High-Order Space-Time Methods for Conservation Laws

H.T. Huynh
Glenn Research Center, Cleveland, Ohio
Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.

- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include creating custom thesauri, building customized databases, organizing and publishing research results.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at [http://www.sti.nasa.gov](http://www.sti.nasa.gov)
- E-mail your question to help@sti.nasa.gov
- Fax your question to the NASA STI Information Desk at 443–757–5803
- Phone the NASA STI Information Desk at 443–757–5802
- Write to:
  STI Information Desk
  NASA Center for AeroSpace Information
  7115 Standard Drive
  Hanover, MD 21076–1320
High-Order Space-Time Methods for Conservation Laws

H.T. Huynh
Glenn Research Center, Cleveland, Ohio
Acknowledgments

This work was supported by the Aeronautical Sciences Project of NASA. The author wishes to thank Drs. James DeBonis and Seth Spiegel for their thorough reviews and numerous valuable suggestions.

Level of Review: This material has been technically reviewed by technical management.

Available from

NASA Center for Aerospace Information
7115 Standard Drive
Hanover, MD 21076–1320

National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312

Available electronically at http://www.sti.nasa.gov
High-Order Space-Time Methods for Conservation Laws

H.T. Huynh
National Aeronautics and Space Administration
Glenn Research Center
Cleveland, Ohio 44135

Abstract. Current high-order methods such as discontinuous Galerkin and/or flux reconstruction can provide effective discretization for the spatial derivatives. Together with a time discretization, such methods result in either too small a time step size in the case of an explicit scheme or a very large system in the case of an implicit one. To tackle these problems, two new high-order space-time schemes for conservation laws are introduced: the first is explicit and the second, implicit. The explicit method here, also called the moment scheme, achieves a CFL (Courant-Friedrichs-Lewy) condition of 1 for the case of one-spatial dimension regardless of the degree of the polynomial approximation. (For standard explicit methods, if the spatial approximation is of degree $p$, then the time step sizes are typically proportional to $1/p^2$.) Fourier analyses for the one and two-dimensional cases are carried out. The property of super accuracy (or super convergence) is discussed. The implicit method is a simplified but optimal version of the discontinuous Galerkin scheme applied to time. It reduces to a collocation implicit Runge-Kutta (RK) method for ordinary differential equations (ODE) called Radau IIA. The explicit and implicit schemes are closely related since they employ the same intermediate time levels, and the former can serve as a key building block in an iterative procedure for the latter. A limiting technique for the piecewise linear scheme is also discussed. The technique can suppress oscillations near a discontinuity while preserving accuracy near extrema. Preliminary numerical results are shown.

1. Introduction

The discontinuous Galerkin (DG) method is currently among the most widely used high-order numerical methods for solving the compressible Navier-Stokes equations on unstructured meshes. It was introduced for the neutron transport equation by Reed and Hill (1973), analyzed by LaSaint and Raviart (1974) and developed and made popular for fluid dynamics equations by Cockburn, Shu, Bassi, Rebay, and others (see e.g., Bassi and Rebay 1997a,b, Cockburn, Karniadakis, and Shu 2000, Cockburn and Shu 2005, 2009, Shu 2012, and the references therein). Efficient DG schemes using the Lagrange polynomials for nodal points as basis functions known as nodal DG methods can be found in (Hesthaven and Warburton 2008). Alternative methods employing the differential form (as opposed to DG, which uses the integral form) have been proposed. Kopriva and Kolias (1996) pioneered this approach with the staggered-grid method on quadrilateral meshes. The method was extended to triangular meshes by Liu, Vinokur, and Wang (2006) and named spectral difference (SD). Another class of schemes called spectral volume (SV) was presented by Wang, Zhang, and Liu (2004).

Recently, an approach to high-order accuracy with the advantage of simplicity and economy called flux reconstruction (FR) was introduced in (Huynh 2007, 2009a). The approach employs the differential form and results in several new schemes with favorable properties. In addition, by recovering DG, SD, and SV, it unifies existing methods. Extensions of the FR method to 2D unstructured meshes were presented in (Wang and Gao 2009) and (Gao and Wang 2009) and was named lifting collocation penalty or LCP schemes. The involved authors later combined FR and LCP and called them correction procedure via reconstruction or CPR methods. Extension to 3D was presented in (Haga, Gao, and Wang 2010, 2011).
and (Wang, Gao, and Haga 2011). A mathematical foundation for FR was provided via energy
stability proofs: for an SD scheme identified via FR by Jameson (2010), for a one-parameter class of FR
schemes in 1D by Vincent, Castonguay, and Jameson (2011), for 2D triangular grids by Castonguay,
Vincent and Jameson (2012), for linear-advection-diffusion problem in 1D by Castonguay et al. (2013)
and for advection-diffusion in 2D, by Williams et al. (2011). Extension for a particularly simple FR
scheme called $g_2$ to the Navier-Stokes equations was presented in (Liang, Miyaji, and Zhang 2013).
Applications of FR/CPR methods to turbulent internal flows for turbomachinaries were elaborated in (Lu,
Yuan, and Dawes 2012 and Lu, Liu, and Dawes 2013). Interface elements dealing with non-conforming
polynomials together with p-adaptation for viscous flow simulations were presented in (Cagnone and
Nadarajah 2012, Cagnone, Vermeire, and Nadarajah 2013), and an implicit large eddy simulation (ILES)
solver was developed for FR/CPR schemes by Vermeire, Cagnone, and Nadarajah (2013).

The high-order methods discussed can provide effective discretization for the spatial derivatives.
Together with a time discretization, however, such methods result in either too small a time step size in
the case of an explicit scheme or a very large system in the case of an implicit one. Efforts to improve
time stepping have been proposed by several authors.

Concerning explicit schemes, a space-time DG method with a favorable CFL (Courant-Friedrichs-
Lewy) condition employing a staggered mesh was introduced in Lowrie et. al (1995). For the case of a
semi-discrete formulation, Warburton and Hagstrom (2008) presented a staggered-mesh scheme where
the solution is projected onto a mesh staggered from the original one and then projected back onto the
original mesh resulting in a solution ‘smoother’ than the original. They obtain a CFL condition
proportional to $1/p$ as opposed to $1/p^2$ where $p$ is the degree of the approximating polynomials.
However, multi-dimensional extensions of these staggered-mesh approaches are both complex and costly.
Additional efforts can be found under the heading “Taming the CFL number” of (Hesthaven and
Warburton 2008) and the references therein. In addition to these efforts, Dumbser and Munz (2005)
developed ADER-DG using the ADER finite-volume scheme (Arbitrary order using derivatives). Space-
time expansion DG (or STE-DG) as a modification of ADER-DG was introduced in (Lörcher, Gassner,
and Munz 2007 and Gassner, Lörcher, and Munz 2008), but these methods have highly restricted CFL
limits.

Researchers in the finite-volume community have devised several explicit methods with a CFL
condition of 1 for the case of one spatial dimension. In fact, in 1977, Van Leer presented a series of five
schemes for convection with this property. Among the five, scheme I is the least accurate but becomes the
most popular and, in the literature, is generally the scheme implied when the MUSCL approach is
referred. Concerning the more accurate schemes such as III and V, the problem is, as stated in (Van Leer
and Nomura 2005), “When trying to extend these schemes beyond advection, viz., to a nonlinear
hyperbolic system like the Euler equations, the first author ran into insuperable difficulties because the
exact shift operator no longer applies, and he abandoned the idea”.

For scheme III, which can be considered as a piecewise-linear space-time DG method for convection,
this difficulty was overcome in (Huynh 2006) where the exact shift operator was replaced by (a) an
integration by parts (Gauss theorem), (b) the observation that the space-time fluxes contain the
information provided by the shift operator, and (c) a successive procedure of updating the moments. The
resulting method is called the upwind moment scheme. The approach was further analyzed and applied to
nonlinear hyperbolic equations by Suzuki and Van Leer (2007). As briefly discussed in (Huynh 2007), the
moment scheme can be extended to arbitrary order. Such an extension was independently obtained by Lo
and was studied in combination with Van Leer’s recovery scheme for diffusion (2011). The description of
the extension here and that of Lo thus have considerable similarities (the differences will also be
discussed). Extensions for scheme V can be found in (Eymann and Roe 2013).

Concerning implicit schemes, it was shown by LaSaint and Raviart (1974) via a quadrature of left
Radau type (Hildebrand 1987) that discretizing an ODE via discontinuous Galerkin results in an implicit
Runge-Kutta (RK) method. In addition, the scheme is accurate to order $2p + 1$ if the solution is approximated by a polynomial of degree $p$ in time. Implicit space-time DG schemes for conservation laws have been studied by several authors (e.g., Bar-Yoseph and Elata 1990, Van der Vegt and Van der Ven 2002, Pesch and Van der Vegt 2008). The common difficulty is that the implicit system is large. Recent efforts include the large eddy simulation solver by Vermeire, Cagnone, and Nadarajah (2013).

Another aspect of time stepping methods concerns their performance for large scale calculations using massively parallel supercomputers or graphics processing units (GPU). Here, the DG or FR/CPR spatial discretization is highly parallelizable. It is well known that this ease of parallelization is retained with an explicit time stepping, e.g., RK, but not an implicit one.

In this paper, to tackle the time stepping problem, two new high-order space-time schemes for conservation laws are introduced: the first is explicit and the second, implicit. The explicit method, called the moment scheme, achieves a CFL condition of 1 for the case of one-spatial dimension regardless of the degree of approximation. Compared with standard approaches, the key difference is the built-in physics of advection. Such a CFL condition is a significant gain compared with a typical time step size proportional to $1/p^2$. For the case of multiple-dimensions, however, there is a reduction in the CFL limits. The current extension differs from that of Lo in (a) the case of two-spatial dimensions and its Fourier analysis are presented and (b) additional details, observations, and the relation with the implicit time stepping are discussed. It will be shown via Fourier analysis that for both one and two-dimensional cases, the moment schemes are accurate to order $2p + 1$ if the solution is approximated by polynomials of degree $p$.

The implicit method here is a simplified but optimal DG scheme applied to time. It employs the FR approach and simplification is made possible by the fact that time always moves forward (as opposed to space where waves can travel in all directions). The method reduces to an implicit RK scheme using the right Radau points as collocation points called the Radau IIA (Lambert 1991, Hairer, Norsett, and Wanner 1991, Hairer and Wanner 1993). It is also accurate to order $2p + 1$ if polynomial of degree $p$ is employed to approximate the solution. A key advantage of using the right Radau points (as opposed to the left Radau points employed by LaSaint and Raviart 1974) is that the method is simpler and involves fewer interpolations in the solution procedure. More importantly, the explicit and implicit schemes here are closely related since they use the same intermediate time levels, and the former can serve as a key building block in an iterative procedure for the latter.

The DG derivation here also differs from those in the literature: the standard DG formulation is highly algebraic, and the behavior of the solution is hidden in a system of equations; the current formulation is geometric, and the behavior of the solution can be observed.

In addition to the time stepping methods, since the numerical tests have discontinuities, a limiting technique that can suppress oscillations near discontinuities while preserving accuracy near extrema is discussed. Finally, preliminary numerical results are shown.

This paper is essentially self-contained and organized as follows. In Section 2, we review Van Leer’s scheme III except that our derivation is for arbitrary order. The moment schemes, which have a CFL condition of 1 and can be considered as explicit space-time DG methods, are presented in Section 3. Extension to the case of two spatial dimensions is discussed in Section 4. Fourier analyses are carried out in Section 5 for the one- as well as two-dimensional cases. DG time discretization in the FR/CPR framework is presented in Section 6. Section 7 deals with limiting. Preliminary numerical solutions are shown in Section 8. Finally, conclusions and discussions can be found in Section 9.

