Abstract

A new high-resolution and genuinely multidimensional numerical method for solving conservation laws is being developed. It was designed to avoid the limitations of the traditional methods, and was built from ground zero with extensive physics considerations. Nevertheless, its foundation is mathematically simple enough that one can build from it a coherent, robust, efficient and accurate numerical framework.

Two basic beliefs that set the new method apart from the established methods are at the core of its development. The first belief is that, in order to capture physics more efficiently and realistically, the modeling focus should be placed on the original integral form of the physical conservation laws, rather than the differential form. The latter form follows from the integral form under the additional assumption that the physical solution is smooth, *an assumption that is difficult to realize numerically in a region of rapid change, such as a boundary layer or a shock*. The second belief is that, with proper modeling of the integral and differential forms themselves, the resulting numerical solution should automatically be consistent with the properties derived from the integral and differential forms, e.g., the jump conditions across a shock and the properties of characteristics. Therefore a much simpler and more robust method can be developed by not using the above derived properties explicitly.

Specifically, to capture physics as fully as possible, the method requires that: (i) space and time be unified and treated as a single entity; (ii) both local and global flux conservation in space and time be enforced; and (iii) a multidimensional scheme be constructed without using the dimensional-splitting approach, such that multidimensional effects and source terms (which are scalars) can be modeled more realistically.

To simplify mathematics and broaden its applicability as much as possible, the method attempts to use the simplest logical structures and approximation techniques. Specifically, (i) it uses a staggered space-time mesh such that flux at any interface separating two conservation elements can be evaluated internally in a simpler and more consistent manner, without using a separate flux model; (ii) it does not use many well-established techniques such as Riemann solvers, flux splittings and monotonicity constraints such that the limitations and complications associated with them can be avoided; and (iii) it does not use special techniques that are not applicable to more general problems.

Furthermore, triangles in 2D space and tetrahedrons in 3D space are used as the basic building blocks of the spatial meshes, such that the method (i) can be used to construct 2D and 3D non-dissipative schemes in a natural manner; and (ii) is compatible with the simplest unstructured meshes.

Note that while numerical dissipation is required for shock capturing, it may also result in annihilation of small disturbances such as sound waves and, in the case of flow with a large Reynolds number, may overwhelm physical dissipation. To overcome this difficulty, two different and mutually complementary types of adjustable numerical dissipation are introduced in the present development.

1. Introduction

Since its inception in 1991 [1], the space-time conservation element and solution element method [1–32] has been used to obtain highly accurate numerical solutions for flow problems involving shocks, rarefaction waves, acoustic waves, vortices, ZND detonation waves, shock/acoustic waves/vortices interactions, dam-break and hydraulic jump. This article is the first of a series of papers that will provide a systematic and up-to-date description of this new method (hereafter it may be referred to abbreviatedly as the space-time CE/SE method or simply as the CE/SE method). To answer frequently-asked questions and clarify possible misconceptions, we shall begin this paper with (i) an overall view of the CE/SE method and its capabilities, and (ii) an extensive comparison of the basic concepts used by the CE/SE method with those used by other methods.

Currently, the field of computational fluid dynamics (CFD) represents a diverse collection of numerical methods, with each of them having its own limitations. Generally speaking, these methods were originally introduced to solve special classes of flow problems. Development of the CE/SE method is motivated by a desire to build a brand new, more general and coherent numerical framework that avoids the limitations of the traditional methods.

The new method is built on a set of design principles given in [2]. They include: (i) To enforce both local and global flux conservation in space and time, with flux evaluation at an interface being an integral part of the solution procedure and requiring no interpolation or extrapolation; (ii) To unify space and time and treat them as a single entity; (iii) To consider mesh values of dependent variables and their derivatives as independent variables, to be solved for simultaneously; (iv) To use only local discrete variables rather than global variables like the expansion coefficients used in spectral methods; (v) To define conservation elements and solution elements such that the simplest stencil will result; (vi) To require that, as much as possible, a numerical analogue be constructed so as to share with the corresponding physical equations the same space-time invariant properties, such that numerical dissipation can be minimized [5,10,24]; (vii) *To exclude the use of characteristics-based techniques (such as Riemann solvers)*; and (viii) To avoid the use of ad hoc techniques as much as possible.

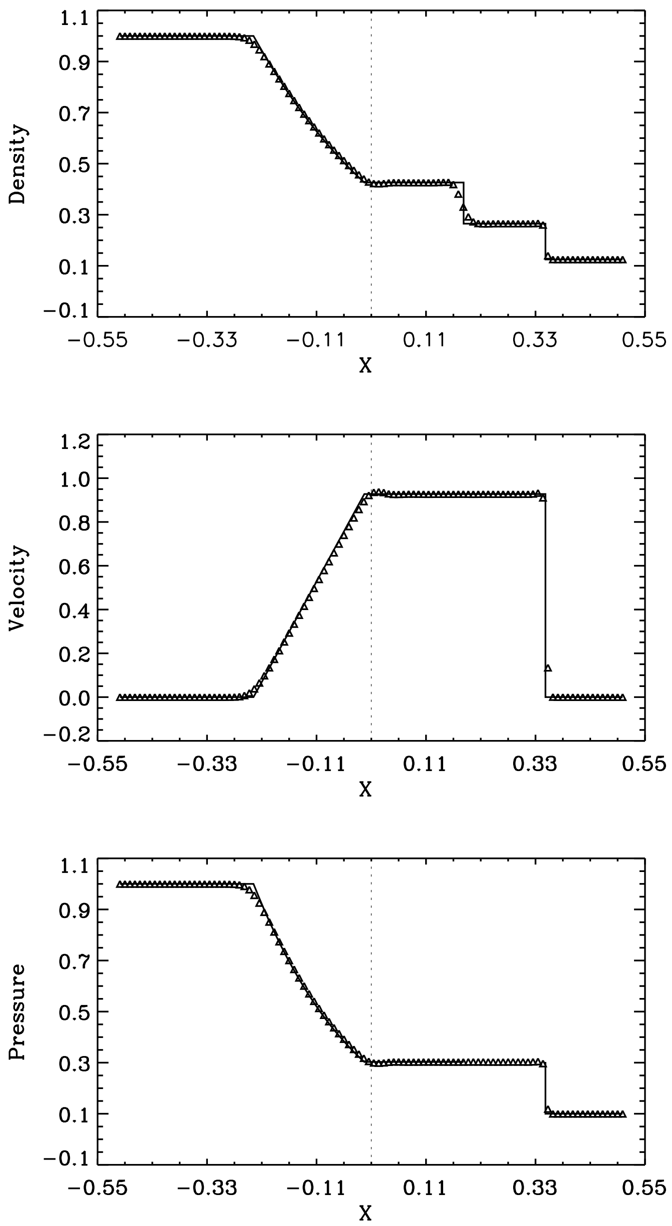
Moreover, the development of the CE/SE method is also guided by two basic beliefs that set it apart from the established methods. The first belief is that, in order to capture physics more efficiently and realistically, the modeling focus should be placed on the original integral form of the physical conservation laws, rather than the differential form. The latter form follows from the integral form under the additional assumption that the physical solution is smooth, *an assumption that is difficult to realize numerically in a region of rapid change, such as a boundary layer or a shock*. The second belief is that, with proper modeling of the integral and differential forms themselves, the resulting numerical solution should automatically be consistent with the properties derived from the integral and differential forms, e.g., the jump conditions across a shock and the properties of characteristics. In other words, a much simpler and more robust method can be developed by not using the above derived properties explicitly.

With the exception of the Navier-Stokes solver, all the 1D schemes described in [2] have been extended to become their 2D counterparts [9–11,14]. A more complete account of these new 2D schemes and their applications will be given in this and the following papers [3,4]. It will be shown in Sec. 3 that the spatial meshes used in these schemes are built from triangles—*in such a manner that the resulting meshes are completely different from those used in the finite element method.* As a result, these schemes are (i) compatible with the simplest unstructured meshes [31], and (ii) constructed *without using the dimensional-splitting approach, i.e., without applying a 1D scheme in each coordinate direction.* The dimensional-splitting approach is widely used in the construction of multidimensional upwind schemes. Unfortunately, this approach is flawed in several respects [33]. In particular, because a source term is not aligned with a special direction, it is difficult to imagine how this dimensional-splitting approach, *in a logically consistent manner*, can be used to solve a multidimensional problem involving source terms, such as those modeling chemical energy release.

Moreover, as will be shown shortly, because the CE/SE 2D schemes share with their 1D versions the same design principles, not only is the extension to 2D a straightforward matter, each of the new 2D schemes also shares with its 1D version virtually identical fundamental characteristics.

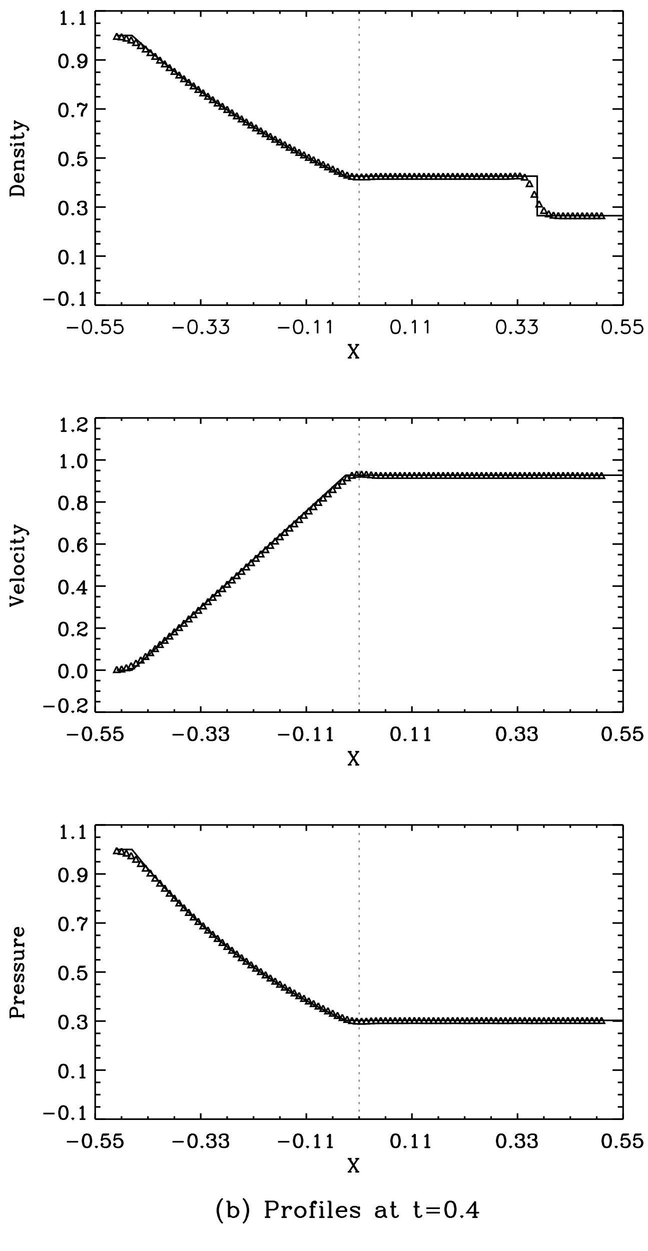
At this juncture, note that monotonicity conditions are not observed by general flow fields, e.g., those involving ZND detonation waves [21]. As a result, techniques involving monotonicity constraints are not used in the present development.

To give the reader, in advance, a concrete example that demonstrate the validity of the two basic beliefs referred to earlier, a self-contained Fortran program is listed in Appendix A. It is a CE/SE solver [23] for an extended Sod's shock tube problem that is the original Sod's problem [38] with the additional complication of imposing a non-reflecting boundary condition at each end of the computational domain. Note that the flow under consideration contains discontinuities and, *relative to the computational frame*, is subsonic throughout. It is well known that implementing a non-reflecting boundary condition for a subsonic flow is much more difficult than doing the same for a supersonic flow. This difficulty is further exacerbated by the fact that the traditional non-reflecting boundary conditions, e.g., the characteristic, the radiation (asymptotic), the buffer-zone, and the absorbing boundary conditions [39-44] are all based on an assumption that is not valid for the present case, i.e., that the flow is continuous. In spite of the fact that solving the present extended Sod's problem is substantially more difficult than the original Sod's problem, the main loop in the program listed herein contains only 39 Fortran statements. Not only is it very small in size, this program has a very simple logical structure. With the exception of a single "if" statement used to identify the time levels at which the non-reflecting boundary conditions must be imposed, it contains no conditional Fortran statements or functions such as "if", "amax", or "amin" that are often used in programs implementing high-resolution upwind methods. The small size of the listed program reflects the simplicity of the techniques employed by the CE/SE method to capture shock waves. It also results from the fact that the non-reflecting boundary conditions used in the present solver are the simple extrapolation conditions Eqs. (2.66) and (2.67) given in Sec. 2. They are much simpler than the traditional non-reflecting boundary conditions. On the other hand, the absence of Fortran conditional

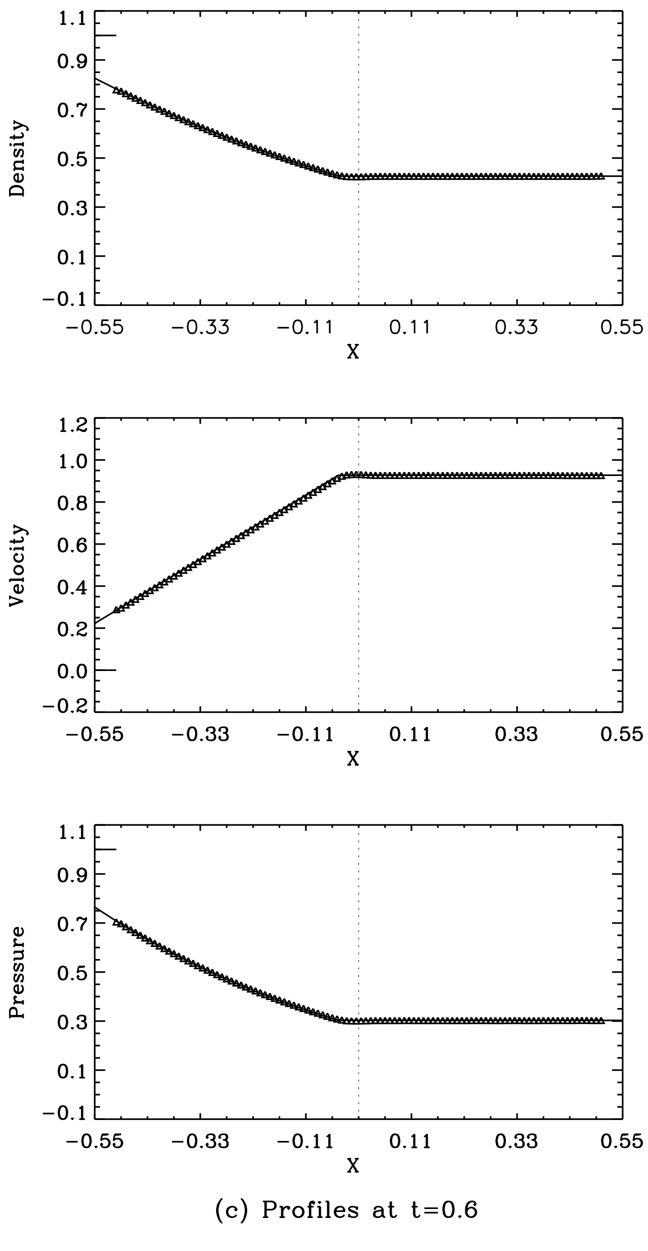


(a) Profiles at t=0.2

Figure 1: The CE/SE solution of the extended Sod's problem using the boundary conditions Eqs. (2.66) and (2.67) ($\Delta t = 0.004$, $\Delta x = 0.01$, CFL$\approx$0.88, $\epsilon = 0.5$, $\alpha = 1$).



(b) Profiles at t=0.4



(c) Profiles at t=0.6

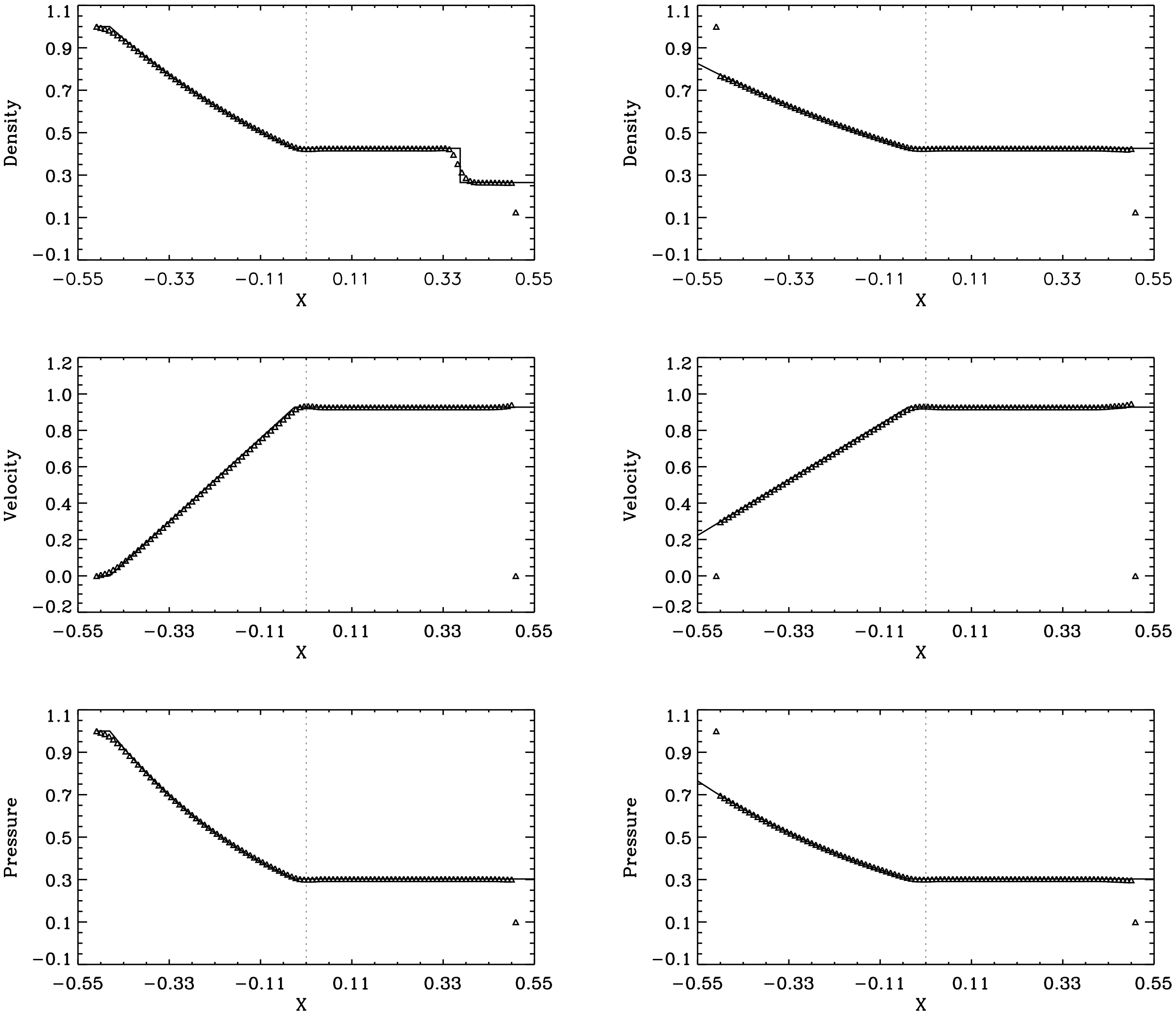
Figure 1: (continued).

statements is a result of avoiding the use of ad hoc techniques. In spite of its small size and simple logical structure, according to the numerical results generated by the listed program (presented here as Figs. 1(a)–(c), with the numerical results and the exact solutions denoted by triangles and solid lines, respectively; see also [23]), the present solver is capable of generating nearly perfect non-reflecting solutions using the same time-step size from $t = 0$. Note that, at $t = 10$, all the waves have exited the computational domain, i.e., the exact solution is constant within it. The theoretical values of density, velocity, and pressure are approximately 0.4262000, 0.9277462 and 0.3030000, respectively. The maximum magnitudes of the errors in the numerically computed values of density, velocity, and pressure at $t = 10$ are approximately 0.0004, 0.0007, and 0.0004, respectively.

Note that Eqs. (2.66) and (2.67) represent only one of many sets of simple and robust non-reflecting boundary conditions developed especially for the CE/SE method [23]. Behind this development is *a radical new concept based entirely on an assumption about the space-time flux distribution in the neighborhood of a spatial boundary.* As a result, implementation of these CE/SE non-reflecting boundary conditions does not require the use of characteristics-based techniques.

To further demonstrate the nontraditional nature of the CE/SE method, the numerical results generated using the *steady-state* non-reflecting boundary conditions that were introduced and rigorously justified in [23] will also be presented here. Consider an alternate CE/SE solver that differs from the above CE/SE solver only in the fact that the steady-state boundary conditions Eq. (2.68) given in Sec. 2 are now taking the place of Eqs. (2.66) and (2.67). At $t = 0.2$, the waves generated in the interior of the computational domain have not yet reached the boundaries. In this case, with the given initial conditions (i.e., two different uniform states separated by a discontinuity located at the dead center of the domain), each of the above two solvers yield the same uniform solution in the vincinity of the right or left boundary. As a result, at $t = 0.2$, the numerical results generated by the alternate solver are identical to those shown in Fig. 1(a). The numerical results of the alternate solver at $t = 0.4$ are shown in Fig. 2(a). It is seen that, by this time, the shock wave has passed cleanly through the right boundary. There is good agreement between the numerical solution and the exact solution everywhere *in the interior* except for a slight disagreement in the vicinity of the right boundary. Note that the right boundary values, *which do not vary with time*, are discontinuous with respect to the neighboring interior values. The numerical results at $t = 0.6$ are shown in Fig. 2(b). As seen from the density profile, by this time, the contact discontinuity has also passed through the right boundary. Agreement between the numerical solution and the exact solution continue to be good in the interior. However, both left and right boundary values are now discontinuous with respect to the neighboring interior values.

Note that several recent applications [13,16,17,26,28] of the CE/SE method to 2D aeroacoustics problems reveal that: (i) the trivial nature of implementing CE/SE non-reflecting boundary conditions is manifested even for 2D problems; (ii) accuracy of the numerical results for *nonlinear* Euler problems is comparable to that of a 4-6th order compact difference scheme, even though nominally the CE/SE solver used is only of 2nd-order accuracy; and (iii) most importantly, the CE/SE method is capable of accurately modeling both small disturbances and strong shocks, and thus provides a unique tool for solving flow problems



(a) Profiles at t=0.4 (b) Profiles at t=0.6

Figure 2: The CE/SE solution of the extended Sod's problem using the boundary condition Eq. (2.68) ($\Delta t = 0.004$, $\Delta x = 0.01$, CFL$\approx$0.88, $\epsilon = 0.5$, $\alpha = 1$).

where the interactions between sound waves and shocks are important, such as the noise field around a supersonic over- and under-expanded jet. The fact listed in item (i) is more fundamental in nature, and will be further discussed in a separate paper. The following comments pertain to items (ii) and (iii):

(a) Assuming the same order of accuracy, generally speaking, the accuracy of a scheme that enforces the space-time flux-conservation property is higher than that of a scheme that does not. A compact scheme generally does not enforce the flux-conservation property of the nonlinear Euler equations. On the contrary, not only is the present scheme flux-conserving, its accuracy in nonlinear calculations is enhanced by its surprisingly small dispersive errors [2,8,13,16,17]. Moreover, the nominal order of accuracy of an Euler solver is determined assuming a linearized form of the Euler equations. Thus its significance with respect to a *highly* nonlinear solution of the Euler equations may be questionable.

(b) *while numerical dissipation is required for shock resolution, it may also result in annihilation of small disturbances such as sound waves.* Thus, a solver that can handle both small disturbances and strong shocks must be able to overcome this difficulty. It will be explained shortly that the CE/SE method is intrinsically endowed with this capability. On the other hand, a high-resolution upwind scheme that focuses only on shock resolution may introduce too much numerical dissipation [45].

Next we shall review briefly the inviscid version of the a-μ scheme described in [2]. In addition to providing a historical perspective, the review will remove, once and for all, any lingering doubt from the reader's mind that the CE/SE method indeed differs substantially in both concept and methodology from the well-established methods. In particular, it will give in advance answers to questions such as: (i) is there any difference between the space-time elements used here and those used in the finite element method? and (ii) what are the key differences between the CE/SE method and other finite volume methods?

To proceed, consider an initial-value problem involving the PDE

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0 \tag{1.1}$$

where the convection speed a is a constant. The exact solution to any such problem has three fundamental properties: (i) it does not dissipate with time; (ii) its value at a spatial point at a later time has a finite domain of dependence (a point) at an earlier time; and (iii) it is completely determined by the initial data at a given time. Ideally, a numerical solution for Eq. (1.1) should also possess the same three properties. Because (i) a solution of a *dissipative* numerical scheme will dissipate with time, (ii) the value of a solution of an *implicit* scheme at any point (x, t) is dependent on all initial data, and all the boundary data up to the time t, and (iii) the unique determination of a solution by a scheme involving more than two time levels requires the specification of the data at at least the first two time levels, an ideal solver must be a *two-level, explicit, and non-dissipative (i.e., neutrally stable)* *scheme*. In 1991, the first solver known to the authors that satisfies the above conditions was reported in [1]. Because this new solver models Eq. (1.1) which is characterized by the parameter a, it is referred to as the a scheme. The a scheme is non-dissipative if the Courant number is less than unity.

