth part of the solution we compare the Taylor expansions of both solutions. We will explain this procedure for the conservation of mass in 1-D only and give the results for the 1-D and the 2-D case.

We then briefly introduce a numerical algorithm to solve these equations efficiently and to high order of accuracy. In some numerical experiments at the end we verify the theoretical results. Convergence histories for smooth solutions in one and two space dimensions illustrate the influence of the second order correction terms. A solution of a Mach 10 flow indicates the robustness of the method even for strong shocks.

2 Linearisation of the Equations in 1-D

The one-dimensional compressible Euler equations in conservation form can be written as

$$\frac{\partial}{\partial t} \mathbf{U} + \frac{\partial}{\partial x} \mathbf{F}(\mathbf{U}) = 0 \quad (1)$$

where \mathbf{U} is the state vector of the conserved quantities and $\mathbf{F}(\mathbf{U})$ is the flux given by

$$\mathbf{U} = \begin{pmatrix} \rho \\ m \\ E \end{pmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{pmatrix} m \\ \rho u^2 + p \\ u(E + p) \end{pmatrix}.$$

Here, ρ is the mass density, $m = \rho u$ is the momentum, E is the total energy, u is the velocity and p is the pressure related to \mathbf{U} by the equation of state

$$p = (\gamma - 1)\left(E - \rho \frac{u^2}{2}\right).$$

The ration of the specific heat capacities γ takes the value 1.4 for air.

The simplest approach to solve the non-linear system is to use the quasi-linear form of (1), i.e.

$$\mathbf{U}_t + \frac{\partial \mathbf{F}}{\partial \mathbf{U}} \mathbf{U}_x =: \mathbf{U}_t + \mathbf{A} \mathbf{U}_x = 0.$$

Since equation (1) is hyperbolic, the matrix \mathbf{A} has only real eigenvalues and a full set of eigenvectors. Thus, the Jacobian matrix \mathbf{A} can be diagonalized by the matrix \mathbf{R} of right eigenvectors. We have the relation

$$\mathbf{A} = \mathbf{R} \mathbf{\Lambda} \mathbf{R}^{-1} \quad \text{or} \quad \mathbf{\Lambda} = \mathbf{R}^{-1} \mathbf{A} \mathbf{R},$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \lambda_3) = \text{diag}(u + c, u, u - c)$. Freezing the matrix \mathbf{A} locally, the transformation $\mathbf{W} = \mathbf{R}^{-1} \mathbf{U}$ leads to an approximation of (1), that decouples the equations, i.e.

$$\mathbf{R}^{-1} \mathbf{U}_t + \mathbf{\Lambda} \mathbf{R}^{-1} \mathbf{U}_x \approx \mathbf{W}_t + \mathbf{\Lambda} \mathbf{W}_x = 0.$$

This is called the local characteristic approach (for further references, see [19]).

A less obvious linearization of the equations is used in the flux-vector splitting in Steger and Warming [16] and in the flux-difference splitting introduced by Roe [14]. In both approaches the homogeneity of the Euler-equations, i.e. the flux can be written as

$$\mathbf{F}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} \mathbf{U} =: \mathbf{A} \mathbf{U}, \quad (2)$$

for any state vector \mathbf{U} , is used. We can write (2) as

$$\mathbf{F}(\mathbf{U}) = \mathbf{R} \mathbf{\Lambda} \mathbf{R}^{-1} \mathbf{U} = \sum_{i=1}^3 (\alpha_i \mathbf{r}_i) \lambda_i, \quad (3)$$

where $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ is the matrix of right eigenvectors of \mathbf{A} . The vector $(\alpha_1, \alpha_2, \alpha_3)^T := \mathbf{R}^{-1} \mathbf{U}$ and c is the speed of sound, given by $c^2 = \gamma p / \rho$. Using (3) and the fact that

$$\mathbf{U} = \mathbf{I} \mathbf{U} = \mathbf{R} \mathbf{R}^{-1} \mathbf{U} = \sum_{i=1}^3 (\alpha_i \mathbf{r}_i) \quad (4)$$

can be decomposed into the same vectors $(\alpha_i \mathbf{r}_i)$ as the flux, (1) becomes

$$\sum_{i=1}^3 \left(\frac{\partial}{\partial t} (\alpha_i \mathbf{r}_i) + \frac{\partial}{\partial x} ((\alpha_i \mathbf{r}_i) \lambda_i) \right) = 0. \quad (5)$$

The non-linear system (5) can locally be approximated by three linear systems or a total of nine scalar advection equations of the form

$$\frac{\partial}{\partial t} \omega + \frac{\partial}{\partial x} (a(x) \omega) = 0, \quad (6)$$

where ω is one of the components of $\alpha_i \mathbf{r}_i$, $i = 1, 2, 3$ and a is the corresponding characteristic speed λ_i , $i = 1, 2, 3$ which in this process becomes a function of space only. The resulting numerical scheme is consistent since the sum of all equations in (6) gives (5). After solving the decoupled scalar equations for a small time step Δt an approximation of the solution of the non-linear system is obtained by adding up the linear solutions. Iteration of this propagation step with updated values of \mathbf{r}_i and λ_i yields the numerical scheme. This can be interpreted as a Heugens principal for short times, i.e. interaction between different 'waves' are neglected. The non-linear coupling takes place during the averaging process in the finite volume discretization.

