Efficient Characteristic Projection in Upwind Difference Schemes for Hyperbolic Systems (The Complementary Projection Method)

Ronald P. Fedkiw
Barry Merriman
Stanley Osher *

October 20, 1997

Abstract

The standard construction of upwind difference schemes for hyperbolic systems of conservation laws requires the full eigensystem of the Jacobian matrix. This system is used to define the transformation into and out of the characteristic scalar fields, where upwind differencing is meaningful.

When the Jacobian has a repeated eigenvalue, the associated normalized eigenvectors are not uniquely determined, and an arbitrary choice of eigenvectors must be made to span the characteristic subspace. In this report we point out that it is possible to avoid this arbitrary choice entirely. Instead, a complementary projection technique can be used to formulate upwind differencing without specifying a basis.

For systems with eigenvalues of high multiplicity, this approach simplifies the analytical and programming effort and reduces the computational cost. Numerical experiments show no significant difference in computed results between this formulation and the traditional one, and thus we recommend its use for these types of problems.

*Research supported in part by ONR N00014-97-1-0027, ARPA URI-ONR-N00014-92-J-1890, NSF #DMS 94-04942, and ARO DAAH04-95-1-0155

This complementary projection method has other applications. For example, it can be used to extend upwind schemes to some weakly hyperbolic systems. These lack complete eigensystems, so the traditional form of characteristic upwinding is not possible.

1 Introduction

The standard formulation of upwind difference schemes for hyperbolic systems requires finding the Jacobian matrix of the flux function and the associated eigensystem (eigenvalues and left and right eigenvectors). The left eigenvectors define the transformation into the characteristic fields, the associated eigenvalues define upwind directions for these fields, and the right eigenvectors define the transformation back to the primitive variables.

These characteristic upwind schemes are generally considered to give the highest quality numerical results. There is a vast literature on this subject (see e.g. [4] and the references therein). Their only drawback is that they require specifying a complete eigensystem for the problem. In practice, this can involve considerable analytical work, as well as some complications when the eigensystem lacks uniqueness (or even existence). In this paper we point out that, in many cases, the most problematic portion of the eigensystem can be avoided entirely.

As motivation, consider a system which has a repeated eigenvalue (characteristic speed). A common example is the compressible, multi-species, multidimensional Euler equations [3], where the convective flow velocity is an eigenvalue repeated once for each species and each spatial dimension (see section 4.3). In such a system, the distinct eigenvalues have corresponding unique eigenvectors (up to scalar multiples), but the eigenvectors for the repeated eigenvalue are not unique. The eigen-subspace is well defined, but an arbitrary choice of spanning eigenvectors must be made to obtain a complete eigensystem. These arbitrary vectors may form the great majority of the eigensystem.

When designing a numerical method for such a system, various criteria can be applied to help select one eigenbasis from the infinitely many choices. For example, one can look for eigenvectors that are as sparse as possible, in order to save time projecting into and out of characteristic fields. One can also demand that the the left and right eigenvector matrices be numerically well conditioned (i.e. determinant near 1). Still, there is a high degree of arbitrariness left over, and for degenerate systems, there are typically a variety of eigensystems presented in the literature.

Our goal here is to present an alternative approach which eliminates the need to find the ambiguous eigenbasis. The basic idea is to project data directly into the characteristic subspace by using the complement of the projection operator defined by the unambiguous part of the eigensystem. Componentwise upwind differencing can be applied directly to this characteristic *vector* field, in contrast to the usual approach of upwinding characteristic scalar fields.

2 The Complementary Projection Method (CPM)

To describe the complementary projection technique in detail, we will show how it relates to the standard characteristic decomposition used in upwind discretization of a system of n hyperbolic conservation laws in one spatial dimension,

$$\vec{U}_t + [\vec{F}(\vec{U})]_x = 0. \tag{1}$$

Let the Jacobian of the flux function, $\partial\vec{F}(\vec{U})/\partial\vec{U}$, have left and right eigenvectors $\vec{L}^i$ and $\vec{R}^i$, with associated eigenvalues λ^i, $i = 1, \ldots, n$. The left and right eigenvectors are further required to be mutually orthonormal, i.e. $\vec{L}^i \cdot \vec{R}^j = \delta_{ij}$. Equivalently, the row matrix of left eigenvectors , $\mathcal{L}$, and column matrix of right eigenvectors , $\mathcal{R}$, are inverses: $\mathcal{L}\mathcal{R} = \mathcal{R}\mathcal{L} = I$.

Given this complete eigensystem, any upwind difference scheme defined for scalar equations can be extended to the hyperbolic system via a "characteristic decomposition". This can be described fairly generally as follows: the spatial discretization of $[\vec{F}(\vec{U})]_x$ is expressed as a difference of fluxes between two grid cell walls. Thus the essential step is to compute the flux at a grid cell wall, $\vec{F}_w$, given the fluxes, $\vec{F}(\vec{U})$, at the nearby grid cell centers [4].

The first step in defining the flux at a particular cell wall is to project the vector fluxes at each cell center into "scalar fluxes for the i^{th} characteristic field", defined by $f^i = \vec{L}^i_w \cdot \vec{F}(\vec{U})$. Here $\vec{L}^i_w$, $\vec{R}^i_w$ and λ^i_w are used to denote left and right eigenvectors and eigenvalues evaluated at the wall in some fashion. Note the assumed orthonormality implies we can write the original vector flux in terms of these scalar fluxes as

$$\vec{F}(\vec{U}) = f^1\vec{R}^1_w + f^2\vec{R}^2_w + \cdots + f^n\vec{R}^n_w. \tag{2}$$

This shows that we can think of $f^i\vec{R}^i_w$ as the vector contribution to the total flux from the i^{th} characteristic scalar flux, f^i.

Next, for each scalar field i, the cell center characteristic fluxes, f^i, are interpolated to the cell wall of interest in an upwind fashion with the upwind direction defined by the corresponding "characteristic speed" at the wall, λ^i_w. This yields the scalar characteristic wall flux, f^i_w.