At this juncture, the reader may wonder what the merit is of constructing a neutrally stable scheme. After all, it is well known that its nonlinear extensions generally are unstable. To address this question, the significance of constructing such a scheme and the critical role it plays in the development of the CE/SE method will be discussed immediately.

To proceed, note that there are several explicit and implicit extensions [2,12,25] of the a scheme which are solvers for

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} - \mu\frac{\partial^2 u}{\partial x^2} = 0 \tag{1.2}$$

Here the viscosity coefficient $\mu(\geq 0)$ is a constant. Because Eq. (1.2) is characterized by the parameters a and μ, these extensions are referred to as either the explicit a-μ schemes or the implicit a-μ schemes. Each of these schemes reduces to the non-dissipative a scheme when $\mu = 0$. As a result, each of them has the property that *the numerical dissipation of its solutions approaches zero as the physical dissipation approaches zero*.

The above property is important because of the following observation: with a few exceptions, the numerical solution of a time-marching problem generally is contaminated by numerical dissipation. For a nearly inviscid problem, e.g., flow at a large Reynolds number, numerical dissipation may overwhelm physical dissipation and cause a complete distortion of the solution. To avoid such a difficulty, ideally a CE/SE solver for Eq. (1.2) or for the Navier-Stokes equations should possess the above special property. Obviously the development of such a solver must be preceded by that of a neutrally stable solver of Eq. (1.1).

The problem of physical dissipation being overwhelmed by numerical dissipation does not exist for a pure convection problem. However, as explained in the earlier discussion about the delicate nature of simulating small disturbances in the presence of shocks, numerical dissipation must still be handled carefully in this case.

Note that numerical dissipation traditionally is adjusted by varying the magnitude of added artificial dissipation terms. However, *after being stripped of these added artificial dissipation terms, almost every traditional scheme (such as the Lax-Wendroff scheme) is still not free from inherent numerical dissipation. Hence, numerical dissipation generally cannot be avoided completely using the traditional approach.*

This completes the discussion about the roles of non-dissipative schemes in the current development. To proceed further, the construction of the 1D a scheme will be described briefly.

Let $x_1 = x$, and $x_2 = t$ be considered as the coordinates of a two-dimensional Euclidean space E_2. By using Gauss' divergence theorem in the space-time E_2, it can be shown that

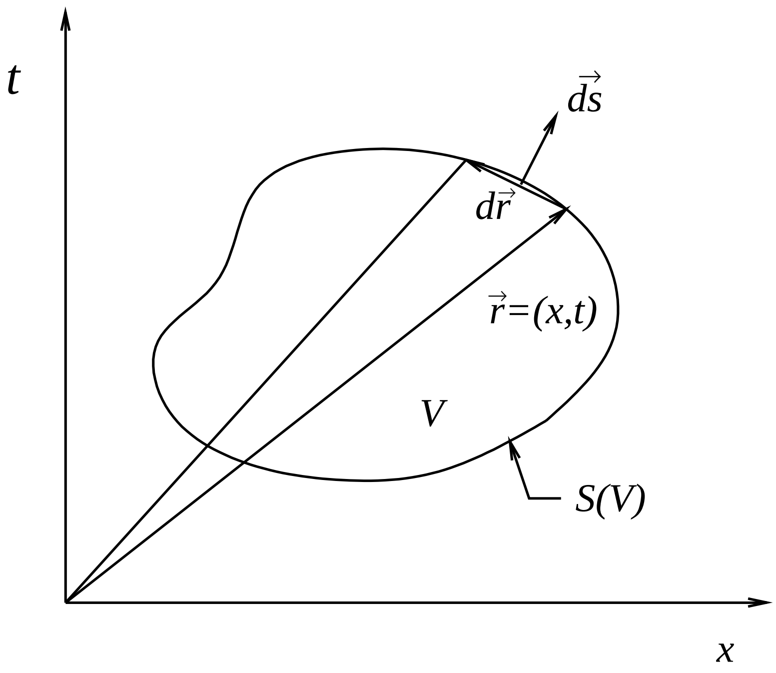


Figure 3: A surface element on the boundary S(V) of an arbitrary space-time region V.

Eq. (1.1) is the differential form of the integral conservation law

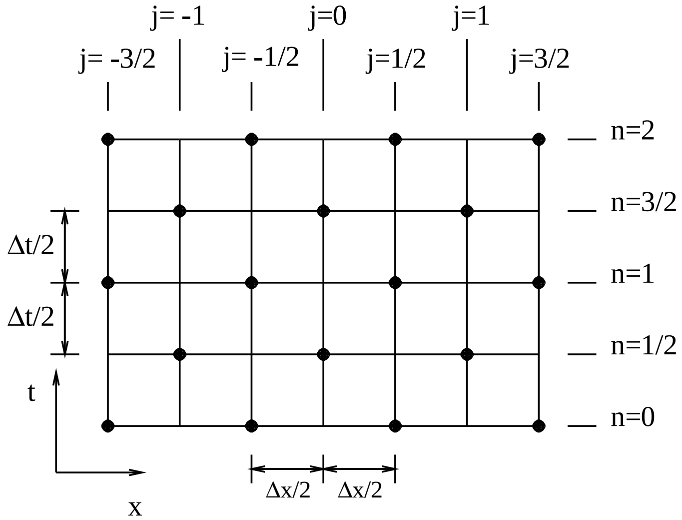
$$\oint_{S(V)} \vec{h} \cdot d\vec{s} = 0 \tag{1.3}$$

As depicted in Fig. 3, here (i) $S(V)$ is the boundary of an arbitrary space-time region V in E_2; (ii) $\vec{h} = (au,\, u)$ is a current density vector in E_2; and (iii) $d\vec{s} = d\sigma\,\vec{n}$ with $d\sigma$ and $\vec{n}$, respectively, being the area and the outward unit normal of a surface element on $S(V)$. Note that (i) $\vec{h} \cdot d\vec{s}$ is the *space-time* flux of $\vec{h}$ leaving the region V through the surface element $d\vec{s}$, and (ii) all mathematical operations can be carried out as though E_2 were an ordinary two-dimensional Euclidean space.

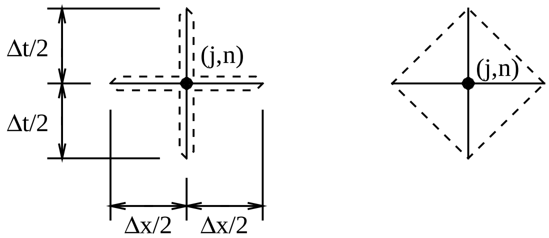
Let Ω denote the set of all mesh points (j, n) in E_2 (dots in Fig. 4(a)) with n being a half or whole integer, and $(j - n)$ being a half integer. For each $(j, n) \in \Omega$, let the solution element $\mathrm{SE}(j, n)$ be the *interior* of the *space-time* region bounded by a dashed curve depicted in Fig. 4(b). It includes a horizontal line segment, a vertical line segment, and their immediate neighborhood. For the discussions given in this paper, the exact size of this neighborhood does not matter. However, in case the conservation law Eq. (1.3) takes a more complicated form in which the right side is a volume integral involving a source term, the SEs must fill the entire computational domain such that the volume integral can be modeled properly [21,22]. A SE that fulfills this requirement is depicted in Fig. 4(c).

For any $(x, t) \in \mathrm{SE}(j, n)$, let $u(x, t)$ and $\vec{h}(x, t)$, respectively, be approximated by $u^*(x, t\,; j, n)$ and $\vec{h}^*(x, t\,; j, n)$ which we shall define shortly. Let

$$u^*(x, t\,; j, n) = u_j^n + (u_x)_j^n (x - x_j) + (u_t)_j^n (t - t^n) \tag{1.4}$$



(a). — The staggered space-time mesh



(b). — SE(j,n) (c). — Alternative SE(j,n)

Figure 4: The SEs and CEs.

where (i) u_j^n, $(u_x)_j^n$, and $(u_t)_j^n$ are constants in SE(j,n), and (ii) (x_j, t^n) are the coordinates of the mesh point (j,n).

We shall require that $u = u^*(x,t\,;j,n)$ satisfy Eq. (1.1) within SE(j,n). As a result,

$$(u_t)_j^n = -a\,(u_x)_j^n \tag{1.5}$$

Combining Eqs. (1.4) and (1.5), one has

$$u^*(x,t\,;j,n) = u_j^n + (u_x)_j^n\,[(x-x_j) - a\,(t-t^n)], \quad (x,t) \in \mathrm{SE}(j,n) \tag{1.6}$$

As a result, there are two independent marching variables u_j^n and $(u_x)_j^n$ associated with each $(j,n) \in \Omega$. Furthermore, because $\vec{h} = (au, u)$, we define

$$\vec{h}^*(x,t\,;j,n) = (au^*(x,t\,;j,n),\, u^*(x,t\,;j,n)) \tag{1.7}$$

Let E_2 be divided into non-overlapping rectangular regions (see Fig. 4(a)) referred to as conservation elements (CEs). As depicted in Figs. 4(d) and 4(e), the CE with its top-right (top-left) vertex being the mesh point $(j,n) \in \Omega$ is denoted by CE$_-(j,n)$ (CE$_+(j,n)$). The discrete approximation of Eq. (1.3) is then

$$\oint_{S(CE_\pm(j,n))} \vec{h}^* \cdot d\vec{s} = 0 \tag{1.8}$$

for all $(j,n) \in \Omega$. At each $(j,n) \in \Omega$, Eq. (1.8) provides the two conditions needed to solve its two independent marching variables. In the following, the manner in which the integrals in Eq. (1.8) should be evaluated will be explained by considering the case that involves CE$_-(j,n)$.

According to Fig. 4(d), $S(CE_-(j,n))$, i.e., the boundary of CE$_-(j,n)$, is formed by four line segments. Among them, $\overline{AB}$ and $\overline{AD}$ lie within SE(j,n). As a result, the flux leaving CE$_-(j,n)$ through these two line segments will be evaluated using Eqs. (1.6) and (1.7) with the assumption that any point (x,t) on them belongs to SE(j,n). On the other hand, because $\overline{CB}$ and $\overline{CD}$ lie within SE$(j-1/2, n-1/2)$, the flux leaving CE$_-(j,n)$ through them will be evaluated assuming any point (x,t) on them belongs to SE$(j-1/2, n-1/2)$.

According to Eq. (1.8), the total flux of $\vec{h}^*$ leaving the boundary of any conservation element is zero. Because the surface integration over any interface separating two neighboring CEs is evaluated using the information from a single SE, obviously the local conservation relation Eq. (1.8) leads to a global flux conservation relation, i.e., *the total flux of $\vec{h}^*$ leaving the boundary of any space-time region that is the union of any combination of CEs will also vanish.*

From the above discussions, it becomes obvious that *the space-time element used in the finite element method differs from the current space-time SE and CE in both concept and the roles they serve.* In particular, the former is not introduced to enforce flux conservation. *In contrast to this, in the CE/SE method, flux conservation transmits information between*

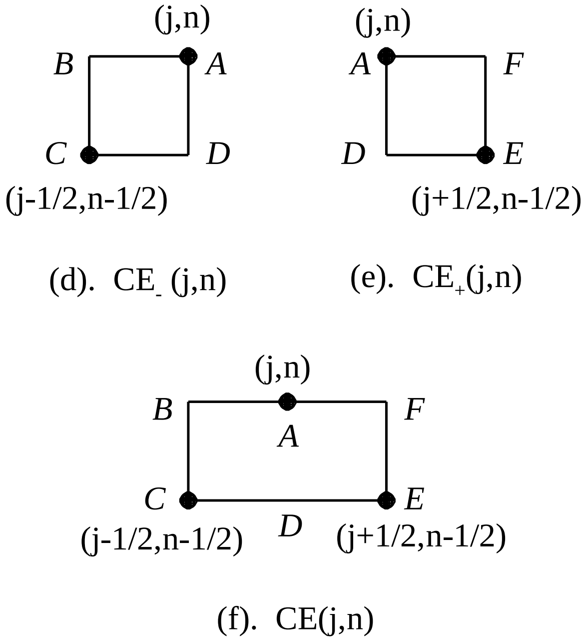


Figure 4: (continued).

neighboring SEs, and no global smoothness requirements are made on the solution to link neighboring SEs. This strategy enables the accurate capturing of traveling multidimensional solution discontinuities, e.g., moving multidimensional shock waves.

Furthermore, the CE/SE method is also fundamentally different from the traditional finite-volume methods such as the high-resolution upwind methods [46,47] and the discontinuous Galerkin method [48] in one important respect, i.e., *because of the space-time staggering nature of its solution elements, the present method has a much simpler and consistent procedure to evaluate the flux at an interface.* The key features of CE/SE flux-evaluation that distinguish it from those of the traditional methods are discussed in the following remarks:

(a) Because an interface separating two neighboring CEs lies within a SE, the flux at this interface is evaluated without interpolation or extrapolation. Furthermore, the SE to which a particular interface belongs is determined by a rule that is independent of the local numerical solution. In other words, *the concept of special upwind treatments and the complications that arise from these treatments are entirely foreign to the CE/SE method.* To be more specific, consider the flux at the interface $\overline{AD}$ depicted in Fig. 4(d). It is completely determined by u_j^n and $(u_x)_j^n$, two numerical variables associated with the predetermined mesh point (j, n), i.e., point A.

(b) Flux evaluation is straightforward and it requires only simple integration involving the first-order Taylor's expansion. *No complicated techniques such as the characteristics-based techniques are ever needed.*

Finally, we also want to emphasize that the concepts used in the construction of the a scheme are fundamentally different from several schemes introduced by Nessyahu and

Tadmor[49], and Sanders and Weiser [50] except that the meshes used by the a scheme and the latter schemes are all staggered in time. The key features of the a scheme that distinguish it from the latter schemes include: (i) the mesh values of both the dependent variable and its spatial derivative are considered as the independent variables, to be solved for simultaneously; and (ii) no interpolation or extrapolation techniques are used in the construction of the a scheme. Note that the differences between the latter schemes and an extension of the a scheme were also clearly spelled out by Huynh [51].

This section is concluded with the following remarks:

(a) The a scheme can be constructed from a different perspective in which both CEs and SEs have the shape of a rhombus [2]. In this alternative construction, the differential condition Eq. (1.5) is not assumed. Instead it becomes a result of a local flux conservation condition and Eq. (1.4). *In other words, the a scheme can be constructed entirely from flux conservation conditions and the assumption that $u^*(x, t\,; j, n)$ is linear in x and t.*

(b) The a scheme has many non-traditional features. They were discussed in great detail in [2].

(c) Because there are two independent marching variables at each mesh point $\in \Omega$, two amplification factors appear in the von Neumann stability analysis of the a scheme [2]. It happens that these two factors are identical to those of the Leapfrog scheme [52] if the latter factors arise from a "correct" von Neumann analysis [2]. Note that the main Leapfrog scheme (excluding its starting scheme which relates the mesh variables at the first two time levels), the Lax scheme [52], and the main DuFort-Frankel scheme [52] share one special property, i.e., a solution to any one of these schemes is formed by two decoupled solutions. Traditionally the von Neumann analysis for these schemes is performed without taking into account this decoupled nature. It is explained in [2] why such an erroneous analysis will result in a *dispersive property prediction that makes the dispersion appear worse than it really is.* Moreover, because (i) the a scheme and the Leapfrog scheme share the same amplification factors, and (ii) the a scheme is a two-level scheme while the Leapfrog scheme is a three-level scheme, *the a scheme can be considered as a more advanced and compact Leapfrog scheme.*

The fact that the amplification factors of the a scheme are related to those of a celebrated classical scheme is only one among a string of similar unexpected coincidences encountered during the development of the CE/SE method. As it turns out [2,12,25], the amplification factors of the Lax, the Crank-Nicolson, and the DuFort-Frankel schemes also are related to those of some of the extensions of the a scheme.

2. Review of the 1D Schemes

In this section, we shall (i) review and reformulate the 1D schemes described in [2], and (ii) fill a gap in the derivation of Eq. (4.28) in [2]. Not only does the reformulation enable the reader to see more clearly the structural similarity between the 1D solvers of Eq. (1.1) and their Euler counterparts, it also makes it easier for him to appreciate the consistency between the construction of the 1D CE/SE solvers and that of the 2D solvers to be described in the later sections.

2.1. The a Scheme

As the first step, the marching procedure of the a scheme will be cast into a form slightly different from that given in [2]. To proceed, let the Courant number $\nu \stackrel{def}{=} a\Delta t/\Delta x$. Also let

$$(u_x^+)_j^n \stackrel{def}{=} \frac{\Delta x}{4}(u_x)_j^n \tag{2.1}$$

for any $(j,n) \in \Omega$. Hereafter the superscript symbol "+" is used to denote a *normalized* parameter. Using Eq. (2.1), Eqs. (1.6)–(1.8) imply that

$$\left[(1-\nu)u + (1-\nu^2)u_x^+\right]_j^n = \left[(1-\nu)u - (1-\nu^2)u_x^+\right]_{j+1/2}^{n-1/2} \tag{2.2}$$

and

$$\left[(1+\nu)u - (1-\nu^2)u_x^+\right]_j^n = \left[(1+\nu)u + (1-\nu^2)u_x^+\right]_{j-1/2}^{n-1/2} \tag{2.3}$$

for all $(j,n) \in \Omega$. To simplify notation, in the above and hereafter we adopt a convention that can be explained using the expression on the left side of Eq. (2.2) as an example, i.e.,

$$\left[(1-\nu)u + (1-\nu^2)u_x^+\right]_j^n = (1-\nu)u_j^n + (1-\nu^2)(u_x^+)_j^n$$

Moreover, to streamline the future development, we define

$$(s_+)_{j+1/2}^{n-1/2} \stackrel{def}{=} \left[u - (1+\nu)u_x^+\right]_{j+1/2}^{n-1/2} \tag{2.4}$$

$$(s_-)_{j-1/2}^{n-1/2} \stackrel{def}{=} \left[u + (1-\nu)u_x^+\right]_{j-1/2}^{n-1/2} \tag{2.5}$$

and

$$(u_x^{a+})_j^n \stackrel{def}{=} \frac{1}{2}\left[(s_+)_{j+1/2}^{n-1/2} - (s_-)_{j-1/2}^{n-1/2}\right] \tag{2.6}$$

By adding Eqs. (2.2) and (2.3) together, and using the above definitions, one has

$$u_j^n = \frac{1}{2}\left[(1-\nu)(s_+)_{j+1/2}^{n-1/2} + (1+\nu)(s_-)_{j-1/2}^{n-1/2}\right], \qquad (j,n) \in \Omega \tag{2.7}$$

Let $1-\nu^2 \neq 0$, i.e., $1-\nu \neq 0$ and $1+\nu \neq 0$. Then Eqs. (2.2) and (2.3) can be divided by $(1-\nu)$ and $(1+\nu)$, respectively. By subtracting the resulting equations from each other and using Eqs. (2.4)–(2.6), one has

$$(u_x^+)_j^n = (u_x^{a+})_j^n, \qquad (j,n) \in \Omega \tag{2.8}$$

Because both $(s_+)^{n-1/2}_{j+1/2}$ and $(s_-)^{n-1/2}_{j-1/2}$ are explicit functions of the marching variables at the $(n-1/2)$th time level, Eqs. (2.7) and (2.8) form the explicit marching procedure for the a scheme. Note that these equations can be obtained from the inviscid form of the a-μ scheme, i.e., Eq. (2.14) in [2]. Also note that the superscript symbol "a" in the parameter $(u_x^{a+})^n_j$ is introduced to remind the reader that Eq. (2.8) is valid for the a scheme.

2.2. The a-ϵ Scheme

In the a-ϵ scheme [2], $\mathrm{CE}_+(j,n)$ and $\mathrm{CE}_-(j,n)$, which are depicted in Figs. 4(d) and 4(e), respectively, are not considered as conservation elements, i.e., Eq. (1.8) is no longer applicable. Instead, one assumes that

$$\oint_{S(CE(j,n))} \vec{h}^* \cdot d\vec{s} = 0, \qquad (j,n) \in \Omega \tag{2.9}$$

where $\mathrm{CE}(j,n)$ is the union of $\mathrm{CE}_+(j,n)$ and $\mathrm{CE}_-(j,n)$ (see Fig. 4(f)). In other words, $\mathrm{CE}(j,n)$ is a conservation element in the a-ϵ scheme. Again the local conservation condition Eq. (2.9) leads to a global conservation condition [2], i.e., *the total flux of $\vec{h}^*$ leaving the boundary of any space-time region that is the union of any combination of new CEs will also vanish.*

It was explained in [2] that Eq. (2.7) follows directly from Eq. (2.9). As a result, the former is also valid in the a-ϵ scheme. The a-ϵ scheme is formed by Eq. (2.7) and another equation that differs from Eq. (2.8) only in the expression on the right side. To show more clearly the similarity of the 1D schemes and their 2D versions to be described in the later sections, in the following, the counterpart of Eq. (2.8) in the a-ϵ scheme will be rederived from a perspective different from that presented in [2].

Let $(j,n) \in \Omega$. Then $(j \pm 1/2, n-1/2) \in \Omega$. Let

$$u'^n_{j\pm 1/2} \stackrel{def}{=} u^{n-1/2}_{j\pm 1/2} + (\Delta t/2)(u_t)^{n-1/2}_{j\pm 1/2} \tag{2.10}$$

Substituting Eqs. (1.5) and (2.1) into Eq. (2.10) and using the definition $\nu = a\Delta t/\Delta x$, one has

$$u'^n_{j\pm 1/2} = \left[u - 2\nu\, u_x^+\right]^{n-1/2}_{j\pm 1/2} \tag{2.11}$$

Note that, by definition, $(j \pm 1/2, n) \notin \Omega$ if $(j,n) \in \Omega$. Thus $u'^n_{j\pm 1/2}$ is associated with a mesh point $\notin \Omega$. The reader is warned that similar situations may occur in the rest of this paper.

According to Eq. (2.10), $u'^n_{j\pm 1/2}$ can be interpreted as a first-order Taylor's approximation of u at $(j \pm 1/2, n)$. Thus

$$(u_x^{c+})^n_j \stackrel{def}{=} \frac{u'^n_{j+1/2} - u'^n_{j-1/2}}{4} = \frac{\Delta x}{4}\left(\frac{u'^n_{j+1/2} - u'^n_{j-1/2}}{\Delta x}\right) \tag{2.12}$$

is a central-difference approximation of $\partial u/\partial x$ at (j,n), normalized by the same factor $\Delta x/4$ that appears in Eq. (2.1). Note that the superscript "c" is used to remind the reader of the central-difference nature of the term $(u_x^{c+})^n_j$. In the a-ϵ scheme, Eq. (2.8) is replaced by

$$(u_x^+)^n_j = (u_x^{a+})^n_j + 2\epsilon(u_x^{c+} - u_x^{a+})^n_j \tag{2.13}$$

where ϵ is a real number.

At this juncture, note that, at each mesh point $(j,n) \in \Omega$, Eqs. (2.7) and (2.8) are the results of two conservation conditions given in Eq. (1.8). Because Eq. (2.13) does not reduce to Eq. (2.8) except in the special case $\epsilon = 0$, at each mesh point $(j,n) \in \Omega$, generally the a-ϵ scheme satisfies only the single conservation condition Eq. (2.9) rather than the two consevation conditions Eq. (1.8). However, because $(u_x^{a+})_j^n$ generally is present on the right side of Eq. (2.13), the a-ϵ scheme generally will still be burdened with the cost of solving two conservation conditions at each mesh point. *The exception occurs only for the special case $\epsilon = 1/2$, under which Eq. (2.13) reduces to $(u_x^+)_j^n = (u_x^{c+})_j^n$.*

Note that the first part of the expression on the right side of Eq. (2.13), i.e., $(u_x^{a+})_j^n$, emerges from the development of the non-dissipative a scheme. As a result, it is the non-dissipative part. On the other hand, the second part, whose magnitude can be adjusted by the parameter ϵ, represents numerical dissipation introduced by the difference between the central difference term $(u_x^{c+})_j^n$ and the non-dissipative term $(u_x^{a+})_j^n$. Thus one may heuristically conclude that the numerical dissipation associated with the a-ϵ scheme can be increased by increasing the value of ϵ. It was shown in [2] that this conclusion is indeed valid in the stability domain of the a-ϵ scheme, i.e.,

$$0 \le \epsilon \le 1, \qquad \text{and} \qquad \nu^2 < 1 \tag{2.14}$$

According to Eqs. (2.4)–(2.6), (2.11) and (2.12), both $(u_x^{c+})_j^n$ and $(u_x^{a+})_j^n$ are explicitly dependent on ν (and therefore explicitly dependent on Δt). However, $(u_x^{c+} - u_x^{a+})_j^n$ is not dependent on ν. As a matter of fact, it can be shown that

$$(u_x^{c+} - u_x^{a+})_j^n = \frac{1}{2}\left[(u_x^+)_{j+1/2}^{n-1/2} + (u_x^+)_{j-1/2}^{n-1/2}\right] - \frac{1}{4}\left(u_{j+1/2}^{n-1/2} - u_{j-1/2}^{n-1/2}\right) \tag{2.15}$$

Let $(du_x)_j^n$ be the parameter defined by Eq. (3.2) in [2]. Then it can be shown that

$$(u_x^{c+} - u_x^{a+})_j^n = \frac{\Delta x}{4}(du_x)_j^n \tag{2.16}$$

Note that, in the original development [2], $(du_x)_j^n$ was introduced to break the symmetry of the stencil of the a scheme with respect to space-time inversion. This symmetry breaking results in the a-ϵ scheme that was originally defined by the matrix equation Eq. (3.6) of [2]. Its two component equations are Eq. (2.7) and

$$(u_x^+)_j^n = (u_x^{a+})_j^n + \epsilon\left[(u_x^+)_{j+1/2}^{n-1/2} + (u_x^+)_{j-1/2}^{n-1/2} - \frac{1}{2}\left(u_{j+1/2}^{n-1/2} - u_{j-1/2}^{n-1/2}\right)\right] \tag{2.17}$$

with the latter being equivalent to Eq. (2.13). *It should be emphasized that the fact that $(u_x^+)_j^n = (u_x^{c+})_j^n$ when $\epsilon = 1/2$, and that therefore the a-ϵ scheme can be considered as a central-difference scheme in this special case, was a later accidental discovery.*

2.3. The Euler a Scheme

For a reason that will soon become obvious to the reader, reformulation of the inviscid ($\mu = 0$) version of the Navier-Stokes solver described in Section 5 of [2] will precede that of the Euler solvers described in Section 4 of [2]. Because the inviscid version is also an Euler solver and, like the a scheme, is free of numerical dissipation if it is stable, it will be referred to as the Euler a scheme.