3 Linearization of the 2-D Euler-equations

Unfortunately, the above approach is not possible in several space dimensions. In the 2-D case, the equations have the form

$$\frac{\partial}{\partial t} \mathbf{U} + \frac{\partial}{\partial x} \mathbf{F}_1(\mathbf{U}) + \frac{\partial}{\partial y} \mathbf{F}_2(\mathbf{U}) = 0, \quad (7)$$

with

$$\mathbf{U} = \begin{pmatrix} \rho \\ m \\ n \\ E \end{pmatrix}, \quad \mathbf{F}_1 = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(E + p) \end{pmatrix}, \quad \mathbf{F}_2 = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(E + p) \end{pmatrix}.$$

Here, $\mathbf{m} = (m, n)^T$ is the momentum and $\mathbf{u} = (u, v)^T = (m/\rho, n/\rho)^T$ is the velocity. The pressure p is given by

$$p = (\gamma - 1) \left(E - \rho \frac{u^2 + v^2}{2} \right).$$

The equations are still hyperbolic and a linearization of the form

$$\mathbf{U}_t + \frac{\partial \mathbf{F}_1}{\partial \mathbf{U}} \mathbf{U}_x + \frac{\partial \mathbf{F}_2}{\partial \mathbf{U}} \mathbf{U}_y = 0$$

is still possible, but the Jacobian matrices of \mathbf{F}_1 and \mathbf{F}_2 cannot be diagonalized simultaneously, i.e. there is no matrix \mathbf{R} . Hyperbolicity allows the diagonalization of each Jacobian matrix and any linear combination of both. Thus a number of methods use the one-dimensional decomposition in (5) in each space direction. Let $\Lambda = \mathbf{R}^{-1} \mathbf{A} \mathbf{R}$ be the decomposition of the flux in x -direction with eigenvalues $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_4)$ and $\tilde{\Lambda} = \mathbf{S}^{-1} \mathbf{B} \mathbf{S}$ the decomposition of the flux in y -direction with eigenvalues $\tilde{\Lambda} = \text{diag}(\mu_1, \dots, \mu_4)$. With $\mathbf{R} = (\mathbf{r}_1, \dots, \mathbf{r}_4)$, $\mathbf{R}^{-1} \mathbf{U} = (\alpha_1, \dots, \alpha_4)^T$, $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_4)$ and $\mathbf{S}^{-1} \mathbf{U} = (\beta_1, \dots, \beta_4)^T$ we can decompose the one-dimensional problems into

$$\begin{aligned} \sum_{i=1}^4 ((\alpha_i \mathbf{r}_i)_t + (\lambda_i (\alpha_i \mathbf{r}_i))_x) &= \mathbf{U}_t + \mathbf{F}_1(\mathbf{U})_x, \\ \sum_{i=1}^4 ((\beta_i \mathbf{s}_i)_t + (\mu_i (\beta_i \mathbf{s}_i))_y) &= \mathbf{U}_t + \mathbf{F}_2(\mathbf{U})_y. \end{aligned} \tag{8}$$

In this representation we get twice the amount of the conserved quantity \mathbf{U} . We need to modify the decomposition:

$$\begin{aligned} \frac{1}{2} \left(\sum_{i=1}^4 (\alpha_i \mathbf{r}_i)_t + (2\lambda_i (\alpha_i \mathbf{r}_i))_x \right) &= \frac{1}{2} \mathbf{U}_t + \mathbf{F}_1(\mathbf{U})_x, \\ \frac{1}{2} \left(\sum_{i=1}^4 (\beta_i \mathbf{s}_i)_t + (2\mu_i (\beta_i \mathbf{s}_i))_y \right) &= \frac{1}{2} \mathbf{U}_t + \mathbf{F}_2(\mathbf{U})_y, \end{aligned}$$

so that the sum gives the original system of the Euler-equations. Hence, the propagation velocity of the linearized operators seems to be twice as large as in the 1-D case, which leads to the well known restriction of the CFL-number to 1/2 in the 2-D case and to 1/3 in the 3-D case, respectively. This analysis shows that we get a stability problem with the above approach (c.f. [12]).

On the other hand, an operator splitting approach, i.e. using the one-dimensional operators in (8) one after the other, leads to the well known restriction of the order to at most two [17].

The MoT provides the necessary information to find a linearization in the multi-dimensional case. We are not interested in a direction wise decomposition which directly leads to the dimensional splitting approach. Instead,

we are seeking for an approximation of the non-linear system by a set of linear but multi-dimensional advection equations equivalent to the 1-D case. Before we briefly introduce the idea of the MoT, we rewrite (7) in a more convenient form. Let

$$\mathbf{F}(\mathbf{U}) = (\mathbf{F}_1, \mathbf{F}_2) = \mathbf{U}\mathbf{u}^T + \begin{pmatrix} \mathbf{0}^T \\ \mathbf{I} \\ \mathbf{u}^T \end{pmatrix} p \quad (9)$$

be the $(N + 2) \times N$ matrix representing the multi-dimensional flux. N denotes the dimension of the space, \mathbf{I} is the $N \times N$ identity matrix and $\mathbf{0}$ is the N -dimensional vector of zeros. Then (7) becomes

$$\mathbf{U}_t + \nabla \cdot (\mathbf{F}(\mathbf{U})) = \mathbf{0} \quad (10)$$

where the divergence acts on the rows of \mathbf{F} .

The MoT described in [4] is a multi-dimensional generalization of the flux-vector decomposition in (3). The vectors $\alpha_i \mathbf{r}_i$ are moving with the characteristic speeds λ_i . In several space dimensions an equivalent quantity propagates along the characteristic surfaces called Monge cone. The derivation of the MoT shows that three quantities are necessary and sufficient to obtain a consistent method. We define

$$\mathbf{R}_1(\mathbf{U}) := \frac{1}{\gamma} \begin{pmatrix} \rho \\ \rho \mathbf{u} \\ \rho H \end{pmatrix}, \quad \mathbf{R}_2(\mathbf{U}) := \frac{\gamma - 1}{\gamma} \begin{pmatrix} \rho \\ \rho \mathbf{u} \\ \rho \mathbf{u}^2 / 2 \end{pmatrix}, \quad \mathbf{L}(\mathbf{U}) := \frac{\rho c}{\gamma} \begin{pmatrix} \mathbf{0}^T \\ \mathbf{I} \\ \mathbf{u}^T \end{pmatrix}.$$

Here $H = (E + p)/\rho$ denotes the total enthalpy and c is the speed of sound. Notice the structure of \mathbf{L} , which is similar to the part in front of the pressure in the multi-dimensional flux (9). In the 1-D case, \mathbf{R}_1 , \mathbf{R}_2 and \mathbf{L} are related to the vectors $\alpha_i \mathbf{r}_i$ as follows:

$$\mathbf{R}_1 = \alpha_1 \mathbf{r}_1 + \alpha_3 \mathbf{r}_3, \quad \mathbf{R}_2 = \alpha_2 \mathbf{r}_2, \quad \mathbf{L} = \alpha_1 \mathbf{r}_1 - \alpha_3 \mathbf{r}_3.$$

The propagation of these quantities along the characteristic directions, i.e. the flow direction \mathbf{u} for \mathbf{R}_2 and the Monge cone for \mathbf{R}_1 and \mathbf{L} leads to a consistent numerical method. Notice, that this approach introduces infinitely many propagation directions into the numerical scheme.