2. Exact Time Evolution and Projection

In the fourth installment of a celebrated series of five papers entitled “Towards the ultimate conservative difference scheme”, Van Leer (1977) introduced five schemes for advection; among them, the first three are piecewise linear and, the last two, piecewise parabolic. Concerning the piecewise linear
methods, scheme I is the least accurate but, due to its ease of extension to systems, becomes the most popular and is typically known as the MUSCL scheme (monotone upstream-centered schemes for conservation laws). As shown in (Huynh 2003), scheme II formulated for a staggered mesh results in the CE/SE scheme with $\varepsilon = 1/2$ (conservation element/solution element, Chang 1995). Scheme III is remarkable since it is third-order accurate in spite of being piecewise linear. The property of having higher than expected order of accuracy is called super accuracy or super convergence. The last two schemes are piecewise parabolic. Here, scheme IV leads to the standard PPM (piecewise parabolic method). Scheme V stores and updates both the cell averages and interface values. Schemes III and V have the same error, which is only $1/9$ that of scheme IV. As mentioned above, when trying to extend the more accurate schemes to the Euler equations, Van Leer “ran into insuperable difficulties”. However, scheme III plays a critical role in serving as a guide for the moment scheme (Huynh 2006), which can be extended to systems with relative ease. Below, we review Van Leer’s approach for scheme III, which amounts to a projection after an exact shift to account for time evolution. Our presentation of the scheme is for arbitrary order, not just piecewise linear.

Consider the scalar advection equation

$$u_t + au_x = 0 \tag{2.1}$$

with initial condition

$$u(x, 0) = u_0(x) \tag{2.2}$$

where $t$ is time, $x$ space, and $a$ the advection speed, a positive constant. By assuming that $u_0$ is periodic or of compact support, boundary conditions are trivial and are omitted. The exact solution at time $t$ is obtained by shifting the data curve to the right a distance $at$,

$$u_{\text{exact}}(x, t) = u_0(x - at). \tag{2.3}$$

**Legendre Polynomials.** To approximate the solution, we need some preparations. Let $k$ be a nonnegative integer. Denote by $P_k$ the space of polynomials of degree $\leq k$. As is standard in the finite-volume community, for the reference interval, we use $I = [-1/2, 1/2]$. Let the inner product of any two functions $v_1$ and $v_2$ on $I$ be defined by

$$(v_1, v_2) = \int_{-1/2}^{1/2} v_1(\xi) v_2(\xi) \, d\xi, \tag{2.4}$$

and the $L_2$ norm of a function $v$ by

$$\|v\| = \left(\int_{-1/2}^{1/2} (v(\xi))^2 \, d\xi\right)^{1/2}. \tag{2.5}$$

Let $L_k$ be the Legendre polynomial of degree $k$ defined by applying the Gram-Schmidt process to orthogonalize the monomials $1, \xi, \xi^2, \xi^3, \ldots$ together with the condition that $L_k(1/2) = 1$. The first few Legendre polynomials on $I = [-1/2, 1/2]$ are

$$L_0(\xi) = 1, \quad L_1(\xi) = 2\xi, \quad L_2(\xi) = 6\xi^2 - 1/2, \quad L_3(\xi) = 20\xi^3 - 3\xi, \quad L_4(\xi) = 70\xi^4 - 15\xi^2 + 3/8. \tag{2.6}$$
Their plots are shown in Fig. 2.1. Note that compared with the standard Legendre polynomials $P_k$ defined on $[-1, 1]$, for $\xi$ in $[-1/2, 1/2]$, $L_k(\xi) = P_k(2\xi)$.

![Fig. 2.1 Legendre Polynomials](image)

**Discretization.** Let the domain of calculation be divided into non-overlapping cells or elements denoted by $E_j = [x_{j-1/2}, x_{j+1/2}]$ of centers $x_j$ and widths $h_j$. For each $E_j$, let the local coordinate $\xi$ on $I = [-1/2, 1/2]$ be

$$\xi = (x - x_j)/h_j. \quad (2.7)$$

Next, let $p \geq 0$ be an integer. On each cell $E_j$, we wish to approximate the solution by a polynomial of degree $p$ denoted by $w_j$. To this end, we need to define the projection operator. Consider the $p + 1$ Legendre polynomials $L_k, k = 0, \ldots, p$. With $t$ fixed, the solution $u(x, t)$ can be approximated on $E_j$ by its projection onto $P_p$:

$$w_j(\xi) = \sum_{k=0}^{p} u_{j,k} L_k(\xi) \quad (2.8)$$

where the coefficients are given by

$$u_{j,k} = \int_{-1/2}^{1/2} u(x_j + \xi h_j, t) L_k(\xi) d\xi / \|L_k\|^2. \quad (2.9)$$

Thus, on $E_j$, $w_j$ is the least square fit of degree $p$ for $u$, and since the projection does not increase the norm, we have $\|w_j\| \leq \|u\|$. The collection of all $w_j$ as $j$ varies forms a piecewise polynomial function denoted by, in the physical coordinate, $w(x, t)$. The function $w(x, t)$ is generally discontinuous at the cell interfaces (see Fig. 2.2(a)).

Note that $u_{j,0}$ is the cell average value on $E_j$, and $u_{j,1}$ represents half of the normalized slope so that for $p = 1$, the value at the right interface is $u_{j,0} + u_{j,1}$. In fact, for all $p$, the value at the right interface is $u_{j,0} + u_{j,1} + \cdots + u_{j,p}$, i.e., each $u_{j,k}$ results in a $k$-th degree approximation to the right interface value.
For simplicity but not necessity, assume that the mesh is uniform: \( h_j = h \). Denote the time step by \( \Delta t \) and the CFL number \( a \Delta t / h \) by \( \sigma \). Then, since \( a \geq 0 \), the CFL condition takes the form

\[
0 \leq \sigma \leq 1.
\] (2.10)

That is, in one time step, the wave advects a distance no more than one cell width. It is this property that we wish to preserve when extending the method to one-dimensional systems of equations.

Let the projection of the initial data be carried out and denoted by, on \( E_j \),

\[
w_j^0(\xi, 0) = \sum_{k=0}^{p} u_{j,k}^0 L_k(\xi).
\]

**Exact Time Evolution and Projection.** Assume that at time \( t^n \), the projection \( w_j(\xi, t^n) \) is known; i.e., \( u_{j,k}^n = u_{j,k} \) are known for all \( j \) and all \( k \leq p \) (the superscript \( n \) is understood). We wish to calculate all \( u_{j,k}^{n+1} \) at time \( t^{n+1} \) (the superscript \( n+1 \) is retained).

Given the piecewise polynomial data \( w(x, t^n) = \{ w_j(x) \} \), the exact solution at time \( t^{n+1} \) for this data is obtained by shifting the curve \( w(x) \) to the right a distance \( a \Delta t \) as shown in Fig. 2.2(b):

\[
v(x) = w(x - a \Delta t).
\]

That is, on \( E_j \), with the local coordinate \( \xi \) given by (2.7),

\[
v_j(\xi) = \begin{cases} w_{j-1}(\xi - \sigma + 1), & \text{if } \xi < -1/2 + \sigma \\
w_j(\xi - \sigma), & \text{if } \xi \geq -1/2 + \sigma 
\end{cases}
\] (2.11)

To evaluate the solution, we split the integral into two parts (again see Fig. 2.2(b)); with \( k = 0, \ldots , p \),

\[
u_{j,k}^{n+1} = \left[ \int_{-1/2}^{-1/2+\sigma} w_{j-1}(\xi - \sigma + 1) L_k(\xi) d\xi + \int_{-1/2+\sigma}^{1/2} w_j(\xi - \sigma) L_k(\xi) d\xi \right] / \| L_k \|^2.
\] (2.12)

Such a solution can easily be evaluated by a mathematical programming package such as Mathematica.

For the case \( p = 0 \) or piecewise constant data, scheme III reduces to the first-order upwind scheme.

For the case \( p = 1 \) or piecewise linear, the solution is

\[
u_{j,0}^{n+1} = u_{j,k} + \sigma(u_{j-1,0} + u_{j-1,1} - u_{j,0} - u_{j,1}) + \sigma^2(-u_{j-1,1} + u_{j,1}),
\] (2.13)

\[
u_{j,1}^{n+1} = u_{j,1} + \sigma(-3u_{j-1,0} - 3u_{j-1,1} - 3u_{j,0} - 3u_{j,1}) + \sigma^2(3u_{j-1,0} + 6u_{j-1,1} - 3u_{j,0} + 3u_{j,1}) + \sigma^3(-2u_{j-1,1} + 2u_{j,1}).
\] (2.14)

In matrix form,
\[
\begin{pmatrix}
\sigma^{n+1} \\
\sigma^{n+1}
\end{pmatrix} = \begin{pmatrix}
-3\sigma(1-\sigma) & \sigma(1-\sigma) \\
1-\sigma & -\sigma(1-\sigma)
\end{pmatrix} \begin{pmatrix}
\sigma^{n-1} \\
\sigma^{n-1}
\end{pmatrix} + \begin{pmatrix}
\frac{1}{3\sigma(1-\sigma)} & \sigma(1-\sigma) \\
3\sigma(1-\sigma) & (1-\sigma)(1-2\sigma-2\sigma^2)
\end{pmatrix} \begin{pmatrix}
\sigma^{n} \\
\sigma^{n}
\end{pmatrix}
\] (2.15)

Fig. 2.2 Scheme III (a) Piecewise linear data; (b) Solution obtained by shifting the data a distance \(a\Delta t\) and calculating the moments of the discontinuous function.

Note the term \(\sigma^3\) above for the case \(p = 1\). For a general \(p\), the function \(w_j(\xi - \sigma) L_p(\xi)\) is of degree \(2p\) in \(\xi\); as a result, the term of highest degree in the solution is \(\sigma^{2p+1}\). It will be shown by Fourier analysis that the method is accurate to order \(2p + 1\) and is stable for \(0 \leq \sigma \leq 1\).

Also note that the method is clearly energy-stable: the \(L^2\)-norm of the solution at time \(t^{n+1}\) is bounded by that of the data at time \(t^n\) since the solution is obtained by an orthogonal projection.

As the final remark of this section, we discuss the extension of this approach. For the advection case, the discontinuity at an interface evolves via the exact shift operator. In the case of systems such as the Euler equations, in one spatial dimension, these discontinuities result in some combination of a shock, a contact, or a fan. Tracking these waves accurately is extremely difficult. Resolving these waves for the multi-dimensional cases appears to be an impossible task.