To proceed, consider the Euler equations [2]

$$\frac{\partial u_m}{\partial t} + \frac{\partial f_m}{\partial x} = 0, \qquad m = 1,2,3 \tag{2.18}$$

where (i) u_m, $m = 1,2,3$, are the independent flow variables to be solved for, and (ii) f_m, $m = 1,2,3$, are known functions [2] of u_m, $m = 1,2,3$. Assuming that the physical solution is smooth, Eq. (2.18) is a result of the more fundamental space-time flux conservation laws

$$\oint_{S(V)} \vec{h}_m \cdot d\vec{s} = 0, \qquad m = 1,2,3 \tag{2.19}$$

where $\vec{h}_m = (f_m,\, u_m)$, $m = 1,2,3$.

To proceed, let (i)

$$f_{m,k} \stackrel{def}{=} \partial f_m / \partial u_k, \qquad m,k = 1,2,3 \tag{2.20}$$

and (ii) F^+ be the 3×3 matrix formed by $(\Delta t/\Delta x) f_{m,k}$, $m,k = 1,2,3$. Note that, as a result of (ii), $F^+ = (\Delta t/\Delta x)F$ where F is the matrix that appears in Eq. (4.8) in [2]. Also let $(u_m)_j^n$ be the numerical version of u_m at any $(j,n) \in \Omega$. Because f_m and $f_{m,k}$ are functions of u_m, for any $(j,n) \in \Omega$, we can define $(f_m)_j^n$ and $(f_{m,k})_j^n$ to be the values of f_m and $f_{m,k}$, respectively, when u_m, $m = 1,2,3$, respectively, assume the values of $(u_m)_j^n$, $m = 1,2,3$. Furthermore, because f_m, $m = 1,2,3$, are homogeneous functions of degree 1 [53, p. 11] in the variables u_m, $m = 1,2,3$, we have

$$(f_m)_j^n = \sum_{k=1}^{3} (f_{m,k})_j^n (u_k)_j^n \tag{2.21}$$

Note that Eq. (2.21) is not essential in the development of the 1D CE/SE Euler solvers. However, in some instances, it is used to recast some equations into more convenient forms.

For any $(x,t) \in \mathrm{SE}(j,n)$, $u_m(x,t)$, $f_m(x,t)$ and $\vec{h}_m(x,t)$ are approximated by

$$u_m^*(x,t\,;j,n) \stackrel{def}{=} (u_m)_j^n + (u_{mx})_j^n (x - x_j) + (u_{mt})_j^n (t - t^n) \tag{2.22}$$

$$f_m^*(x,t\,;j,n) = (f_m)_j^n + (f_{mx})_j^n (x - x_j) + (f_{mt})_j^n (t - t^n) \tag{2.23}$$

and

$$\vec{h}_m^*(x,t\,;j,n) = (f_m^*(x,t\,;j,n),\, u_m^*(x,t\,;j,n)) \tag{2.24}$$

respectively [2]. Here (i) $(u_m)_j^n$ and $(u_{mx})_j^n$ are the independent marching variables to be solved for, and (ii) $(f_{mx})_j^n$, $(f_{mt})_j^n$, and $(u_{mt})_j^n$ are the functions of $(u_m)_j^n$ and $(u_{mx})_j^n$, $m = 1,2,3$, defined by Eqs. (4.10), (4.11), and (4.17) in [2].

For all $(j,n) \in \Omega$, we assume that

$$\oint_{S(CE_{\pm}(j,n))} \vec{h}_m^* \cdot d\vec{s} = 0, \qquad m = 1,2,3 \tag{2.25}$$

Note that Eqs. (2.18), (2.19) and (2.25) are the Euler counterparts of Eqs. (1.1), (1.3) and (1.8), respectively. With the aid of Eqs. (2.22)–(2.24), Eq. (2.25) implies that, for all $(j,n) \in \Omega$,

$$\begin{aligned}(u_m)_j^n - (u_m)_{j\pm 1/2}^{n-1/2} \pm \frac{\Delta x}{4}\left[(u_{mx})_{j\pm 1/2}^{n-1/2} + (u_{mx})_j^n\right] \\ \pm \frac{\Delta t}{\Delta x}\left[(f_m)_{j\pm 1/2}^{n-1/2} - (f_m)_j^n\right] \pm \frac{(\Delta t)^2}{4\Delta x}\left[(f_{mt})_{j\pm 1/2}^{n-1/2} + (f_{mt})_j^n\right] = 0.\end{aligned} \tag{2.26}$$

Eq. (2.26) is the inviscid version of the Navier-Stokes marching scheme originally given in Eq. (5.19) of [2].

For each $(j,n) \in \Omega$, let (i)

$$(u_{mx}^+)_j^n \stackrel{def}{=} \frac{\Delta x}{4}(u_{mx})_j^n, \qquad m = 1,2,3 \tag{2.27}$$

(ii) $\vec{u}_j^n$ and $(\vec{u}_x^+)_j^n$, respectively, be the 3×1 column matrices formed by $(u_m)_j^n$ and $(u_{mx}^+)_j^n$, $m = 1,2,3$, and (iii) $(F^+)_j^n$ be the 3×3 matrix formed by $(\Delta t/\Delta x)(f_{m,k})_j^n$, $m,k = 1,2,3$. Then with the aid of Eqs. (4.10), (4.11) and (4.17) in [2], and Eq. (2.21), one can rewrite Eq. (2.26) as a pair of matrix equations, i.e. for any $(j,n) \in \Omega$,

$$\left[(I - F^+)\vec{u} + \left(I - (F^+)^2\right)\vec{u}_x^+\right]_j^n = \left[(I - F^+)\vec{u} - \left(I - (F^+)^2\right)\vec{u}_x^+\right]_{j+1/2}^{n-1/2} \tag{2.28}$$

and

$$\left[(I + F^+)\vec{u} - \left(I - (F^+)^2\right)\vec{u}_x^+\right]_j^n = \left[(I + F^+)\vec{u} + \left(I - (F^+)^2\right)\vec{u}_x^+\right]_{j-1/2}^{n-1/2} \tag{2.29}$$

where I is the 3×3 identity matrix.

Note that the flux conservation conditions Eqs. (2.2) and (2.3), and its Euler counterparts, i.e., Eqs. (2.28) and (2.29) share the same algebraic structure. As a matter of fact, the former pair will become the latter pair if the symbols 1, ν, u *and* u_x^+ *are replaced by* I, F^+, $\vec{u}$ *and* $\vec{u}_x^+$, *respectively.* As a result, Eqs. (2.28) and (2.29) will be solved by a procedure similar to that used earlier to extract Eqs. (2.7) and (2.8) from Eqs. (2.2) and (2.3). However, because (i) matrix multiplication is not commutative and (ii) the matrix $(F^+)_j^n$ is a function of $(u_m)_j^n$, $m = 1,2,3$, while ν is a simple constant, as will be shown shortly, the algebraic structure of the solution to Eqs. (2.28) and (2.29) is more complex than that of Eqs. (2.7) and (2.8).

To proceed, let $(j,n) \in \Omega$ and

$$(\vec{s}_+)_{j+1/2}^{n-1/2} \stackrel{def}{=} \left[\vec{u} - (I + F^+)\vec{u}_x^+\right]_{j+1/2}^{n-1/2} \tag{2.30}$$

and

$$(\vec{s}_-)_{j-1/2}^{n-1/2} \stackrel{def}{=} \left[\vec{u} + (I - F^+)\vec{u}_x^+\right]_{j-1/2}^{n-1/2} \tag{2.31}$$

Then the addition of Eqs. (2.28) and (2.29) implies that

$$\vec{u}_j^n = \frac{1}{2}\left\{\left[(I - F^+)\vec{s}_+\right]_{j+1/2}^{n-1/2} + \left[(I + F^+)\vec{s}_-\right]_{j-1/2}^{n-1/2}\right\} \tag{2.32}$$

Note that: (i) Eq. (2.32) is equivalent to Eq. (4.24) in [2]; and (ii) Eqs. (2.30)–(2.32) are the Euler counterparts of Eqs. (2.4), (2.5) and (2.7), respectively.

Equation (2.32) represents the first part of the solution to Eqs. (2.28) and (2.29). To obtain the second part, one must assume the existence of the inverse of the matrix $[I-(F^+)^2]_j^n$ for all $(j,n) \in \Omega$. In the following, we shall briefly discuss the significance of this assumption.

Let v and c be the fluid speed and sonic speed, respectively. They are known functions of u_m, $m = 1,2,3$ [2]. For each $(j,n) \in \Omega$, let v_j^n and c_j^n, respectively, denote the values of v and c when u_m, $m = 1,2,3$, respectively, assume the values of $(u_m)_j^n$, $m = 1,2,3$. Let

$$(\nu_1)_j^n \stackrel{def}{=} \frac{\Delta t}{\Delta x}(v_j^n - c_j^n), \quad (\nu_2)_j^n \stackrel{def}{=} \frac{\Delta t}{\Delta x}v_j^n, \quad (\nu_3)_j^n \stackrel{def}{=} \frac{\Delta t}{\Delta x}(v_j^n + c_j^n) \tag{2.33}$$

Then, by using (i) the relation $F^+ = (\Delta t/\Delta x)F$, (ii) the fact that the eigenvalues of the matrix F are $v-c$, v and $v+c$ (see Eq. (4.8) in [2]), and (iii) the fact that the eigenvalues of $f(A)$ are $f(\lambda_1)$, $f(\lambda_2)$, $f(\lambda_3)$, ..., $f(\lambda_n)$ if the eigenvalues of a matrix A are λ_1, λ_2, λ_3, ..., λ_n and $f(A)$ is a polynomial of A, one concludes that the eigenvalues of $[I-(F^+)^2]_j^n$ are $[1-((\nu_\ell)_j^n)^2]$, $\ell = 1,2,3$. Because any square matrix is nonsingular (and therefore its inverse exists) if and only if all its eigenvalues are nonzero [54, p.14], one concludes that the inverse of $[I-(F^+)^2]_j^n$ exists if and only if

$$[(\nu_\ell)_j^n]^2 \neq 1, \qquad \ell = 1,2,3 \tag{2.34}$$

In this paper, we shall assume a more restrictive condition than Eq. (2.34), i.e., for all $(j,n) \in \Omega$, the local Courant number $\nu_j^n < 1$. Here

$$\nu_j^n \stackrel{def}{=} \max\{|(\nu_1)_j^n|, |(\nu_2)_j^n|, |(\nu_3)_j^n|\} \tag{2.35}$$

Note that, because

$$(I - F^+)(I + F^+) = (I + F^+)(I - F^+) = I - (F^+)^2 \tag{2.36}$$

the inverse of $[I \pm (F^+)]_j^n$ exists if the inverse of $[I-(F^+)^2]_j^n$ exists.

Let $(j,n) \in \Omega$. Let the marching variables at the $(n-1/2)$th time level be given. Then $\vec{u}_j^n$ can be evaluated using Eq. (2.32). Because $[I \pm F^+]_j^n$ is a function of $\vec{u}_j^n$, it follows that

$$(\vec{S}_+)_j^n \stackrel{def}{=} \left[(I - F^+)_j^n\right]^{-1}\left[(I - F^+)\vec{u} - \left(I - (F^+)^2\right)\vec{u}_x^+\right]_{j+1/2}^{n-1/2} \tag{2.37}$$

$$(\vec{S}_-)^n_j \stackrel{def}{=} \left[(I+F^+)^n_j\right]^{-1} \left[(I+F^+)\vec{u} + \left(I-(F^+)^2\right)\vec{u}^+_x\right]^{n-1/2}_{j-1/2} \tag{2.38}$$

and

$$(\vec{u}^{a+}_x)^n_j \stackrel{def}{=} \frac{1}{2}(\vec{S}_+ - \vec{S}_-)^n_j \tag{2.39}$$

can also be evaluated. Note that, in the above and hereafter, the inverse of a matrix A is denoted by A^{-1}.

To obtain the second part of the solution to Eqs. (2.28) and (2.29), they are multiplied from the left by

$$\left[(I-F^+)^n_j\right]^{-1} \quad \text{and} \quad \left[(I+F^+)^n_j\right]^{-1}$$

respectively. Let the resulting expressions be subtracted from each other. Then, with the aid of Eq. (2.36), one obtains

$$(\vec{u}^+_x)^n_j = (\vec{u}^{a+}_x)^n_j, \qquad (j,n) \in \Omega \tag{2.40}$$

Equations (2.32) and (2.40) define the marching procedure of the Euler a scheme. Note that the superscript symbol "a" in $(\vec{u}^{a+}_x)^n_j$ is intoduced to remind the reader that Eq. (2.40) is valid for the Euler a scheme.

It has been shown by numerical experiments that the Euler a scheme is neutrally stable in the interior of the computational domain up to at least a thousand time steps when $\nu^n_j < 1$ for all $(j,n) \in \Omega$. In these numerical experiments involving a shock-tube problem, the computational domain was allowed to grow with time, so that the undisturbed fluid state could always be prescribed at the computational boundaries as the exact solution. As a matter of fact, by using an analysis similar to that given at the end of Sec. 6 in [7], one can show that the linearized form of the Euler a scheme is neutrally stable when $\nu^n_j < 1$ for all $(j,n) \in \Omega$.

The parameters $(\vec{S}_+)^n_j$ and $(\vec{S}_-)^n_j$ can be evaluated by using Eqs. (2.37) and (2.38) directly. This direct evaluation involves inverting two 3×3 matrices which is computationally costly. In the following, we shall describe a more efficient approach.

According to Eqs. (2.37) and (2.38), $(\vec{S}_+)^n_j$ and $(\vec{S}_-)^n_j$ are the solutions to

$$(I-F^+)^n_j(\vec{S}_+)^n_j = \left[(I-F^+)\vec{u} - \left(I-(F^+)^2\right)\vec{u}^+_x\right]^{n-1/2}_{j+1/2} \tag{2.41}$$

and

$$(I+F^+)^n_j(\vec{S}_-)^n_j = \left[(I+F^+)\vec{u} + \left(I-(F^+)^2\right)\vec{u}^+_x\right]^{n-1/2}_{j-1/2} \tag{2.42}$$

respectively. Note that: (i) each of Eqs. (2.41) and (2.42) represents a system of three scalar equations; (ii) because of the reason given in the paragraph preceding Eq. (2.37), the coefficients of both systems are known if the marching variables at the $(n-1/2)$th time level are given, i.e., both systems can be considered as linear; and (iii) because of the assumption $\nu^n_j < 1$, each system has a unique solution. As a result of (i)–(iii), both $(\vec{S}_+)^n_j$ and $(\vec{S}_-)^n_j$ can be solved efficiently by using the Gaussian elimination method.

2.4. The Simplified Euler a scheme

In implementing the Euler a scheme, two systems of linear equations must be solved for each $(j,n) \in \Omega$. As a result, the Euler a scheme is *locally implicit* [1, p.22]. In this subsection we shall develop a simplified version that is completely explicit.

To proceed, the expressions

$$\left[(I - F^+)_j^n\right]^{-1} \quad \text{and} \quad \left[(I + F^+)_j^n\right]^{-1}$$

in Eqs. (2.37) and (2.38) are approximated by

$$\left[(I - F^+)_{j+1/2}^{n-1/2}\right]^{-1} \quad \text{and} \quad \left[(I + F^+)_{j-1/2}^{n-1/2}\right]^{-1}$$

respectively. As a result, one has

$$(\vec{S}_+)_j^n \approx (\vec{s}_+)_{j+1/2}^{n-1/2} \quad \text{and} \quad (\vec{S}_-)_j^n \approx (\vec{s}_-)_{j-1/2}^{n-1/2} \tag{2.43}$$

where $(\vec{s}_+)_{j+1/2}^{n-1/2}$ and $(\vec{s}_-)_{j-1/2}^{n-1/2}$ are defined in Eqs. (2.30) and (2.31), respectively. Let

$$(\vec{u}_x^{a'+})_j^n \stackrel{def}{=} \frac{1}{2}\left[(\vec{s}_+)_{j+1/2}^{n-1/2} - (\vec{s}_-)_{j-1/2}^{n-1/2}\right] \tag{2.44}$$

Then (i) $(\vec{u}_x^{a'+})_j^n$ can be evaluated explicitly, and (ii) as a result of Eqs. (2.39) and (2.43), Eq. (2.40) can be approximated by

$$(\vec{u}_x^+)_j^n = (\vec{u}_x^{a'+})_j^n, \qquad (j,n) \in \Omega \tag{2.45}$$

The marching procedure defined by Eqs. (2.32) and (2.45) is referred to as the simplified Euler a scheme. Note that the superscript symbol "a'" in $(\vec{u}_x^{a'+})_j^n$ is introduced to remind the reader that Eq. (2.45) is valid for the *simplified* Euler a scheme.

Generally $\mathrm{CE}_\pm(j,n)$, $(j,n) \in \Omega$, are not conservation elements in the simplified scheme. However, because Eq. (2.32) is equivalent to the conservation condition [2]

$$\oint_{S(CE(j,n))} \vec{h}_m^* \cdot d\vec{s} = 0, \qquad (j,n) \in \Omega \quad \text{and} \quad m = 1,2,3 \tag{2.46}$$

$\mathrm{CE}(j,n)$, $(j,n) \in \Omega$, are the conservation elements in the simplified scheme.

Note that by replacing the symbols s_+, s_-, u_x^{a+}, u, u_x^+, 1 and ν in Eqs. (2.4)–(2.8) by $\vec{s}_+$, $\vec{s}_-$, $\vec{u}_x^{a'+}$, $\vec{u}$, $\vec{u}_x^+$, I and F^+, respectively, these equations will become Eqs. (2.30), (2.31), (2.44), (2.32) and (2.45), respectively. In other words, the a scheme and the simplified Euler a scheme share the same algebraic structure.

The simplified Euler a scheme generally is unstable. However, as will be shown shortly, this scheme can be extended to become the simplified Euler a-ϵ scheme which does have a large stability domain.

2.5. The Euler a-ϵ Scheme

The process by which the a-ϵ scheme was constructed from the a scheme will be used to construct the Euler a-ϵ scheme from the Euler a scheme.

In the Euler a-ϵ scheme, the conservation conditions given in Eq. (2.46) are assumed. Because Eq. (2.32) is equivalent to Eq. (2.46), the former is also a part of of the Euler a-ϵ scheme. The Euler a-ϵ scheme is formed by Eq. (2.32) and another equation that differs from Eq. (2.40) only in the expression on the right side.

To proceed, let $(j,n) \in \Omega$ and

$$\vec{u}'^{\,n}_{j\pm1/2} \stackrel{def}{=} \vec{u}^{\,n-1/2}_{j\pm1/2} + (\Delta t/2)(\vec{u}_t)^{n-1/2}_{j\pm1/2} \tag{2.47}$$

where $(\vec{u}_t)^{n-1/2}_{j\pm1/2}$ is the column matrix formed by $(u_{mt})^{n-1/2}_{j\pm1/2}$, $m = 1,2,3$. With the aid of Eqs. (4.10) and (4.17) in [2], Eq. (2.47) implies that

$$\vec{u}'^{\,n}_{j\pm1/2} = (\vec{u} - 2F^{+}\vec{u}^{\,+}_{x})^{n-1/2}_{j\pm1/2} \tag{2.48}$$

Let

$$(\vec{u}^{\,c+}_{x})^{n}_{j} \stackrel{def}{=} \frac{\vec{u}'^{\,n}_{j+1/2} - \vec{u}'^{\,n}_{j-1/2}}{4} \tag{2.49}$$

Then the Euler a-ϵ scheme is formed by Eq. (2.32) and

$$(\vec{u}^{\,+}_{x})^{n}_{j} = (\vec{u}^{\,a+}_{x})^{n}_{j} + 2\epsilon(\vec{u}^{\,c+}_{x} - \vec{u}^{\,a+}_{x})^{n}_{j} \tag{2.50}$$

where ϵ is a real number. Obviously Eq. (2.50) reduces to (i) Eq. (2.40) when $\epsilon = 0$, and (ii) $(\vec{u}^{\,+}_{x})^{n}_{j} = (\vec{u}^{\,c+}_{x})^{n}_{j}$ when $\epsilon = 1/2$. Also it has been shown numerically that (i) the Euler a-ϵ scheme generally is stable if

$$0 \le \epsilon \le 1, \qquad \text{and} \qquad \nu^{n}_{j} < 1 \quad \text{for all} \quad (j,n) \in \Omega \tag{2.51}$$

and (ii) the numerical dissipation associated with the scheme increases as the value of ϵ increases. Note that Eqs. (2.47)–(2.50) are the Euler counterparts of Eqs. (2.10)–(2.13), respectively.

2.6. The Simplified Euler a-ϵ Scheme

According to Eq. (2.50), *excluding the special case* $\epsilon = 1/2$, implementation of the Euler a-ϵ scheme also requires the evaluation of $(\vec{u}^{\,a+}_{x})^{n}_{j}$ and therefore (see Eqs. (2.37)–(2.39)) the solution of Eqs. (2.41) and (2.42). Thus the Euler a-ϵ scheme is locally implicit if $\epsilon \neq 1/2$. A totally explicit variant, referred to as the simplified Euler a-ϵ scheme, is defined by Eq. (2.32) (or, equivalently, Eq. (2.46)) and

$$(\vec{u}^{\,+}_{x})^{n}_{j} = (\vec{u}^{\,a'+}_{x})^{n}_{j} + 2\epsilon(\vec{u}^{\,c+}_{x} - \vec{u}^{\,a'+}_{x})^{n}_{j} \tag{2.52}$$

Obviously the simplified Euler a-ϵ scheme (i) reduces to the simplified Euler a scheme when $\epsilon = 0$, and (ii) *is identical to the Euler a-ϵ scheme when* $\epsilon = 1/2$.

Note that by replacing the symbols s_+, s_-, u_x^{a+}, u, u_x^+, u', u_x^{c+}, 1 and ν in Eqs. (2.4)–(2.7) and (2.11)–(2.13) by $\vec{s}_+$, $\vec{s}_-$, $\vec{u}_x^{a'+}$, $\vec{u}$, $\vec{u}_x^+$, $\vec{u}'$, $\vec{u}_x^{c+}$, I and F^+, respectively, these equations will become Eqs. (2.30), (2.31), (2.44), (2.32), (2.48), (2.49) and (2.52) respectively. In other words, the a-ϵ scheme and the simplified Euler a-ϵ scheme share the same algebraic structure.

It has been shown numerically that the simplified Euler a-ϵ scheme is stable if

$$0.03 \le \epsilon \le 1, \qquad \text{and} \qquad \nu_j^n < 1 \quad \text{for all} \quad (j,n) \in \Omega \tag{2.53}$$

A comparison between Eqs. (2.51) and (2.53) reveals that the simplified version is only slightly less stable than the original version.

According to Eqs. (2.30), (2.31), (2.44), (2.48) and (2.49), both $(\vec{u}_x^{c+})_j^n$ and $(\vec{u}_x^{a'+})_j^n$ are explicitly dependent on the the matrices $(F^+)_{j+1/2}^{n-1/2}$ and $(F^+)_{j-1/2}^{n-1/2}$ (and therefore explicitly dependent on Δt). However, $(\vec{u}_x^{c+} - \vec{u}_x^{a'+})_j^n$ is free from this dependency. Let (i) $(du_{mx})_j^n$ be the parameter defined by Eq. (4.26) in [2], and (ii) $(d\vec{u}_x)_j^n$ be the column matrix formed by $(du_{mx})_j^n$, $m = 1, 2, 3$. Then it can be shown that

$$(\vec{u}_x^{c+} - \vec{u}_x^{a'+})_j^n = \frac{1}{2}\left[(\vec{u}_x^+)_{j+1/2}^{n-1/2} + (\vec{u}_x^+)_{j-1/2}^{n-1/2}\right] - \frac{1}{4}\left(\vec{u}_{j+1/2}^{n-1/2} - \vec{u}_{j-1/2}^{n-1/2}\right) = \frac{\Delta x}{4}(d\vec{u}_x)_j^n \tag{2.54}$$

With the above preliminaries, we are now ready to provide a proof for Eq. (4.28) in [2]. Note that the last equation was introduced in [2] simply as a "natural generalization" of Eq. (3.10) in [2].