It turns out that a finite number of directions is sufficient in terms of accuracy and consistency of the method. This approximation corresponds to the use of an integration rule for the surface integral over the Monge cone. It is not related to the projection onto some arbitrary coordinate lines as in the dimensional splitting case.

For given directions \mathbf{n}_i , $i = 1, \dots, k$ we want to decompose the state vector \mathbf{U} into

$$\mathbf{U} = \mathbf{R}_1 + \mathbf{R}_2 + \frac{1}{k} \sum_{i=1}^k \mathbf{L} \cdot \mathbf{n}_i, \quad (11)$$

which is equivalent to (4) in the 1-D case. Equation (11) leads to the following consistency constraints on the directions \mathbf{n}_i :

$$\sum_{i=1}^k \mathbf{n}_i = 0, \quad (12)$$

$$\sum_{i=1}^k n_{ij} n_{ik} = k \delta_{jk} \quad \text{or} \quad \sum_{i=1}^k \mathbf{n}_i \mathbf{n}_i^T = k \mathbf{I}, \quad (13)$$

These conditions lead to a consistent decomposition of the flux, too. We can rewrite the Euler equations (10) as

$$\begin{aligned} \mathbf{U}_t + \nabla \cdot (\mathbf{F}(\mathbf{U})) &= (\mathbf{R}_2)_t + \nabla \cdot (\mathbf{R}_2 \mathbf{u}^T) \\ &+ \frac{1}{4} \sum_{i=1}^4 \left((\mathbf{R}_1 + \mathbf{L} \cdot \mathbf{n}_i)_t + \nabla \cdot ((\mathbf{R}_1 + \mathbf{L} \cdot \mathbf{n}_i) (\mathbf{u} + \mathbf{n}_i c)^T) \right) = 0. \end{aligned}$$

Reordering of terms and the use of (11) shows that

$$\begin{aligned} \mathbf{U}_t &= \left(\mathbf{R}_2 + \mathbf{R}_1 + \frac{1}{4} \sum_{i=1}^4 \mathbf{L} \cdot \mathbf{n}_i \right)_t, \\ \nabla \cdot (\mathbf{U} \mathbf{u}^T) &= \nabla \cdot \left(\left(\mathbf{R}_2 + \mathbf{R}_1 + \frac{1}{4} \sum_{i=1}^4 \mathbf{L} \cdot \mathbf{n}_i \right) \mathbf{u}^T \right). \end{aligned}$$

The remaining linear terms in \mathbf{n}_i vanish because of (12). For the quadratic term, we get

$$\frac{1}{4} c \mathbf{L} \sum_{i=1}^4 \mathbf{n}_i \mathbf{n}_i^T = \frac{1}{4} 4 c \mathbf{L} = \begin{pmatrix} \mathbf{0}^T \\ \mathbf{I} \\ \mathbf{u}^T \end{pmatrix} p$$

because of (13) and the definition of \mathbf{L} . This proves that the decomposition is valid and that the linearization

$$\begin{aligned} (\mathbf{R}_2)_t + \nabla \cdot (\mathbf{R}_2 \mathbf{u}^T) &= 0 \\ \frac{1}{4} \left\{ (\mathbf{R}_1 + \mathbf{L} \cdot \mathbf{n}_i)_t + \nabla \cdot ((\mathbf{R}_1 + \mathbf{L} \cdot \mathbf{n}_i) (\mathbf{u} + \mathbf{n}_i c)^T) \right\} &= 0, \quad i = 1, \dots, 4 \end{aligned} \quad (14)$$

is locally a first order approximation of (10). The solution of the non-linear system (10) can be replaced by solving 20 scalar advection equations of the form

$$h_t + \nabla \cdot (h \mathbf{a}) = 0 \quad (15)$$

where h is one of the components in \mathbf{R}_2 or $\mathbf{R}_1 + \mathbf{L} \cdot \mathbf{n}_i$ and \mathbf{a} is \mathbf{u} or $\mathbf{u} + \mathbf{n}_i c$, respectively.

In the following, we choose

$$\mathbf{n}_i \in \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\} \quad (16)$$

as directions, since for this choice the resulting method has a simple form and (12) and (13) are valid. A more general choice for the \mathbf{n}_i and the resulting consistency equations are shown in [6]. Notice that the choice of vectors along the coordinate axes in (16) will not lead to the dimensional splitting. In this case one use a different integration rule for the Monge cone. The propagation directions are in general not aligned with the coordinate directions.

4 Error Analysis

To obtain the error due to the linearization process we compare the Taylor expansions of the solution in (10) with the sum of solutions in (14) after time Δt . In principal there is no difference between the expansions in one and two dimensions and therefore we illustrate this process for the equation of mass conservation in the 1-D case only.