3. Moment Scheme for the Case of One Spatial Dimension

**Formulation.** To avoid the exact shift operator—which makes extensions difficult or impossible—we take a different path by using a space-time domain and integration by parts together with the observation that the space-time fluxes contain the information provided by the shift operator. Consider the scalar conservation law

\[
u_t + f_x = 0
\] (3.1)
where $t$ is time, $x$ spatial coordinate, and $f = f(u(x, t))$ the flux. With the wave speed $a(u) = df/du$, (3.1) implies

$$u_t + au_x = 0.$$  

(3.2)

On the cell $E = E_j$, let $\phi$ be a test function, i.e., a polynomial of degree $p$. The conservation law in weak form is

$$\int_E u_t(x, t) \phi(x) \, dx + \int_E f_x \phi \, dx = 0.$$  

(3.3)

We will set $\phi = L_k$, the Legendre basis function, later. Thus, the weak form amounts to projecting onto $P_p$. As in the DG formulation, by applying integration by parts to the second term,

$$\frac{\partial}{\partial t} \int_E u(x, t) \phi(x) \, dx + \int_E [f \phi]_{x_{j+1/2}}^{x_{j-1/2}} - \int_E f_x \phi \, dx = 0.$$  

(3.4)

The fluxes $f$ for the boundary term $[f \phi]_{x_{j-1/2}}$ are chosen to be the upwind ones. At each interface, this flux is common for the two adjacent cells and allows the data in these cells to interact. Without this common flux, there is no interaction, and the resulting method is inconsistent with physics.

Integrating in time from $t^n$ to $t^*$ ($t^*$ will be set to the intermediate time levels as well as the final time $t^{n+1}$ later), we obtain

$$\int_E u(x, t^*) \phi(x) \, dx - \int_E u(x, t^n) \phi(x) \, dx + \int_{t^n}^{t^*} [f_{upw} \phi]_{x_{j-1/2}}^{x_{j+1/2}} \, dt$$

$$- \int_{t^n}^{t^*} f \phi_x \, dx \, dt = 0.$$  

(3.5)

Denote the surface and volume integral by, respectively,

$$SI = -\int_{t^n}^{t^*} [f_{upw} \phi]_{x_{j-1/2}}^{x_{j+1/2}} \, dt$$  

(3.6)

and

$$VI = \int_{t^n}^{t^*} \int_E f \phi_x \, dx \, dt.$$  

(3.7)

The problem reduces to the evaluation of $SI$ and $VI$,

$$\int_E u(x, t^*) \phi(x) \, dx = \int_E u(x, t^n) \phi(x) \, dx + SI + VI.$$  

(3.8)

Here, for $\phi = L_k$ and $t^* = t^{n+1}$, the left hand side above in the local coordinate yields the solution $h_j ||L_k||^2 u_{j,k}^{n+1}$ whereas the first term on right hand side, $h_j ||L_k||^2 u_{j,k}^n = h_j ||L_k||^2 u_{j,k}^n$.

Consider now the case of advection. So far, the current approach is equivalent to the exact evolution approach of the previous section. Indeed, with the piecewise polynomial data $w_j(x)$ at time $t^n$, the exact
space-time solution is \( w(x, t) = w_j(x - at) \), which has discontinuities along the characteristic lines joining the points \((x_{j-1/2}, t^n)\) and \((x_{j-1/2} + a\Delta t, t^{n+1})\) as shown in Fig. 3.1. The two formulations, one by projecting \( w(x, t^{n+1}) \), and the other by integrating on the space-time cell \( E_j \times [t^n, t^{n+1}] \) yield identical solutions, which are exact for the data. At first glance, it appears that we have made the problem more difficult since the volume integral takes the form

\[
VI = \int_{t^n}^{t^{n+1}} \int_E au \phi_x \, dx \, dt. \tag{3.9}
\]

In the previous section, we only deal with a discontinuity at one point in space; here, we have to deal with a discontinuity along a line in a space-time domain. However, as will be shown, the term \( \phi_x \) comes to the rescue in a procedure which ignores the discontinuity.

![Exact Time Evolution](image)

**Fig. 3.1.** Space-time cells \( E_j \times [t^n, t^{n+1}] \) and the paths of discontinuities (dark line segments) for the case of advection

**Time Integration.** We proceed by discussing time integration. Due to the term \( \sigma^{2p+1} \) in the solution as noted in the previous section, we need a quadrature with a degree of precision of \( 2p \) in time, i.e., it is exact for polynomials of degree \( 2p \) or less (note that we only need a degree of precision of \( 2p \) since the vertical edge of \( E_j \times [t^n, t^{n+1}] \) accounts for the additional power in \( \sigma^{2p+1} \)). Here, we use the right Radau quadrature with \( p + 1 \) evaluation points—consistent with the spatial degree of freedom. Another reason for using this quadrature is that the DG method applied to time can be reduced to the Radau IIA method (Huynh 2009b), which is a collocation method corresponding to this quadrature. In other words, the moment scheme formulated here has the same intermediate time levels and is closely related to the implicit Runge-Kutta (RK) scheme resulting from a space-time DG discretization (see §6).

A brief review of the right Radau quadrature and the associated Butcher matrix is in order (we can also employ the left Radau quadrature as in LaSaint and Raviart (1974); the resulting method is, however, more complex). Set \( s = p + 1 \) where \( s \) is the number of stages. Using notations from ODE, on the interval \([0, 1]\), let the \( s \) right Radau points be denoted by \( c_l, l = 1, \ldots, s \) where \( c_s = 1 \). Recall that these points result in a quadrature formula on \([0,1]\) that is exact for polynomials of degree \( 2(s - 1) \) or less. For each \( l \), let the corresponding Lagrange polynomials on \([0,1]\) be denoted by \( L_{a_l} \) and defined by the condition that \( L_{a_l}(c_m) = 0 \) for all \( m \neq l \) and \( L_{a_l}(c_l) = 1 \):
The plots of these functions for the case $s = 2$ and $s = 3$ (where $s = p + 1$) are shown in Fig. 3.2.

\[
\text{La}_l(\tau) = \prod_{\substack{m=1 \atop m \neq l}}^{s} \frac{\tau - c_m}{c_l - c_m},
\]

(3.10)

Fig. 3.2 Lagrange polynomials for the right Radau points: (a) Two points $c_1 = 1/3$, $c_2 = 1$, (b) Three points $c_1 = 0.155051$, $c_2 = 0.644949$, and $c_3 = 1$.

Next, suppose the values of a function $v$ are known at these Radau points: $v_m = v(c_m)$, $m = 1, \ldots, s$. We wish to obtain a quadrature formula for $\int_0^{c_l} v(\tau) d\tau$, $l = 1, \ldots, s$. To this end, the polynomial of degree $p = s - 1$ interpolating $v_m$, $m = 1, \ldots, s$, can be expressed as

\[
v(\tau) = \sum_{m=1}^{s} v_m \text{La}_m(\tau).
\]

(3.11)

Integrating from 0 to $c_l$, we obtain

\[
\int_0^{c_l} v(\tau) d\tau = \sum_{m=1}^{s} v_m \int_0^{c_l} \text{La}_m(\tau) \, d\tau.
\]

(3.12)

Denote the Butcher matrix of dimension $s \times s$ by $A$. The entries $a_{lm}$ of $A$ are given by

\[
a_{lm} = \int_0^{c_l} \text{La}_m(\tau) \, d\tau.
\]

(3.13)

These $a_{lm}$ can easily be obtained by using a software package such as Mathematica. The above two equations imply, for each $l$,

\[
\int_0^{c_l} v(\tau) d\tau = \sum_{m=1}^{s} a_{lm} v_m.
\]

(3.14)

Note that for each row $l$, setting $v = 1$, we obtain $c_l = \sum_{m=1}^{s} a_{lm}$. Also note that for $l = 1, \ldots, p$, the above quadrature has degree of precision of only $p$. However, for $l = s = p + 1$, i.e., for the integral from
0 to 1, due to the definition of the Radau points, the corresponding quadrature has a degree of precision of 2p. The column vector $C = \{c_i\}$ and the matrix $A$ for the case $p = 1$ and $p = 2$ are, respectively

$$
\begin{bmatrix}
1/3 \\
5/12 \\
1/4 \\
3/4
\end{bmatrix}
\begin{bmatrix}
1 \\
-1/12 \\
1/4
\end{bmatrix}
\text{ and }
\begin{bmatrix}
0.155051 \\
0.196815 \\
0.644949
\end{bmatrix}
\begin{bmatrix}
0.394424 \\
0.993835 \\
0.376403
\end{bmatrix}
\begin{bmatrix}
-0.0655354 \\
0.292073 \\
0.512486
\end{bmatrix}
\begin{bmatrix}
0.023771 \\
-0.0415488 \\
0.111111
\end{bmatrix}.
$$

(3.15)

Returning to the conservation law, for the space-time domain $E_j \times [t^n, t^{n+1}]$, denote the time level for each stage $l$ of the $s$ stages of the above right Radau quadrature by

$$
t^{n,l} = t^n + c_l \Delta t.
$$

(3.16)

These are the stages of a Runge-Kutta type scheme where $t^{n,s} = t^{n+1}$. See Fig. 3.3(a) and (b).

**Space-Time Solution with No Interaction.** Next, we discuss an approximation for the solution on the space-time cell $E_j \times [t^n, t^{n+1}]$ with no interaction (NI) i.e., no upwinding. This approximation denoted by $w_j^{NI}(x, t)$ is a polynomial of degree $p$, which can be a Taylor series around $(x_j, t^n)$. The case of $p = 1$ is trivial: $w_j(x)$ provides $u_x$ and (3.2), $u_t$. For a general $p$, in each cell $E_j$, at time $t^n$, the solution is known and is a polynomial of degree $p$ in $x$,

$$
w_j(x) = \sum_{k=0}^{p} u_{j,k} L_k(x).
$$

(3.17)

From (3.2), we can calculate all the time as well as mixed space-time derivatives up to degree $p$, i.e., $u_t$, $u_{xt}$, $u_{tt}$, ... at $(x_j, t^n)$ via the Cauchy-Kovalevsky procedure (Harten, Osher, Engquist, and Chakravarthy 1987). The space-time Taylor series expansion around $(x_j, t^n)$ up to degree $p$ then yields $w_j^{NI}(x, t)$ on $E_j \times [t^n, t^{n+1}]$.

As an alternative to the above Taylor series expansion, we can use the nodal Lobatto points together with a RK time stepping with no interaction. For example, consider the case $p = 3$. See Fig. 3.3(b). On $E_j$, at time $t^n$, since the cubic $w_j(x)$ is known, we can obtain the values at the four Lobatto points via a $(4 \times 4)$ matrix multiplication. Next, applying the standard 4-stage RK scheme to (3.2) with no interaction, we can evaluate the nodal values at time $t^{n+1}$. From the known nodal values for $u$ at $t^n$ and $t^{n+1}$, we can calculate the corresponding $u_t$ again with no interaction by (3.2). Then, using a Hermite cubic interpolation with the two ends at $t^n$ and $t^{n+1}$, we can evaluate the nodal values accurate to degree $p = 3$ for all intermediate time levels $t^{n,l} = t^n + c_l \Delta t$, $l = 1, 2, 3$. Note that this step is inexpensive and the equations for either primitive or conservative variables can be employed with the former more economical than the latter.

An observation concerning the case of advection with $a \geq 0$ is in order. We focus, for the moment, on only one cell, say, $E_j$ with the polynomial data $w_j$. For the space-time domain $E_j \times [t^n, t^{n+1}]$, along the two interfaces $(x_{j-1/2}, t)$ and $(x_{j+1/2}, t)$, the fluxes are identical to those by method of characteristics applied to $w_j$ assuming the domain for $w_j(x)$ is $(-\infty, \infty)$. Indeed, both the method of characteristics and space-time Taylor series produce the exact solution for the polynomial data. However, when interaction comes into play, whereas along the right interface, the upwind fluxes reduces to the result by method of characteristics for cell $j$, along the left interface, the values by $w_j$ at $x_{j-1/2} = a c_l \Delta t$, namely, $aw_j^{NI}(x_{j-1/2}, t^{n,l}) = w_j(x_{j-1/2} - a c_l \Delta t)$, are from outside the cell $E_j$ and will be discarded by the
upwind step. In other words, along \((x_{j-1/2}, t)\), the upwind procedure sorts out and employs the values by \(w_{j-1}\).