To proceed, note that Eq. (2.47) is the matrix form of Eq. (4.27) in [2], i.e., $\vec{u}'^n_{j\pm1/2}$ is the column matrix formed by $(u'_m)_{j\pm1/2}^n$, $m = 1, 2, 3$, which were introduced in the latter equation. As a result, with the aid of Eqs. (2.27), (2.49) and (2.54), Eq. (2.52) can be rewritten as

$$(u_{mx})_j^n = \left[(u'_m)_{j+1/2}^n - (u'_m)_{j-1/2}^n\right]/\Delta x + (2\epsilon - 1)(du_{mx})_j^n \tag{2.55}$$

i.e., Eq. (4.28) in [2].

Because Eqs. (4.24) in [2] are equivalent to Eq. (2.32), the Euler scheme defined by Eqs. (4.24) and (4.28) in [2] is identical to the simplified Euler a-ϵ scheme.

2.7. The a-ϵ-α-β Scheme and Its Euler Versions

Consider the a-ϵ scheme defined by Eqs. (2.7) and (2.13). If discontinuities are present in a numerical solution, the above scheme is not equipped to suppress numerical wiggles that generally appear near these discontinuities. In the following, we shall describe a remedy for this deficiency.

Let

$$(u_{x\pm}^{c+})_j^n \stackrel{def}{=} \pm\frac{1}{2}(u'^n_{j\pm1/2} - u_j^n) \tag{2.56}$$

Then it can be shown that

$$(u_x^{c+})_j^n = \frac{1}{2}\left[(u_{x+}^{c+})_j^n + (u_{x-}^{c+})_j^n\right] \tag{2.57}$$

i.e., $(u_x^{c+})_j^n$ is the simple average of $(u_{x+}^{c+})_j^n$ and $(u_{x-}^{c+})_j^n$. Next, let the function W_o be defined by (i) $W_o(0,0,\alpha)=0$ and (ii)

$$W_o(x_-,x_+;\alpha)=\frac{|x_+|^\alpha x_- + |x_-|^\alpha x_+}{|x_+|^\alpha + |x_-|^\alpha}, \qquad (|x_+|+|x_-|>0) \tag{2.58}$$

where x_+, x_- and $\alpha \geq 0$ are real variables. Note that (i) to avoid dividing by zero, in practice a small positive number such as 10^{-60} is added to the denominator in Eq. (2.58); and (ii) $W_o(x_-,x_+;\alpha)$, a nonlinear weighted average of x_- and x_+, becomes their simple average if $\alpha=0$ or $|x_-|=|x_+|$. Furthermore, let

$$(u_x^{w+})_j^n \stackrel{def}{=} W_o\left((u_{x+}^{c+})_j^n,(u_{x-}^{c+})_j^n;\alpha\right) \tag{2.59}$$

Note that the superscript "w" is used to remind the reader of the weighted-average nature of the term $(u_x^{w+})_j^n$. With the aid of the above definitions, a more advanced scheme, referred to as the a-ϵ-α-β scheme, can be defined by Eq. (2.7) and

$$(u_x^{+})_j^n=(u_x^{a+})_j^n+2\epsilon(u_x^{c+}-u_x^{a+})_j^n+\beta(u_x^{w+}-u_x^{c+})_j^n \tag{2.60}$$

Here $\beta \geq 0$ is another adjustable constant. Note that Eq. (2.60) can be rewritten as

$$(u_x^{+})_j^n=\beta W_o\left((u_{x+}^{c+})_j^n,(u_{x-}^{c+})_j^n;\alpha\right)+(1-\beta)(u_x^{c+})_j^n+(2\epsilon-1)(u_x^{c+}-u_x^{a+})_j^n \tag{2.61}$$

It can be shown easily that the a-ϵ-α-β scheme reduces to the a-ϵ scheme if $\alpha=0$ or $\beta=0$.

The expression on the right side of Eq. (2.60) contains three parts. The first part is a non-dissipative term $(u_x^{a+})_j^n$. The second part is the product of 2ϵ and the difference between the central difference term $(u_x^{c+})_j^n$ and the non-dissipative term $(u_x^{a+})_j^n$. The third part is the product of β and the difference between a weighted average of $(u_{x+}^{c+})_j^n$ and $(u_{x-}^{c+})_j^n$ and their simple average. Numerical dissipation introduced by the second part generally is effective in damping out numerical instabilities that arise from the smooth region of a solution. But it is less effective in suppressing numerical wiggles that often occur near a discontinuity. On the other hand, numerical dissipation introduced by the third part is very effective in suppressing numerical wiggles. Moreover, because the condition $|(u_{x+}^{c+})_j^n|=|(u_{x-}^{c+})_j^n|$ more or less prevails and thus the weighted average is nearly equal to the simple average (see the comment given immediately following Eq. (2.58)) in the smooth region of the the solution, numerical dissipation introduced by the third part has very slight effect in the smooth region.

From the above disscusion, one concludes that there are two different types of numerical dissipation associated with the a-ϵ-α-β scheme and they complement each other. As a result, the a-ϵ-α-β scheme can handle both small disturbances and sharp discontinuies simultaneously if the values of ϵ, α and β are chosen properly (usually $\epsilon=1/2$, $\alpha=1,2$ and $\beta=1$). Also note that, to give the CE/SE method more flexibility in controlling local numerical dissipation, the parameters ϵ and β can even be considered as functions of local dynamical variables and mesh parameters (see Sec. 8).

Similarly, the Euler a-ϵ scheme and the simplified Euler a-ϵ scheme can be modified to become the Euler a-ϵ-α-β scheme and the simplified Euler a-ϵ-α-β scheme, respectively, by simply replacing Eqs. (2.50) and (2.52) with

$$(\vec{u}_x^+)_j^n = (\vec{u}_x^{a+})_j^n + 2\epsilon(\vec{u}_x^{c+} - \vec{u}_x^{a+})_j^n + \beta(\vec{u}_x^{w+} - \vec{u}_x^{c+})_j^n \tag{2.62}$$

and

$$(\vec{u}_x^+)_j^n = (\vec{u}_x^{a'+})_j^n + 2\epsilon(\vec{u}_x^{c+} - \vec{u}_x^{a'+})_j^n + \beta(\vec{u}_x^{w+} - \vec{u}_x^{c+})_j^n \tag{2.63}$$

respectively. Here $(\vec{u}_x^{w+})_j^n$ is the 3×1 column matrix formed by

$$W_o\left((u_{mx+}^{c+})_j^n, (u_{mx-}^{c+})_j^n; \alpha\right), \qquad m = 1, 2, 3$$

where

$$(u_{mx\pm}^{c+})_j^n \stackrel{def}{=} \pm\frac{1}{2}((u'_m)_{j\pm1/2}^n - (u_m)_j^n) \tag{2.64}$$

with $(u'_m)_{j\pm1/2}^n$ and $(u_m)_j^n$ being the mth components of $\vec{u}'^n_{j\pm1/2}$ and $\vec{u}_j^n$, respectively.

2.8. The 1D CE/SE Shock-Capturing Scheme

Let $\epsilon = 1/2$ and $\beta = 1$. Then the Euler a-ϵ-α-β scheme and the simplified Euler a-ϵ-α-β scheme reduce to the same scheme. The reduced scheme is defined by Eq. (2.32) and

$$(u_{mx}^+)_j^n = W_o\left((u_{mx+}^{c+})_j^n, (u_{mx-}^{c+})_j^n; \alpha\right), \qquad m = 1, 2, 3 \tag{2.65}$$

where $(j, n) \in \Omega$.

The above scheme is one of the simplest among the Euler solvers known to the authors. *The value of α is the only adustable parameter allowed in this scheme.* Because it is totally explicit and has the simplest stencil, the scheme is also highly compatible with parallel computing. Furthermore, it has been shown that the scheme can accurately capture shocks and contact discontinuities with high resolution and no numerical oscillations. For these distinctive features and for convenience of future reference, the reduced scheme will be given a special name, i.e., the 1D CE/SE shock-capturing scheme. Note that this scheme with $\alpha = 1$ is implemented in the two shock-tube solvers referred to in Sec. 1. Consider only the case that all spatial boundary points $(j, n) \in \Omega$ are at the time levels $n = 0, 1, 2, \ldots$ (see Fig. 4(a)). The non-reflecting boundary conditions used in the first solver, i.e., the one listed in Appendix A, are: (i)

$$\vec{u}_j^n = \vec{u}_{j-1/2}^{n-1/2} \qquad \text{and} \qquad (\vec{u}_x^+)_j^n = (\vec{u}_x^+)_{j-1/2}^{n-1/2}, \qquad n = 1, 2, 3, \ldots \tag{2.66}$$

if (j, n) is a mesh point on the right spatial boundary; and (ii)

$$\vec{u}_j^n = \vec{u}_{j+1/2}^{n-1/2} \qquad \text{and} \qquad (\vec{u}_x^+)_j^n = (\vec{u}_x^+)_{j+1/2}^{n-1/2}, \qquad n = 1, 2, 3, \ldots \tag{2.67}$$

if (j, n) is a mesh point on the left spatial boundary. On the other hand, for the alternate solver, the steady-state boundary conditions

$$\vec{u}_j^n = \vec{u}_j^0 \qquad \text{and} \qquad (\vec{u}_x^+)_j^n = (\vec{u}_x^+)_j^0, \qquad n = 1, 2, 3, \ldots \tag{2.68}$$

is imposed at any mesh point (j, n) on the right or left spatial boundary.

3. Geometrical Description of Conservation Elements in Two Spatial Dimensions

In Sec. 2, it was established that, for each 1D CE/SE solver, there were $2M$ independent marching variables per mesh point with M being the number of conservation laws to be solved. Because M conservation conditions are imposed over each CE, two CEs were introduced at each mesh point such that both the 1D a scheme and the 1D Euler a scheme can be constructed by solving, at each mesh point $(j,n) \in \Omega$, for the $2M$ variables using the $2M$ conservation conditions imposed over $\mathrm{CE}_-(j,n)$ and $\mathrm{CE}_+(j,n)$.

As will be shown in the following sections, for each 2D CE/SE solver, there are $3M$ independent marching variables per mesh point. As a result, construction of the 2D a scheme and the 2D Euler a scheme demands that three CEs be defined at each mesh point. In this section, only the basic geometric structures of these CEs will be described.

Consider a spatial domain formed by congruent triangles (see Fig. 5). The center of each triangle is marked by either a hollow circle or a solid circle. The distribution of these hollow and solid circles is such that if the center of a triangle is marked by a solid (hollow) circle, then the centers of the three neighboring triangles with which the first triangle shares a side are marked by hollow (solid) circles. As an example, point G , the center of the triangle ΔBDF, is marked by a solid circle while points A, C and E, the centers of the triangles ΔFMB, ΔBJD and ΔDLF, respectively, are marked by hollow circles. These centers are the spatial projections of the space-time mesh points used in the 2D CE/SE solvers.

To specify the exact locations of the mesh points in space-time, one must also specify their temporal coordinates. In the 2D CE/SE development, again we assume that the mesh points are located at the time levels $n = 0, \pm 1/2, \pm 1, \pm 3/2, \ldots$ with $t = n\,\Delta t$ at the nth time level. Furthermore, we assume that the spatial projections of the mesh points at a whole-integer (half-integer) time level are the points marked by hollow (solid) circles in Fig. 5.

Let the triangles depicted in Fig. 5 lie on the time level $n = 0$. Then those points marked by hollow circles are the mesh points at this time level. On the other hand, those points marked by solid circles are not the mesh points at the time level $n = 0$. They are the spatial projections of the mesh points at half-integer time levels.

Points A, C and E, which are depicted in Figs. 5 and 6(a), are three mesh points at the time level $n = 0$. Point G', which is depicted in Fig. 6(a), is a mesh point at the time level $n = 1/2$. Its spatial projection at the time level $n = 0$ is point G. Because point G is not a mesh point, it is not marked by a circle in the space-time plots given in Figs. 6(a) and 6(c). Hereafter, only a mesh point, e.g., point G', will be marked by a solid or hollow circle in a space-time plot.

The conservation elements associated with point G' are defined to be the space-time quadrilateral cylinders $GFABG'F'A'B'$, $GBCDG'B'C'D'$, and $GDEFG'D'E'F'$ that are depicted in Fig. 6(a). Here (i) points B, D and F are the vertices of the triangle with point G as its center (centroid) (see also Fig. 5), and (ii) points A', B', C', D', E' and F' are on the time level $n = 1/2$ with their spatial projections on the time level $n = 0$ being points A,

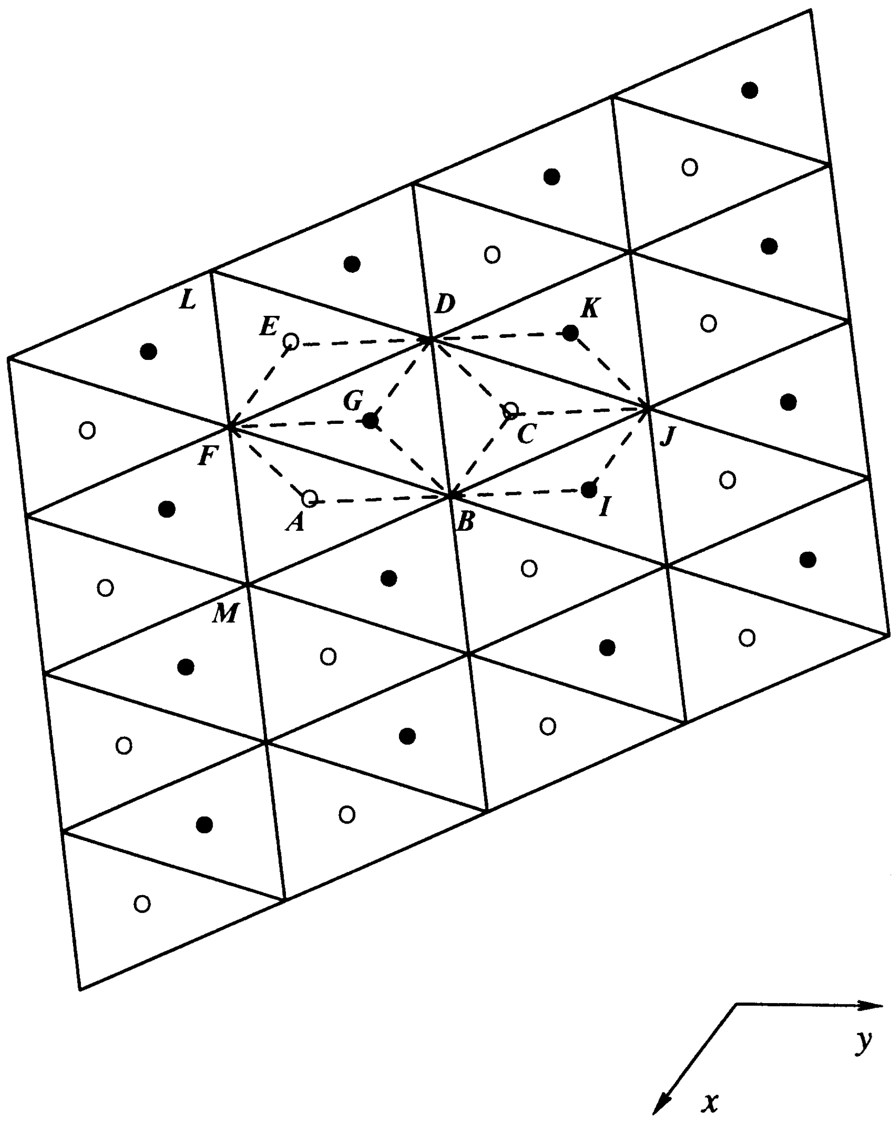


Figure 5: A spatial domain formed from congruent triangles, showing the spatial projections of the mesh points.

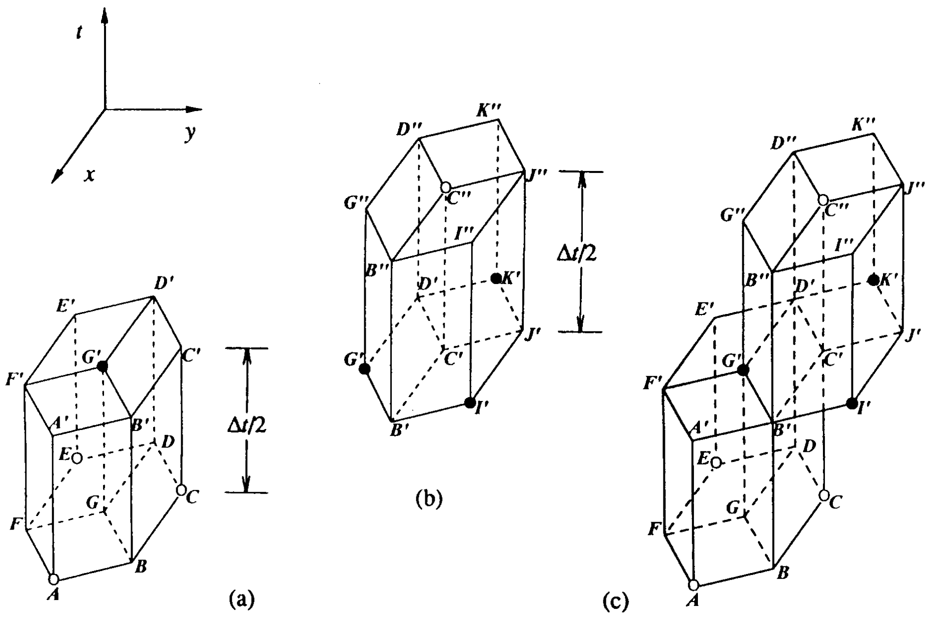


Figure 6: (a) The CEs associated with G'. (b) the CEs associated with C''. (c) The relative positions of the CEs of successive time steps.

B, C, D, E and F, respectively.

Point G' is a mesh point at a half-integer time level. For a mesh point at a whole-integer time-level, the conservation elements associated with it can be constructed in a similar fashion. As an example, consider Fig. 6(b). Here points $B'\,(B'')$, $I'\,(I'')$, $J'\,(J'')$, $K'\,(K'')$, $D'\,(D'')$, $G'\,(G'')$ and $C'\,(C'')$ are on the time level $n = 1/2$ $(n = 1)$ with their spatial projections on the time level $n = 0$, respectively, being the points B, I, J, K, D, G and C that are depicted in Fig. 5. Point C'' is a mesh point at the time level $n = 1$. By definition, the conservation elements associated with point C'' are the quadrilateral cylinders $C'J'K'D'C''J''K''D''$, $C'D'G'B'C''D''G''B''$ and $C'B'I'J'C''B''I''J''$. The relative space-time positions of the six CEs associated with mesh points G' and C'' are depicted in Fig. 6(c).

Recall that, in the development of the 1D a scheme, a pair of diagonally opposite vertices of each $\mathrm{CE}_{\pm}(j,n)$ (see Figs. 4(d) and 4(e)) are assigned as mesh points. Furthermore, the boundary of each $\mathrm{CE}_{\pm}(j,n)$ is a subset of the union of the SEs associated with the two diagonally opposite mesh points of this CE. In the 2D development, as seen from Figs. 6(a)–(c), two diagonally opposite vertices of each CE are also assigned as mesh points. In Sec. 4, we shall define the SEs such that even in the 2D case, the boundary of a CE is again a subset of the union of the SEs associated with the two diagonally opposite mesh points of this CE.

As a preliminary to the derivation of several equations to be given in Sec. 4, this section is concluded with a discussion of several geometric relations involving point G and the vertices of the hexagon $ABCDEF$ that are depicted in Fig. 5. By using the facts that (i) points A, C, E and G are the geometric centers of four neighboring congruent triangles $\triangle FMB$,

$\triangle BJD$, $\triangle DLF$ and $\triangle BDF$, respectively; and (ii) any two of the above four triangles form a parallelogram (note: two congruent triangles sharing one side may not form a parallelogram), it can be shown that:

(a) $\overline{CD}$, $\overline{GE}$, $\overline{BG}$ and $\overline{AF}$ are parallel line segments of equal length.

(b) $\overline{AB}$, $\overline{GC}$, $\overline{FG}$ and $\overline{ED}$ are parallel line segments of equal length.

(c) $\overline{BC}$, $\overline{GD}$, $\overline{AG}$ and $\overline{FE}$ are parallel line segments of equal length.

(d) Point G is the geometric center of the hexagon $ABCDEF$ and the triangle ACE.

Note that the line segments $\overline{GA}$, $\overline{GC}$, $\overline{GE}$ $\overline{AC}$, $\overline{CE}$ and $\overline{EA}$ are not shown in Fig. 5. Also note that, because the hexagon $BIJKDG$ (depicted in Fig. 5) is congruent to the hexagon $ABCDEF$, a set of geometric relations similar to those listed above also exists for the vertices and the center of the hexagon $BIJKDG$.

4. The 2D a Scheme

In this section, we consider a dimensionless form of the 2D convection equation, i.e.,

$$\frac{\partial u}{\partial t} + a_x \frac{\partial u}{\partial x} + a_y \frac{\partial u}{\partial y} = 0 \tag{4.1}$$

where a_x, and a_y are constants. Let $x_1 = x$, $x_2 = y$, and $x_3 = t$ be the coordinates of a three-dimensional Euclidean space E_3. By using Gauss' divergence theorem in the space-time E_3, it can be shown that Eq. (4.1) is the differential form of the integral conservation law

$$\oint_{S(V)} \vec{h} \cdot d\vec{s} = 0 \tag{4.2}$$

Here (i) $S(V)$ is the boundary of an arbitrary space-time region V in E_3, (ii)

$$\vec{h} \stackrel{def}{=} (a_x u,\ a_y u,\ u) \tag{4.3}$$

is a current density vector in E_3, and (iii) $d\vec{s} = d\sigma\, \vec{n}$ with $d\sigma$ and $\vec{n}$, respectively, being the area and the outward unit normal of a surface element on $S(V)$. It was shown in Sec. 3, that E_3 can be divided into nonoverlapping space-time regions referred to as conservation elements (CEs).

In the following analysis, the nontraditional space-time mesh that was sketched in Sec. 3 will be rigorously defined. To proceed, the spatial projections of the mesh points depicted in Fig. 5 are reproduced in Fig. 7. Note that the dashed lines that appear in Fig. 7 are the spatial projections of the vertical interfaces (see Fig. 6(a)-(c)) that separate different CEs. Also note that, as a result of the geometric relations listed at the end of Sec. 3, any dashed line can point only in one of three different fixed directions. We assume that the congruent triangles depicted in Fig. 5 are aligned such that one of the above fixed directions is the x-direction.

Each mesh point marked by a solid or hollow circle is assigned a pair of spatial indices (j, k) according to the location of its spatial projection. Obviously, a mesh point can be uniquely identified by its spatial indices (j, k) and the time level n where it resides. According to Figs. 8 and 9, the spatial projections of the mesh points that share the same value of j (k) lie on a straight line on the x-y plane with this straight line pointing in the direction of the k- $(j$-$)$ mesh axis.

Let

$$t^n \stackrel{def}{=} n\Delta t, \qquad n = 0,\ \pm 1/2,\ \pm 1,\ \pm 3/2, \ldots \tag{4.4}$$

Let j and k be spatial mesh indices with $j, k = 0,\ \pm 1/3,\ \pm 2/3,\ \pm 1, \ldots$. Let Ω_1 denote the set of mesh points (j, k, n) with $j, k = 0, \pm 1, \pm 2, \ldots$, and $n = \pm 1/2,\ \pm 3/2,\ \pm 5/2, \ldots$. These mesh points are marked by solid circles. Let Ω_2 denote the set of mesh points (j, k, n) with $j, k = 1/3,\ 1/3 \pm 1,\ 1/3 \pm 2, \ldots$, and $n = 0, \pm 1, \pm 2, \ldots$. These mesh points are marked by hollow circles. The union of Ω_1 and Ω_2 will be denoted by Ω. *Note that the same symbol Ω was also used to denote the set of mesh points used in the 1D solvers (see Sec.2). Hereafter, unless specified otherwise, the new definition of Ω is assumed.*

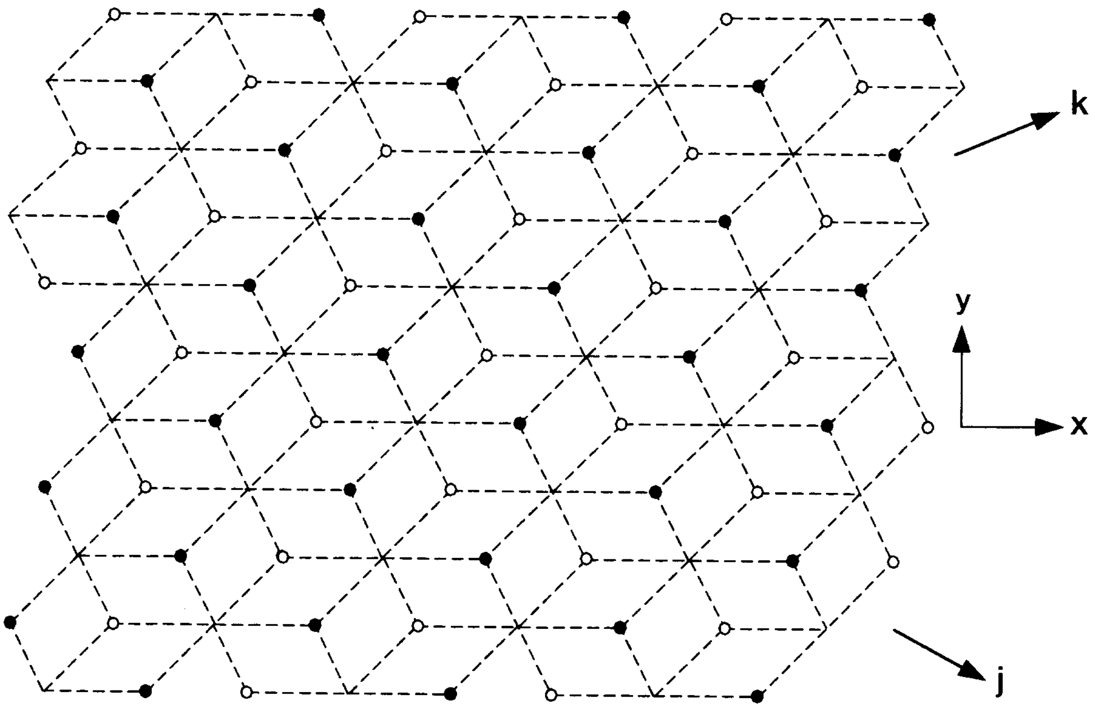


Figure 7: The relative spatial positions of the mesh points $\in \Omega_1$ and the mesh points $\in \Omega_2$ (dash lines are spatial boundaries of the conservation elements depicted in figs 10(a) and 11(a)).