For the density after time Δt we get

$$\rho(x, t_0 + \Delta t) = \rho(x, t_0) + \Delta t \rho_t(x, t_0) + \frac{\Delta t^2}{2} \rho_{tt}(x, t_0) + O(\Delta t^3)$$

and with the Euler equations the time derivatives can be replaced by spatial derivatives, i.e.

$$\begin{aligned} \rho_t &= -(\rho u)_x, \\ \rho_{tt} &= (-m_x)_t = (\rho u^2 + p)_{xx} = \left(\rho \left(u^2 + \frac{c^2}{\gamma} \right) \right)_{xx}. \end{aligned}$$

To advance the solution in terms of the linearized equations we first decompose the density at time t_0 into

$$\rho(x, t_0) = \rho_1(x, t_0) + \rho_2(x, t_0) + \rho_3(x, t_0) \quad \text{with} \quad \rho_1 = \rho_3 = \frac{1}{2\gamma} \rho \quad \text{and} \quad \rho_2 = \frac{\gamma - 1}{\gamma} \rho.$$

For each part ρ_i we consider the advection equations

$$(\rho_{1/3})_t + ((u \pm c)\rho_{1/3})_x = 0 \quad \text{and} \quad (\rho_2)_t + (u\rho_2)_x = 0,$$

which leads to

$$\begin{aligned} (\rho_{1/3})_t &= -((u \pm c)\rho_{1/3})_x, \\ (\rho_{1/3})_{tt} &= ((u \pm c)((u \pm c)\rho_{1/3})_x)_x, \\ (\rho_2)_t &= -(u\rho_2)_x, \\ (\rho_2)_{tt} &= (u(u\rho_2)_x)_x. \end{aligned}$$

Since $(\rho_1 + \rho_2 + \rho_3)_t = -(u(\rho_1 + \rho_2 + \rho_3))_x = -(u\rho)_x$, the first order part is equal in both cases which proves consistency. In the second order term we get

$$\rho_{tt} - (\rho_1 + \rho_2 + \rho_3)_{tt} = -\frac{\rho}{2}(\gamma u u_x + c c_x) \neq 0 \quad (17)$$

which shows accuracy of first order only.

A closer look at the structure of the coefficients suggests the use of correction terms in the advection equations. The choice of

$$\hat{\rho}_{1/3} = \rho_{1/3} \pm k^\rho \quad \text{with} \quad k^\rho = -\frac{\Delta t \rho}{2c}(\gamma u u_x + c c_x)$$

instead of $\rho_{1/3}$ eliminates the error in (17). With

$$\rho(x, t_0 + \Delta t) - (\hat{\rho}_1(x, t_0 + \Delta t) + \rho_2(x, t_0 + \Delta t) + \hat{\rho}_3(x, t_0 + \Delta t)) = O(\Delta t^3),$$

we get a second order approximation in smooth regions of the solution. The same analysis can be done for the other components of the state vector in one dimension and also in the 2-D case. The resulting correction terms are

$$\begin{aligned} k^\rho &= -\frac{\Delta t \rho}{2c}(\gamma u u_x + c c_x), \\ k^m &= -\frac{\Delta t}{2}\rho((\gamma - 2)c u_x + u c_x) + u k^\rho, \\ k^E &= -\frac{\Delta t \rho c}{2\gamma(\gamma - 1)}(u u_x - c c_x) + u k^m - \frac{u^2}{2} k^\rho \end{aligned}$$

in the 1-D case and

$$\begin{aligned}
k_1^\rho &= -\frac{\Delta t \rho}{2 c \gamma} (\gamma(v u_y + u u_x) + c c_x), \\
k_2^\rho &= -\frac{\Delta t \rho}{2 c \gamma} (\gamma(u v_x + v v_y) + c c_y), \\
k_1^m &= -\frac{\Delta t \rho}{2 \gamma} (c(\gamma - 1)(v_y + u_x) + v c_y + u c_x - c u_x) + u k_1^\rho, \\
k_2^m &= \frac{\Delta t \rho}{2 \gamma} (c u_y) + u k_2^\rho, \\
k_1^n &= \frac{\Delta t \rho}{2 \gamma} (c v_x) + v k_1^\rho, \\
k_2^n &= -\frac{\Delta t \rho}{2 \gamma} (c(\gamma - 1)(u_x + v_y) + u c_x + v c_y - c v_y) + v k_2^\rho, \\
k_1^E &= -\frac{\Delta t \rho c}{2 \gamma(\gamma - 1)} (u u_x + v u_y - c c_x) + u k_1^m + v k_1^n - \frac{u^2 + v^2}{2} k_1^\rho, \\
k_2^E &= -\frac{\Delta t \rho c}{2 \gamma(\gamma - 1)} (v v_y + u v_x - c c_y) + u k_2^m + v k_2^n - \frac{u^2 + v^2}{2} k_2^\rho
\end{aligned}$$

in the 2-D case with the special choice (16) for the \mathbf{n}_i . In the linearization (14) the matrix \mathbf{L} has to be replaced by $\mathbf{L} + \mathbf{K}$ where $\mathbf{K} = (\mathbf{k}_1, \mathbf{k}_2)$.

5 Numerical solution of the scalar equations

From the previous considerations, we will focus on the linear advection equation in the form (15). Without loss of generality, we can restrict ourselves to two space dimensions. We will use the notation

$$h_t + (a(x, y) h)_x + (b(x, y) h)_y = 0 \quad (18)$$

where h is the unknown solution and $a, b : \mathbb{R}^2 \rightarrow \mathbb{R}$ are given real functions in space only.

In this section we will briefly describe the main idea of transport as proposed in [7]. The space is discretized with a Cartesian mesh of step size Δx and Δy in the x - and y -direction, respectively. The center of the finite volumes denoted by (x_i, y_j) are located at $x_i = i \Delta x$, $y_j = j \Delta y$ with the cell interfaces at $x_{i \pm 1/2} = (i \pm 1/2) \Delta x$ and $y_{j \pm 1/2} = (j \pm 1/2) \Delta y$. Using the average values of the function h over each cell as the dependent variables, i.e.