Thus, for general conservation laws, via one of the above two procedures (Cauchy-Kovalevsky or explicit RK), for all intermediate time levels \(t^{n,l}\) of the right Radau type, at each cell interface \(x_{j+1/2}\), we have two values \(u_L\) and \(u_R\) (no interaction). See Fig. 3.3. These values in turn yield the upwind flux \(f_{upw}(x_{j+1/2}, t^{n,l})\).

### Fig. 3.3
Space-time nodal points with space via Lobatto and time via right Radau: (a) \(p = 1\), and (b) \(p = 3\). The number of spatial points, \(p + 1\), is the same as the number of Legendre basis functions. Along each interface, at each Radau intermediate time level, we can obtain the values \(u_L\) and \(u_R\), which in turn yield the upwind flux.

#### Moment Solution
We can now discuss the moment solution. On the space-time domain \(E_j \times [t^n, t^{n+1}]\), with \(t^* = t^{n,l}\), switching to the local coordinate \(\xi\) on \(I = [-1/2, 1/2]\), since \(dx = h_j d\xi\), (3.8) yields

\[
\frac{h_j}{2} L_k \|u_{j,k}^{n,l}\|^2 = \frac{h_j}{2} L_k \|u_{j,k}\|^2 - \int_{t^n}^{t^{n,l}} [f_{upw} L_{k,1/2} - f_{upw}(x_{j-1/2}, t^{n,m}) L_k(x_{j-1/2})] d\xi dt.
\]

For the above last two terms, integrate in (normalized) time from 0 to \(c_t\) via the right Radau quadrature (3.14), we obtain the following surface and volume integrals: with Legendre index \(k\) and time index \(l\),

\[
SI_{k}^l = -\Delta t \sum_{m=1}^{s} a_{lm} [f_{upw}(x_{j+1/2}, t^{n,m}) L_k(x_{j+1/2}) - f_{upw}(x_{j-1/2}, t^{n,m}) L_k(x_{j-1/2})]
\]

and

\[
VI_{k}^l = \Delta t \sum_{m=1}^{s} a_{lm} \int I f(\xi, t^{n,m})(L_k)\xi d\xi.
\]

That is,

\[
\frac{h_j}{2} L_k \|u_{j,k}^{n,l}\|^2 = \frac{h_j}{2} L_k \|u_{j,k}\|^2 - SI_{k}^l + VI_{k}^l.
\]
The solution we are seeking corresponds to time level \( l = s \) and Legendre indices \( k = 0, \ldots, p \).

The surface integral is straightforward. The volume integral is where the difficulty lies. This integral must be carried out in a manner that two contrasting requirements hold: (a) the procedure can be extended with relative ease to the case of systems but (b) the space-time discontinuity in Fig. 3.1 is somehow accounted for.

What comes to the rescue here is the fact that the degree of \( (L_k)_{\xi} \) is one lower than that of \( L_k \); as a result, we can successively calculate \( u_{j,0}^{n,l}, u_{j,1}^{n,l}, u_{j,2}^{n,l}, \ldots, u_{j,p}^{n,l} \) as follows.

First, for \( k = 0 \), the volume integral vanishes since \( (L_0)_{\xi} = 0 \). This fact and (3.21) yield \( u_{j,0}^{n,l} \) for all \( l \).

With \( u_{j,0}^{n,l} \) known, we can calculate \( u_{j,1}^{n,l} \) where \( k = 1 \) as follows. Since (2.6) implies \( (L_1)_{\xi} = 2 \), the volume integral \( \mathcal{V}_k^{l} \) of (3.20) involves only the projection of \( f \) onto \( 2L_0 \). Thus, we need to approximate \( f \) up to \( L_0 \) but no higher, which is given by \( f(u_{j,0}^{n,l}) \). The solution \( u_{j,1}^{n,l} \) can therefore be obtained for all \( l \).

Next, with \( u_{j,0}^{n,l} \) and \( u_{j,1}^{n,l} \) known, we can calculate \( u_{j,2}^{n,l} \) where \( k = 2 \) as follows. Since (2.6) implies \( (L_2)_{\xi} = 12\xi = 6L_1 \), the volume integral \( \mathcal{V}_k^{l} \) of (3.20) involves the projection of \( f \) onto \( 6L_1 \). Thus, we only need to approximate \( f \) up to \( L_1 \). With \( u_{j,0}^{n,l} \) and \( u_{j,1}^{n,l} \) known, we can approximate \( u \) by \( u_{j,0}^{n,l} L_0(\xi) + u_{j,1}^{n,l} L_1(\xi) \), and \( f \) by \( f_{j,0}^{n,l} L_0(\xi) + f_{j,1}^{n,l} L_1(\xi) \) (by whatever procedure that provides \( f \) to the same degree as \( u \)). The solution \( u_{j,2}^{n,l} \) can then be obtained for all \( l \).

Assuming that \( u_{j,0}^{n,l}, \ldots, u_{j,k-1}^{n,l} \) are known, we now calculate \( u_{j,k}^{n,l} \). Since \( (L_k)_{\xi} \) is of degree \( k - 1 \), it can be expressed as a linear combination of \( L_0, \ldots, L_{k-1} \). In fact, in general,

\[
(L_{2m+1})_{\xi} = 2 \sum_{l=0}^{m} L_{2l} \quad \text{and} \quad (L_{2m+2})_{\xi} = 2 \sum_{l=0}^{m} L_{2l+1}, \tag{3.22}
\]

i.e.,

\[
(L_{2m+1})_{\xi} = 2L_{2m} + 2L_{2m-2} + \ldots \quad \text{and} \quad (L_{2m+2})_{\xi} = 2L_{2m+1} + 2L_{2m-1} + \ldots
\]

Thus, the volume integral \( \mathcal{V}_k^{l} \) by (3.20) involves projecting \( f \) onto \( P_{k-1} \), and we only need to approximate \( f \) up to \( L_{k-1} \). Since the solution can be approximated by \( u_{j,0}^{n,l} L_0(\xi) + \cdots + u_{j,k-1}^{n,l} L_{k-1}(\xi) \), the flux \( f \) can be approximated by \( f_{j,0}^{n,l} L_0(\xi) + \cdots + f_{j,k-1}^{n,l} L_{k-1}(\xi) \) (again, by whatever procedure that provides \( f \) to the same degree as \( u \)) and the solution \( u_{j,k}^{n,l} \) follows for all \( l \). This completes the algorithm.

Note that in the calculation for \( u_{j,p}^{n,l} \), namely \( k = p \), the last Legendre component \( k \) for the above successive calculation, we only need to evaluate \( u_{j,p}^{n,s} \); the quantities \( u_{j,p}^{n,l} \) where \( l = 1, \ldots, p \) are not needed (we would need them if we use the left Radau or Gauss quadrature or if we use the current procedure as a building block in an iterative procedure for the implicit right Radau method).

Also note that if \( u \) is approximated by \( \psi_j = u_{j,0} L_0(\xi) + \cdots + u_{j,k-1} L_{k-1}(\xi) \), there are several ways to approximate \( f \) up to \( L_{k-1} \). One way is by evaluating \( f_m = f(\psi_j(\xi_m)) \) where \( \xi_m \) are the \( k \) Gauss points and then projecting the polynomial interpolating \( \{f_m\} \) onto \( P_{k-1} \).

For the case of advection, the current moment approach is equivalent to the exact evolution approach of the previous section. This fact has also been verified by the author via a Mathematica program.
For the piecewise linear case, the moment scheme is simple: \( w_{j}^{NL}(x, t) \) is a linear function where, with \( \xi \) on \([-1/2, 1/2]\) and \( \tau \) on \([0, 1]\),
\[
  w_{j}^{NL}(\xi, \tau) = u_{j, 0} + 2u_{j, 1}\xi - 2u_{j, 1}\varpi\tau.
\]
The extension of the piecewise linear case to the Euler equations is also simple (Huynh 2006).

**Algorithm.** The moment scheme is summarized below. Assume that the data \( u_{j,k} \) at time \( t^{n} \) are given for all \( j \) and \( k = 0, \ldots, p \).

1. In each space-time cell \( E_{j} \times [t^{n}, t^{n+1}] \), obtain \( w_{j}^{NL} \) with no interaction by a Cauchy-Kovalevsky procedure or an explicit RK with no interaction.
2. At each interface \( x_{j-1/2} \), for each right Radau intermediate time level, calculate the left and right values using \( w_{j-1}^{NL} \) and \( w_{j}^{NL} \). Obtain upwind fluxes from these two values.
3. Successively evaluate \( u_{j,0}^{n,1}, u_{j,1}^{n,1}, u_{j,2}^{n,1}, \ldots, \) and \( u_{j,p}^{n,1} \) via (3.21) by calculating the surface integral via (3.19) and volume integral via (3.20). This completes the algorithm.

We conclude this section by the following remarks.

If the upwind fluxes at the interfaces are given (e.g., in an iterative procedure for the implicit right Radau time stepping), the above algorithm provides the solution for the space-time cell \( E_{j} \times [t^{n}, t^{n+1}] \). This observation shows that the moment scheme can serve as a key building block in an iterative procedure for the implicit time-stepping method.

One way to apply the moment approach to the nodal type of schemes is to transform the nodal values into the modal (Legendre) components, carry out the moment calculations, and then transfer back. The various transferring, however, involve some costs. The question is whether it is possible to use an iteration procedure in the nodal framework with no transferring.

**4 Moment Scheme for the Case of Two Spatial Dimensions**

For the case of two spatial dimensions, the approach of exact time evolution and projection of Section 2 can be extended in a manner that if the flow is along the diagonal direction, say, from southwest to northeast, then the solution recovers the exact data of the southwest cell for the CFL number \((\sigma_{x}, \sigma_{y}) = (1,1)\). Such an approach must allow for interaction among cells that share a common edge as well as cells that share a common corner. Its key drawback is that extension to systems is extremely difficult. Therefore, the extension of the moment scheme here allows interaction only among immediate neighbors, i.e., cells that share a common edge. The price to pay for this simplification is a reduction in CFL condition when the flow is not along the edge direction.

To proceed, with \( u = u(x,y,t) \), the 2D conservation law takes the form
\[
  u_{t} + f_{x} + g_{y} = 0.
\]
Let initial condition be \( u(x,y,0) = u_{\text{init}}(x,y) \). For the Euler equations, \( u, f, \) and \( g \) are vectors of four entries. With no loss of generality, we consider only the scalar case. Assume that the solution \( u \) is periodic or of compact support so that boundary conditions are trivial.
Denote \( f = (f, g) \) and \( \nabla = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) \). Then \( \nabla \cdot f = f_x + g_y \), and the above can be written as

\[
u_t + \nabla \cdot f = 0. \tag{4.1}\]

Let the domain of calculation be divided into a non-overlapping unstructured mesh of cells \( E_j \) which are quadrilaterals or triangles. On each cell \( E_j \), let the solution \( u \) be approximated by \( u_j \) of degree \( p \). If the cell is a quadrilateral, we can use either a polynomial in \( \xi \) and \( \eta \) of degree \( p \) or less, or we can use tensor products of degree \( p \times p \). For example, if \( p = 1 \), we can use either a linear polynomial \( a + b\xi + c\eta \), or for the case of tensor products, a bilinear polynomial \( a + b\xi + c\eta + d\xi\eta \).