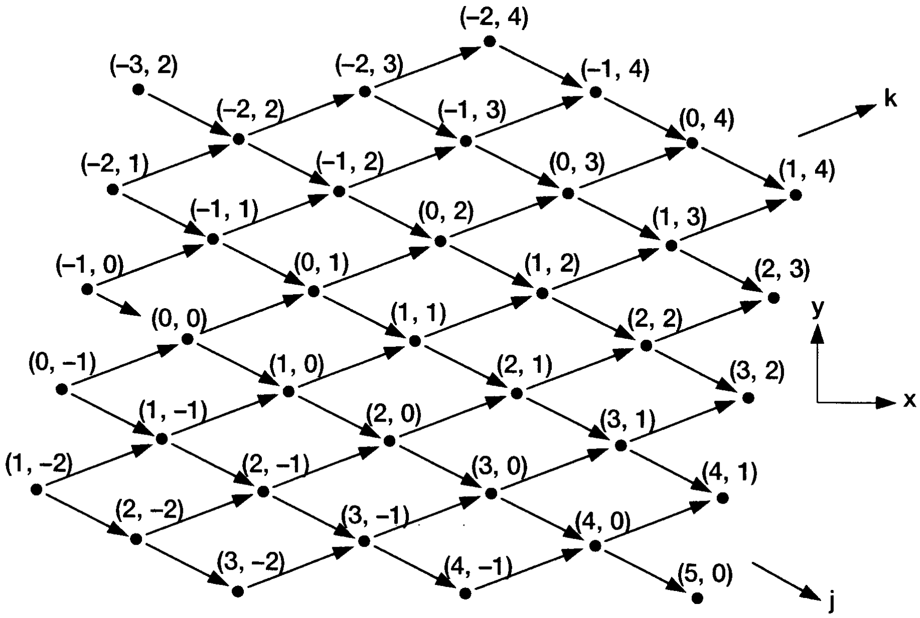


Figure 8: The spatial mesh indices (j,k) of the mesh points $\in \Omega_1$
($n = \pm 1/2,\ \pm 3/2,\ \pm 5/2, \ldots$).

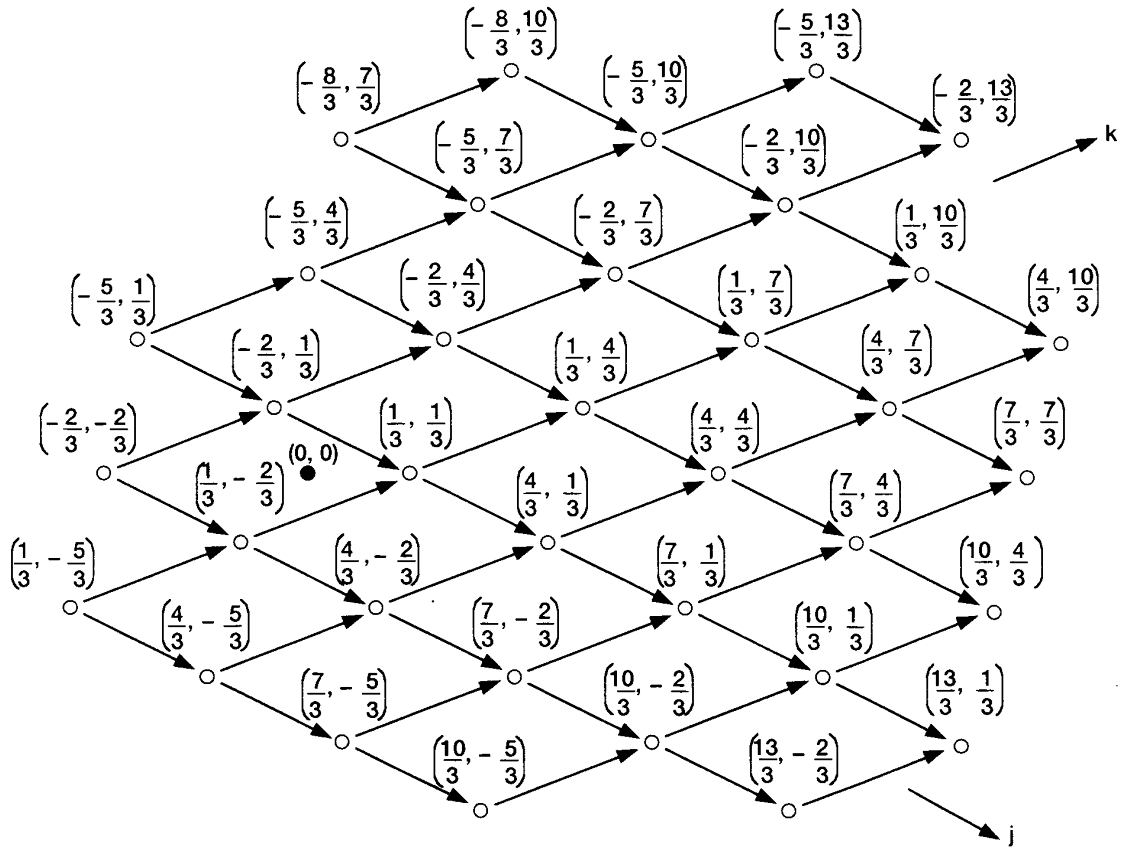


Figure 9: The spatial mesh indices (j,k) of the mesh points $\in \Omega_2$ ($n = 0, \pm 1, \pm 2, \ldots$).

Each mesh point $(j,k,n) \in \Omega$ is associated with (i) three conservation elements (CEs), denoted by $\mathrm{CE}_r(j,k,n)$, $r=1,2,3$ (see Figs. 10(a) and 11(a)); and (ii) a solution element (SE), denoted by $\mathrm{SE}(j,k,n)$ (see Figs. 10(b) and 11(b)). Each CE is a quadrilateral cylinder in space-time while each SE is the union of three vertical planes, a horizontal plane, and their immediate neighborhoods. *Note that the CEs and the SE associated with a mesh point* $(j,k,n) \in \Omega_1$ *differ from those associated with a mesh point* $(j,k,n) \in \Omega_2$ *in their space-time orientations.*

By using the geometric relations listed at the end of Sec. 3, one can conclude that the spatial coordinates of the vertices of the hexagon $ABCDEF$, which appears in both Figs. 10(a) and 11(a), are uniquely determined by three positive parameters w, b and h (see Fig. 12(a)) if (i) one assumes that $\overline{DA}$ is aligned with the x-direction, and (ii) the spatial coordinates of point G (the centroid of the hexagon) are given. Note that w, b and h, respectively, are the lengths of the line segments $\overline{DM}$, $\overline{MH}$ and $\overline{BH}$ with (i) $\overline{DM}$ being a median of the triangle $\triangle BDF$, and (ii) points G, M and H being on the line segment $\overline{DA}$. Also note that a dashed line in Fig. 7 may appear in other figures as a solid line.

According to Fig. 7, E_3 can be filled with the CEs defined above. Moreover, it is seen from Figs. 10(a), 10(b), 11(a), and 11(b) that *the boundary of a CE is formed by the subsets of two neighboring SEs.*

Let the space-time mesh be uniform, i.e., the parameters Δt, w, b, and h are constants. Let $x_{j,k}$ and $y_{j,k}$ be the x- and y- coordinates of any mesh points $(j,k,n) \in \Omega$. Let $x_{0,0}=0$ and $y_{0,0}=0$. Then information provided by Figs. 12(a) and 12(b) implies that

$$x_{j,k} = (j+k)w + (k-j)b, \quad y_{j,k} = (k-j)h \tag{4.5}$$

Let $\vec{n}_1$, $\vec{n}_2$, $\vec{n}_3$, $\vec{n}_4$, $\vec{n}_5$, and $\vec{n}_6$ be the vectors depicted in Fig. 12(a). They lie on the x-y plane and are the outward unit normals to $\overline{AB}$, $\overline{BC}$, $\overline{CD}$, $\overline{DE}$, $\overline{EF}$, and $\overline{FA}$, respectively. It can be shown that

$$\vec{n}_1 = \frac{(h, -b+w/3, 0)}{\sqrt{h^2+(b-w/3)^2}}, \qquad \vec{n}_4 = -\vec{n}_1 \tag{4.6a}$$

$$\vec{n}_2 = (0,1,0), \qquad \vec{n}_5 = -\vec{n}_2 \tag{4.6b}$$

and

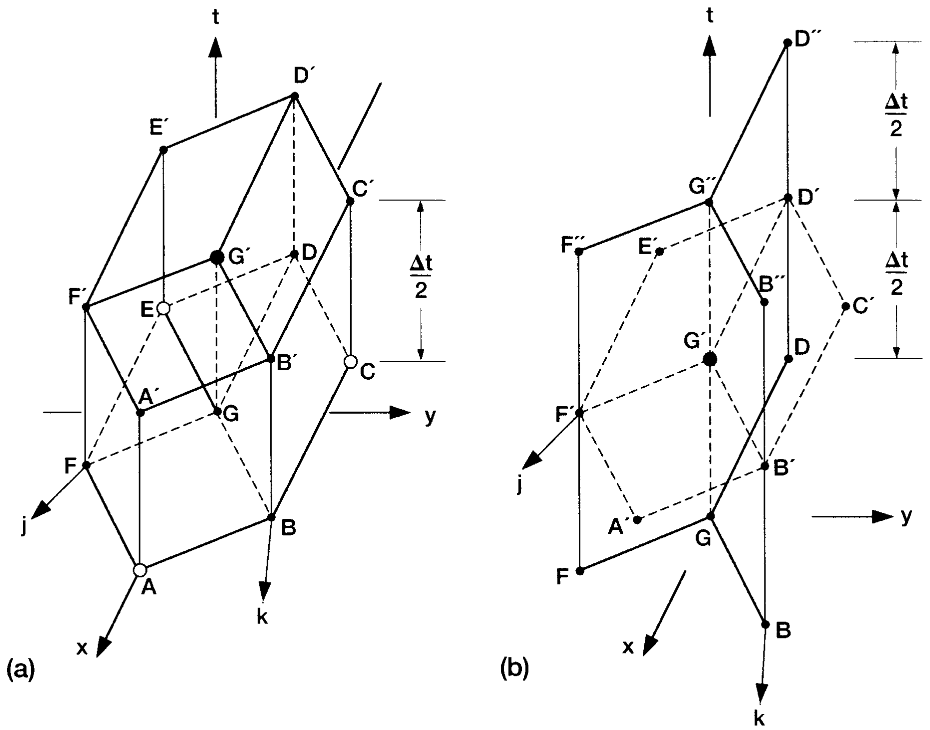
$$\vec{n}_3 = \frac{(-h, b+w/3, 0)}{\sqrt{h^2+(b+w/3)^2}}, \qquad \vec{n}_6 = -\vec{n}_3 \tag{4.6c}$$

For any $(x,y,t) \in \mathrm{SE}(j,k,n)$, $u(x,y,t)$ and $\vec{h}(x,y,t)$, respectively, are approximated by

$$u^*(x,y,t;j,k,n) \stackrel{def}{=} u_{j,k}^n + (u_x)_{j,k}^n(x-x_{j,k}) + (u_y)_{j,k}^n(y-y_{j,k}) + (u_t)_{j,k}^n(t-t^n) \tag{4.7}$$

and

$$\vec{h}^*(x,y,t;j,k,n) \stackrel{def}{=} [a_x u^*(x,y,t;j,k,n), a_y u^*(x,y,t;j,k,n), u^*(x,y,t;j,k,n)] \tag{4.8}$$



CE_1 (j, k, n) = box GFABG′F′A′B′
CE_2 (j, k, n) = box GBCDG′B′C′D′
CE_3 (j, k, n) = box GDEFG′D′E′F′

SE (j, k, n) = the union of four planes A′B′C′D′E′F′, GBB″G″, GDD″G″, and GG″F″F and their immediate neighborhoods.

$$G' = (j, k, n) \in \Omega_1,$$

$$A' = \left(j+\frac{1}{3}, k+\frac{1}{3}, n\right), \quad B' = \left(j-\frac{1}{3}, k+\frac{2}{3}, n\right), \quad C' = \left(j-\frac{2}{3}, k+\frac{1}{3}, n\right),$$

$$D' = \left(j-\frac{1}{3}, k-\frac{1}{3}, n\right), \quad E' = \left(j+\frac{1}{3}, k-\frac{2}{3}, n\right), \quad F' = \left(j+\frac{2}{3}, k-\frac{1}{3}, n\right)$$

Figure 10: (a) Conservation elements $\mathrm{CE}_r(j,k,n)$, $r=1,2,3$ for any $(j,k,n)\in\Omega_1$. (b) Solution element $\mathrm{SE}(j,k,n)$ for any $(j,k,n)\in\Omega_1$.

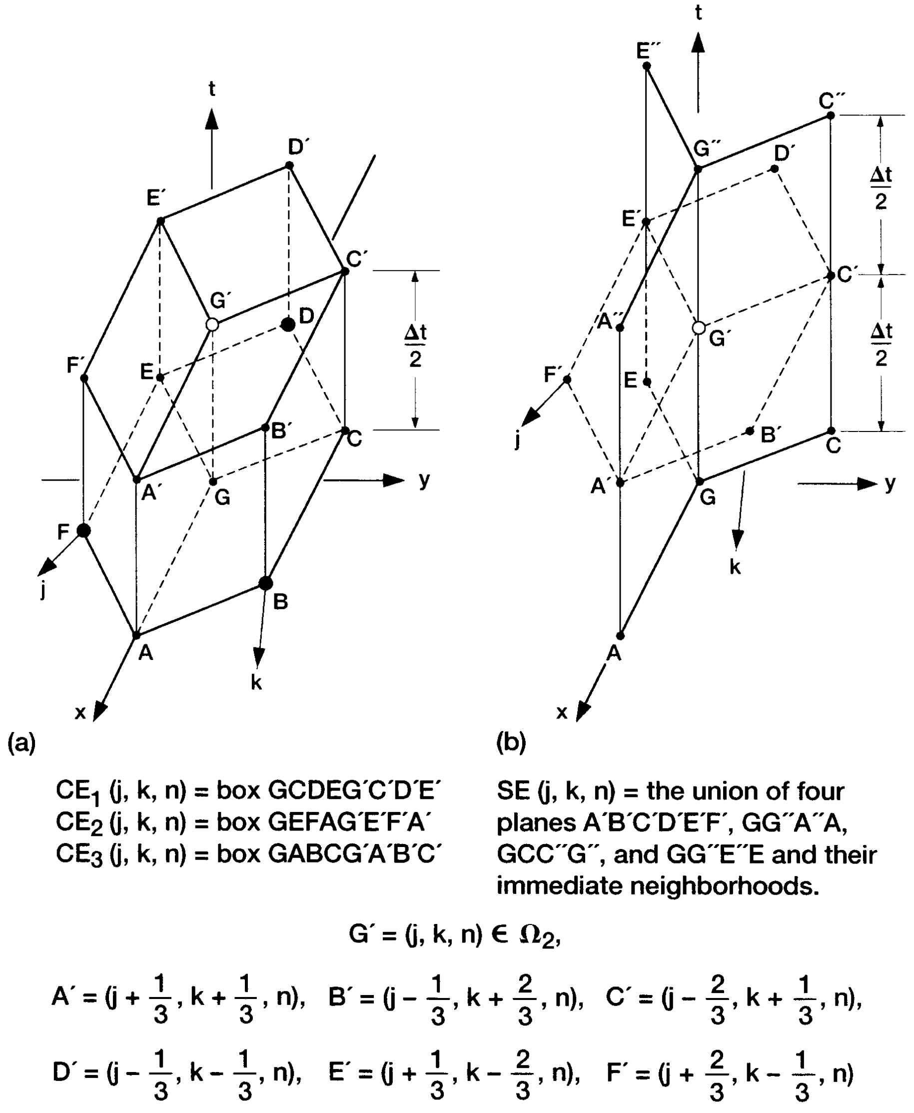


Figure 11: (a) Conservation elements $\mathrm{CE}_r(j,k,n)$, $r=1,2,3$ for any $(j,k,n)\in\Omega_2$. (b) Solution element $\mathrm{SE}(j,k,n)$ for any $(j,k,n)\in\Omega_2$.

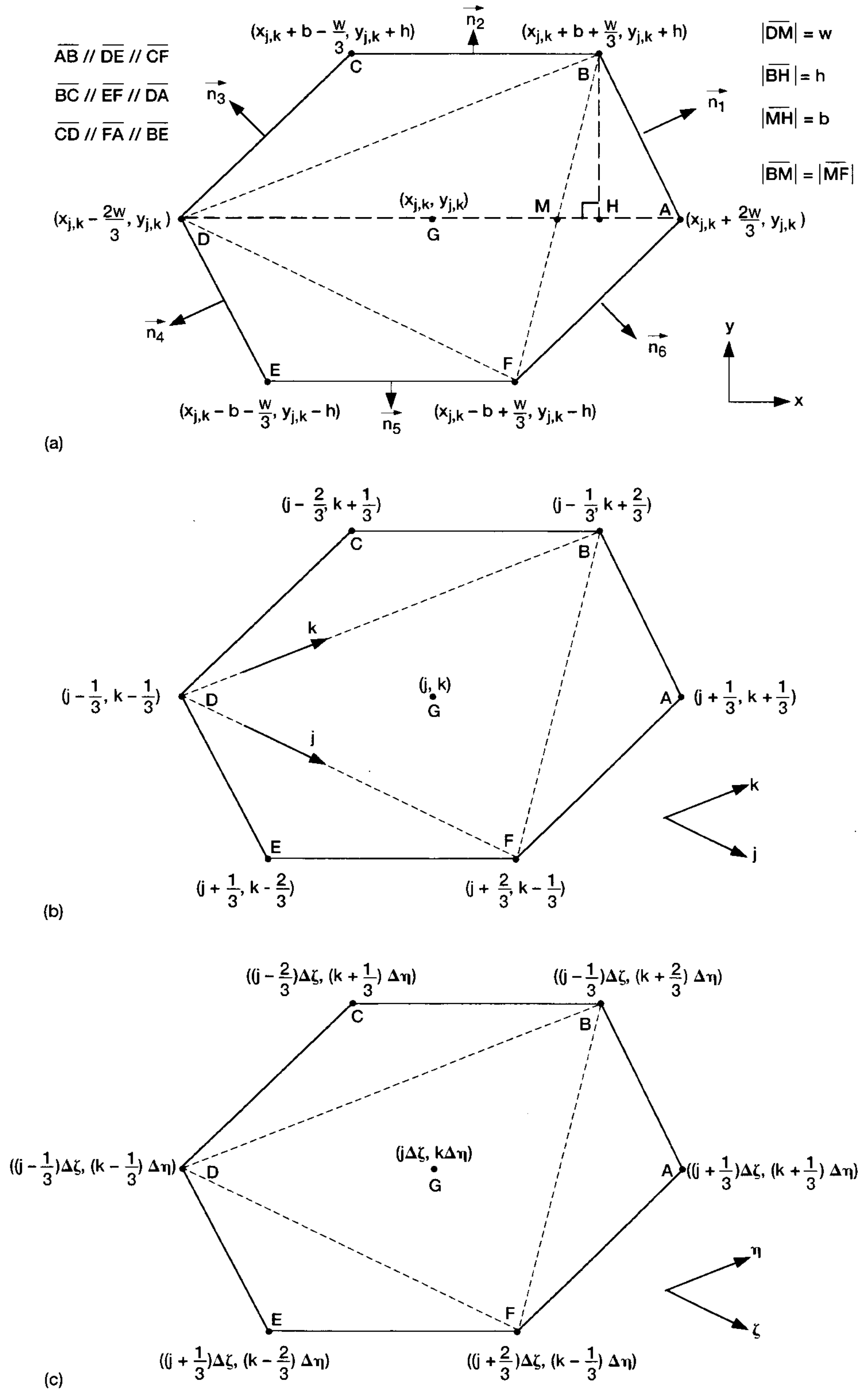


Figure 12: Geometry of the hexagon ABCDEF. (a) Relative positions of the vertices in terms of (x,y). (b) Relative positions of the vertices in terms of (j,k). (c) Relative positions of the vertices in terms of (ζ, η).

where $u^n_{j,k}$, $(u_x)^n_{j,k}$, $(u_y)^n_{j,k}$, and $(u_t)^n_{j,k}$ are constants within $\text{SE}(j,k,n)$. The last four coefficients, respectively, can be considered as the numerical analogues of the values of u, $\partial u/\partial x$, $\partial u/\partial y$, and $\partial u/\partial t$ at $(x_{j,k}, y_{j,k}, t^n)$. As a result, the expression on the right side of Eq. (4.7) can be considered as the first-order Taylor's expansion of $u(x,y,t)$ at $(x_{j,k}, y_{j,k}, t^n)$. Also note that Eq. (4.8) is the numerical analogue of Eq. (4.3).

We shall require that $u = u^*(x,y,t;j,k,n)$ satisfy Eq. (4.1) within $\text{SE}(j,k,n)$. As a result,

$$(u_t)^n_{j,k} = -\left[a_x(u_x)^n_{j,k} + a_y(u_y)^n_{j,k}\right] \tag{4.9}$$

Substituting Eq. (4.9) into Eq. (4.7), one has

$$\begin{aligned} u^*(x,y,t;j,k,n) &= u^n_{j,k} + (u_x)^n_{j,k}\left[(x - x_{j,k}) - a_x(t - t^n)\right] \\ &+ (u_y)^n_{j,k}\left[(y - y_{j,k}) - a_y(t - t^n)\right]. \end{aligned} \tag{4.10}$$

Thus there are three independent marching variables, i.e., $u^n_{j,k}$, $(u_x)^n_{j,k}$, and $(u_y)^n_{j,k}$ associated with a mesh point $(j,k,n) \in \Omega$. For any $(j,k,n) \in \Omega_1$, these variables will be determined in terms of those associated with the mesh points $(j+1/3, k+1/3, n-1/2)$, $(j-2/3, k+1/3, n-1/2)$, and $(j+1/3, k-2/3, n-1/2)$ (see Fig. 13(a)) by using the three flux conservation relations

$$\oint_{S(CE_r(j,k,n))} \vec{h}^* \cdot d\vec{s} = 0, \qquad r = 1,2,3 \tag{4.11}$$

Similarly, the marching variables at any $(j,k,n) \in \Omega_2$ are determined in terms of those associated with the mesh points $(j-1/3, k-1/3, n-1/2)$, $(j+2/3, k-1/3, n-1/2)$, and $(j-1/3, k+2/3, n-1/2)$ (see Fig. 13(b)) by using the three flux conservation relations Eq. (4.11). Obviously, Eq. (4.11) is the numerical analogue of Eq. (4.2).

As a result of Eq. (4.11), the total flux leaving the boundary of any CE is zero. Because the flux at any interface separating two neighboring CEs is calculated using the information from a single SE, the flux entering one of these CEs is equal to that leaving another. It follows that the local conservation conditions Eq. (4.11) will lead to a global conservation condition, i.e., *the total flux leaving the boundary of any space-time region that is the union of any combination of CEs will also vanish.*

In the following, several preliminaries will be given prior to the evaluation of Eq. (4.11). To proceed, note that a mesh line with j and n being constant or a mesh line with k and n being constant is not aligned with the x-axis or the y-axis. We shall introduce a new spatial coordinate system (ζ, η) with its axes aligned with the above mesh lines (see Fig. 12(c)).