Finally, the desired total wall flux vector is defined as the sum of all the characteristic vector contributions,

$$\vec{F}_w = f^1_w\vec{R}^1_w + f^2_w\vec{R}^2_w + \cdots + f^n_w\vec{R}^n_w. \tag{3}$$

To introduce the alternative approach, suppose that from the n eigenvalues we have a p-fold repeated eigenvalue. Without loss of generality, we will assume that the first p eigenvalues, $\lambda_w^1 = \lambda_w^2 = \cdots = \lambda_w^p$, are repeated. The corresponding p dimensional characteristic subspace is the span of $\{\vec{L}_w^1, \ldots, \vec{L}_w^p\}$. The part of the original cell center flux vector $\vec{F}(\vec{U})$ that lies in this characteristic subspace is

$$\vec{\mathcal{F}} \;=\; f^1\vec{R}_w^1 + f^2\vec{R}_w^2 + \cdots + f^p\vec{R}_w^p. \tag{4}$$

Note that all of the characteristic fields contributing to $\vec{\mathcal{F}}$ have the same upwind direction for interpolation, since their characteristic speeds (eigenvalues) are identical.

Since $\vec{\mathcal{F}}$ has a well-defined upwind direction, upwind differencing is possible without decomposing $\vec{\mathcal{F}}$ further into the individual scalar fluxes. Instead, we can directly apply upwind interpolation to the cell center values of the vector $\vec{\mathcal{F}}$, in a component by component fashion. Let $\vec{\mathcal{F}}_w$ denote the resulting flux value interpolated to the cell wall of interest. Then, the net cell wall flux required in the numerical method can be defined via the "partially decomposed" form

$$\vec{F}_w = \vec{\mathcal{F}}_w + f_w^{p+1}\vec{R}_w^{p+1} + f_w^{p+2}\vec{R}_w^{p+2} + \cdots + f_w^n\vec{R}_w^n, \tag{5}$$

instead of the fully decomposed form in equation 3.

So far there is no obvious benefit to this formulation. The critical observation that makes this partial decomposition useful is that we can compute $\vec{\mathcal{F}}$ without knowing the basis of left and right eigenvectors used to define it in equation 4. Instead, it is simply the complement of the remaining part of the decomposition, i.e.

$$\vec{\mathcal{F}} \;=\; \vec{F}(\vec{U}) - \left(f^{p+1}\vec{R}_w^{p+1} + f^{p+2}\vec{R}_w^{p+2} + \cdots + f^n\vec{R}_w^n\right). \tag{6}$$

Thus, in order to apply a fully upwind scheme to a problem where one characteristic subspace has a repeated eigenvalue, all that is required are the left and right eigenvectors corresponding to complementary subspace. In practice, we simply define the cell wall flux via equation 5 and compute $\vec{\mathcal{F}}_w$ from $\vec{\mathcal{F}}$ as calculated in equation 6, which requires only the left and right eigenvectors associated with $\{\lambda_w^{p+1}, \ldots, \lambda_w^n\}$. There is never any need to choose a basis for—or characterize in any direct way—the subspace associated with the repeated eigenvalue.

The basic method of complementary projection is exceedingly simple. In the following remarks, we elaborate on its properties.

3 Remarks

Remark 1 For a system with a p-fold repeated eigenvalue, the above argument shows the entire vector field $\vec{\mathcal{F}}$ has not only a definite upwind direction, it actually has a well-defined characteristic speed. Thus, further decomposition into scalar characteristic fields does not provide any greater insight into the time evolution of the data. Instead, it is simply an arbitrary decomposition into scalars that have no greater significance than the scalar components of $\vec{\mathcal{F}}$ itself.

Thus it seems that if we consider only the quality of the computed solution, there is no motivation for further decomposition of $\vec{\mathcal{F}}$. Our numerical experiments on standard test problems confirm this—i.e. there is no significant difference between solutions computed using full or complementary projection.

Moreover, by not decomposing $\vec{\mathcal{F}}$ we can avoid the arbitrary selection of spanning left and right eigenvectors for the degenerate subspace. This represents a reduction in the need for tedious analysis, programming, and publishing, and can also noticeably reduce computational costs.

Based on these factors, we strongly encourage practitioners to use the complementary projection formulation for systems with a repeated eigenvalue.

Remark 2 When applied to a system with a repeated eigenvalue having a large multiplicity, complementary projection may require fewer operations and therefore result in a faster code. Let us compare the computational costs of full projection versus complementary projection in detail.

We will express the cost as a function of the dimension of the undecomposed subspace, p, and the overall system size n. We will compare only the cost of the portion of the problem that is treated differently in each method, i.e. the cost of treating the fields with the p-fold repeated eigenvalue.

The computational cost of a full decomposition into the p scalar fields is $pW_1 + pW_0$, where W_1 is the average cost of projecting into and out of a field, and W_0 is the average cost of doing a scalar upwind interpolation. The field projections require computing $\vec{L}_w^i \cdot \vec{F}(\vec{U})$ and $f_w^i \vec{R}_w^i$. These are operations on n-vectors, so the cost is proportional to n, and $W_1 = \alpha n$. The cost of a scalar interpolation, W_0, has no dependence on system size n or multiplicity p. Thus the total cost of the standard decomposition has the form $\alpha pn + pW_0$.

In the complementary projection method, the computational cost is nW_2 $+nW_0$, where W_2 is the work per component required to compute $\vec{\mathcal{F}}$ via equation 6. This is proportional to the number of terms which is $n-p$, and $W_2 = \beta(n-p)$. Thus the overall cost of complementary projection takes the form $\beta(n-p)n + nW_0$.

In the limit of a large system with a large multiplicity, the cost of the traditional method scales like pn, while the new method scales like $(n-p)n$. If we further assume that the repeated eigenvalue dominates the system, so that p dominates $n-p$ (e.g. in the equations for multi-species flow, $n-p=2$ as $p, n \to \infty$), then the complementary projection method is asymptotically less costly than the traditional approach.