$$h_{i,j}^n = \frac{1}{|V_{i,j}|} \int_{V_{i,j}} h(x, y, n \Delta t) dx dy, \quad (19)$$

with $V_{i,j} = [x_{i-1/2}, x_{i+1/2}] \times [y_{j-1/2}, y_{j+1/2}]$ the domain of cell (i, j) . Integrating the conservation law (18) over space and time

$$\int_t^{t+\Delta t} \int_{V_{i,j}} h_t(x, y, t) dx dy dt + \int_t^{t+\Delta t} \int_{V_{i,j}} \operatorname{div} \left(h(x, y, t) \begin{pmatrix} a(x, y) \\ b(x, y) \end{pmatrix} \right) dx dy dt = 0$$

and using Gauss theorem and the definition in (19), we get

$$|V_{i,j}| (h_{i,j}^{n+1} - h_{i,j}^n) + \int_t^{t+\Delta t} \int_{\partial V_{i,j}} h(x, y, t) \begin{pmatrix} a(x, y) \\ b(x, y) \end{pmatrix} \mathbf{n} dO dt = 0,$$

where $|V_{i,j}|$ is the volume and $\partial V_{i,j}$ the boundary of the domain $V_{i,j}$. For the Cartesian grid we can write

$$\begin{aligned} |V_{i,j}| (h_{i,j}^{n+1} - h_{i,j}^n) + \int_t^{t+\Delta t} \int_{y_{j-1/2}}^{y_{j+1/2}} f(h, x_{i+1/2}, y) - f(h, x_{i-1/2}, y) dy dt + \\ \int_t^{t+\Delta t} \int_{x_{i-1/2}}^{x_{i+1/2}} g(h, x, y_{j+1/2}) - g(h, x, y_{j-1/2}) dx dt = 0 \end{aligned} \quad (20)$$

with the flux $f(h, x, y) = h(x, y, t) a(x, y)$ and $g(h, x, y) = h(x, y, t) b(x, y)$ in x - and y -direction, respectively.

Since the function h is not known for a time larger than t the second integral can only be computed approximately. The simplest choice is to keep h , a and b constant in each cell. With the assumption that $a, b > 0$, we obtain the following numerical method

$$\begin{aligned} |V_{i,j}| (h_{i,j}^{n+1} - h_{i,j}^n) + \Delta t (\Delta y (f(h_{i,j}^n, x_i, y_j) - f(h_{i-1,j}^n, x_{i-1}, y_j)) + \\ \Delta x (g(h_{i,j}^n, x_i, y_j) - g(h_{i,j-1}^n, x_i, y_{j-1}))) = 0. \end{aligned} \quad (21)$$

in conservation form. In the special case of the linear equation (18) the functions a and b are known everywhere and can be evaluated at the cell interfaces. In view of solving a non-linear system of equations the method (21) is more appropriate since the velocities a and b are functions of the solution. The approach (21) introduces the cell normals into the numerical method as shown in Figure 1. This leads to a strong influence of the underlying grid on the numerical solution. It is the aim of this paper to replace

these fluxes in normal direction by a procedure more related to the 'physics' of the problem.

The theory of characteristics use a different property of (18) to advance the solution in time. Along the integral curves $(x(t), y(t))^T$ given by

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} a(x, y, t) \\ b(x, y, t) \end{pmatrix} =: \mathbf{u} \quad \text{with} \quad \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \quad (22)$$

the solution $h(x(t), y(t), t)$ of (18) reduces to the ODE

$$\frac{d}{dt}h(x(t), y(t), t) = -h(x(t), y(t), t) \left(\frac{\partial}{\partial x}a(x(t), y(t)) + \frac{\partial}{\partial y}b(x(t), y(t)) \right) \quad (23)$$

with initial conditions $h(x_0, y_0, 0) = h_0(x_0, y_0)$.

Independent of the space dimension there is only one propagation direction for each point. Therefore, in the scalar case this transformation always reduces the problem to the integration of an ODE. For the linear equation (18) and smooth functions a and b , (22) can always be integrated numerically and the solution of h at any time t can be computed in one step. This idea can also be generalized to non-linear scalar equations (c.f. [1]). In the case of a non-linear system, the equations no longer reduce to an ODE since the characteristic of different families interact. The treatment of occurring shocks is more difficult.

To achieve a mesh-independent solver both ideas are combined. The finite volume part is taken to capture shocks and the characteristic part is taken to use the physical propagation directions instead of coordinate axes imposed by the grid. The characteristic curves (22) are used to trace the transport of quantities, in the case of (18) the value of h . This gives a mapping between the domain of influence (Figure 2) at the new time step and the cell itself as the domain of dependence.

We can describe this behavior analytically using Dirac's delta function. Let h^u be an approximation of $h(x, y, t + \Delta t)$ using (23). For small time steps Δt we can replace the integration in (22) by a simple Euler step and get

$$h^u(\mathbf{x}, t, \Delta t) = \int_{\mathbb{R}^N} h(\mathbf{y}, t) \delta(\mathbf{x} - (\mathbf{y} + \Delta t \mathbf{u}(\mathbf{y}, t))) d\mathbf{y} \quad (24)$$

up to first order. The δ -function "searches" backward for the point \mathbf{y} that determines the value of h at $(\mathbf{x}, t + \Delta t)$. For simplicity, we drop the indices i and j since all actions are the same on each cell. We define $\Omega_0 = V_{i,j}$, the

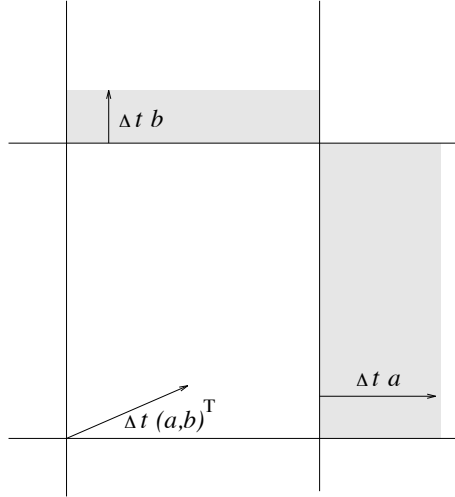


Figure 1: Flux in dimensional splitting.