Let $\vec{e}_x$ and $\vec{e}_y$ be the unit vectors in the x- and the y- directions, respectively. Let $\vec{e}_\zeta$ and $\vec{e}_\eta$ be the unit vectors in the directions of $\overrightarrow{DF}$ and $\overrightarrow{DB}$ (i.e., the j- and the k- directions–see Figs. 12(a)-(c)), respectively. It can be shown that

$$\vec{e}_\zeta = \left[(w-b)\vec{e}_x - h\vec{e}_y\right]/\Delta\zeta \tag{4.12}$$

and

$$\vec{e}_\eta = \left[(w+b)\vec{e}_x + h\vec{e}_y\right]/\Delta\eta \tag{4.13}$$

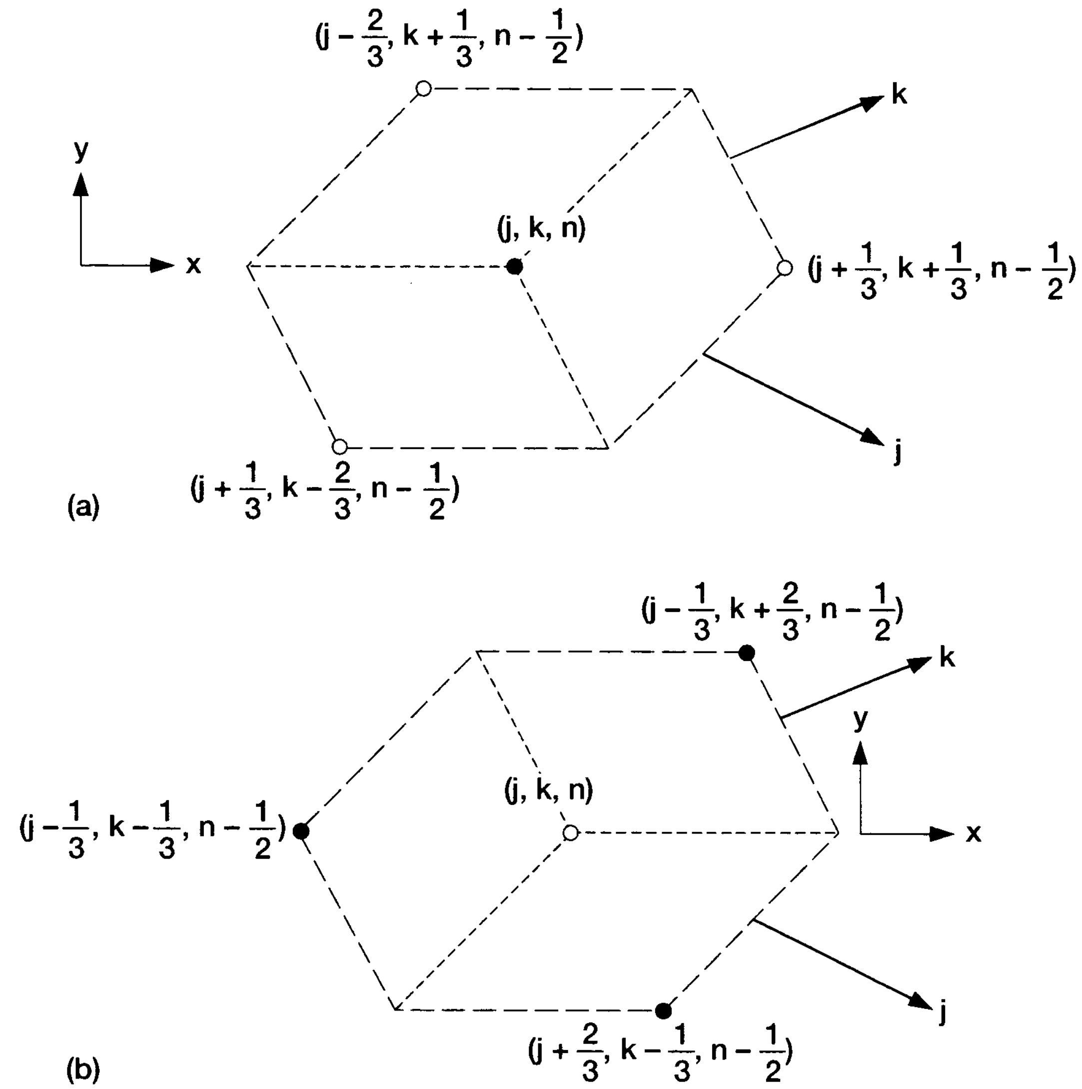


Figure 13: (a) The mesh points (j, k, n), $(j + 1/3, k + 1/3, n - 1/2)$, $(j - 2/3, k + 1/3, n - 1/2)$, and $(j + 1/3, k - 2/3, n - 1/2)$ that belongs to $\in \Omega_1$. (b) The mesh points (j, k, n), $(j - 1/3, k - 1/3, n - 1/2)$, $(j + 2/3, k - 1/3, n - 1/2)$, and $(j - 1/3, k + 2/3, n - 1/2)$ that belongs to $\in \Omega_2$.

where

$$\Delta\zeta \stackrel{def}{=} |\overline{DF}| = \sqrt{(w-b)^2 + h^2} \tag{4.14}$$

and

$$\Delta\eta \stackrel{def}{=} |\overline{DB}| = \sqrt{(w+b)^2 + h^2} \tag{4.15}$$

Let the origin of (x, y) also be that of (ζ, η). Then, at any point in E_3, the coordinates (ζ, η) are defined in terms of (x, y) using the relation

$$\zeta\, \vec{e}_\zeta + \eta\, \vec{e}_\eta = x\, \vec{e}_x + y\, \vec{e}_y \tag{4.16}$$

Substituting Eqs. (4.12) and (4.13) into Eq. (4.16), one has

$$\begin{pmatrix} x \\ y \end{pmatrix} = T \begin{pmatrix} \zeta \\ \eta \end{pmatrix} \tag{4.17}$$

and

$$\begin{pmatrix} \zeta \\ \eta \end{pmatrix} = T^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \tag{4.18}$$

Here

$$T \stackrel{def}{=} \begin{pmatrix} \dfrac{w-b}{\Delta\zeta} & \dfrac{w+b}{\Delta\eta} \\ \\ -\dfrac{h}{\Delta\zeta} & \dfrac{h}{\Delta\eta} \end{pmatrix} \tag{4.19}$$

and

$$T^{-1} \stackrel{def}{=} \begin{pmatrix} \dfrac{\Delta\zeta}{2w} & -\dfrac{(w+b)\Delta\zeta}{2wh} \\ \\ \dfrac{\Delta\eta}{2w} & \dfrac{(w-b)\Delta\eta}{2wh} \end{pmatrix} \tag{4.20}$$

Note that the existence of T^{-1}, the inverse of T, is assured if $wh \neq 0$.

With the aid of Eqs. (4.5), (4.18), and (4.20), it can be shown that the coordinates (ζ, η) of any mesh point $(j, k, n) \in \Omega$ are given by

$$\zeta = j\,\Delta\zeta, \qquad \text{and} \qquad \eta = k\,\Delta\eta \tag{4.21}$$

i.e., $\Delta\zeta$ and $\Delta\eta$ are the mesh intervals in the ζ- and the η- directions, respectively.

Next we shall introduce several coefficients that are tied to the coordinate system (ζ, η). Let

$$\begin{pmatrix} a_\zeta \\ a_\eta \end{pmatrix} \stackrel{def}{=} T^{-1} \begin{pmatrix} a_x \\ a_y \end{pmatrix} \tag{4.22}$$

Also, for any $(j,k,n) \in \Omega$, let

$$\begin{pmatrix} (u_\zeta)^n_{j,k} \\ \\ (u_\eta)^n_{j,k} \end{pmatrix} \stackrel{def}{=} T^t \begin{pmatrix} (u_x)^n_{j,k} \\ \\ (u_y)^n_{j,k} \end{pmatrix} \tag{4.23}$$

where T^t is the transpose of T. For those who are familiar with tensor analysis [55], the following comments will clarify the meaning of the above definitions:

(a) (a_ζ, a_η) are the *contravariant* components with respect to the coordinates (ζ,η) for the spatial vector whose x- and y- components are a_x and a_y, respectively.

(b) $((u_\zeta)^n_{j,k}, (u_\eta)^n_{j,k})$ are the *covariant* components with respect to the coordinates (ζ,η) for the spatial vector whose x- and y- components are $(u_x)^n_{j,k}$ and $(u_y)^n_{j,k}$, respectively.

(c) Because the contraction of the contravariant components of a vector and the covariant components of another is a scalar, Eq. (4.9) can be rewritten as

$$(u_t)^n_{j,k} = -\left[a_\zeta (u_\zeta)^n_{j,k} + a_\eta (u_\eta)^n_{j,k}\right] \tag{4.24}$$

(d) Under the *linear* coordinate transformation defined by Eqs. (4.17) and (4.18), $(\zeta - j\Delta\zeta, \eta - k\Delta\eta)$ are the contravariant components with respect to the coordinates (ζ,η) for the spatial vector whose x- and y- components are $x - x_{j,k}$ and $y - y_{j,k}$, respectively. Using the same reason given in (c), Eq. (4.10) implies that

$$u^\star(x,y,t;j,k,n) = u^\star(\zeta,\eta,t;j,k,n) \tag{4.25}$$

where

$$\begin{aligned} u^\star(\zeta,\eta,t;j,k,n) \stackrel{def}{=} u^n_{j,k} \quad &+ \quad (u_\zeta)^n_{j,k}\left[(\zeta - j\Delta\zeta) - a_\zeta(t-t^n)\right] \\ &+ \quad (u_\eta)^n_{j,k}\left[(\eta - k\Delta\eta) - a_\eta(t-t^n)\right] \end{aligned} \tag{4.26}$$

Note that Eqs. (4.24) and (4.25) can also be verified directly using Eqs. (4.18), (4.20), (4.22), and (4.23).

Next, let (i)

$$\nu_\zeta \stackrel{def}{=} \frac{3\Delta t}{2\Delta\zeta} a_\zeta, \qquad \text{and} \qquad \nu_\eta \stackrel{def}{=} \frac{3\Delta t}{2\Delta\eta} a_\eta \tag{4.27}$$

and (ii)

$$(u^+_\zeta)^n_{j,k} \stackrel{def}{=} \frac{\Delta\zeta}{6}(u_\zeta)^n_{j,k}, \qquad \text{and} \qquad (u^+_\eta)^n_{j,k} \stackrel{def}{=} \frac{\Delta\eta}{6}(u_\eta)^n_{j,k} \tag{4.28}$$

The coefficients defined in Eqs. (4.27) and (4.28) can be considered as the *normalized* counterparts of those defined in Eqs. (4.22) and (4.23). Furthermore, because $\Delta\zeta$ and $\Delta\eta$ are the mesh intervals in the ζ- and η- directions, respectively, Eq. (4.27) implies that $(2/3)\nu_\zeta$ and $(2/3)\nu_\eta$, respectively, are equal to the Courant numbers in the ζ- and η- directions, respectively.

Furthermore, let

$$\sigma_{11}^{(1)\pm} \stackrel{def}{=} 1 - \nu_\zeta - \nu_\eta \tag{4.29}$$

$$\sigma_{12}^{(1)\pm} \stackrel{def}{=} \pm(1 - \nu_\zeta - \nu_\eta)(1 + \nu_\zeta) \tag{4.30}$$

$$\sigma_{13}^{(1)\pm} \stackrel{def}{=} \pm(1 - \nu_\zeta - \nu_\eta)(1 + \nu_\eta) \tag{4.31}$$

$$\sigma_{21}^{(1)\pm} \stackrel{def}{=} 1 + \nu_\zeta \tag{4.32}$$

$$\sigma_{22}^{(1)\pm} \stackrel{def}{=} \mp(1 + \nu_\zeta)(2 - \nu_\zeta) \tag{4.33}$$

$$\sigma_{23}^{(1)\pm} \stackrel{def}{=} \pm(1 + \nu_\zeta)(1 + \nu_\eta) \tag{4.34}$$

$$\sigma_{31}^{(1)\pm} \stackrel{def}{=} 1 + \nu_\eta \tag{4.35}$$

$$\sigma_{32}^{(1)\pm} \stackrel{def}{=} \pm(1 + \nu_\eta)(1 + \nu_\zeta) \tag{4.36}$$

$$\sigma_{33}^{(1)\pm} \stackrel{def}{=} \mp(1 + \nu_\eta)(2 - \nu_\eta) \tag{4.37}$$

$$\sigma_{11}^{(2)\pm} \stackrel{def}{=} 1 + \nu_\zeta + \nu_\eta \tag{4.38}$$

$$\sigma_{12}^{(2)\pm} \stackrel{def}{=} \mp(1 + \nu_\zeta + \nu_\eta)(1 - \nu_\zeta) \tag{4.39}$$

$$\sigma_{13}^{(2)\pm} \stackrel{def}{=} \mp(1 + \nu_\zeta + \nu_\eta)(1 - \nu_\eta) \tag{4.40}$$

$$\sigma_{21}^{(2)\pm} \stackrel{def}{=} 1 - \nu_\zeta \tag{4.41}$$

$$\sigma_{22}^{(2)\pm} \stackrel{def}{=} \pm(1 - \nu_\zeta)(2 + \nu_\zeta) \tag{4.42}$$

$$\sigma_{23}^{(2)\pm} \stackrel{def}{=} \mp(1 - \nu_\zeta)(1 - \nu_\eta) \tag{4.43}$$

$$\sigma_{31}^{(2)\pm} \stackrel{def}{=} 1 - \nu_\eta \tag{4.44}$$

$$\sigma_{32}^{(2)\pm} \stackrel{def}{=} \mp(1 - \nu_\eta)(1 - \nu_\zeta) \tag{4.45}$$

and

$$\sigma_{33}^{(2)\pm} \stackrel{def}{=} \pm(1 - \nu_\eta)(2 + \nu_\eta) \tag{4.46}$$

Note that:

(a) Each of Eqs. (4.29)–(4.46) represents two equations. One corresponds to the upper signs while the other, to the lower signs.

(b) The definitions given in Eqs. (4.29)–(4.37) will be used in the first marching step of the 2D *a* scheme; while those given in Eqs. (4.38)–(4.46) will be used in the second marching step. It is seen that the expressions on the right sides of the former can be converted to those of the latter, respectively, by reversing the "+" and "−" signs. Moreover, for every pair of r and s $(r, s = 1, 2, 3)$, $\sigma_{rs}^{(1)-}$ and $\sigma_{rs}^{(2)-}$ are converted to $\sigma_{rs}^{(2)+}$ and $\sigma_{rs}^{(1)+}$, respectively, if ν_ζ, and ν_η are replaced by $-\nu_\zeta$, and $-\nu_\eta$, respectively.

(c) We have

$$\sigma_{11}^{(q)\pm} + \sigma_{21}^{(q)\pm} + \sigma_{31}^{(q)\pm} = 3, \qquad q = 1, 2 \tag{4.47}$$

and

$$\begin{aligned} &\sigma_{12}^{(q)\pm} + \sigma_{22}^{(q)\pm} + \sigma_{32}^{(q)\pm} \\ &= \sigma_{13}^{(q)\pm} + \sigma_{23}^{(q)\pm} + \sigma_{33}^{(q)\pm} = 0, \qquad q = 1, 2 \end{aligned} \tag{4.48}$$

To simplify the following development, let

$$(j, k; 1, 1) \stackrel{def}{=} j + 1/3, k + 1/3 \tag{4.49a}$$

$$(j, k; 1, 2) \stackrel{def}{=} j - 2/3, k + 1/3 \tag{4.49b}$$

$$(j, k; 1, 3) \stackrel{def}{=} j + 1/3, k - 2/3 \tag{4.49c}$$

$$(j, k; 2, 1) \stackrel{def}{=} j - 1/3, k - 1/3 \tag{4.50a}$$

$$(j, k; 2, 2) \stackrel{def}{=} j + 2/3, k - 1/3 \tag{4.50b}$$

$$(j, k; 2, 3) \stackrel{def}{=} j - 1/3, k + 2/3 \tag{4.50c}$$

Note that (i) $(j, k; 1, r)$, $r = 1, 2, 3$, are the spatial mesh indices of points A, C, and E depicted in Fig. 10(a), respectively, (ii) $(j, k; 2, r)$, $r = 1, 2, 3$, are the spatial mesh indices of points D, F, and B depicted in Fig. 11(a), respectively, and (iii) the mesh indices on the right sides of Eqs. (4.49a,b,c) can be converted to those in Eqs. (4.50a,b,c), respectively, by reversing the "+" and "−" signs.

Equation (4.11) is evaluated in Appendix B. Let $(j, k, n) \in \Omega_q$ with $q = 1, 2$. Then, for any $r = 1, 2, 3$, the result of evaluation can be expressed as:

$$\left[\sigma_{r1}^{(q)+} u + \sigma_{r2}^{(q)+} u_\zeta^+ + \sigma_{r3}^{(q)+} u_\eta^+\right]_{j,k}^n = \left[\sigma_{r1}^{(q)-} u + \sigma_{r2}^{(q)-} u_\zeta^+ + \sigma_{r3}^{(q)-} u_\eta^+\right]_{(j,k;q,r)}^{n-1/2} \tag{4.51}$$

According to Eqs. (4.29)–(4.31), $\sigma_{11}^{(1)\pm}$, $\sigma_{12}^{(1)\pm}$, and $\sigma_{13}^{(1)\pm}$ contain a common factor $(1 - \nu_\zeta - \nu_\eta)$. Similarly, each of three consecutive pairs of coefficients defined in Eqs. (4.32)–(4.46) also contain a common factor. As a result, if one assumes that (i) $1 - \nu_\zeta - \nu_\eta \neq 0$, (ii) $1 + \nu_\zeta \neq 0$, (iii) $1 + \nu_\eta \neq 0$, (iv) $1 + \nu_\zeta + \nu_\eta \neq 0$, (v) $1 - \nu_\zeta \neq 0$ and (vi) $1 - \nu_\eta \neq 0$, i.e.,

$$\left[1 - (\nu_\zeta + \nu_\eta)^2\right] \left(1 - \nu_\zeta^2\right) \left(1 - \nu_\eta^2\right) \neq 0 \tag{4.52}$$

then the six equations ($q = 1, 2$ and $r = 1, 2, 3$) given in Eq. (4.51) can be simplified as

$$\left[u + (1 + \nu_\zeta) u_\zeta^+ + (1 + \nu_\eta) u_\eta^+\right]_{j,k}^n = s_1^{(1)}, \qquad (j, k, n) \in \Omega_1 \tag{4.53}$$

$$\left[u - (2 - \nu_\zeta) u_\zeta^+ + (1 + \nu_\eta) u_\eta^+\right]_{j,k}^n = s_2^{(1)}, \qquad (j, k, n) \in \Omega_1 \tag{4.54}$$

$$\left[u + (1+\nu_\zeta)u_\zeta^+ - (2-\nu_\eta)u_\eta^+\right]_{j,k}^n = s_3^{(1)}, \qquad (j,k,n) \in \Omega_1 \tag{4.55}$$

$$\left[u - (1-\nu_\zeta)u_\zeta^+ - (1-\nu_\eta)u_\eta^+\right]_{j,k}^n = s_1^{(2)}, \qquad (j,k,n) \in \Omega_2 \tag{4.56}$$

$$\left[u + (2+\nu_\zeta)u_\zeta^+ - (1-\nu_\eta)u_\eta^+\right]_{j,k}^n = s_2^{(2)}, \qquad (j,k,n) \in \Omega_2 \tag{4.57}$$

and

$$\left[u - (1-\nu_\zeta)u_\zeta^+ + (2+\nu_\eta)u_\eta^+\right]_{j,k}^n = s_3^{(2)}, \qquad (j,k,n) \in \Omega_2 \tag{4.58}$$

respectively. Here

$$s_1^{(1)} \overset{def}{=} \left[u - (1+\nu_\zeta)u_\zeta^+ - (1+\nu_\eta)u_\eta^+\right]_{(j,k;1,1)}^{n-1/2}, \qquad (j,k,n) \in \Omega_1 \tag{4.59}$$

$$s_2^{(1)} \overset{def}{=} \left[u + (2-\nu_\zeta)u_\zeta^+ - (1+\nu_\eta)u_\eta^+\right]_{(j,k;1,2)}^{n-1/2}, \qquad (j,k,n) \in \Omega_1 \tag{4.60}$$

$$s_3^{(1)} \overset{def}{=} \left[u - (1+\nu_\zeta)u_\zeta^+ + (2-\nu_\eta)u_\eta^+\right]_{(j,k;1,3)}^{n-1/2}, \qquad (j,k,n) \in \Omega_1 \tag{4.61}$$

$$s_1^{(2)} \overset{def}{=} \left[u + (1-\nu_\zeta)u_\zeta^+ + (1-\nu_\eta)u_\eta^+\right]_{(j,k;2,1)}^{n-1/2}, \qquad (j,k,n) \in \Omega_2 \tag{4.62}$$

$$s_2^{(2)} \overset{def}{=} \left[u - (2+\nu_\zeta)u_\zeta^+ + (1-\nu_\eta)u_\eta^+\right]_{(j,k;2,2)}^{n-1/2}, \qquad (j,k,n) \in \Omega_2 \tag{4.63}$$

and

$$s_3^{(2)} \overset{def}{=} \left[u + (1-\nu_\zeta)u_\zeta^+ - (2+\nu_\eta)u_\eta^+\right]_{(j,k;2,3)}^{n-1/2}, \qquad (j,k,n) \in \Omega_2 \tag{4.64}$$

The current 2D a scheme will be constructed using Eqs. (4.53)–(4.58) without assuming Eq. (4.52). Note that Eqs. (4.53)–(4.58) imply Eq. (4.51) for *any* ν_ζ *and* ν_η. However, the reverse is false unless Eq. (4.52) is assumed.

Note that the expressions within the brackets in Eqs. (4.53)–(4.55) and (4.59)–(4.61), respectively, can be converted to those in Eqs. (4.56)–(4.58) and (4.62)–(4.64) by reversing the "+" and "−" signs.

It can be shown that Eqs. (4.53)–(4.55) are equivalent to

$$u_{j,k}^n = \frac{1}{3}\left[(1-\nu_\zeta-\nu_\eta)s_1^{(1)} + (1+\nu_\zeta)s_2^{(1)} + (1+\nu_\eta)s_3^{(1)}\right] \tag{4.65}$$

$$(u_\zeta^+)_{j,k}^n = (u_\zeta^{a+})_{j,k}^n \overset{def}{=} \frac{1}{3}\left(s_1^{(1)} - s_2^{(1)}\right) \tag{4.66}$$

and

$$(u_\eta^+)_{j,k}^n = (u_\eta^{a+})_{j,k}^n \overset{def}{=} \frac{1}{3}\left(s_1^{(1)} - s_3^{(1)}\right) \tag{4.67}$$

where $(j,k,n) \in \Omega_1$. Also Eqs. (4.56)–(4.58) are equivalent to

$$u_{j,k}^n = \frac{1}{3}\left[(1+\nu_\zeta+\nu_\eta)s_1^{(2)} + (1-\nu_\zeta)s_2^{(2)} + (1-\nu_\eta)s_3^{(2)}\right] \tag{4.68}$$

$$(u_\zeta^+)_{j,k}^n = (u_\zeta^{a+})_{j,k}^n \stackrel{def}{=} \frac{1}{3}\left(s_2^{(2)} - s_1^{(2)}\right) \tag{4.69}$$

and

$$(u_\eta^+)_{j,k}^n = (u_\eta^{a+})_{j,k}^n \stackrel{def}{=} \frac{1}{3}\left(s_3^{(2)} - s_1^{(2)}\right) \tag{4.70}$$

where $(j,k,n) \in \Omega_2$.

At this juncture, it should be emphasized that Eqs. (4.65) and (4.68) can be derived directly from Eq. (4.51). As a matter of fact, with the aid of Eqs. (4.47) and (4.48), we can obtain Eq. (4.65) (Eq. (4.68)) by summing over the three equations with $q = 1$ ($q = 2$) and $r = 1, 2, 3$ in Eq. (4.51).

The 2D a scheme is formed by repeatedly applying the two marching steps defined by Eqs. (4.65)–(4.67) and Eqs. (4.68)–(4.70), respectively. It has been shown numerically that it is of second order in accuracy for $u_{j,k}^n$, $(u_\zeta)_{j,k}^n$ and $(u_\eta)_{j,k}^n$ assuming that ν_ζ and ν_η are held constant in the process of mesh refinement (note: as a result of Eq. (4.28), the 2D a scheme is third order accurate for $(u_\zeta^+)_{j,k}^n$ and $(u_\eta^+)_{j,k}^n$). Note that the superscript symbol "a" in $(u_\zeta^{a+})_{j,k}^n$ and $(u_\eta^{a+})_{j,k}^n$ is introduced to remind the reader that Eqs. (4.66), (4.67), (4.69) and (4.70) are valid for the 2D a scheme. Although the 2D a scheme is constructed using a procedure very much parallel to that used to construct the 1D a scheme, the former is more complex than the latter in many aspects. *One key difference between these two schemes is that the 2D a scheme is formed by two distinctly different marching steps while the 1D a scheme is formed by repeatedly applying the same marching step defined by the inviscid version of Eq. (2.14) in [2].* It is this difference that, in the 2D case, makes it necessary to divide the mesh points into two sets Ω_1 and Ω_2.