Using \( u_j \), the fluxes \( f \) and \( g \) can be approximated by \( f_j \) and \( g_j \) of the same degree and type as \( u_j \). The piecewise polynomial functions \( \{ u_j \}, \{ f_j \}, \) and \( \{ g_j \} \) can be and usually are discontinuous across cell interfaces.

From here on, unless otherwise stated, we focus only on the approximate polynomials. For simplicity of notation, the subscript \( j \) is understood: \( u_j, f_j \), and \( E_j \) are abbreviated to \( u, f, \) and \( E \) when there is no ambiguity.

Let \( \phi \) be a test function on \( E \), i.e., a polynomial of the same type as \( u \). We require the solution to satisfy, for the moment, formally,

\[
\int_E u_t \phi \, dE + \int_E (\nabla \cdot f) \phi \, dE = 0.
\]

As in the 1D case, integrating by parts and applying the upwind flux for the boundary term to account for data interaction, the solution is required to satisfy

\[
\int_E u_t \phi \, dE + \int_{\partial E} (f \cdot n)_{\text{upw}} \phi \, ds - \int_E \nabla \phi \cdot f \, dE = 0. \tag{4.2}
\]

Integrating in time from \( t^n \) to \( t^* = t^{n+1} \),

\[
\int_E u(x, y, t^*) \phi \, dE - \int_E u(x, y, t^n) \phi \, dE - \text{SI} - \text{VI} = 0 \tag{4.3}
\]

where the surface and volume integral are, respectively,

\[
\text{SI} = - \int_{t^n}^{t^*} \int_{\partial E} (f \cdot n)_{\text{upw}} \phi \, ds \, dt \tag{4.4}
\]

and

\[
\text{VI} = \int_{t^n}^{t^*} \int_E \nabla \phi \cdot f \, dE \, dt. \tag{4.5}
\]

The problem reduces to the evaluation of SI and VI.

To this end, as in the 1D case, we first obtain \( w_{ji}^{NI} \) with no interaction by a Cauchy-Kovalevsky procedure or an explicit RK with no interaction up to appropriate order. For the case of tensor products, we would need to derive time derivatives of degree up to \( 2p \) if the Cauchy-Kovalevsky procedure is
employed. Next, along each edge, for each right Radau intermediate time level, calculate the left and right
values using appropriate $w_j^{nl}$. Obtain normal upwind fluxes from these two values.

The surface integral again poses no difficulty. As for the volume integral, we make use the fact that
both components of $\nabla \Phi$, namely, $\phi_x$ and $\phi_y$ are of at least one degree lower than $\Phi$. As in the 1D case,
we can evaluate the volume integral and successively calculate the Legendre components $u_j^{n,1}$, $u_j^{n,1}$,
$u_j^{n,0,1}$, $u_j^{n,2,0}$, $u_j^{n,1,1}$, $u_j^{n,0,2}$, ... This completes the algorithm.

5 Fourier Analyses

One-Dimensional Case. For the 1D case, consider the advection equation with positive speed $a$,

$$u_t + au_x = 0.$$ 

The cells are $E_j = [j - 1/2, j + 1/2]$. Denote by $U_j$ the column vector of $K = p + 1$ components,

$$U_j = (u_{j,0}, u_{j,1}, ..., u_{j,p})^T$$

where $T$ represents the transpose. Let $\Delta t$ be the time step and set $\sigma = a\Delta t$. Assume that the data $U_j^n = U_j$ are known. The solution $U_j^{n+1}$ can be expressed as

$$U_j^{n+1} = C_0U_j + C_{-1}U_{j-1}$$

(5.1)

where $C_0$ and $C_{-1}$ are $K \times K$ matrices. As examples, for $p = 1$, i.e., the linear case,

$$C_{-1} = \begin{pmatrix}
\sigma & \sigma(1-\sigma) \\
-3\sigma(1-\sigma) & -\sigma(3-6\sigma+2\sigma^2)
\end{pmatrix}, \quad C_0 = \begin{pmatrix}
1-\sigma & \sigma(1-\sigma) \\
-3\sigma(1-\sigma) & (1-\sigma)(-1+2\sigma+2\sigma^2)
\end{pmatrix}.$$ 

and, for $p = 2$ or the parabolic case,

$$C_{-1} = \begin{pmatrix}
\sigma & \sigma(1-\sigma) & \sigma(1-\sigma)(1-2\sigma) \\
-3\sigma(1-\sigma) & -\sigma(3-6\sigma+2\sigma^2) & -3\sigma(1-\sigma)(1-3\sigma+\sigma^2) \\
5\sigma(1-\sigma)(1-2\sigma) & 5\sigma(1-\sigma)(1-3\sigma+\sigma^2) & \sigma(5-30\sigma+50\sigma^2-30\sigma^3+6\sigma^4)
\end{pmatrix}.$$ 

and

$$C_0 = \begin{pmatrix}
1-\sigma & -\sigma(1-\sigma) & -\sigma(1-\sigma)(1-2\sigma) \\
3\sigma(1-\sigma) & (1-\sigma)(1-2\sigma-2\sigma^2) & -3\sigma(1-\sigma)(1-\sigma-\sigma^2) \\
-5\sigma(1-\sigma)(1-2\sigma) & 5\sigma(1-\sigma)(1-\sigma-\sigma^2) & (1-\sigma)(1-4\sigma-4\sigma^2+6\sigma^3+6\sigma^4)
\end{pmatrix}.$$ 

Note the highest degree term for $p = 1$ is $\sigma^3$ and, that for $p = 2$, $\sigma^5$.

Stability. Let the wave number be denoted by $w$, $-\pi < w \leq \pi$, and the imaginary number, by $I$ instead of the standard notation $i$ to avoid confusion with the cell index $i$ later. Assuming the solution is a harmonic such that $U_{j-1} = e^{-iw}U_j$, we obtain the amplification matrix

$$A = C_0 + e^{-iw}C_{-1}$$

(5.2)

which satisfies, by (5.1),
With each value of $\sigma$ and $w$, the corresponding matrix $A$ has $p + 1 = K$ complex eigenvalues. For stability, these eigenvalues must have magnitude no larger than 1. Numerical calculations for these eigenvalues for all $p$ tested by this author result in stable schemes provided that $0 \leq \sigma \leq 1$.

Figures 5.1 show the magnitude of the amplification factors (or eigenvalues) for the case $p = 1$ and Figs 5.2, the case $p = 2$.

**Accuracy.** To calculate the order of accuracy, denote the principal eigenvalue of $A$ by $e_p = e_p(\sigma, w)$. This value approximates $e^{-I\omega}$. A scheme is accurate to order $m$ if, for a fixed $\sigma$, $e_p$ approximates $e^{-I\omega}$ to $O(w^{m+1})$ for small $w$,

$$e_p - e^{-I\omega} = O(w^{m+1}). \quad (5.4)$$
In practice, it is difficult if not impossible to derive a Taylor series expression for $e_p$ when $p > 1$. Therefore, we obtain the order of accuracy of a scheme by a numerical calculation. First, set $\sigma = \sigma_0 = 0.8$, say, and $w = w_c = \pi/4$. We can calculate the coarse mesh error

$$er_c = e_p(\sigma_0, w_c) - e^{-i\sigma_0 w_c}.$$  

(5.5)

By halving the wave number, $w_f = w_c/2 = \pi/8$ (equivalent to doubling the number of mesh points), the corresponding fine mesh error is

$$er_f = e_p(\sigma_0, w_f) - e^{-i\sigma_0 w_f}.$$  

(5.6)

For a scheme to be $m$-th order accurate, after one time step,

$$\frac{|er_c|}{er_f} \approx 2^{m+1}.  \quad (5.7)$$

That is,

$$m \approx \frac{\log \left( \frac{|er_c|}{er_f} \right)}{\log(2)} - 1.  \quad (5.8)$$

Thus, for a scheme to be of order $m$, when we march to a certain final time, doubling the mesh results in an error smaller by a factor of $2^m$.

It can be verified, as will be shown in the more involved 2D cases, that a moment scheme using polynomials of degree $p$ is accurate to order $2p + 1$.

Two-Dimensional Case. We next briefly describe Fourier (Von Neumann) analysis for the 2D case. On the domain $(-\infty, \infty) \times (-\infty, \infty)$, consider the advection equation

$$u_t + au_x + bu_y = 0.$$ 

The cells are the squares $E_{i,j} = [i - 1/2, i + 1/2] \times [j - 1/2, j + 1/2]$ centered at $(i,j)$. For the case of polynomials of degree $p$, the number of Legendre basis functions is $K_p = (p + 1)(p + 2)/2$. For the case of tensor products, the number of Legendre basis functions again denoted by $K_p$, is $K_p = (p + 1)^2$. In addition, denote by $U_{i,j}$ the column vector of $K_p$ components,

$$U_{i,j} = \begin{pmatrix} u_{i,j,0,0} & u_{i,j,1,0} & u_{i,j,0,1} & \cdots \end{pmatrix}^T  \quad (5.9)$$

As an example, for $p = 1$,

$$U_{i,j} = \begin{pmatrix} u_{i,j,0,0} & u_{i,j,1,0} & u_{i,j,0,1} \end{pmatrix}^T.$$ 

Or, for the case of tensor products,

$$U_{i,j} = \begin{pmatrix} u_{i,j,0,0} & u_{i,j,1,0} & u_{i,j,0,1} & u_{i,j,1,1} \end{pmatrix}^T.$$
Set \( \sigma_x = a \Delta t \) and \( \sigma_y = b \Delta t \). The 2D moment solution can be written as

\[
U_{i,j}^{n+1} = C_{0,0} U_{i,j} + C_{-1,0} U_{i-1,j} + C_{0,-1} U_{i,j-1}.
\]  

(5.10)

Here, \( C_{0,0}, C_{-1,0}, \) and \( C_{0,-1} \) are \( K_p \times K_p \) matrices whose entries involve \( \sigma_x \) and \( \sigma_y \). As an example, for the linear case (5.9),

\[
C_{-1,0} = \begin{pmatrix}
\sigma_x & \sigma_x(1-\sigma_x) & -\sigma_x\sigma_y \\
-3\sigma_x(1-\sigma_x) & -\sigma_x(3-6\sigma_x + 2\sigma_y^2) & \sigma_x\sigma_y(3-2\sigma_x) \\
3\sigma_x\sigma_y & \sigma_x\sigma_y(3-2\sigma_x) & \sigma_x(1-2\sigma_y^2)
\end{pmatrix},
\]

and

\[
C_{0,-1} = \begin{pmatrix}
\sigma_y & -\sigma_x\sigma_y & \sigma_y(1-\sigma_x) \\
3\sigma_x\sigma_y & \sigma_y(1-2\sigma_x^2) & \sigma_x\sigma_y(3-2\sigma_y) \\
-3\sigma_y(1-\sigma_y) & \sigma_x\sigma_y(3-2\sigma_y) & -\sigma_y(3-6\sigma_y + 2\sigma_x^2)
\end{pmatrix},
\]

Stability. Let the wave number in the \( x \)-direction be denoted by \( \omega_x \) and in the \( y \)-direction, \( \omega_y \) (involving no partial derivative), and let the imaginary number be denoted by \( I \). Similar to the 1D case, replacing \( U_{i-1,j} \) and \( U_{i,j-1} \) respectively by \( e^{-I \omega_x} U_{i,j} \) and \( e^{-I \omega_y} U_{i,j} \), we obtain the amplification matrix

\[
A = C_{0,0} + e^{-I \omega_x} C_{-1,0} + e^{-I \omega_y} C_{0,-1}.
\]  

(5.11)

Consider a fixed value \((\sigma_x, \sigma_y)\). For each wave number \((\omega_x, \omega_y)\) where \(-\pi < \omega_x \leq \pi\) and \(-\pi < \omega_y \leq \pi\), the corresponding matrix \( A \) has \( K_p \) eigenvalues. For stability, these eigenvalues must have magnitude no larger than 1 for all wave numbers. The domain of all values \((\sigma_x, \sigma_y)\) satisfying this stability condition is called the stability region of the scheme.