This analysis makes it clear that complementary projection carries out more upwind interpolations than the traditional approach (always n, instead of p), but it can save even more work by avoiding p scalar field projections. However, use of a vectorizing or parallel computer could potentially alter this conclusion (e.g. by reducing the cost of the vector inner products used for full projection).

Also note that it is possible to minimize W_1 by making the eigensystems $\mathcal{L}$ and $\mathcal{R}$ collectively as sparse as possible. For example, consider multi-species flow $n-p=2$, and thus the complementary projection method scales like n. If the eigensystem was dense, then the full projection method scales like n^2, while the sparse eigensystem chosen in [3] yields a full projection method which scales like n.

Remark 3 Consider this projection technique on a more abstract level. We are able to project onto the target subspace (and define $\vec{\mathcal{F}}$) without a basis because we know the complementary projection explicitly. That is, $\vec{\mathcal{F}} = (I-P)\vec{F}$, where P is the projection defined explicitly by the known part of the eigensystem. Since we have all the information needed to perform P, we can perform the complement, $I-P$, with no additional information.

This algebraic trick can only be used to define a single basis-free projection operator: we can project onto a subspace S_1 without a basis for it, given a basis for its complement. But if we need projection operators for two linearly independent subspaces S_1 and S_2, it is clear that we must select a basis for at least one of them.

For example, this means that if the eigensystem of a flux function $\vec{F}(\vec{U})$ has two distinct, repeated eigenvalues, it is not possible to separately upwind each associated characteristic subspace without finding a basis for *either one*. An eigenbasis must be selected for one of the subspaces, and then the other can be treated without a basis.

Remark 4 In contrast to Remark 3, there is a special situation in which multiple complementary projections can be used efficiently within a single decomposition. If the flux Jacobian matrix has a block diagonal structure, it is possible to apply complementary projection separately within each block. In particular, within each major block it is possible treat a single repeated eigenvalue without ever constructing an eigenbasis for the associated characteristic subspace.

To clarify the procedure in this case, let B_1 and B_2 be the image spaces in R^n associated with two distinct blocks in the diagonal of the Jacobian. Consider subspaces $S_1 \subset B_1$ and $S_2 \subset B_2$. We will show it is possible to define the projections onto S_1 and S_2 without specifying a basis for either one.

Let P_i be the projection onto the complement of S_i in B_i. Construction of P_i requires knowing only a basis in R^n for the complement of S_i in B_i—which does not require choosing a basis for the other subspace, S_j. Then, projection onto S_i is defined in complementary fashion as $Q_i - P_i$, where Q_i is the projection from R^n onto B_i. Note that the $n \times n$ matrix Q_i is trivial, since it is simply an identity matrix where the corresponding block, B_i, in the Jacobian is located, and zero elsewhere.

Remark 5 Another important situation where this complementary projection can be of use is the upwind discretization of a weakly hyperbolic system. These systems have characteristic subspaces that lack a basis of eigenvectors. The simplest example of such a system is

$$u_t + au_x + v_x = 0 \tag{7}$$

$$v_t + av_x = 0, \tag{8}$$

where a is a real constant. The Jacobian is an irreducible Jordan block; it has repeated eigenvalue a, but only a one dimensional family of eigenvectors spanned by $(1, 0)$. The traditional upwind technique requires a full eigensystem, and so it does not even apply. However, this system can be upwinded with componentwise a-upwind differencing and special techniques for weakly hyperbolic systems which damp out the unwanted linear growth.

More generally, a subsystem locally equivalent can occur as a block inside a larger hyperbolic system. The traditional upwind technique requiring a full eigensystem again does not apply. Still, as long as there is an eigenbasis for the other characteristic fields, these fields can be upwinded in the standard way and the complement, $\mathcal{F}$, can be solved componentwise, using special techniques for weakly hyperbolic systems. Note that the standard alternative is to treat the entire system with the weakly hyperbolic solver and thus degrade the quality of the solution in the fields which are not weakly hyperbolic. For an example of a system of practical interest, where this technique can be applied, see [2].

In practice, a complicated hyperbolic system may develop a repeated eigenvalue or become weakly hyperbolic (eigenvectors become dependent) *transiently* during a calculation. A full characteristic decomposition is appropriate as the primary numerical method, but some special "back-up" treatment is required when these degenerate cases arise. The method of complementary characteristic projection provides a convenient "back-up" formula for the flux in these circumstances.

Remark 6 Complementary projection can be used to upwind difference a characteristic subspace composed of characteristic fields moving with *different speeds*, as long as they all have the same upwind direction. I.e., equations 5 and 6 provide a stable upwind differencing of the system as long as $\lambda_w^1, \ldots, \lambda_w^p$ are all of the same sign.

For an extreme example, one could lump together all the positive speed fields and apply componentwise upwinding with *no decomposition*, knowing only a basis for the negative speed fields (which would in contrast be treated by standard decomposition into scalar fields). If it so happened that all the fields were positive at some cell wall, upwind differencing could be applied in a componentwise fashion to compute the cell wall flux $\vec{F}_w$, with no characteristic field projections at all (the $p = n$ case).

However, lumping together fields moving at different speeds into a single undecomposed subspace is not as attractive as it is for the case of a repeated eigenvalue. The repeated eigenvalue case is free of any negative consequences, while the more general application of complementary projection has several deficiencies.

One major deficit is that there is no savings in analytical work—formulas for the entire eigensystem must be available. To see why, note that since the characteristic speeds are different they will not *always* have the same upwind direction. Under the right conditions they will differ in sign, and the associated fields cannot be lumped into a subspace with a single upwind direction. Since one must be prepared for this to occur, the characteristic scalar decomposition must be available as an option for all fields, and so the associated eigenvectors must be known even if they are seldom used. Still, lumping together different fields can give a major savings in computational work, because we only need this information when eigenvalues change sign.