As a preliminary for the stability analysis of the 2D a scheme given in Sec. 6, for any $(j,k,n) \in \Omega$, let

$$\vec{q}(j,k,n) \stackrel{def}{=} \begin{pmatrix} u \\ \\ u_\zeta^+ \\ \\ u_\eta^+ \end{pmatrix}_{j,k}^n \tag{4.71}$$

Furthermore, let the six 3×3 matrices $Q_r^{(q)}$, $q = 1, 2$, and $r = 1, 2, 3$, respectively, be the special cases of those defined in Eqs. (5.18)–(5.23) (see Sec. 5) with $\epsilon = 0$. Then Eqs. (4.65)–(4.70) can be expressed as

$$\vec{q}(j,k,n) = \sum_{r=1}^{3} Q_r^{(q)} \vec{q}((j,k;q,r), n - 1/2), \qquad (j,k,n) \in \Omega_q \tag{4.72}$$

Combining Eqs. (4.72) and (4.49a)–(4.50c), one has (i)

$$\begin{aligned} \vec{q}(j,k,n) &= Q_1^{(1)} Q_2^{(2)} \vec{q}(j+1,k,n-1) + Q_1^{(1)} Q_3^{(2)} \vec{q}(j,k+1,n-1) \\ &+ Q_2^{(1)} Q_1^{(2)} \vec{q}(j-1,k,n-1) + Q_2^{(1)} Q_3^{(2)} \vec{q}(j-1,k+1,n-1) \end{aligned}$$

$$
\begin{aligned}
&+ \quad Q_3^{(1)}Q_1^{(2)}\vec{q}(j,k-1,n-1) + Q_3^{(1)}Q_2^{(2)}\vec{q}(j+1,k-1,n-1) \\
&+ \quad (Q_1^{(1)}Q_1^{(2)} + Q_2^{(1)}Q_2^{(2)} + Q_3^{(1)}Q_3^{(2)})\,\vec{q}(j,k,n-1) \qquad (4.73)
\end{aligned}
$$

where $(j,k,n) \in \Omega_1$; and (ii)

$$
\begin{aligned}
\vec{q}(j,k,n) \quad = \quad & Q_1^{(2)}Q_2^{(1)}\vec{q}(j-1,k,n-1) + Q_1^{(2)}Q_3^{(1)}\vec{q}(j,k-1,n-1) \\
+ \quad & Q_2^{(2)}Q_1^{(1)}\vec{q}(j+1,k,n-1) + Q_2^{(2)}Q_3^{(1)}\vec{q}(j+1,k-1,n-1) \\
+ \quad & Q_3^{(2)}Q_1^{(1)}\vec{q}(j,k+1,n-1) + Q_3^{(2)}Q_2^{(1)}\vec{q}(j-1,k+1,n-1) \\
+ \quad & (Q_1^{(2)}Q_1^{(1)} + Q_2^{(2)}Q_2^{(1)} + Q_3^{(2)}Q_3^{(1)})\,\vec{q}(j,k,n-1) \qquad (4.74)
\end{aligned}
$$

where $(j,k,n) \in \Omega_2$. Note that (i) Eq. (4.73) relates the marching variables at two adjacent half-integer time levels; and (ii) Eq. (4.74) relates the marching variables at two adjacent whole-integer time levels.

The 2D a scheme has several nontraditional features. They are summarized in the following comments:

(a) As in the case of the 1D a scheme, the 2D a scheme also has the simplest stencil possible, i.e., in each of their two marching steps, the stencil is a tetrahedron in 3D space-time with one vertex at the upper time level and the other three vertices at the lower time level.

(b) As in the case of the 1D a scheme, each of the six flux conservation conditions associated with the 2D a scheme., i.e., those given in Eq. (4.51) ($q = 1,2$ and $r = 1,2,3$), represents a relation among the marching variables associated with *only two neighboring SEs*.

(c) As in the case of the 1D a scheme, the 2D a scheme also is non-dissipative if it is stable. It is shown in Sec. 7 that the 2D a scheme is *neutrally stable* if

$$
|\nu_\zeta| < 1.5, \quad |\nu_\eta| < 1.5, \quad \text{and} \quad |\nu_\zeta + \nu_\eta| < 1.5 \qquad (4.75)
$$

As depicted in Fig. 14, the domain of stability defined by Eq. (4.75) is a hexagonal region in the ν_ζ-ν_η plane. Moreover, it will also be shown later that *Eq. (4.75) can be interpreted as the requirement that the physical domain of dependence of Eq. (4.1) be within the numerical domain of dependence*. Note that the points on the ν_ζ-ν_η plane that *violate* Eq. (4.52) form the boundary of a hexagonal region which is entirely within the stability domain defined in Eq. (4.75). As was emphasized earlier, the 2D a scheme applies even at these points.

(d) It is shown in [9] that the 2D a scheme has the following property, i.e., for any $(j,k,n) \in \Omega$,

$$
\vec{q}(j,k,n+1) \rightarrow \vec{q}(j,k,n) \quad \text{as} \quad \Delta t \rightarrow 0 \qquad (4.76)
$$

if a_x, a_y, w, b, and h are held constant. The 1D a scheme also possesses a similar property, i.e., Eq. (2.19) in [2]. The above property usually is not shared by other schemes that use a mesh that is staggered in time, e.g., the Lax scheme [52].

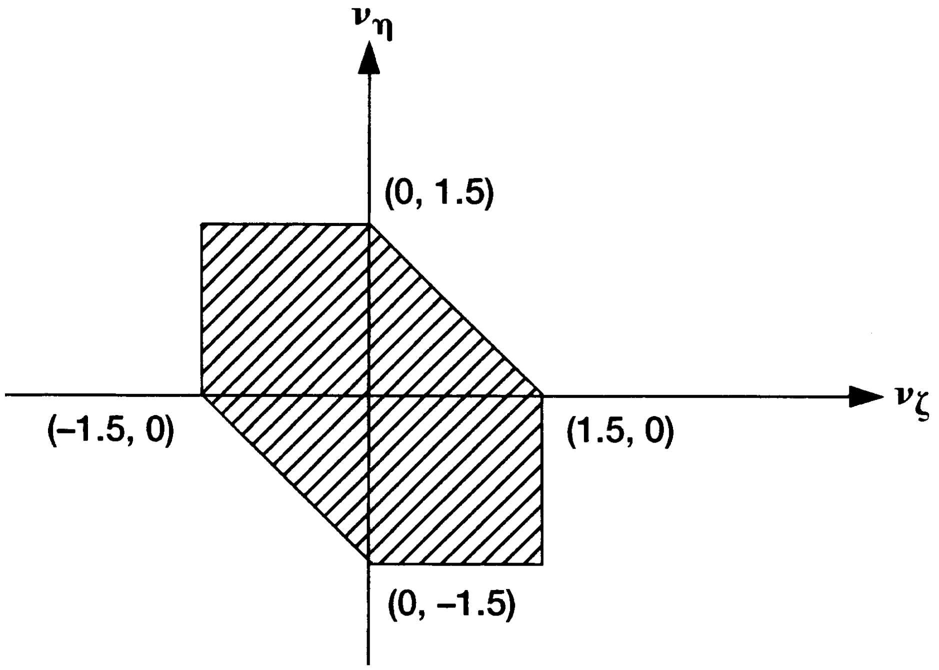


Figure 14: The stability domain of the 2D a-scheme.

(e) As in the case of the 1D a scheme, the 2D a scheme is also a two-way marching scheme. In other words, Eqs. (4.53)–(4.58) can also be used to construct the backward time-marching version of the 2D a scheme. More discussions on this subject are given in [9].

This section is concluded with the following remarks:

(a) the 2D a scheme is only a special case of the 2D a-μ scheme described in [9]. It is a solver for the 2D convection-diffusion equation

$$\frac{\partial u}{\partial t} + a_x \frac{\partial u}{\partial x} + a_y \frac{\partial u}{\partial y} - \mu \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) = 0 \tag{4.77}$$

where a_x, a_y, and μ (≥ 0) are constants. Note that this solver, as in the case of its 1D counterpart, is unconditionally stable if $a_x = a_y = 0$.

(b) It should be emphasized that, with the aid of Eqs. (4.17)–(4.20), (4.22), and (4.23), the 2D a scheme can also be expressed in terms of the marching variables and the coefficients tied to the coordinates (x, y). In other words, the coordinates (ζ, η) are introduced solely for the purpose of simplifying the current development. *The essence of the 2D a scheme, and the schemes to be introduced in the following sections, is not dependent on the choice of the coordinates in terms of which these schemes are expressed.*

5. The 2D a-ϵ and a-ϵ-α-β Schemes

The 2D a scheme is non-dissipative and reversible in time. It is well known that a non-dissipative numerical analogue of Eq. (4.1) generally becomes unstable or highly dispersive when it is extended to model the 2D unsteady Euler equations. It is also obvious that a scheme that is reversible in time cannot model a physical problem that is irreversible in time, e.g., an inviscid flow problem involving shocks. As a result, the 2D a scheme will be extended to become the dissipative 2D a-ϵ and a-ϵ-α-β scheme before it is extended to model the Euler equations. As will be shown, the 2D extensions are carried out in a fashion completely parallel to their 1D counterparts.

5.1. The 2D a-ϵ Scheme

To proceed, note that the CEs for the 2D a-ϵ scheme generally are not those associated with the 2D a scheme. Here only a single CE is associated with a mesh point $(j,k,n) \in \Omega$. This CE, denoted by CE(j,k,n), is the union of CE$_r(j,k,n)$, $r = 1,2,3$. In other words,

$$CE(j,k,n) \stackrel{def}{=} [CE_1(j,k,n)] \cup [CE_2(j,k,n)] \cup [CE_3(j,k,n)] \tag{5.1}$$

Instead of Eq. (4.11), here we assume the less stringent conservation condition

$$\oint_{S(CE(j,k,n))} \vec{h}^* \cdot d\vec{s} = 0 \tag{5.2}$$

Obviously, (i) E_3 can be filled with the new CEs, and (ii) the total flux leaving the boundary of any space-time region that is the union of any new CEs will also vanish.

Moreover, because of Eq. (5.1), Eq. (5.2) must be true if Eq. (4.11) is assumed. As a matter of fact, a direct evaluation of Eq. (5.2) reveals that it is equivalent to Eq. (4.65) (Eq. (4.68)) if $(j,k,n) \in \Omega_1$ ($(j,k,n) \in \Omega_2$). As a result, Eqs. (4.65) and (4.68) are shared by the 2D a scheme and 2D a-ϵ scheme. Recall that Eq. (2.7) is also shared by the 1D a and a-ϵ schemes. In this section, using a procedure similar to that which was used to extend the 1D a scheme to become the 1D a-ϵ scheme, the two marching steps that form the 2D a-ϵ scheme will be constructed by modifying the other equations in the 2D a scheme, i.e., Eqs. (4.66), (4.67), (4.69), and (4.70). As a prerequisite, first we shall provide a geometric interpretation of the procedure by which the second equation of the 1D a scheme, i.e., Eq. (2.8), was extended to become the second equation of the 1D a-ϵ scheme, i.e., Eq. (2.13).

The key step in extending the 1D a scheme to the 1D a-ϵ scheme is the construction of a central difference approximation of $\partial u/\partial x$ at the mesh point (j,n). The approximation is given as the fraction within the parentheses on the extreme right side of Eq. (2.12). Consider a line segment in the x-u space joining the two points $(x_{j-1/2}, u'^n_{j-1/2})$ and $(x_{j+1/2}, u'^n_{j+1/2})$. It is obvious that the above central-difference approximation is the value of the slope du/dx of this line segment. In the following modification, instead of considering a line segment in the x-u space joining two points, we begin with the construction of a plane in the ζ-η-u space that intersects three given points.

To proceed, for any $(j,k,n) \in \Omega_q$, $q=1,2$, let

$$u'^{n}_{(j,k;q,r)} \stackrel{def}{=} \left(u + \frac{\Delta t}{2} u_t\right)^{n-1/2}_{(j,k;q,r)}, \qquad r = 1,2,3 \tag{5.3}$$

By its definition, $u'^{n}_{(j,k;q,r)}$ is a finite-difference approximation of u at $((j,k;q,r),n)$. With the aid of Eqs. (4.24), (4.27) and (4.28), Eq. (5.3) implies that

$$u'^{n}_{(j,k;q,r)} = \left[u - 2\left(\nu_\zeta u^+_\zeta + \nu_\eta u^+_\eta\right)\right]^{n-1/2}_{(j,k;q,r)} \tag{5.4}$$

For both the case $q=1$ (see Fig. 15(a)) and the case $q=2$ (see Fig. 15(b)), let P, Q, and R be the three points in the ζ-η-u space with their (i) ζ- and η-coordinates being those of the mesh points $((j,k;q,r),n-1/2)$, $r=1,2,3$, respectively, and (ii) their u-coordinates being $u'^{n}_{(j,k;q,r)}$, $r=1,2,3$, respectively. It can be shown that the plane in the ζ-η-u space that intersects the above three points is represented by

$$u = (u^c_\zeta)^n_{j,k}(\zeta - j\Delta\zeta) + (u^c_\eta)^n_{j,k}(\eta - k\Delta\eta) + (u^c)^n_{j,k} \tag{5.5}$$

where

$$(u^c)^n_{j,k} \stackrel{def}{=} \frac{1}{3}\sum_{r=1}^{3} u'^{n}_{(j,k;q,r)} \tag{5.6}$$

$$(u^c_\zeta)^n_{j,k} \stackrel{def}{=} (-1)^q \left(u'^{n}_{(j,k;q,2)} - u'^{n}_{(j,k;q,1)}\right)/\Delta\zeta \tag{5.7}$$

and

$$(u^c_\eta)^n_{j,k} \stackrel{def}{=} (-1)^q \left(u'^{n}_{(j,k;q,3)} - u'^{n}_{(j,k;q,1)}\right)/\Delta\eta \tag{5.8}$$

The coordinates of the points O and O_c depicted in both Fig. 15(a) and Fig. 15(b) are $(j\Delta\zeta, k\Delta\eta, u^n_{j,k})$ and $(j\Delta\zeta, k\Delta\eta, (u^c)^n_{j,k})$, respectively. Here $u^n_{j,k}$ is evaluated using (i) Eq. (4.65) if $q=1$ and (ii) Eq. (4.68) if $q=2$. Equation (5.5) implies that point O_c is on the same plane that contains points P, Q, and R. Because generally $u^n_{j,k} \neq (u^c)^n_{j,k}$, points O, P, Q and R generally are not on the same plane. Moreover, for every point on the plane represented by Eq. (5.5),

$$\left(\frac{\partial u}{\partial \zeta}\right)_\eta = (u^c_\zeta)^n_{j,k}, \quad \text{and} \quad \left(\frac{\partial u}{\partial \eta}\right)_\zeta = (u^c_\eta)^n_{j,k} \tag{5.9}$$

As a result of the above considerations, and the fact that the spatial projection of the mesh point $(j,k,n) \in \Omega_q$ on the $(n-1/2)$th time level is the centroid of the triangle formed with the mesh points $((j,k;q,r),n-1/2)$, $r=1,2,3$, one concludes that $(u^c)^n_{j,k}$, $(u^c_\zeta)^n_{j,k}$, and $(u^c_\eta)^n_{j,k}$ are central-difference approximations of u, $\partial u/\partial\zeta$, and $\partial u/\partial\eta$, respectively, at the mesh point (j,k,n).

To proceed, for any $(j,k,n) \in \Omega$, let

$$(u^{c+}_\zeta)^n_{j,k} \stackrel{def}{=} \frac{\Delta\zeta}{6}(u^c_\zeta)^n_{j,k} \quad \text{and} \quad (u^{c+}_\eta)^n_{j,k} \stackrel{def}{=} \frac{\Delta\eta}{6}(u^c_\eta)^n_{j,k} \tag{5.10}$$

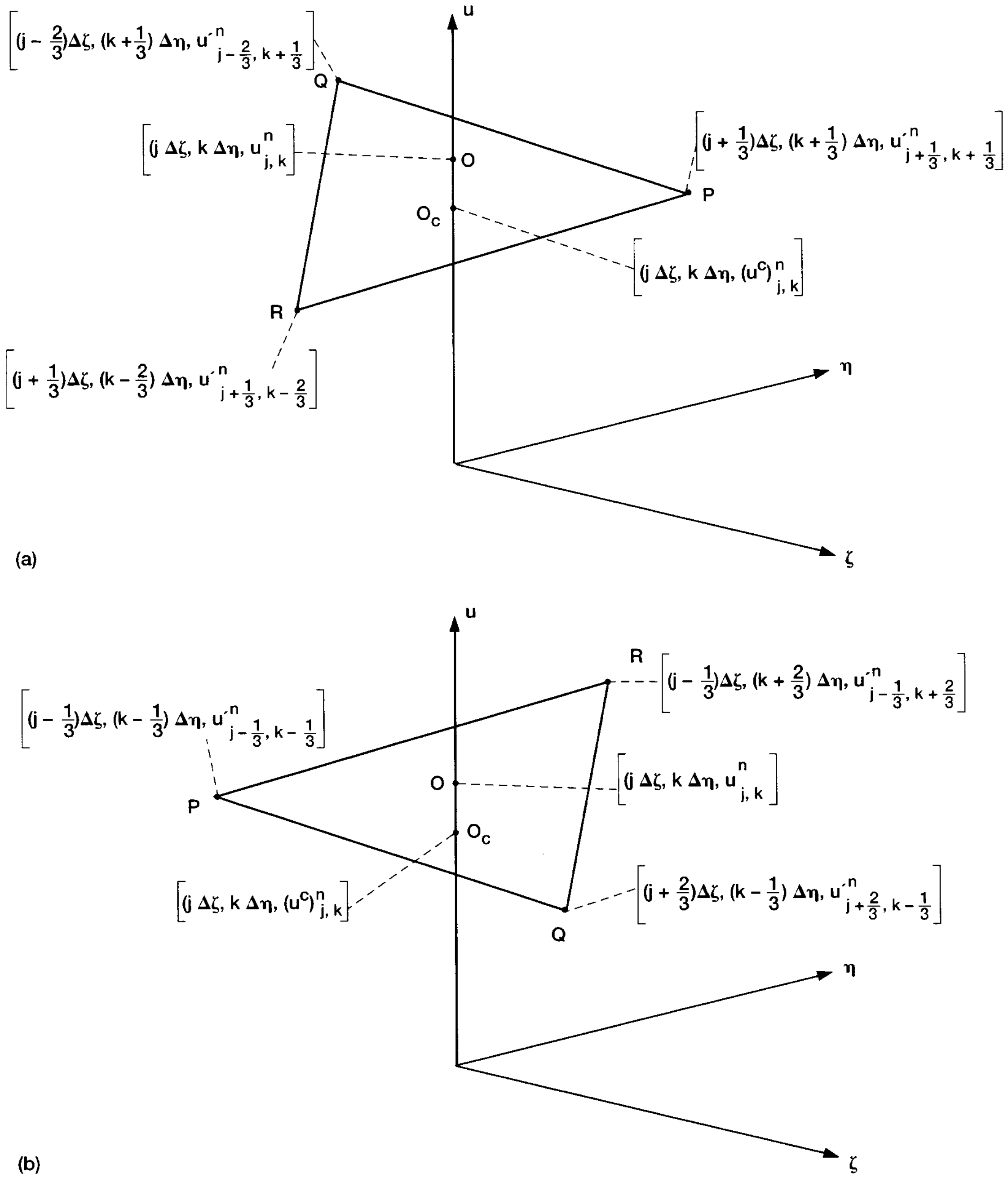


Figure 15: Construction of the 2D a-ϵ and a-ϵ-α-β schemes. (a) $(j, k, n) \in \Omega_1$. (b) $(j, k, n) \in \Omega_2$.

Then the 2D a-ϵ scheme can be defined as follows: For any $(j,k,n) \in \Omega_1$, we assume Eq. (4.65) and

$$(u_\zeta^+)_{j,k}^n = (u_\zeta^{a+})_{j,k}^n + 2\epsilon\left(u_\zeta^{c+} - u_\zeta^{a+}\right)_{j,k}^n \tag{5.11}$$

and

$$(u_\eta^+)_{j,k}^n = (u_\eta^{a+})_{j,k}^n + 2\epsilon\left(u_\eta^{c+} - u_\eta^{a+}\right)_{j,k}^n \tag{5.12}$$

with *the understanding that* $(u_\zeta^{a+})_{j,k}^n$ *and* $(u_\eta^{a+})_{j,k}^n$ *are those defined in Eqs. (4.66) and (4.67).* On the other hand, for any $(j,k,n) \in \Omega_2$, we assume Eqs. (4.68), (5.11) and (5.12) with *the understanding that* $(u_\zeta^{a+})_{j,k}^n$ *and* $(u_\eta^{a+})_{j,k}^n$ *are those defined in Eqs. (4.69) and (4.70).*

With the aid of Eqs. (5.4), (5.7), (5.8), (5.10), (4.66), (4.67), (4.69) and (4.70), it can be shown that (i)

$$(u_\zeta^{c+} - u_\zeta^{a+})_{j,k}^n = \frac{1}{6}\left[\left(u + 4u_\zeta^+ - 2u_\eta^+\right)_{(j,k;1,2)}^{n-1/2} - \left(u - 2u_\zeta^+ - 2u_\eta^+\right)_{(j,k;1,1)}^{n-1/2}\right] \tag{5.13}$$

and

$$(u_\eta^{c+} - u_\eta^{a+})_{j,k}^n = \frac{1}{6}\left[\left(u - 2u_\zeta^+ + 4u_\eta^+\right)_{(j,k;1,3)}^{n-1/2} - \left(u - 2u_\zeta^+ - 2u_\eta^+\right)_{(j,k;1,1)}^{n-1/2}\right] \tag{5.14}$$

if $(j,k,n) \in \Omega_1$; and (ii)

$$(u_\zeta^{c+} - u_\zeta^{a+})_{j,k}^n = \frac{1}{6}\left[\left(u + 2u_\zeta^+ + 2u_\eta^+\right)_{(j,k;2,1)}^{n-1/2} - \left(u - 4u_\zeta^+ + 2u_\eta^+\right)_{(j,k;2,2)}^{n-1/2}\right] \tag{5.15}$$

and

$$(u_\eta^{c+} - u_\eta^{a+})_{j,k}^n = \frac{1}{6}\left[\left(u + 2u_\zeta^+ + 2u_\eta^+\right)_{(j,k;2,1)}^{n-1/2} - \left(u + 2u_\zeta^+ - 4u_\eta^+\right)_{(j,k;2,3)}^{n-1/2}\right] \tag{5.16}$$

if $(j,k,n) \in \Omega_2$. Note that $(u_\zeta^{c+})_{j,k}^n$, $(u_\zeta^{a+})_{j,k}^n$, $(u_\eta^{c+})_{j,k}^n$ and $(u_\eta^{a+})_{j,k}^n$ are explicitly dependent on ν_ζ and ν_η (and therefore explicitly dependent on Δt). However, according to Eqs. (5.13)–(5.16), $(u_\zeta^{c+} - u_\zeta^{a+})_{j,k}^n$ and $(u_\eta^{c+} - u_\eta^{a+})_{j,k}^n$ are free from this depenency. Note that a similar occurrence was encountered in the construction of the 1D a-ϵ scheme (see the comment given following Eq. (2.14)).

At this juncture, note that:

(a) The 2D a-ϵ scheme becomes the 2D a scheme when $\epsilon = 0$.

(b) For the special case with $\epsilon = 1/2$, Eqs. (5.11) and (5.12) reduce to $(u_\zeta^+)_{j,k}^n = (u_\zeta^{c+})_{j,k}^n$ and $(u_\eta^+)_{j,k}^n = (u_\eta^{c+})_{j,k}^n$, respectively.

(c) Using the same reason given in the paragraph preceding Eq. (2.14), one may conclude that numerical dissipation in the 2D a-ϵ scheme may be controlled by varying the value of ϵ. In fact, it will be shown in Sec. 7 that (i) the 2D a-ϵ scheme is unstable if $\epsilon < 0$ or $\epsilon > 1$, and (ii) numerical diffusion indeed increases as ϵ increases, at least in the range of $0 \le \epsilon \le 0.7$.

(d) Consider the case $(j, k, n) \in \Omega_1$. Then, with the aid of Eqs. (4.28) and (5.13), Eq. (5.11) can be rewritten as:

$$(u_\zeta)^n_{j,k} = \frac{6}{\Delta\zeta}(u^{a+}_\zeta)^n_{j,k}$$
$$+\frac{\epsilon}{3}\left[\left(\frac{6u}{\Delta\zeta} + 4u_\zeta - \frac{2\Delta\eta}{\Delta\zeta}u_\eta\right)^{n-1/2}_{(j,k;1,2)} - \left(\frac{6u}{\Delta\zeta} - 2u_\zeta - \frac{2\Delta\eta}{\Delta\zeta}u_\eta\right)^{n-1/2}_{(j,k;1,1)}\right] \tag{5.17}$$

Let (i) $u^{n-1/2}_{(j,k;1,2)}$, $(u_\zeta)^{n-1/2}_{(j,k;1,2)}$ and $(u_\eta)^{n-1/2}_{(j,k;1,2)}$ be identified with the values of u, $\partial u/\partial\zeta$ and $\partial u/\partial\eta$ at the mesh point $((j, k; 1, 2), n - 1/2)$, respectively; and (ii) $u^{n-1/2}_{(j,k;1,1)}$, $(u_\zeta)^{n-1/2}_{(j,k;1,1)}$ and $(u_\eta)^{n-1/2}_{(j,k;1,1)}$ be identified with the values of u, $\partial u/\partial\zeta$ and $\partial u/\partial\eta$ at the mesh point $((j, k; 1, 1), n - 1/2)$, respectively. Then it can be shown that the expression within the brackets on the right side of Eq. (5.17) is $O(\Delta\zeta, \Delta\eta)$. Furthermore, because Eq. (4.26) is applicable only for those points $(\zeta, \eta, t) \in \mathrm{SE}(j, k, n)$ only (see Figs. 10(b) and 11(b)), the expression enclosed within the first bracket on the right side of Eq. (4.26) is $O(\Delta\zeta, \Delta t)$. From the above considerations, one concludes that the error of $u^\star(\zeta, \eta, t; j, k, n)$ introduced by adding the extra term involving ϵ on the right side of Eq. (5.17) is second order in $\Delta\zeta$, $\Delta\eta$, and Δt. In other words, addition of the term involving ϵ results in lowering the order of accuracy of $(u_\zeta)^n_{j,k}$ but not that of $u^n_{j,k}$. A similar conclusion is also applicable to Eq. (5.11) for $(j, k, n) \in \Omega_2$ and to Eq. (5.12) for either $(j, k, n) \in \Omega_1$ or $(j, k, n) \in \Omega_2$.