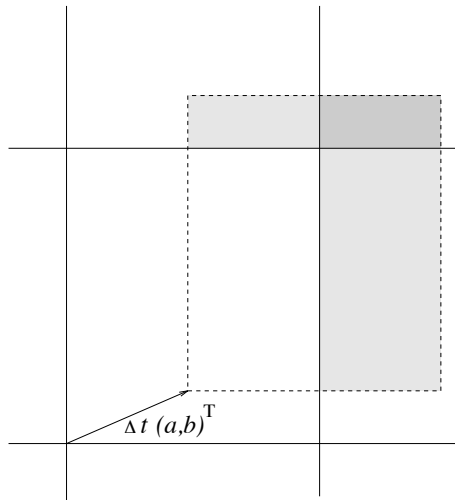


Figure 2: Movement of all points with characteristic speed.

domain of the center cell, and Ω_i , $i = 1, \dots, 8$, the neighboring cells. If we now restrict the domain of dependence to Ω_0 , the function

$$h_{\Omega_0}^u(\mathbf{x}, t, \Delta t) = \int_{\Omega_0} h(\mathbf{y}, t) \delta(\mathbf{x} - (\mathbf{y} + \Delta t \mathbf{u}(\mathbf{y}, t))) d\mathbf{y} \quad (25)$$

represents the distribution or motion of the quantity h in Ω_0 after time Δt . The transport from cell Ω_0 to any other domain Ω_i can be computed as

$$f_{\Omega_0\Omega_i}^u := \int_{\Omega_i} h_{\Omega_0}^u(\mathbf{x}, t, \Delta t) d\mathbf{x}. \quad (26)$$

The dashed line in Figure 2 shows the support of the function $h_{\Omega_0}^u$ for a constant velocity in the cell. The dark grey domain in the upper right corner indicates the contribution to the diagonal cell that means a contribution to a domain which has no finite boundary with the cell.

In a finite volume discretization, the update of the mean value in a cell at time $t + \Delta t$ can be done by adding all in- and outgoing fluxes:

$$h_{\Omega_0}^{n+1} = h_{\Omega_0}^n - \frac{1}{|\Omega_0|} \sum_{j=1}^k (f_{\Omega_0\Omega_j}^u - f_{\Omega_j\Omega_0}^u) = \frac{1}{|\Omega_0|} \sum_{j=0}^k f_{\Omega_j\Omega_0}^u. \quad (27)$$

Taking into account only the nearest neighbors on a Cartesian mesh, the upper limit is $3^N - 1$ for N space dimensions.

The two approaches described in (21) and (22,23), respectively, differ in the calculation of the net difference between in- and outgoing quantity h . In (21) the flux across a cell interface during time Δt is approximated, i.e. how many 'mass' points propagates across the cell surface. The second approach (22,23) tries to detect where the 'mass' points, initially in one cell, travel during time Δt , possibly crossing several cell interfaces.

We emphasize that during this transport process the conservation property of the equation is still satisfied, because of

$$\begin{aligned} \int_{\mathbb{R}^N} h_{\Omega_0}^u(\mathbf{x}, t, \Delta t) d\mathbf{x} &= \int_{\mathbb{R}^N} \int_{\Omega_0} h(\mathbf{y}, t) \delta(\mathbf{x} - (\mathbf{y} + \Delta t \mathbf{u}(\mathbf{y}, t))) d\mathbf{y} d\mathbf{x} \\ &= \int_{\Omega_0} h(\mathbf{y}, t) \int_{\mathbb{R}^N} \delta(\mathbf{x} - (\mathbf{y} + \Delta t \mathbf{u}(\mathbf{y}, t))) d\mathbf{x} d\mathbf{y} \quad (28) \\ &= \int_{\Omega_0} h(\mathbf{y}, t) d\mathbf{y}, \end{aligned}$$

i.e. the integration of $h_{\Omega_0}^u$ gives the 'amount' of h within Ω_0 . Thus, we enforce the conservation property already on the level of 'wave' propagation, rather than on the level of fluxes or contributions.

In the report [7] the convergence of this type of scheme for scalar conservation laws is proven. A similar idea, in the context of kinetic theory can be found in [8]. The transport collapse operator described by K. W. Morton and P. N. Childs in [2] and the rotated Riemann solver proposed by R. LeVeque [10] as well as the multidimensional method introduced by P. Collela [3] lead in first order to the same result if applied to a scalar equation.

The approach described above can be extended naturally to higher order. A proper reconstruction of the function h (and a and b) from the average values h_{ij} and the integration of (22) and (23) with sufficient accuracy leads to the numerical method of desired order independent of the dimension of space.

6 Numerical Example

The first set of examples illustrates the influence of the correction terms on the accuracy of the solution. The central part of the code is the solution of the scalar advection equation of the form

$$u_t + (a(x)u)_x = 0$$

in 1-D. Table 1 shows the results for smooth initial values $u(x,0) = e^{-x^2}$ and $a(x) = -a \tan(x)$. Notice that the function $a(x)$ has a sign change in the region of interest $x \in [-2, 2]$. This is most like the case for the non-linear system. The symbol C_∞ denotes the maximum error and L_1 the integral error of the numerical solution. V1 and V2 are different limiter functions with only minor influence in the smooth regions of the solutions. Because of the sign change of a , the order of convergence in the maximum norm is only one. This can be verified analytically. The integral error shows the predicted second order convergence.

We use this method to solve the resulting linear equations in (14) for each component. As initial values we use the constant states $\mathbf{U}_L = (3/4, 1, 7/3)^T$ and $\mathbf{U}_R = (1, 1, 3)^T$ from a steady shock for $|x| > 1$. For the values of $|x| < 1$ we pick $u(x,0)$ such that $u(x,0) \in C^2$, i.e. the initial data is smooth enough. Integration is stopped before the formation of the shock. Table 2 shows the result for the Euler equations. Here, V1 denotes the calculation with correction terms and V2 without them. The estimated error is more than two orders of magnitude smaller for V1 than V2. The L_1 error for V1 shows a nice

second order convergence and even the maximum error decays faster than first order. For the steady shock, only one of the characteristic velocities change sign. Thus only one of the waves has the observed degeneracy of the maximum error as shown above. The other waves have second order convergence in the maximum norm, too. Without the correction terms, the solution is only of first order in any norm.