Figure 5.3 shows the stability region for a square mesh and polynomials of degree \( p = 1, \ldots, 4 \). As is well-known, the CFL limits along the diagonal for the first-order upwind scheme \((p = 0)\) is \( \sqrt{2}/2 \approx 0.7071 \). In the case of the moment schemes, for \( p = 1, \ldots, 4 \), they are, respectively,

\[0.4714, \ 0.2859, \ 0.2190, \ \text{and} \ 0.1654.\]
Figure 5.3. Stability regions for schemes using polynomials of degree $p$. Note that for $p = 3$ and 4, the stability conditions along the $x$ and $y$ directions are 1, but these conditions reduce very quickly when the flow slightly deviates from the axes directions.

Figure 5.4 shows the stability region for a square mesh and polynomials via tensor products of degree $p \times p$ where $p = 1, 2, 3,$ and 4. The CFL limits along the diagonal for $p = 1, \ldots, 4$ are respectively $0.4714, 0.2834, 0.2526,$ and $0.1702$. 
Figure 5.4. Stability regions for schemes using polynomials via tensor products of degree $p \times p$. Again, for $p = 3$ and 4, the stability conditions along the $x$ and $y$ directions are 1, but these conditions reduce very quickly when the flow slightly deviates from the axes directions.

Thus, the reduction in the CFL condition for flows along the diagonal direction is significant. The main reason is that information does not directly communicate among cells that share only one corner. That is, for a flow from, say, southwest to northeast, information flows to the cells to the east and the north before it can reach the cell at the northeast corner. This difficulty concerning information flow can be remedied by using the moment scheme as a procedure to update the solution in an iterative process for an implicit space-time DG scheme as will be discussed in the next section.
Accuracy. To calculate the order of accuracy, using the notation of the 1D case, denote the principal eigenvalue of the amplification matrix $\mathbf{A}$ defined in (5.11) by $e_p = e_p(\sigma_x, \sigma_y, w_x, w_y)$. This principal eigenvalue approximates $e^{-I(\sigma_x w_x + \sigma_y w_y)}$. To find the order of accuracy of the scheme, first, fix $(\sigma_x, \sigma_y)$, say, $(\sigma_x, \sigma_y) = (0.2, 0.1)$. For the coarse mesh, set $(w_x, w_y) = (\pi/4, \pi/6)$. We can calculate the coarse mesh error

$$e_r = e_p(\sigma_x, \sigma_y, w_x, w_y) - e^{-I(\sigma_x w_x + \sigma_y w_y)}.$$  

(5.12)

Next, halving the wave number, set $(w_x, w_y) = (\pi/8, \pi/12)$. A formula similar to the above yields the corresponding fine mesh error $e_f$. For a scheme to be $m$-th order accurate, (after one time step)

$$\frac{\log(e_r/e_f)}{\log(2)} - 1 \approx m. \quad (5.8)$$

As can be seen in Tables 5.1 and 5.2, the 2D moment schemes using polynomials either of degree $p$ or via tensor products of degree $p \times p$ are accurate to order $2p + 1$.

<table>
<thead>
<tr>
<th>Polynomial Degree</th>
<th>Coarse Mesh Error</th>
<th>Fine Mesh Error</th>
<th>Order of Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1.88 \times 10^{-3}$</td>
<td>$1.28 \times 10^{-4}$</td>
<td>2.87</td>
</tr>
<tr>
<td>2</td>
<td>$2.58 \times 10^{-5}$</td>
<td>$4.26 \times 10^{-7}$</td>
<td>4.92</td>
</tr>
<tr>
<td>3</td>
<td>$1.83 \times 10^{-7}$</td>
<td>$7.47 \times 10^{-10}$</td>
<td>6.94</td>
</tr>
<tr>
<td>4</td>
<td>$7.22 \times 10^{-10}$</td>
<td>$7.49 \times 10^{-13}$</td>
<td>8.91</td>
</tr>
</tbody>
</table>

Table 5.1. Errors and order of accuracy of 2D moment schemes using polynomials of degree $p$; $(\sigma_x, \sigma_y) = (0.2, 0.1)$, coarse mesh $(w_x, w_y) = (\pi/4, \pi/6)$, and fine mesh $(w_x, w_y) = (\pi/8, \pi/12)$.

<table>
<thead>
<tr>
<th>Polynomial Degree</th>
<th>Coarse Mesh Error</th>
<th>Fine Mesh Error</th>
<th>Order of Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1×1</td>
<td>$8.3 \times 10^{-4}$</td>
<td>$5.42 \times 10^{-5}$</td>
<td>2.94</td>
</tr>
<tr>
<td>2×2</td>
<td>$3.09 \times 10^{-6}$</td>
<td>$4.83 \times 10^{-8}$</td>
<td>5.</td>
</tr>
<tr>
<td>3×3</td>
<td>$9.47 \times 10^{-9}$</td>
<td>$3.81 \times 10^{-11}$</td>
<td>6.96</td>
</tr>
<tr>
<td>4×4</td>
<td>$1.14 \times 10^{-11}$</td>
<td>$1.41 \times 10^{-14}$</td>
<td>8.66</td>
</tr>
</tbody>
</table>

Table 5.2. Errors and order of accuracy of 2D moment schemes using tensor products (degree $p \times p$); $(\sigma_x, \sigma_y) = (0.2, 0.1)$, coarse mesh $(w_x, w_y) = (\pi/4, \pi/6)$, and fine mesh $(w_x, w_y) = (\pi/8, \pi/12)$. Due to the larger degree of freedom, errors for schemes via tensor products are smaller than the corresponding errors for schemes via polynomials of degree $p$.  

NASA/TM—2013-218077 22
6 Implicit Space-Time DG Schemes

The moment scheme discussed in Section 3 is closely related to the implicit space-time DG method of this section. When applying the DG approach to discretize time, as shown by La Saint and Raviart (1974) who employed a quadrature of left Radau type, the result is an implicit Runge-Kutta scheme. It was later shown in (Huynh 2009b) that using the FR framework, the right Radau points are a natural choice for quadrature points. The resulting DG method is simpler and is identical to a collocation method called Radau IIA (Lambert 1991, Hairer, Norsett, and Wanner 1991, Hairer and Wanner 1993). We review the DG method and apply the FR approach to time marching to show the close relation between the resulting implicit space-time method and the moment scheme.

Consider the ODE

\[ u'(t) = f(u, t), \quad (6.1a) \]

with initial condition

\[ u(t^0) = u^0. \quad (6.1b) \]

Here, some of the notations are from ODE; as such, the above \( f \) is not related to the flux \( f \) of Section 4. Next, let \( \Delta t \) be the time step size and set \( t^n = n\Delta t, n = 0, 1, 2, \ldots; \) the assumption of constant step size is only for convenience; in practice, time step sizes can vary. When dealing with time, we use \( I = [0, 1] \) and \( t = t^n + \tau \Delta t \) where \( \tau \) varies on \( I \); thus, \( t \) varies on \( I^n = [t^n, t^{n+1}] \).

Assume that the data \( u^n \) at time \( t^n \) is known, we wish to calculate the solution \( u^{n+1} \) at \( t^{n+1} \). For \( n = 0, u^n \) is the initial condition \( u^0 \) in (6.1b).

The DG method seeks to approximate the solution for (6.1) by a function denoted by \( u_h \) which can be and usually is discontinuous across each \( t^n \). On \( I^n \), this function is a polynomial of degree \( s - 1 \), which is defined by its values at \( s \) points at locations corresponding to the stages. Note that when applying these time-stepping methods to conservation laws, if \( p \) is the degree of the spatial polynomial, the number of stages in time is

\[ s = K = p + 1. \]

Due to the property of super accuracy in time discussed later, we could combine parabolic approximation in space or \( p = 2 \) with linear approximation in time or \( s = 2 \). Such combinations remain to be explored.

Assume that the calculation for \( I^{n-1} = [t^{n-1}, t^n] \) is finished, and we have the solution \( u_h \) on that interval. Thus, the value just to the left of \( t^n \) is known, namely \( u_h(t^n^-) \). Recall that the solution can be and usually is discontinuous across each \( t^n \). We assume time marches forward, thus, the value at \( t^n \) is that from the left (small square of Fig. 6.1),

\[ u^n = u_h(t^n^-). \quad (6.2) \]

We wish to calculate \( u_h \) on \( I^n = [t^n, t^{n+1}] \), i.e., to obtain

\[ u_h(t^{n+1^-}) = u^{n+1}. \quad (6.3) \]
Fig. 6.1. DG solution $u_h$ can be and usually is discontinuous across each $t^n$. The value at $t^n$ is chosen to be that from the left (small squares): $u^n = u_h(t^n^-)$.

**DG formulation.** On $I^n = [t^n, t^{n+1}]$, let $\phi_k$ be a set of basis functions for $P_K^q$, $k = 1, \ldots, K$. The solution is approximated by

$$u_h = \sum_{k=1}^{K} u_k \phi_k.$$

Let $\phi$ be a test function, which is in $P_K^q$; typically, $\phi$ is one of the basis functions $\phi_k$, $k = 1, \ldots, K$. Again on $I^n$, the weak form for $u' = f$ is, formally (the correct equation is (6.6) or (6.7)),

$$(u', \phi) = (f, \phi).$$

Let $f_h$ be the projection of $f$ onto $P_K^q$. Since $\phi$ varies on $P_K^q$, we have $(f, \phi) = (f_h, \phi)$, i.e., we can project first, then apply the inner product. The above implies

$$(u', \phi) = (f_h, \phi).$$

We wish to find $u_h$ in $P_K^q$ such that for all $\phi$ in $P_K^q$, again formally,

$$(u_h', \phi) = (f_h, \phi). \quad (6.4)$$

For the DG method, $u_h(t^{n+})$ is allowed to be and is generally different from $u_h(t^{n-}) = u^n$. By applying integration by parts to the left hand side above,

$$[u_h \phi]_t^{n+1} - (u^n, \phi') = (f_h, \phi).$$

Concerning the boundary term, there is no ambiguity with test function $\phi$, but for $u_h$, since it is discontinuous across $t^n$ and $t^{n+1}$, a more precise expression is

$$[u_h \phi]_t^{n+1} = u_h(t^{n+1-})\phi(t^{n+1}) - u_h(t^{n+})\phi(t^n).$$

To involve the initial condition, however, we replace $u_h(t^{n+})$ in the boundary term by $u_h(t^{n-}) = u^n$ in (6.2). Thus, the DG solution $u_h$ is required to satisfy, for all $\phi$ in $P_K^q$,

$$u_h(t^{n+1})\phi(t^{n+1}) - u_h(t^{n-})\phi(t^n) - (u_h, \phi') = (f_h, \phi). \quad (6.6)$$
Applying integration by parts again to \((u_h, \phi')\),

\[-[u_h(t^n-) - u_h(t^{n+})] \phi(t^n) + (u'_h, \phi) = (f_h, \phi).\]  

(6.7)

Note that this time we use \(u_h(t^{n+})\) for the boundary term \(u_h(t^n)\).