There is another complication which can make complementary projection undesirable, in this non-repeated eigenvalue case. If fields moving at different speeds are lumped into a single subspace, there is the potential for a loss of resolution, when two discontinuities propagating in different fields at different speeds move close together. In each individual characteristic scalar field, there is only an isolated discontinuity; this will be resolved to the extent possible by the chosen upwind scheme for all time. However, in a vector mixture of two discontinuous fields, both discontinuities could be present in the same vector component. Since they move with different speeds, the faster discontinuity could overtake the slower one. No matter how fine the grid, as the discontinuities pass through each other there will be a temporary loss of resolution. The resulting errors—which are avoided in the full decomposition—can seriously corrupt the calculation.

Remark 7 In contrast to the loss of resolution difficulties mentioned in Remark 6, such problems do not arise during calculations in the repeated eigenvalue case. Even if multiple discontinuities are present in different degenerate fields, because they move at the same speed, they cannot merge. A high accuracy upwind scheme will maintain resolution as long as the initial data was resolved by the grid. Further, even when it is possible in principle, there is no practical way to isolate the discontinuities by projecting them into different degenerate scalar fields. This is because there is no simple way to determine which of the infinitely many distinct decompositions will yield the desired separation of features.

Returning to the considerations in Remark 1, note that this reasoning does suggest one possible accuracy-related motivation for performing a full characteristic decomposition in the repeated eigenvalue case. Namely, the possibility that one of the non-unique decompositions might yield a smoother set of scalar fields for scalar upwind differencing than those provided by the components of the vector data, $\vec{\mathcal{F}}$. However, there does not seem to be any practical, general way of determining which of the infinitely many possible decompositions would yield the smoothest set of scalar fields. In the absence of such knowledge, complementary projection remains our recommended method for treating systems with repeated eigenvalues.

4 Examples

We illustrate this approach by considering a few common hyperbolic systems of equations. All calculations were carried out using the ENO method described in [4], though complementary projection can be used with any characteristic upwinding scheme. (The eigenvalues and eigenvectors are all evaluated at cell walls. In what follows, we will assume that this is given and drop the subscript 'w' as a notational change only.)

4.1 1D Euler Equations

This simple system provides a clear illustration of the operational differences between full decomposition and complementary projection.

The 1D Euler equations are

$$\vec{U}_t + [\vec{F}(\vec{U})]_x = 0, \tag{9}$$

$$\vec{U} = \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix}, \quad \vec{F}(\vec{U}) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ (E+p)u \end{pmatrix}, \tag{10}$$

where

$$E = -p + \frac{\rho u^2}{2} + \rho h, \qquad h(T) = h^f + \int_0^T c_p(s)ds. \tag{11}$$

Here t is time, x is the spatial dimension, ρ is the density, u is the velocity, E is the energy per unit volume, h is enthalpy per unit mass, h^f is the heat of formation or enthalpy at $0K$, c_p is the specific heat at constant pressure, and p is the pressure [3].

We assume pressure is a function (or table look-up) of the density and internal energy per unit mass, $p = p(\rho, e)$, and denote its corresponding partial derivatives by p_ρ and p_e. The Jacobian matrix of $\vec{F}(\vec{U})$ has eigenvalues

$$\lambda^1 = u - c, \quad \lambda^2 = u, \quad \lambda^3 = u + c, \tag{12}$$

and eigenvectors

$$\vec{L}^1 = \left(\frac{b_2}{2} + \frac{u}{2c}, \frac{-b_1 u}{2} - \frac{1}{2c}, \frac{b_1}{2}\right), \tag{13}$$

$$\vec{L}^2 = (1 - b_2, b_1 u, -b_1)\,, \tag{14}$$

$$\vec{L}^3 = \left(\frac{b_2}{2} - \frac{u}{2c}, \frac{-b_1 u}{2} + \frac{1}{2c}, \frac{b_1}{2}\right), \tag{15}$$

$$\vec{R}^1 = \begin{pmatrix} 1 \\ u - c \\ H - uc \end{pmatrix}, \quad \vec{R}^2 = \begin{pmatrix} 1 \\ u \\ H - \frac{1}{b_1} \end{pmatrix}, \quad \vec{R}^3 = \begin{pmatrix} 1 \\ u + c \\ H + uc \end{pmatrix}, \tag{16}$$

where

$$c = \sqrt{p_\rho + \frac{p p_e}{\rho^2}}, \qquad H = \frac{E + p}{\rho}, \tag{17}$$

$$b_1 = \frac{p_e}{\rho c^2}, \qquad b_2 = 1 + b_1 u^2 - b_1 H. \tag{18}$$

Since all the eigenvalues are distinct, the above eigensystem is unique (up to scalar multiples) and provides a good reference for comparison of full projection and complementary projection methods. We will use complementary projection to avoid decomposing the characteristic field moving with the flow velocity u (the 2nd field, or u-field).

The vector flux contributions from the 1st and 3rd fields are computed in the usual way, using eigenvector projection. Next we form

$$\vec{\mathcal{F}} = \vec{F}(\vec{U}) - \vec{L}^1 \vec{F}(\vec{U})\vec{R}^1 - \vec{L}^3 \vec{F}(\vec{U})\vec{R}^3. \tag{19}$$

Note that $\vec{\mathcal{F}}$ is precisely the unprojected 2nd field $\vec{L}^2\vec{F}(\vec{U})\vec{R}^2$, yet it is obtained without use of $\vec{L}^2$ or $\vec{R}^2$. We apply componentwise upwinding to $\vec{\mathcal{F}}$, in the u-upwind direction. Since $\vec{\mathcal{F}}$ is a 3 dimensional vector, 3 upwind interpolations are required. The resulting vector flux is combined with the contributions from the 1st and 3rd fields to get the total flux.

In contrast, the standard method would project the 3 dimensional $\vec{\mathcal{F}}$ into the 1 dimensional scalar u-field and apply the upwind interpolation only once. Thus, the complementary projection method is more costly in this case.

In numerical experiments, we have noticed no difference between the complementary calculations in the case of the 1D Euler equations, except that they run slower (as predicted since the savings occurs as the number of repeated eigenvalues increases). Even in the case of two shocks intersecting [1] —which causes a transient loss of resolution and is therefore more sensitive to different schemes—the numerical results agree quite nicely. Neither scheme seems to have an advantage over the other as far as accuracy or quality of the computed solutions are concerned.