scalar equation								
n	C_∞ -error		L_1 -error		order			
	V1	V2	V1	V2	C_∞		L_1	
10	1.1e-1	7.0e-2	2.8e-2	1.7e-2	V1	V2	V1	V2
20	9.0e-2	8.4e-2	1.1e-2	9.3e-3	0.296	< 0	1.298	0.832
40	4.1e-2	3.8e-2	2.8e-3	3.0e-3	1.117	1.149	1.983	1.650
80	1.0e-2	1.0e-2	7.2e-4	7.4e-4	1.995	1.873	1.987	2.008
160	3.4e-3	3.3e-3	1.8e-4	1.8e-4	1.612	1.643	2.026	2.064
320	1.6e-3	1.4e-3	4.4e-5	4.4e-5	1.132	1.264	1.993	2.016
640	8.6e-4	6.4e-4	1.1e-5	1.1e-5	0.854	1.092	1.984	1.976
1280	4.0e-4	3.4e-4	2.8e-6	2.9e-6	1.093	0.939	1.991	1.921

Table 1: Convergence history for a scalar equation. V1 and V2 differ in the choice of the limiter function.

Euler equation								
n	C_∞ -error		L_1 -error		order			
	V1	V2	V1	V2	C_∞		L_1	
20	5.1e-3	1.1e-2	7.9e-4	1.4e-3	V1	V2	V1	V2
40	2.5e-3	3.8e-3	2.5e-4	5.4e-4	1.041	1.521	1.634	1.354
80	8.4e-4	1.7e-3	5.4e-5	2.6e-4	1.570	1.155	2.228	1.057
160	2.2e-4	9.5e-4	1.0e-5	1.3e-4	1.909	0.845	2.386	0.998
320	6.9e-5	4.9e-4	2.4e-6	6.6e-5	1.700	0.936	2.124	0.977
640	2.2e-5	2.5e-4	5.8e-7	3.3e-5	1.647	0.972	2.026	0.988
1280	7.1e-6	1.3e-4	1.5e-7	1.7e-5	1.622	0.986	1.989	0.992

Table 2: Convergence history for the Euler equations. V1 with, V2 without correction terms.

We get analogous results in the 2-D case. Table 3 shows the convergence history for the scalar equation

$$u_t + (-y u)_x + (x u)_y = 0.$$

The initial values are $u(x, y, 0) = \exp(-((x - 1/2)^2 + y^2))$.

scalar equation 2-D				
n	C_∞ -error	L_1 -error	order	
5	1.0e-1	1.5e-2	C_∞	L_1
10	3.8e-2	5.0e-3	1.450	1.571
20	7.6e-3	1.0e-3	2.308	2.267
40	1.8e-3	2.1e-4	2.100	2.285
80	4.2e-4	4.7e-5	2.086	2.175
160	1.0e-4	1.1e-5	2.018	2.074
320	2.6e-5	2.7e-6	2.012	2.031
640	6.3e-6	6.8e-7	2.007	2.012

Table 3: Convergence history for a scalar equation with smooth data in two space dimensions.

Euler equation 2-D				
n	C_∞ -error	L_1 -error	order	
18			C_∞	L_1
36	3.8e-2	5.0e-3		
72	2.6e-2	8.3e-4		
144	8.0e-3	1.8e-4	1.704	2.165
288	2.2e-3	2.4e-5	1.879	2.445
576	6.4e-4	6.8e-6	1.768	2.320
1152				

Table 4: Convergence history for the Euler equation with smooth data in two space dimensions.

The first 2-D example for the Euler-equations consists of smooth perturbations in the density, velocity and pressure. We use $\rho_1 = 1, \rho_2 = 1.1$ for the density perturbation, $u_1 = 0, u_2 = 0.1$ in the x -velocity, $v_1 = 0, v_2 = -0.1$ in the y -velocity and $p_1 = 1.0, p_2 = 0.9$ in the pressure. In the computational domain $[-2, 2]^2$ we used a radial symmetric function to connect the two values. The center of the perturbations is located at $(1/2, 1/2)^T, (1/2, -1/2)^T, (-1/2, -1/2)^T$ and $(-1/2, 1/2)^T$ for density, x -velocity, y -velocity and pressure, respectively. Table 4 shows the convergence results.

The last example addresses the robustness of the method. We compute the solution of a Richtmyer-Meshkov instability. A strong Mach 10 shock hits a sinusoidal shaped contact surface that generates vorticity. Figure 3 shows the results for different times. In the first picture the incident shock has already passed the contact surface. The numerical method in this ex-

ample is exactly the same as in the convergence tests. The limiter function used is a version of the monotized central-differences but it is extended beyond the TVD-region proposed by Sweby [18] to achieve better accuracy in smooth regions of the flow. Even the contact surfaces from the Mach stem of the reflected shocks are visible. The disadvantages are small oscillations behind the shocks.

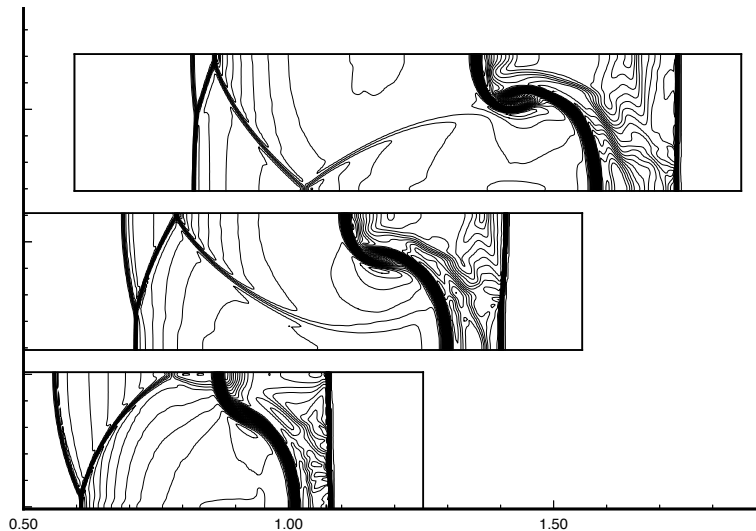


Figure 3: *Density contour lines at three different times for the Richtmyer-Meshkov instability with an incident Mach 10 shock.*

7 Conclusions

The multi-dimensional linearization of the Euler equations offers a simple way to obtain high order methods that are independent of the space dimensions. The solution of a non-linear system of equations is reduced to a finite number of linear problems. This reduction is independent of the numerical method and the discretization used. The numerical implementation needs to meet the requirement of high accuracy for the scalar case.