Also note that the above, the result of integration by parts twice, is the ‘strong form’ whereas (6.6) is the ‘weak form’.

**Flux Reconstruction formulation.** Our goal is to eliminate the test function in (6.7) so that the integral formulation results in a differential one. To this end, we raise the following question: can we find a polynomial \(q_\alpha = q_\alpha(x)\) on \(I^n\) which possesses the property that, for any \(\phi\) of degree \(K-1\) or less,

\[-\phi(t^n) = (g', \phi).\]  

(6.8)

If such a polynomial exists, we can combine \((u'_h, \phi)\) in (6.7) with \((g', \phi)\) and eliminate the test function. From (6.7), since \(f_h\) is of degree \(K-1\), we require \(g'\) to also be of degree \(K-1\); as a result, \(g\) is of degree \(K\). Applying integration by parts to the right hand side of (6.8), we have

\[-\phi(t^n) = g(t^{n+1}) \phi(t^{n+1}) - g(t^n) \phi(t^n) - (g, \phi').\]  

(6.9)

The above holds if \(g\) satisfies,

\[g(t^n) = 1, \quad g(t^{n+1}) = 0,\]  

(6.10)

and, for all \(\phi\) in \(P_{K-1}\),

\[(g, \phi') = 0.\]  

(6.11)

Since \(\phi\) is of degree \(K-1\), \(\phi'\) is of degree \(K-2\); moreover, \(\phi'\) spans \(P_{K-2}\) as \(\phi\) spans \(P_{K-1}\). Eq. (6.11) implies that \(g\) is orthogonal to \(P_{K-2}\), i.e., for any polynomial \(\varphi\) of degree \(K-2\),

\[(g, \varphi) = 0.\]  

(6.12)

Orthogonality to \(P_{K-2}\) provides \(K-1\) conditions; (6.10) provides the other two.

Let the right Radau polynomial of degree \(K\) be denoted by \(R_{R,K}\) and defined by the conditions that \(R_{R,K}(t^n) = 1\) and \(R_{R,K}\) vanishes at the \(K\) right Radau points on \(I^n\). See Fig. 6.2(a). We wish to show that \(g = R_{R,K}\).

That is, (6.10) and (6.12) hold with \(g\) replaced by \(R_{R,K}\). Indeed, (6.10) is immediate by the definition of \(R_{R,K}\). Next, since \(R_{R,K}\) is of degree \(K\), for any \(\varphi\) of degree \(K-2\), \(R_{R,K}\varphi\) is of degree \(2(K-1)\), and \(R_{R,K}\varphi\) vanishes at the \(K\) right Radau points. In addition, the right Radau quadrature is exact for \(R_{R,K}\varphi\). Consequently, (6.12) holds, and (6.13) follows.

The answer to the question posed for (6.8) is thus positive. Note that \(g = g_{LB}\), which corrects for the jump at the left boundary, is the right Radau polynomial (vanishing at \(t^{n+1}\)). Loosely put, the condition \(g(t^n) = 1\) deals with the jump at the left boundary whereas (6.12) together with \(g(t^{n+1}) = 0\) implies that away from the left boundary, \(g\) approximates the zero function.
We now return to our task of eliminating $\phi$. With $g$ defined above, $(g', \phi) = -\phi(t^n)$. Consequently, by (6.7),

$$[u_h(t^n^-) - u_h(t^n^+)] (g', \phi) + (u_h', \phi) = (f_h, \phi). \tag{6.14}$$

What is crucial here is that $\phi$ can be factored out and canceled and, since $u^n = u_h(t^n^-)$,

$$([u^n - u_h(t^n^+)] g + u_h) = f_h. \tag{6.15}$$

Therefore, $[u^n - u_h(t^n^+)] g + u_h$ is a polynomial of degree $K$ whose value at $t^n = t^{n+1}$ is, by (6.10),

$$[u_h(t^n^-) - u_h(t^n^+)] g(t^n) + u_h(t^n^+) = u_h(t^n^-) = u^n. \tag{6.16}$$

Thus, to deal with the jump $[u_h(t^n^-) - u_h(t^n^+)]$ = $u^n - u_h(t^n^+)$ at $t^n$, we set

$$U = [u_h(t^n^-) - u_h(t^n^+)] g + u_h. \tag{6.17}$$

See Fig. 6.2(b). Differentiating $U$ and employing (6.15), we obtain

$$U' (t) = f_h(t, u_h(t)). \tag{6.18}$$

The function $U$ is the key idea of FR. On $[t^n, t^{n+1}]$, assume that $u_h$ which does not match the boundary condition $u^n$ is given. We can add the correction term $[u^n - u_h(t^n^+)] g$ to $u_h$ as in (6.17) so that the resulting $U$ takes on the left boundary value $u^n$, does not alter the value $u_h(t^{n+1}) = u^{n+1}$, and retains the property of ‘best possible’ approximation to $u_h$. The function $U$ is continuous across cell boundaries (as opposed to $u_h$ which is discontinuous). To put it differently, $u_h$ is of degree $p$; $U$ is of degree $s = p + 1$ determined by $p + 2$ conditions; two conditions are known, namely, $U(t^n) = u^n$ and $U(t^{n+1}) = u_h(t^{n+1})$; the remaining $p = s - 1$ conditions are given by the requirement that the projection of $U$ and $u_h$ onto $P_{p-1}$ are the same. This last requirement is the result of the fact that $g$ is orthogonal to $P_{p-1}$ by (6.12).

Denoting the $K$ right Radau points on $I^n$ by $t^{n,k}$, $k = 1, ..., K$ where $t^{n,K} = t^{n,s} = t^{n+1}$. See also Fig. 3.3. Since $g$ vanishes at these points, (6.17) implies

$$U(t^{n,k}) = u_h(t^{n,k}). \tag{6.19}$$

By employing these Radau points as collocation points (or quadrature points), (6.18) leads to

$$U'(t^{n,k}) = f(t^{n,k}, U^{n,k}). \tag{6.20}$$

Here, we approximate $f_h$, which is of degree $p = K - 1$, by using the $K$ values $f(t^{n,k}, U^{n,k})$, $k = 1, ..., K$.

In summary, to calculate the DG solution, we first obtain the polynomial $U$ by solving the implicit equations (6.20); here, $U$ is of degree $K = s$ interpolating the $K + 1$ values at $t^n$ and at the $K$ right Radau points: $U(t^n) = u^n$ and $U(t^{n,k}) = U^{n,k}$, $k = 1, ..., K$. The DG solution $u_h$ of degree $K - 1$ is defined by the values $U$ at these $K$ right Radau points.

The scheme defined by (6.20) is identical to an implicit collocation Runge-Kutta scheme using the right Radau points and accurate to order $2p + 1 = 2K - 1$ called Radau IIA (Lambert 1991, Hairer, Norsett, and Wanner 1991).
Figure 6.2. (a) The correction function $g$, i.e., the right Radau polynomial of degree $s = 3$ on $l = [0,1]$; (b) $u_h$ is of degree $s - 1 = 2$ and discontinuous across $t^n$; $U$ is of degree $s = 3$, continuous across $t^n$, and defined by: $U(t^n) = u^n$, $U(t^{n+1}) = u_h(t^{n+1})$, and the projections of $U$ and $u_h$ onto $P_{s-2}$ (providing $s - 1$ conditions) are the same.

We now discuss the relation between the current implicit time stepping and the moment scheme (explicit time stepping) of Section 3. For the conservation law $u_t + f_x = 0$, with the same spatial discretization using the Legendre polynomial, the key difference between the two time stepping methods is the following. For the moment scheme, the Cauchy-Kovalevsky procedure, i.e., the space-time Taylor series expansion with no interaction denoted by $w_j^{NL}$ applies. For the implicit scheme, this expansion must be replaced by the solution $u_j^{n+1}$ itself; recall that the Legendre components correspond to $k = 0, ..., p$, and the time stages, $l = 1, ..., s$. As such, the moment method of updating the solution can be employed in an iterative manner to calculate the implicit solution. This subject, however, is beyond the scope of the current paper.

7. A Limiting Technique

As is well-known, high-order methods generate oscillations near a shock or a discontinuity. One way to avoid oscillations is to limit the data so that it is not ‘too steep’ before updating. The problem with limiting is that it can cause a loss of accuracy near extrema. We explore ideas that that can suppress oscillations near shocks while preserving accuracy near extrema here. The limiting procedure below is designed for the piecewise linear case with one spatial dimension. The procedure employs the techniques introduced in (Huynh 1995) and (Suresh and Huynh 1997) but has a stencil of only three cells and is simplified by requiring that the magnitude of the slope lies in a certain range.

To proceed, assume that $u_{j,0}$ and $u_{j,1}$ are given for all $j$. We abbreviate the cell average values:

$$u_{j,0} = u_j.$$

As discussed after (2.9), the first moment $u_{j,1}$ represents half of the normalized slope so that the value at the right interface of cell $j$ is $u_{j,0} + u_{j,1}$. Next, set

$$s_L = u_j - u_{j-1}, \quad s_R = u_{j+1} - u_j,$$

and

$$u_{j,0} = u_j.$$
The limit $s_{\text{Lim1}}$ for the normalized slope is similar to Van Leer's limit (1977) except that we use the magnitudes. It can suppress oscillations near a discontinuity, but has the drawback of causing a loss of accuracy near extrema. With $2|u_{j,1}|$ representing the normalized slope, if

$$2|u_{j,1}| \leq s_{\text{Lim1}},$$

limiting does not alter the slope, and we move on to the next $j$ index. At most of the smooth regions of the flow, the above condition is satisfied; thus, limiting is typically needed only for a small number of cells.

If (7.3) does not hold, the cell is near a discontinuity or an extremum, and the following calculations are carried out. Our goal is to enlarge the limit in a manner that near a shock, it reduces to (7.1), but near an extremum, it is larger and provides enough ‘room’ so that the original slope is not altered. To put it differently, at smooth regions, the results by the moment or DG schemes are highly accurate and should not be altered.

As will be shown by numerical examples, the moment scheme produces, for each cell, a slope that is downwind biased if $\sigma$ is near 0, and upwind biased is $\sigma$ is near 1. Therefore, we take extra care to account for this slope variation.