As a representative example, consider Example 7 in [5] which is the celebrated Woodward and Colella "bang-bang" problem. Using the CPM, the convection step was 23 percent slower (as predicted), although the quality of the solution is the same. In fact the pointwise relative difference between the two solutions is on the order of 10^{-12}.

4.2 2D Euler Equations

This is a common system with a repeated eigenvalue. It also illustrates how complementary projection applies equally well to systems with multiple spatial dimensions.

The 2D Euler equations are

$$\vec{U}_t + [\vec{F}(\vec{U})]_x + [\vec{G}(\vec{U})]_y = 0, \tag{20}$$

$$\vec{U} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ E \end{pmatrix}, \quad \vec{F}(\vec{U}) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho u v \\ (E+p)u \end{pmatrix}, \quad \vec{G}(\vec{U}) = \begin{pmatrix} \rho v \\ \rho u v \\ \rho v^2 + p \\ (E+p)v \end{pmatrix}, \tag{21}$$

where

$$E = -p + \frac{\rho(u^2+v^2)}{2} + \rho h, \qquad h(T) = h^f + \int_0^T c_p(s)ds. \tag{22}$$

Here y is the second spatial dimension, and v is the velocity in that dimension [3]. As in the 1D Euler equations, $p = p(\rho, e)$.

The eigenvalues and (one possible set of) eigenvectors for the Jacobian matrix of $\vec{F}(\vec{U})$ are obtained by setting $A = 1$ and $B = 0$ in the following formulas, while those for the Jacobian of $\vec{G}(\vec{U})$ are obtained with $A = 0$ and $B = 1$.

The eigenvalues are

$$\lambda^1 = \hat{u} - c, \quad \lambda^2 = \lambda^3 = \hat{u}, \quad \lambda^4 = \hat{u} + c, \tag{23}$$

and the eigenvectors are

$$\vec{L}^1 = \left(\frac{b_2}{2} + \frac{\hat{u}}{2c}, -\frac{b_1 u}{2} - \frac{A}{2c}, -\frac{b_1 v}{2} - \frac{B}{2c}, \frac{b_1}{2}\right), \tag{24}$$

$$\vec{L}^2 = \left(\frac{1-b_2}{2} - \frac{\hat{v}}{2c}, \frac{b_1 u}{2} - \frac{B}{2c}, \frac{b_1 v}{2} + \frac{A}{2c}, -\frac{b_1}{2}\right), \tag{25}$$

$$\vec{L}^3 = \left(\frac{1-b_2}{2} + \frac{\hat{v}}{2c}, \frac{b_1 u}{2} + \frac{B}{2c}, \frac{b_1 v}{2} - \frac{A}{2c}, -\frac{b_1}{2}\right), \tag{26}$$

$$\vec{L}^4 = \left(\frac{b_2}{2} - \frac{\hat{u}}{2c}, -\frac{b_1 u}{2} + \frac{A}{2c}, -\frac{b_1 v}{2} + \frac{B}{2c}, \frac{b_1}{2}\right), \tag{27}$$

$$\vec{R}^1 = \begin{pmatrix} 1 \\ u - Ac \\ v - Bc \\ H - \hat{u}c \end{pmatrix}, \quad \vec{R}^2 = \begin{pmatrix} 1 \\ u - Bc \\ v + Ac \\ H - \frac{1}{b_1} + \hat{v}c \end{pmatrix}, \tag{28}$$

$$\vec{R}^3 = \begin{pmatrix} 1 \\ u + Bc \\ v - Ac \\ H - \frac{1}{b_1} - \hat{v}c \end{pmatrix}, \quad \vec{R}^4 = \begin{pmatrix} 1 \\ u + Ac \\ v + Bc \\ H + \hat{u}c \end{pmatrix}, \tag{29}$$

where

$$q^2 = u^2 + v^2, \quad \hat{u} = Au + Bv, \quad \hat{v} = Av - Bu, \tag{30}$$

$$c = \sqrt{p_\rho + \frac{p p_e}{\rho^2}}, \qquad H = \frac{E+p}{\rho}, \tag{31}$$

$$b_1 = \frac{p_e}{\rho c^2}, \qquad b_2 = 1 + b_1 q^2 - b_1 H. \tag{32}$$

Note that the choice of eigenvectors 1 and 4 is unique (up to scalar multiples), but the choice for eigenvectors 2 and 3 is not unique. Any two independent vectors from the spans of eigenvectors 2 and 3 could be used instead.

To avoid choosing any basis for this ambiguous subspace, we apply the standard characteristic scalar projections to the 1st and 4th fields, and then apply complementary projection for the u-fields:

$$\vec{\mathcal{F}} = \vec{F}(\vec{U}) - \vec{L}^1 \vec{F}(\vec{U}) \vec{R}^1 - \vec{L}^4 \vec{F}(\vec{U}) \vec{R}^4. \tag{33}$$

We upwind difference $\vec{\mathcal{F}}$ componentwise in the u-upwind direction. The result is then combined with the flux contributions from the 1st and 4th fields. Note that the eigenvectors for the 2nd and 3rd fields were not needed for the discretization.

Four upwind interpolations are required to compute the contribution from the repeated eigenvalue for the complementary projection method, instead of only 2 upwind interpolations if full projection were used. However, we also save two projections.

For a standard dimension by dimension discretization, the complementary projection method applies independently to the flux for the second spatial dimension. Using the eigenvectors appropriate for $\vec{G}(\vec{U})$, we form

$$\vec{\mathcal{G}} = \vec{G}(\vec{U}) - \vec{L}^1 \vec{G}(\vec{U}) \vec{R}^1 - \vec{L}^4 \vec{G}(\vec{U}) \vec{R}^4 \tag{34}$$

and upwind difference $\vec{\mathcal{G}}$ in the v-upwind direction.

4.3 Multi-species Euler Equations

The multi-species Euler equations provide an important example of a hyperbolic system with an eigenvalue repeated many times. Complementary projection becomes quite attractive for such systems, due to the large analytical and computational savings.