The derivation shows that this linearization can be achieved for any system of conservation laws as long as a decomposition of the state vector and the flux in terms of the MoT is possible. For the shallow water equation there exists such a decomposition as shown in [13] and thus all the results carry over.

References

- [1] Helmut Boeing. *Die numerische Konstruktion der exakten Lösung des skalaren Riemannproblems in zwei Raumdimensionen*. PhD thesis, RWTH Aachen, 1992.
- [2] P. N. Childs and K. W. Morton. Characteristic galerkin methods for scalar conservation laws in one dimension. *SIAM J. Numer. Anal.*, 27(3):533–594, 1990.
- [3] P. Collela. Multidimensional upwind methods for hyperbolic conservation laws. *J. of Comp. Phys.*, 87:171–200, 1990.
- [4] M. Fey. A truly multidimensional method to solve the euler- equations. Technical report, California Inst. of Technology, 1994.
- [5] M. Fey. The Method of Transport for solving the Euler- equations. Technical Report 95-15, Seminar für Angewandte Mathematik, ETH Zürich, 1995.
- [6] M. Fey, R. Jeltsch, and A.-T. Morel. Multidimensional schemes for nonlinear systems of hyperbolic conservation laws. Technical Report 95-11, Seminar für Angewandte Mathematik, ETH Zürich, 1995. to appear in "Proceedings of the 16th Biannual Conference on Numerical Analysis", University of Dundee, June 27-30, 1995, Ed. D. Griffiths, 1996.
- [7] M. Fey and A. Schroll. Monoton split- and unsplit methods for a single conservation law in two space dimensions. Technical Report 57, Inst. für Geometrie und Praktische Mathematik, RWTH Aachen, 1989.
- [8] W. N. G. Hitchon, D.J. Koch, and J. B. Adams. An efficient scheme for convection-dominated transport. *J. of Comp. Phys.*, 83:79–95, 1989.
- [9] Bram van Leer. Flux-vector splitting for the Euler equations. *Lecture Notes in Physics*, 170:507–512, 1982.
- [10] R. J. LeVeque. High resolution finit volume methods on arbitrary grids via wave propagation. *J. Comp. Phys.*, 78(1):36–63, 1988.
- [11] R. J. LeVeque. Simplified multidimensional flux limiter methods. In M.J. Baines and K. W. Morton, editors, *Numerical Methods for Fluid Dynamics*, volume 4. Oxford University Press, 1993.

- [12] R. J. LeVeque. High-resolution conservative Algorithms for Advection in Incompressible Flow. *SIAM J. Numer. Anal.*, to appear.
- [13] A.T. Morel and M. Fey. Multidimensional method of transport for the shallow water equations. submitted to Proceeding of the Fifth International Conference on Hyperbolic Problems, Stony Brook, 1994.
- [14] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comp. Physics*, 43:357–372, 1981.
- [15] P. L. Roe. Linear advection schemes on triangular meshes. Technical Report CoA Rep. No. 8720, Cranfield, 1987.
- [16] J. L. Steger and R. F. Warming. Flux vector splitting of the inviscid gasdynamic equations with application to finite difference methods. *J. Comp. Phys.*, 40(263-293), 1981.
- [17] G. Strang. On the constuction and comparison of difference schemes. *Journal of Numerical Analysis*, 5:506–517, 1968.
- [18] P. K. Sweby. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM J. Numer. Anal.*, 21:995–1011, 1984.
- [19] Helen Yee. Upwind and symmetric shock-capturing schemes. Technical Report TM 89464, NASA, 1987.

Research Reports

No.	Authors	Title
95-14	M. Fey	Decomposition of the multidimensional Euler equations into advection equations
95-13	M.D. Buhmann	Radial Functions on Compact Support
95-12	R. Jeltsch	Stability of time discretization, Hurwitz determinants and order stars
95-11	M. Fey, R. Jeltsch, A.-T. Morel	Multidimensional schemes for nonlinear systems of hyperbolic conservation laws
95-10	T. von Petersdorff, C. Schwab	Boundary Element Methods with Wavelets and Mesh Refinement
95-09	R. Sperb	Some complementary estimates in the Dead Core problem
95-08	T. von Petersdorff, C. Schwab	Fully discrete multiscale Galerkin BEM
95-07	R. Bodenmann	Summation by parts formula for noncentered finite differences
95-06	M.D. Buhmann	Neue und alte These über Wavelets
95-05	M. Fey, A.-T. Morel	Multidimensional method of transport for the shallow water equations
95-04	R. Bodenmann, H.J. Schroll	Compact difference methods applied to initial-boundary value problems for mixed systems
95-03	K. Nipp, D. Stoffer	Invariant manifolds and global error estimates of numerical integration schemes applied to stiff systems of singular perturbation type - Part II: Linear multistep methods
95-02	M.D. Buhmann, F. Derrien, A. Le Méhauté	Spectral Properties and Knot Removal for Interpolation by Pure Radial Sums
95-01	R. Jeltsch, R. Renault, J.H. Smit	An Accuracy Barrier for Stable Three-Time-Level Difference Schemes for Hyperbolic Equations
94-13	J. Waldvogel	Circuits in Power Electronics
94-12	A. Williams, K. Burrage	A parallel implementation of a deflation algorithm for systems of linear equations
94-11	N. Botta, R. Jeltsch	A numerical method for unsteady flows
94-10	M. Rezny	Parallel Implementation of ADVISE on the Intel Paragon
94-09	M.D. Buhmann, A. Le Méhauté	Knot removal with radial function interpolation