For any real numbers $x$ and $y$, let $\text{minmod}(x, y)$ be the median of $x, y,$ and 0,

$$\text{minmod}(x, y) = \frac{1}{2} [\text{sign}(x) + \text{sign}(y)] \min(|x|, |y|).$$

At interface $j - 1/2$, set

$$v_{11} = (u_j + u_{j-1})/2, \quad v_{12} = u_{j-1} + u_{j-1,1} - v_{11}, \quad v_{13} = u_j - u_{j,1} - v_{11}. \quad (7.4)$$

Loosely put, we use $v_{11}$ as a pivot; $v_{12}$ is from the left, and $v_{13}$, the right; typically, near a smooth extremum, $v_{12}$ and $v_{13}$ are of the same sign and relatively close to each other. Set

$$v_{14} = \text{minmod}(v_{12}, v_{13}). \quad (7.5)$$

At interface $j + 1/2$, set

$$v_{21} = (u_j + u_{j+1})/2, \quad v_{22} = u_{j+1} - u_{j+1,1} - v_{21}, \quad v_{23} = u_j + u_{j,1} - v_{21},$$

and

$$v_{24} = \text{minmod}(v_{22}, v_{23}). \quad (7.6)$$

Near a smooth extremum, $v_{14}$ and $v_{24}$ are typically of the same sign and relatively close to each other. For the cell $j$, set

$$v_{\text{mm}} = |\text{minmod}(v_{14}, v_{24})|, \quad v_{\text{mx}} = \max(|v_{14}|, |v_{24}|). \quad (7.7)$$

Thus, $v_{\text{mm}}$ and $v_{\text{mx}}$ are also relatively close to each other near a smooth extremum. Near a discontinuity, however, they are far apart. With ‘ra’ short for ratio and ‘ap’ for accuracy preserving, set

$$s_{\text{Lim1}} = 2 \min(|s_L|, |s_R|). \quad (7.2)$$
\[ v_{ra} = \frac{v_{mm}}{v_{mx} + 10^{-2v}}, \quad v_{ap} = 6v_{mm}v_{ra}. \] (7.7)

Note that \( v_{ra} \) is a non-dimensional quantity. Near a discontinuity, it is close to 0; near an extremum, it is relatively close to 1. Therefore, \( v_{ap} \) provides ‘room’ for the interface value near a smooth extremum. Next, set

\[ s_{\min} = \min(|s_L|, |s_R|), \quad s_{\text{Lim2}} = s_{\text{min}} + 2v_{ap}, \quad s_{\text{Lim3}} = \max(s_{\text{Lim1}}, s_{\text{Lim2}}). \] (7.8)

Note the factor 2 accounts for transferring a quantity of type \( \nu \) to type \( s \). Finally, limit \( u_{j,1} \) by

\[ u_{j,1} \leftarrow \text{sign}(u_{j,1}) \min(|u_{j,1}|, s_{\text{Lim3}}/2). \] (7.9)

Note that if (7.3) holds and we still apply the above limiting procedure, the result recovers the original \( u_{j,1} \). Thus, (7.3) only serves the purpose of saving computing time; whether it is applied or not, the final solution is the same.

8. Numerical Results

The following preliminary numerical tests show the results by the piecewise linear moment (or upwind moment) scheme. Recall that for 1D advection, the scheme is identical to Van Leer’s scheme 3.

The first test is for an advection with constant speed 1. The initial condition is a sine wave on the domain \([0, 1]\) with 6 cells and periodic boundary conditions. The continuous (red) curve represents the exact solution, the dots the cell average numerical solutions, and the line segments, the piecewise linear solutions. Figs. 8.1 and 8.2 show the following behavior of the scheme: for each CFL number, the scheme has a ‘preferred mode’ or eigenvector. This eigenvector corresponds to a piecewise linear function with slopes that are downwind biased if \( \sigma \) is close to 0, and upwind biased if \( \sigma \) is close to 1.

Figs. 8.1 shows the results with \( \sigma = 0.1 \). For small \( \sigma \), the solution ‘likes’ downwind biased slopes in the sense that no matter what kind of slopes we start with, the solution will turn them into downwind biased slopes after a few time steps. In (a), for each cell \( j \), we start with slopes biased in the ‘wrong’ direction, i.e., upwind biased slopes defined by the exact interface value \( u_{j-1/2} \) and exact cell average value \( u_j \). Fig. 8.1(c) shows the solution after 10 time steps; the wave travels a distance of one cell. At the maximum, the slope is now slanted downward, opposite to the initial slope in (a). In fact, in (b), after just 4 time steps, the solution already exhibits this ‘preferred mode’ for the slopes: downwind biased for small \( \sigma \). Fig. 8.1(d) shows the solution after 60 time steps; the wave advects 1 period. Comparing (a) and (d), again note the switch in bias of the slopes.

Figs. 8.2 shows the results with a CFL number of 0.9677, which is the exact opposite compared with Figs. 8.1. For CFL number close to 1, the solution ‘likes’ upwind biased slopes. We again start with slopes biased in the ‘wrong’ direction, i.e., downwind biased slopes defined by the exact cell average value \( u_j \) and exact interface value \( u_{j+1/2} \) as shown in Fig. 8.2(a). After 31 time steps, the wave travels a distance of 30 cells, i.e., the profile advects 5 periods. The solution is shown in Fig. 8.2(b), where the slopes are upwind biased.

The above property of biasing toward the upwind or downwind sides holds true for higher-order as well \((p > 1)\). In that case, the \( p + 1 \) right Radau and left Radau points come into play in a manner similar to the linear case above. (The upwind biased linear functions in Fig.8.1(a) interpolate the values at the two left Radau points for each cell, and the downwind biased ones in Fig.8.2(a), the values at the two right Radau points.)
Fig. 8.1. CFL number $\sigma = 0.1$. The continuous (red) curve represents the exact solution. (a) Piecewise linear data with slopes biased in the ‘wrong’ direction, i.e., upwind biased (slanted upward near the maximum); (b) after just 4 time steps, solution exhibits a ‘preferred mode’ for the slopes, which is downwind biased since $\sigma$ is close to 0; (c) solution after 10 time steps; the wave travels a distance of one cell; and (d) solution after 60 time steps; the wave advects 1 period.

Fig. 8.2. CFL number $\sigma = 0.9677$, number of time steps 31, distance of advection 5 periods. The continuous (red) curve represents the exact solution. (a) Piecewise linear data with slopes biased in the ‘wrong’ direction, i.e., downwind biased, or slanted downward near the maximum; (b) the solution exhibits a ‘preferred mode’ for the slopes, which is upwind biased for $\sigma$ close to 1.
Note that the solution in Fig. 8.1(d) is damped more than that in Fig. 8.2(b) in spite of the fact that the wave travels much further in the latter. The reason is that for Fig. 8.2(b), the error associated with the CFL number of 0.9677 is small (σ = 1 has no error), and the number of time steps is only half that of Fig. 8.1(d). In other words, the errors accumulate more in Fig. 8.1(d) due to the larger number of time steps. Since σ = 1 has no error, for the moment scheme (as opposed to the case of RK time stepping), to minimize error, we take time steps as large as possible.

Note that if we use piecewise linear DG with RK time stepping, the solution behaves in a manner similar to that of Fig. 8.1, i.e., it ‘likes’ downwind biased slopes. The moment scheme, however, has this peculiar property of shifting the slopes it ‘likes’ to upwind biased when σ is close to 1. It appears difficult to devise a semi-discrete formulation with this shifting ability.

The second test is again advection. The initial condition consists of a Gaussian, a rectangular, and a semi-ellipse wave on the domain [0,1] with 100 cells and periodic boundary conditions. The CFL number is .8, and the final time is t = 1, i.e., the profile advects 1 period after 125 time steps. The solutions by the standard upwind scheme using the Van Albada limiter and the upwind moment scheme using the limiting procedure in Section 7 are shown in Fig. 8.3. The dots represent the cell average values. Note that the standard upwind solution in (a) has a loss of accuracy (only first-order accurate) near extrema, and discontinuities are smeared. The upwind moment solution has no loss of accuracy near extrema, and discontinuities are resolved with four points.

![Fig. 8.3. Advection problem; 125 time steps with σ = 0.8; distance travelled: 1 period; (a) solution by standard upwind scheme with Van Albada limiter; (b) solution by the moment scheme with the new limiter.](image)

The next test has both shocks and local extrema (Shu and Osher 1989). On [−1, 1], a moving Mach 3 shock interacts with sine waves in density described by

\[
(\rho, u, p) = \begin{cases} 
(3.857, 2.629, 10.333), & -1 \leq x \leq -0.8 \\
(1 + 0.2 \sin(5\pi x), 0, 1), & -0.8 \leq x \leq 1
\end{cases}
\]  

(7.9)

The final time is \( t = 0.36 \). Fig. 8.4 shows the results with 300 cells by standard upwind (with Van Albada limiter) and the upwind moment scheme. Here, for upwind moment, limiting is applied to the characteristic variables. The solid lines represent the solution with 3600 cells by the standard upwind scheme. Again, the upwind moment solution is highly accurate, and accuracy is preserved near extrema.
Shu and Osher’s problem. The upwind moment solution is highly accurate: the shock is well captured and there is no loss of accuracy near extrema.

The last test is the (2D) oblique shock problem (Yee, Warming, and Harten 1983). The domain is \([0,4] \times [0,1]\). The boundary conditions are: left, inflow with \((\rho, u, v, p) = (1., 2.9, 0., 1/1.4)\); top, \((1.7, 2.6193, -0.5063, 1.5282)\); bottom, solid wall; right, outflow. The rectangular uniform mesh is \(80 \times 20\). Here, the limiter is applied in the \(x\) and \(y\) directions when needed (to the characteristic variables). Figure 8.5 shows (a) pressure distribution result by the upwind moment scheme and (b) pressures along \(y = 0.475\) by the upwind moment and the standard upwind schemes. The solid line represents the exact solution. Again, the upwind moment solution is highly accurate. It appears that the slope update via moment calculation can ‘turn’ faster due to the compact stencil compared to the slope update via standard interpolation using neighboring data. For this case, compared with the standard (simplified) second-order upwind scheme with Van Albada’s limiter, computing cost increases by roughly a factor of five.
4. Conclusions and Discussion

In conclusion, two new high-order space-time schemes for conservation laws were introduced: the first is explicit and the second, implicit. The explicit method called the moment scheme achieves a CFL condition of 1 for the case of one-spatial dimension regardless of the degree of approximation. Fourier analyses for the one and two-dimensional cases were carried out. The moment scheme is accurate to order $2p + 1$ if the solution is approximated by piecewise polynomials of degree $p$. The implicit method is a simplified but optimal version of the discontinuous Galerkin scheme applied to time discretization. It reduces to a collocation implicit Runge-Kutta method using the right Radau points called Radau IIA. The explicit and implicit schemes are closely related since they employ the same intermediate time levels, and the former can serve as a key building block in an iterative procedure for the latter. A limiting technique for the piecewise linear scheme that can suppress oscillations near discontinuities while preserving accuracy near extrema was presented. Preliminary numerical results were shown.

Whereas the current findings are encouraging, further research on both time-stepping and limiting is needed.

References


Current high-order methods such as discontinuous Galerkin and/or flux reconstruction can provide effective discretization for the spatial derivatives. Together with a time discretization, such methods result in either too small a time step size in the case of an explicit scheme or a very large system in the case of an implicit one. To tackle these problems, two new high-order space-time schemes for conservation laws are introduced: the first is explicit and the second, implicit. The explicit method here, also called the moment scheme, achieves a Courant-Friedrichs-Lewy (CFL) condition of 1 for the case of one-spatial dimension regardless of the degree of the polynomial approximation. (For standard explicit methods, if the spatial approximation is of degree $p$, then the time step sizes are typically proportional to $1/p^2$). Fourier analyses for the one and two-dimensional cases are carried out. The property of super accuracy (or super convergence) is discussed. The implicit method is a simplified but optimal version of the discontinuous Galerkin scheme applied to time. It reduces to a collocation implicit Runge-Kutta (RK) method for ordinary differential equations (ODE) called Radau IIA. The explicit and implicit schemes are closely related since they employ the same intermediate time levels, and the former can serve as a key building block in an iterative procedure for the latter. A limiting technique for the piecewise linear scheme is also discussed. The technique can suppress oscillations near a discontinuity while preserving accuracy near extrema. Preliminary numerical results are shown.