The 2D Euler equations for multi-species flow with a total of N species are

$$\vec{U}_t + [\vec{F}(\vec{U})]_x + [\vec{G}(\vec{U})]_y = 0, \tag{35}$$

$$\vec{U} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ E \\ \rho Y_1 \\ \vdots \\ \rho Y_{N-1} \end{pmatrix}, \quad \vec{F}(\vec{U}) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ (E+p)u \\ \rho u Y_1 \\ \vdots \\ \rho u Y_{N-1} \end{pmatrix}, \quad \vec{G}(\vec{U}) = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ (E+p)v \\ \rho v Y_1 \\ \vdots \\ \rho v Y_{N-1}, \end{pmatrix} \tag{36}$$

where

$$E = -p + \frac{\rho(u^2+v^2)}{2} + \rho\left(\sum_{i=1}^{N} Y_i h_i\right), \qquad h_i(T) = h_i^f + \int_0^T c_{p,i}(s)ds.(37)$$

Here, Y_i is the mass fraction of species i, h_i is the enthalpy per unit mass of species i, h_i^f is the heat of formation of species i, and $c_{p,i}$ is the specific heat at constant pressure of species i [3]. Note that $Y_N = 1 - \sum_{i=1}^{N-1} Y_i$.

The pressure is a function of the density, internal energy per unit mass, and the mass fractions, $p = p(\rho, e, Y_1, \cdots, Y_{N-1})$, and the corresponding partial derivatives are denoted by p_ρ, p_e and p_{Y_i}.

The eigenvalues and (one possible set of) eigenvectors for the Jacobian matrix of $\vec{F}(\vec{U})$, are obtained by setting $A = 1$ and $B = 0$ in the following formulas, while those for the Jacobian matrix of $\vec{G}(\vec{U})$ use $A = 0$ and $B = 1$.

The eigenvalues are

$$\lambda^1 = \hat{u} - c, \tag{38}$$

$$\lambda^2 = \cdots = \lambda^{N+2} = \hat{u}, \tag{39}$$

$$\lambda^{N+3} = \hat{u} + c, \tag{40}$$

Note the $(N+1)$-fold repeated eigenvalue.

A particularly sparse choice of left eigenvectors are given by the rows of the matrix

$$\begin{pmatrix} \frac{b_2}{2} + \frac{\hat{u}}{2c} + \frac{b_3}{2} & -\frac{b_1 u}{2} - \frac{A}{2c} & -\frac{b_1 v}{2} - \frac{B}{2c} & \frac{b_1}{2} & \frac{-b_1 z_1}{2} & \cdots & \frac{-b_1 z_{N-1}}{2} \\ 1 - b_2 - b_3 & b_1 u & b_1 v & -b_1 & b_1 z_1 & \cdots & b_1 z_{N-1} \\ \hat{v} & B & -A & 0 & 0 & \cdots & 0 \\ -Y_1 & 0 & 0 & 0 & & & \\ \vdots & \vdots & \vdots & \vdots & & I & \\ -Y_{N-1} & 0 & 0 & 0 & & & \\ \frac{b_2}{2} - \frac{\hat{u}}{2c} + \frac{b_3}{2} & -\frac{b_1 u}{2} + \frac{A}{2c} & -\frac{b_1 v}{2} + \frac{B}{2c} & \frac{b_1}{2} & \frac{-b_1 z_1}{2} & \cdots & \frac{-b_1 z_{N-1}}{2} \end{pmatrix}, \tag{41}$$

and the corresponding sparse choice of right eigenvectors are given by the columns of the matrix

$$\begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 1 \\ u - Ac & u & B & 0 & \cdots & 0 & u + Ac \\ v - Bc & v & -A & 0 & \cdots & 0 & v + Bc \\ H - \hat{u}c & H - \frac{1}{b_1} & -\hat{v} & z_1 & \cdots & z_{N-1} & H + \hat{u}c \\ Y_1 & Y_1 & 0 & & & & Y_1 \\ \vdots & \vdots & \vdots & & I & & \vdots \\ Y_{N-1} & Y_{N-1} & 0 & & & & Y_{N-1} \end{pmatrix}, \tag{42}$$

where I is the $N-1$ by $N-1$ identity matrix and

$$q^2 = u^2 + v^2, \quad \hat{u} = Au + Bv, \quad \hat{v} = Av - Bu, \tag{43}$$

$$c = \sqrt{p_\rho + \frac{pp_e}{\rho^2}}, \qquad H = \frac{E + p}{\rho}, \tag{44}$$

$$b_1 = \frac{p_e}{\rho c^2}, \qquad b_2 = 1 + b_1 q^2 - b_1 H, \tag{45}$$

$$b_3 = b_1 \sum_{i=1}^{N-1} Y_i z_i, \qquad z_i = \frac{-p_{Y_i}}{p_e}. \tag{46}$$

Note that the eigenvectors 2 through $N+2$ are not uniquely determined. Each one could be replaced by an arbitrary linear combination of those shown, as long as linear independence is maintained. This gives an indication of the enormous range of possible eigensystems that could be used, though in practice they would yield similar computed solutions. (The costs may differ, though, depending on sparseness.)

In particular, all the fields in the eigensystem for $\vec{F}(\vec{U})$ have eigenvalue u, except for the first and last. To avoid choosing any eigenbasis for this degenerate subspace, we apply the standard projection method to the first and last fields, and treat all the u-fields by complementary projection,

$$\vec{\mathcal{F}} = \vec{F}(\vec{U}) - \vec{L}^1 \vec{F}(\vec{U}) \vec{R}^1 - \vec{L}^{N+3} \vec{F}(\vec{U}) \vec{R}^{N+3}. \tag{47}$$

We upwind difference $\vec{\mathcal{F}}$ in the u-upwind direction. The resulting cell wall flux is combined with the wall flux contributions from the first and the last fields to yield the net numerical wall flux.

A total of $N + 3$ upwind interpolations are required to compute the contribution from the repeated eigenvalue for the complementary projection method, instead of only $N + 1$ upwind schemes if projection is used. Thus only 2 extra upwind interpolations are needed to eliminate $N + 1$ characteristic projections. Starting at about 4 species, we expect the complementary projection method to be less costly. Moreover, there is no need to ever construct most of the eigensystem shown above. Had this approach been available for previous work, it would have allowed a major savings in analytic work, as well as programming and reporting.

For a dimension by dimension discretization, the same considerations apply to the flux in the other spatial dimension. Using the first and last eigenvectors appropriate for $\vec{G}(\vec{U})$, we form

$$\vec{\mathcal{G}} = \vec{G}(\vec{U}) - \vec{L}^1 \vec{G}(\vec{U}) \vec{R}^1 - \vec{L}^{N+3} \vec{G}(\vec{U}) \vec{R}^{N+3} \tag{48}$$

and upwind $\vec{\mathcal{G}}$ in the v-upwind direction.

Numerical experiments were carried out on examples from [3] and [1]. For the case of nine species, the complementary calculations were faster than the traditional approach, even though the set of eigenvectors for the repeated eigenvalue had been carefully chosen to be as sparse as possible and the implementation took full advantage of the sparseness. As the number of species is increased, the percentage savings in CPU time increases as well.

As a particularly difficult example, we compute example 5.1 from [3] which is a chemically reacting "Sod" shock tube problem. The convection step was 59 percent faster using the CPM, with no degradation in the quality of the solution. We show the solution in figure 1 and the relative difference in figure 2. The differences are on the order of about 1 percent, and only the underresolved species (HO_2 and H_2O_2) differ by as much as 2 percent. The largest differences occur near large gradients in the solution where the two schemes capture discontinuities in slightly different ways. These differences are too small to be seen by the naked eye and have no effect on the size or strength of the discontinuities, only the intermediate points which span the jumps. In fact, both schemes give the result depicted in figure 1.

We note that the standard scheme and the CPM have approximately the same CPU time when the eigenvalue is repeated 4 times. That is, for 4 species (3 mass fraction equations) in 1 spatial dimension, for 3 species (2 mass fraction equations) in 2 spatial dimensions, or for 2 species (1 mass fraction equation) in 3 spatial dimensions. After this point, the CPM is faster with the gains in CPU time proportional to the number of species.

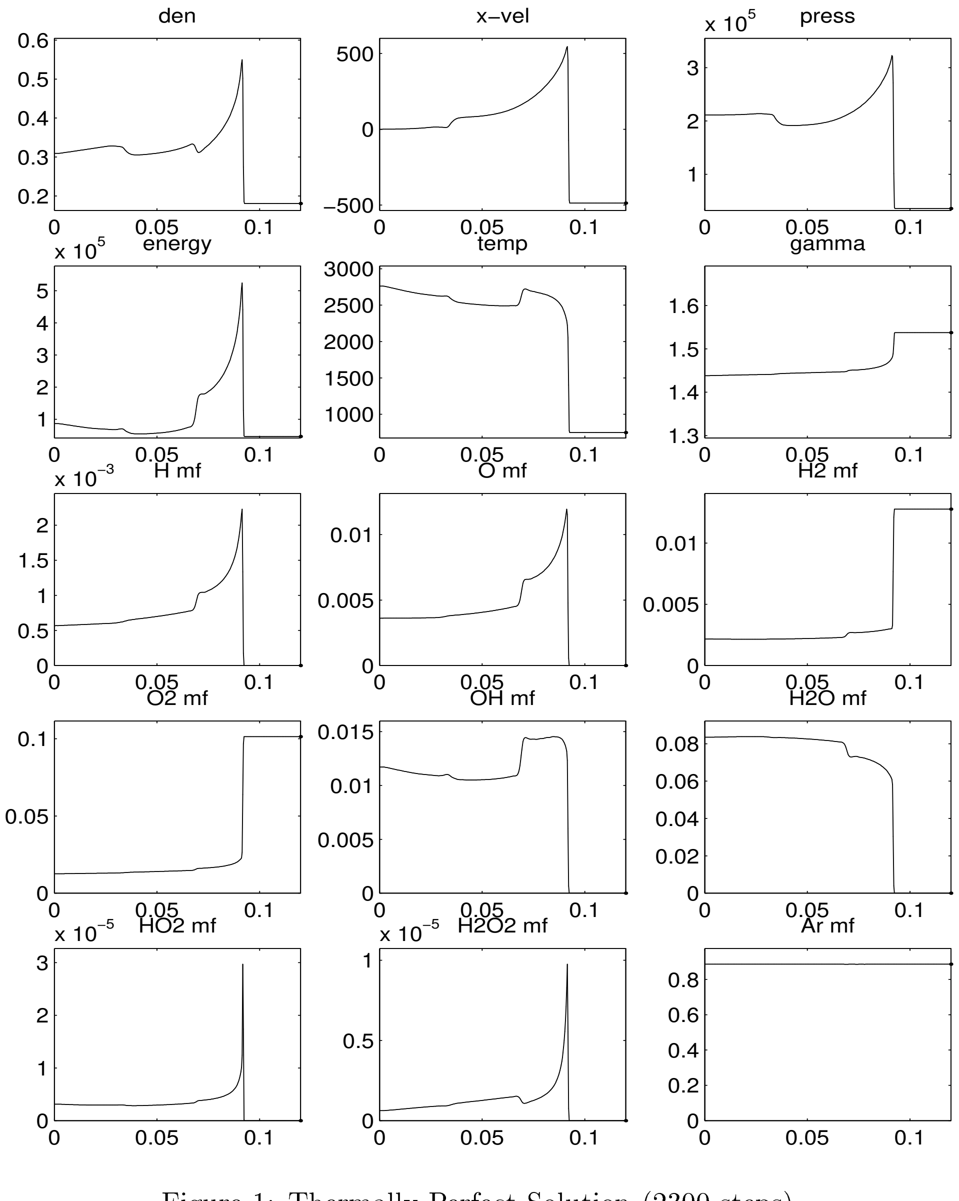

Figure 1: Thermally Perfect Solution (2300 steps)

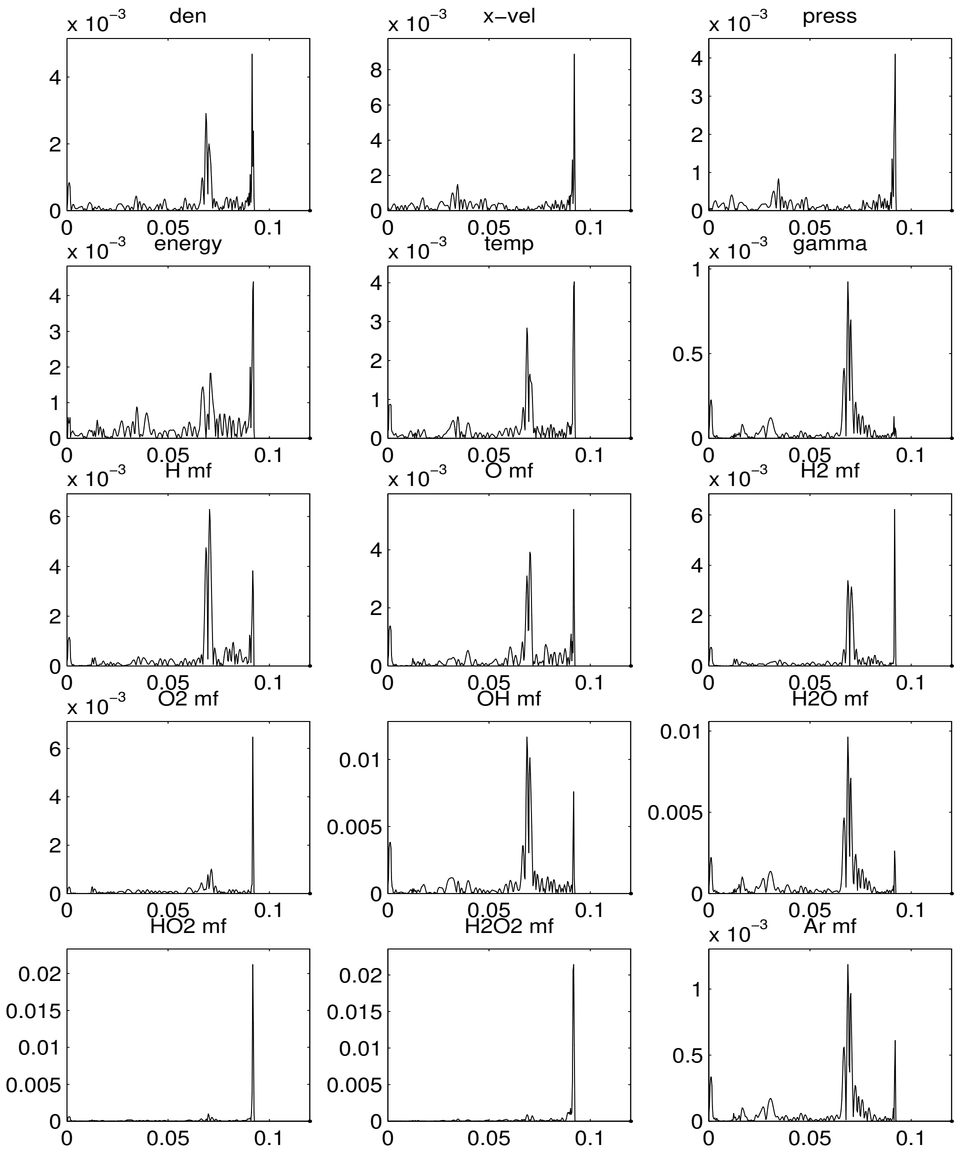

Figure 2: Relative Difference (2300 steps)

5 Conclusions

We have introduced the complementary projection method for use in upwind difference schemes for systems of hyperbolic conservation laws. This approach provides an alternative to full characteristic decomposition of a characteristic subspace, if all associated characteristic speeds are of the same sign. Instead, projection onto the subspace is defined as the complement of the projection onto the remaining characteristic spaces. This allows the application of any upwind method without the need of an eigenbasis for the specified subspace. All that is required is a complete eigenbasis for the complementary subspace.

This has particular application to problems with a repeated eigenvalue. There the eigenspace associated with the repeated eigenvalue does not have a unique eigenbasis. The complementary projection method eliminates the need to construct such a basis, without any negative side effects, reducing the analytical and programming effort required to apply upwind differencing. Our analysis and experiments also show that avoiding the decomposition can save computational time in practical multi-species compressible flow calculations, with no significant change in computed results.

We recommend that in the future, practitioners use the complementary projection method to treat hyperbolic systems with repeated eigenvalues.

This method has other potential applications. The most interesting is formulating upwind difference schemes for weakly hyperbolic systems. For these systems, a complete eigensystem does not exist, and thus traditional upwind characteristic schemes do not apply. In contrast, the complimentary projection method provides a simple way to extend upwind differencing to these systems.

References

[1] Fedkiw, R., *A Survey of Chemically Reacting, Compressible Flows*, UCLA (Dissertation), 1996.

[2] Fedkiw, R., Merriman, B., and Osher, S., *High order numerical methods for weak hyperbolic systems; the BN multiphase explosives model*, UCLA CAM Report (in preparation).

[3] Fedkiw, R., Merriman, B., and Osher, S., *High accuracy numerical methods for thermally perfect gas flows with chemistry*, J. Computational Physics 132, 175-190 (1997).

[4] Fedkiw, R., Merriman, B., Donat, R., and Osher, S., *The Penultimate Scheme for Systems of Conservation Laws: Finite Difference ENO with Marquina's Flux Splitting*, UCLA CAM Report 96-18, July 1996, http://www.math.ucla.edu/applied/cam/.

[5] Shu, C.W. and Osher, S., *Efficient Implementation of Essentially Non-Oscillatory Shock Capturing Schemes II (two)*, J. Computational Physics, v. 83, (1989), pp 32-78.