The 2D a-ϵ scheme can also be expressed in the form of Eq. (4.72) if

$$Q_1^{(1)} \stackrel{def}{=} \frac{1}{3}\begin{pmatrix} 1-\nu_\zeta-\nu_\eta & -(1-\nu_\zeta-\nu_\eta)(1+\nu_\zeta) & -(1-\nu_\zeta-\nu_\eta)(1+\nu_\eta) \\ 1-\epsilon & -(1+\nu_\zeta-2\epsilon) & -(1+\nu_\eta-2\epsilon) \\ 1-\epsilon & -(1+\nu_\zeta-2\epsilon) & -(1+\nu_\eta-2\epsilon) \end{pmatrix} \tag{5.18}$$

$$Q_2^{(1)} \stackrel{def}{=} \frac{1}{3}\begin{pmatrix} 1+\nu_\zeta & (1+\nu_\zeta)(2-\nu_\zeta) & -(1+\nu_\zeta)(1+\nu_\eta) \\ -(1-\epsilon) & -(2-\nu_\zeta-4\epsilon) & 1+\nu_\eta-2\epsilon \\ 0 & 0 & 0 \end{pmatrix} \tag{5.19}$$

$$Q_3^{(1)} \stackrel{def}{=} \frac{1}{3}\begin{pmatrix} 1+\nu_\eta & -(1+\nu_\eta)(1+\nu_\zeta) & (1+\nu_\eta)(2-\nu_\eta) \\ 0 & 0 & 0 \\ -(1-\epsilon) & 1+\nu_\zeta-2\epsilon & -(2-\nu_\eta-4\epsilon) \end{pmatrix} \tag{5.20}$$

$$Q_1^{(2)} \stackrel{def}{=} \frac{1}{3}\begin{pmatrix} 1+\nu_\zeta+\nu_\eta & (1+\nu_\zeta+\nu_\eta)(1-\nu_\zeta) & (1+\nu_\zeta+\nu_\eta)(1-\nu_\eta) \\ -(1-\epsilon) & -(1-\nu_\zeta-2\epsilon) & -(1-\nu_\eta-2\epsilon) \\ -(1-\epsilon) & -(1-\nu_\zeta-2\epsilon) & -(1-\nu_\eta-2\epsilon) \end{pmatrix} \tag{5.21}$$

$$Q_2^{(2)} \stackrel{def}{=} \frac{1}{3}\begin{pmatrix} 1-\nu_\zeta & -(1-\nu_\zeta)(2+\nu_\zeta) & (1-\nu_\zeta)(1-\nu_\eta) \\ 1-\epsilon & -(2+\nu_\zeta-4\epsilon) & 1-\nu_\eta-2\epsilon \\ 0 & 0 & 0 \end{pmatrix} \tag{5.22}$$

and

$$Q_3^{(2)} \stackrel{def}{=} \frac{1}{3}\begin{pmatrix} 1-\nu_\eta & (1-\nu_\eta)(1-\nu_\zeta) & -(1-\nu_\eta)(2+\nu_\eta) \\ 0 & 0 & 0 \\ 1-\epsilon & 1-\nu_\zeta-2\epsilon & -(2+\nu_\eta-4\epsilon) \end{pmatrix} \tag{5.23}$$

Note that, with the above definitions, Eqs. (4.73) and (4.74) are also valid for the 2D a-ϵ scheme.

5.2. The 2D a-ϵ-α-β Scheme

For the same reason that motivates the extension of the 1D a-ϵ scheme to become the 1D a-ϵ-α-β scheme, the 2D a-ϵ scheme will be extended to become the 2D a-ϵ-α-β scheme. As a preliminary for these extensions, first we shall provide a geometric interpretation of the procedure by which the 1D a-ϵ scheme was extended to become the 1D a-ϵ-α-β scheme.

The key step in extending the 1D a-ϵ scheme to 1D a-ϵ-α-β scheme is the construction of a nonlinear weighted average of $(u_{x+}^{c+})_j^n$ and $(u_{x-}^{c+})_j^n$ (see Eqs. (2.56)–(2.61)). Let $P_{j-} = (x_{j-1/2}, u'^n_{j-1/2})$, $P_j = (x_j, u_j^n)$ and $P_{j+} = (x_{j+1/2}, u'^n_{j+1/2})$ be three points in the x-u space. Then according to Eqs. (2.12) and (2.56), $(u_{x-}^{c+})_j^n$, $(u_{x+}^{c+})_j^n$ and $(u_x^{c+})_j^n$, respectively, are equal to the values of the slope du/dx of the three line segments $\overline{P_{j-}P_j}$, $\overline{P_jP_{j+}}$ and $\overline{P_{j-}P_{j+}}$, multiplied by the normalization factor $\Delta x/4$. Equation (2.57) states that $(u_x^{c+})_j^n$ is the simple average of $(u_{x+}^{c+})_j^n$ and $(u_{x-}^{c+})_j^n$. *Thus one can say that the key step in extending the 1D a-ϵ scheme to become the 1D a-ϵ-α-β scheme is the construction of the weighted average of the normalized slopes of $\overline{P_{j-}P_j}$ and $\overline{P_jP_{j+}}$ using the function W_o.* In the construction of the 2D a-ϵ-α-β scheme, paralleling the evaluation of the values of du/dx along the three edges of the triangle $\triangle P_{j-}P_jP_{j+}$ in the x-u space, we shall study the gradient vectors ∇u associated with the four faces of a tetrahedron in the ζ-η-u space. The vertices of the tetrahedron are the points O, P, Q and R depicted in either Fig. 15(a) or Fig. 15(b). The nonlinear weighted average used in the 2D a-ϵ-α-β will be constructed using three of the four gradient vectors referred to above.

To proceed, consider $(j,k,n) \in \Omega_q$. Also let planes #1, #2, and #3, respectively, be the planes containing the following trios of points: (i) points O, Q, and R; (ii) points O, R, and P; and (iii) points O, P, and Q. Then; in general, these planes differ from one another and from the plane that contains points P, Q and R. In the following derivations, first we shall derive the equations representing the former three planes.

As a preliminary for the developments in this and the following sections, for any real numbers s_1, s_2 and s_3, let

$$f_\zeta^{(1)}(s_1,s_2,s_3) \stackrel{def}{=} -(2s_2+s_3)/\Delta\zeta, \qquad f_\eta^{(1)}(s_1,s_2,s_3) \stackrel{def}{=} -(s_2+2s_3)/\Delta\eta \tag{5.24}$$

$$f_\zeta^{(2)}(s_1,s_2,s_3) \stackrel{def}{=} (2s_1+s_3)/\Delta\zeta, \qquad f_\eta^{(2)}(s_1,s_2,s_3) \stackrel{def}{=} (s_1-s_3)/\Delta\eta \tag{5.25}$$

$$f_\zeta^{(3)}(s_1,s_2,s_3) \stackrel{def}{=} (s_1-s_2)/\Delta\zeta, \qquad f_\eta^{(3)}(s_1,s_2,s_3) \stackrel{def}{=} (2s_1+s_2)/\Delta\eta \tag{5.26}$$

$$f_x^{(1)}(s_1,s_2,s_3) \stackrel{def}{=} -\frac{3}{2w}(s_2+s_3), \qquad f_y^{(1)}(s_1,s_2,s_3) \stackrel{def}{=} \frac{(3b+w)s_2+(3b-w)s_3}{2wh} \tag{5.27}$$

$$f_x^{(2)}(s_1,s_2,s_3) \stackrel{def}{=} \frac{3s_1}{2w}, \qquad f_y^{(2)}(s_1,s_2,s_3) \stackrel{def}{=} -\frac{(3b+w)s_1+2ws_3}{2wh} \tag{5.28}$$

$$f_x^{(3)}(s_1,s_2,s_3) \stackrel{def}{=} \frac{3s_1}{2w}, \qquad f_y^{(3)}(s_1,s_2,s_3) \stackrel{def}{=} \frac{(w-3b)s_1+2ws_2}{2wh} \tag{5.29}$$

In the following, consider a mesh point $(j,k,n) \in \Omega_q$ $(q=1,2)$. For any $r=1,2,3$, let

$$x_r \stackrel{def}{=} (-1)^q(u_{j,k}^n - u'^n_{(j,k;q,r)}) \tag{5.30}$$

$$(u_\zeta^{(r)})_{j,k}^n \stackrel{def}{=} f_\zeta^{(r)}(x_1,x_2,x_3), \qquad (u_\eta^{(r)})_{j,k}^n \stackrel{def}{=} f_\eta^{(r)}(x_1,x_2,x_3) \tag{5.31}$$

$$(u_x^{(r)})_{j,k}^n \stackrel{def}{=} f_x^{(r)}(x_1,x_2,x_3), \qquad (u_y^{(r)})_{j,k}^n \stackrel{def}{=} f_y^{(r)}(x_1,x_2,x_3) \tag{5.32}$$

Then it can be shown that, for each $r=1,2,3$, plane $\#r$ is represented by

$$\begin{aligned} u \;=\; & (u_\zeta^{(r)})_{j,k}^n\,(\zeta - j\Delta\zeta) + (u_\eta^{(r)})_{j,k}^n\,(\eta - k\Delta\eta) \\ & + u_{j,k}^n \end{aligned} \tag{5.33}$$

if the coordinates (ζ,η) are used; or by

$$\begin{aligned} u \;=\; & (u_x^{(r)})_{j,k}^n\,(x - x_{j,k}) + (u_y^{(r)})_{j,k}^n\,(y - y_{j,k}) \\ & + u_{j,k}^n \end{aligned} \tag{5.34}$$

if the coordinates (x,y) are used.

Using Eqs. (5.33) and (5.34), one concludes that, at any point on plane $\#r$, $r=1,2,3$, we have

$$\left(\frac{\partial u}{\partial \zeta}\right)_\eta = (u_\zeta^{(r)})_{j,k}^n \quad \text{and} \quad \left(\frac{\partial u}{\partial \eta}\right)_\zeta = (u_\eta^{(r)})_{j,k}^n \tag{5.35}$$

and

$$\left(\frac{\partial u}{\partial x}\right)_y = (u_x^{(r)})_{j,k}^n \quad \text{and} \quad \left(\frac{\partial u}{\partial y}\right)_x = (u_y^{(r)})_{j,k}^n \tag{5.36}$$

As a result of Eqs. (5.35) and (5.36), at any point on plane $\# r$, $r = 1, 2, 3$, $(u_x^{(r)})_{j,k}^n$ and $(u_y^{(r)})_{j,k}^n$ can be considered as the *covariant* components of the vector ∇u with respect to the *Cartesian* coordinates (x, y), while $(u_\zeta^{(r)})_{j,k}^n$ and $(u_\eta^{(r)})_{j,k}^n$ are the *covariant* components of ∇u with respect to the *non-Cartesian* coordinates (ζ, η) [55]. Furthermore, according to Eq. (5.36), at any point on plane $\# r$, $r = 1, 2, 3$, we have

$$|\nabla u| = (\theta_r)_{j,k}^n \stackrel{def}{=} \left[\sqrt{(u_x^{(r)})^2 + (u_y^{(r)})^2}\right]_{j,k}^n \tag{5.37}$$

Note that, by definition, $(\theta_r)_{j,k}^n$, $r = 1, 2, 3$, are scalars. For readers who are not familiar with tensor analysis, it is emphasized that generally $(\theta_r)_{j,k}^n$ would not be a scalar and therefore the first equality sign in Eq. (5.37) would not be valid if $u_x^{(r)}$ and $u_y^{(r)}$ in the same equation, respectively, are replaced by $u_\zeta^{(r)}$ and $u_\eta^{(r)}$.

To proceed further, let

$$(u_\zeta^{(r)+})_{j,k}^n \stackrel{def}{=} \frac{\Delta\zeta}{6}(u_\zeta^{(r)})_{j,k}^n, \qquad (u_\eta^{(r)+})_{j,k}^n \stackrel{def}{=} \frac{\Delta\eta}{6}(u_\eta^{(r)})_{j,k}^n \tag{5.38}$$

Then Eqs. (5.7), (5.8), (5.10), (5.24)–(5.26), (5.30) and (5.31) imply that

$$(u_\zeta^{c+})_{j,k}^n = \frac{1}{3}\left[u_\zeta^{(1)+} + u_\zeta^{(2)+} + u_\zeta^{(3)+}\right]_{j,k}^n \tag{5.39}$$

and

$$(u_\eta^{c+})_{j,k}^n = \frac{1}{3}\left[u_\eta^{(1)+} + u_\eta^{(2)+} + u_\eta^{(3)+}\right]_{j,k}^n \tag{5.40}$$

i.e., (i) u_ζ^{c+} is the simple average of $u_\zeta^{(r)+}$, $r = 1, 2, 3$. and (ii) u_η^{c+} is the simple average of $u_\eta^{(r)+}$, $r = 1, 2, 3$. Equations (5.39) and (5.40) can be considered as the natural extension of Eq. (2.57). *Note that, for simplicity, in the above and hereafter we may suppress the space-time mesh indices if no confusion could occur.*

Note that, as a result of Eq. (5.38), at any point on plane $\# r$, $r = 1, 2, 3$, $(u_\zeta^{(r)+})_{j,k}^n$ and $(u_\eta^{(r)+})_{j,k}^n$ are the *normalized* covariant components of ∇u with respect to the coordinates (ζ, η). On the other hand, as a result of Eqs. (5.9) and (5.10), at any point on the plane that contains the triangle $\triangle PQR$, $(u_\zeta^{c+})_{j,k}^n$ and $(u_\eta^{c+})_{j,k}^n$ are the *normalized* covariant components of ∇u with respect to the same coordinates (ζ, η). Recall that planes #1, #2, and #3, respectively, are the planes that contain the triangles $\triangle OQR$, $\triangle ORP$ and $\triangle OPQ$. The last three triangles and $\triangle PQR$ are the four faces of the tetrahedron $OPQR$. Thus Eqs. (5.39) and (5.40) state that ∇u associated with one face of this tetrahedron is one third of the sum of ∇u associated with the other three faces. This conclusion is true only because the spatial projection of point O on the plane that contains $\triangle PQR$ is the geometric center of $\triangle PQR$.

To proceed further, given any $\alpha \geq 0$, the nonlinear weighted averages $(u_\zeta^{w+})_{j,k}^n$ and $(u_\eta^{w+})_{j,k}^n$ are defined by

$$u_\zeta^{w+} \stackrel{def}{=} \begin{cases} 0, & \text{if } \theta_1 = \theta_2 = \theta_3 = 0 \\ \dfrac{(\theta_2\theta_3)^\alpha u_\zeta^{(1)+} + (\theta_3\theta_1)^\alpha u_\zeta^{(2)+} + (\theta_1\theta_2)^\alpha u_\zeta^{(3)+}}{(\theta_1\theta_2)^\alpha + (\theta_2\theta_3)^\alpha + (\theta_3\theta_1)^\alpha}, & \text{otherwise} \end{cases} \tag{5.41}$$

and

$$u_\eta^{w+} \stackrel{def}{=} \begin{cases} 0, & \text{if } \theta_1 = \theta_2 = \theta_3 = 0 \\ \dfrac{(\theta_2\theta_3)^\alpha u_\eta^{(1)+} + (\theta_3\theta_1)^\alpha u_\eta^{(2)+} + (\theta_1\theta_2)^\alpha u_\eta^{(3)+}}{(\theta_1\theta_2)^\alpha + (\theta_2\theta_3)^\alpha + (\theta_3\theta_1)^\alpha}, & \text{otherwise} \end{cases} \tag{5.42}$$

respectively. To avoid dividing by zero, in practice a small positive number such as 10^{-60} is added to the denominators of the fractions on the right sides of Eqs. (5.41) and (5.42). Note that, in the above weighted averages, the weight assigned to a quantity associated with plane $\# r$ is greater if θ_r is smaller.

Also note that the above denominators vanish if $\alpha > 0$, and any two of θ_1, θ_2, and θ_3 vanish. Thus, consistency of the above definitions requires proof of the proposition: $\theta_1 = \theta_2 = \theta_3 = 0$, *if any two of* θ_1, θ_2, *and* θ_3 *vanish.*

Proof: As an example, let $\theta_1 = \theta_2 = 0$. Then Eq. (5.37) implies that $u_x^{(r)} = u_y^{(r)} = 0$, $r = 1, 2$. In turn, Eqs. (5.27), (5.28) and (5.32) imply that $x_1 = x_2 = x_3 = 0$. $\theta_3 = 0$ now follows from Eqs. (5.29), (5.32) and (5.37). QED.

As a result of Eq. (5.41), we have

$$u_\zeta^{w+} = \begin{cases} u_\zeta^{(1)+}, & \text{if } \theta_1 = 0, \quad \theta_2 > 0, \quad \text{and } \theta_3 > 0 \\ u_\zeta^{(2)+}, & \text{if } \theta_2 = 0, \quad \theta_1 > 0, \quad \text{and } \theta_3 > 0 \\ u_\zeta^{(3)+}, & \text{if } \theta_3 = 0, \quad \theta_1 > 0, \quad \text{and } \theta_2 > 0 \end{cases} \tag{5.43}$$

Assuming $\theta_r > 0$, $r = 1, 2, 3$, we have

$$u_\zeta^{w+} = \frac{(1/\theta_1)^\alpha u_\zeta^{(1)+} + (1/\theta_2)^\alpha u_\zeta^{(2)+} + (1/\theta_3)^\alpha u_\zeta^{(3)+}}{(1/\theta_1)^\alpha + (1/\theta_2)^\alpha + (1/\theta_3)^\alpha} \tag{5.44}$$

Thus the weight assigned to $u_\zeta^{(r)+}$ is proportional to $(1/\theta_r)^\alpha$. By using (i) Eqs. (5.39), (5.41) and (5.44), and (ii) the fact that $u_\zeta^{(r)+} = 0$, $r = 1, 2, 3$, if $\theta_r = 0$, $r = 1, 2, 3$, one arrives at the conclusion that

$$u_\zeta^{w+} = u_\zeta^{c+}, \qquad \text{if} \qquad \theta_1 = \theta_2 = \theta_3 \tag{5.45}$$

Obviously Eqs. (5.43)–(5.45) are still valid if each symbol ζ is replaced by the symbol η.

With the above preliminaries, the 2D a-ϵ-α-β scheme can be defined as follows: For any $(j, k, n) \in \Omega_1$, we assume Eq. (4.65) and

$$(u_\zeta^+)_{j,k}^n = (u_\zeta^{a+})_{j,k}^n + 2\epsilon \left(u_\zeta^{c+} - u_\zeta^{a+}\right)_{j,k}^n + \beta \left(u_\zeta^{w+} - u_\zeta^{c+}\right)_{j,k}^n \tag{5.46}$$

and

$$(u_\eta^+)_{j,k}^n = (u_\eta^{a+})_{j,k}^n + 2\epsilon \left(u_\eta^{c+} - u_\eta^{a+}\right)_{j,k}^n + \beta \left(u_\eta^{w+} - u_\eta^{c+}\right)_{j,k}^n \tag{5.47}$$

with the understanding that $(u_\zeta^{a+})_{j,k}^n$ *and* $(u_\eta^{a+})_{j,k}^n$ *are those defined in Eqs. (4.66) and (4.67).* On the other hand, for any $(j,k,n) \in \Omega_2$, we assume Eqs. (4.68), (5.46) and (5.47) with the understanding that $(u_\zeta^{a+})_{j,k}^n$ *and* $(u_\eta^{a+})_{j,k}^n$ *are those defined in Eqs. (4.69) and (4.70).*

At this juncture, note that, on the smooth part of a solution, θ_1, θ_2, and θ_3 are nearly equal. Thus the weighted averages u_ζ^{w+} and u_η^{w+} are nearly equal to the simple averages u_ζ^{c+}, and u_η^{c+}, respectively (see Eq. (5.45)). As a result, *the effect of weighted-averaging generally is not discernible on the smooth part of a solution.*

Finally note that, according to Eq. (5.37), evaluation of $(\theta_r)^\alpha$ does not involve a fractional power if α is an even integer. Because a fractional power is costly to evaluate, use of the a-ϵ-α-β scheme is less costly when α is an even integer.

6. The Euler Solvers

We consider a dimensionless form of the 2-D unsteady Euler equations of a perfect gas. Let ρ, u, v, p, and γ be the mass density, x-velocity component, y-velocity component, static pressure, and constant specific heat ratio, respectively. Let

$$u_1 = \rho, \qquad u_2 = \rho u, \qquad u_3 = \rho v, \qquad u_4 = p/(\gamma - 1) + \rho(u^2 + v^2)/2 \tag{6.1}$$

$$f_1^x = u_2 \tag{6.2}$$

$$f_2^x = (\gamma - 1)u_4 + (3 - \gamma)(u_2)^2/(2u_1) - (\gamma - 1)(u_3)^2/(2u_1) \tag{6.3}$$

$$f_3^x = u_2\, u_3/u_1 \tag{6.4}$$

$$f_4^x = \gamma u_2 u_4/u_1 - (1/2)(\gamma - 1)u_2 \left[(u_2)^2 + (u_3)^2\right]/(u_1)^2 \tag{6.5}$$

$$f_1^y = u_3 \tag{6.6}$$

$$f_2^y = u_2\, u_3/u_1 \tag{6.7}$$

$$f_3^y = (\gamma - 1)u_4 + (3 - \gamma)(u_3)^2/(2u_1) - (\gamma - 1)(u_2)^2/(2u_1) \tag{6.8}$$

and

$$f_4^y = \gamma u_3 u_4/u_1 - (1/2)(\gamma - 1)u_3 \left[(u_2)^2 + (u_3)^2\right]/(u_1)^2 \tag{6.9}$$

Then the Euler equations can be expressed as

$$\frac{\partial u_m}{\partial t} + \frac{\partial f_m^x}{\partial x} + \frac{\partial f_m^y}{\partial y} = 0, \qquad m = 1, 2, 3, 4 \tag{6.10}$$

Assuming smoothness of the physical solution, Eq. (6.10) is a result of the more fundamental conservation laws

$$\oint_{S(V)} \vec{h}_m \cdot d\vec{s} = 0, \qquad m = 1, 2, 3, 4 \tag{6.11}$$

where

$$\vec{h}_m = \left(f_m^x, f_m^y, u_m\right), \qquad m = 1, 2, 3, 4 \tag{6.12}$$

are the space-time mass, x-momentum component, y-momentum component, and energy current density vectors, respectively.

As a preliminary, let

$$f_{m,\ell}^x \stackrel{def}{=} \partial f_m^x/\partial u_\ell, \qquad \text{and} \qquad f_{m,\ell}^y \stackrel{def}{=} \partial f_m^y/\partial u_\ell, \qquad m, \ell = 1, 2, 3, 4 \tag{6.13}$$

The Jacobian matrices, which are formed by $f_{m,\ell}^x$ and $f_{m,\ell}^y$, $m, \ell = 1, 2, 3, 4$, respectively, are given in [9].

Because f_m^x and f_m^y, $m = 1, 2, 3, 4$, are homogeneous functions of degree 1 [53] in u_1, u_2, u_3, and u_4, we have

$$f_m^x = \sum_{\ell=1}^{4} f_{m,\ell}^x\, u_\ell, \qquad \text{and} \qquad f_m^y = \sum_{\ell=1}^{4} f_{m,\ell}^y\, u_\ell \tag{6.14}$$

Note that Eq. (6.14) is not essential in the development of the CE/SE Euler solvers to be described in the following subsections. However, in certain instances, it will be used to recast some equations into more convenient forms.

6.1. The 2D Euler a Scheme

For any $(x,y,t) \in \mathrm{SE}(j,k,n)$, $u_m(x,y,t)$, $f_m^x(x,y,t)$, $f_m^y(x,y,t)$, and $\vec{h}_m(x,y,t)$, respectively, are approximated by $u_m^*(x,y,t\,;j,k,n)$, $f_m^{x*}(x,y,t\,;j,k,n)$, $f_m^{y*}(x,y,t\,;j,k,n)$, and $\vec{h}_m^*(x,y,t\,;j,k,n)$. They will be defined shortly. Let

$$\begin{aligned} u_m^*(x,y,t\,;j,k,n) \stackrel{def}{=} & (u_m)_{j,k}^n + (u_{mx})_{j,k}^n(x-x_{j,k}) + (u_{my})_{j,k}^n(y-y_{j,k}) \\ & +(u_{mt})_{j,k}^n(t-t^n), \qquad m=1,2,3,4 \end{aligned} \tag{6.15}$$

where $(u_m)_{j,k}^n$, $(u_{mx})_{j,k}^n$, $(u_{my})_{j,k}^n$, and $(u_{mt})_{j,k}^n$ are constants in $\mathrm{SE}(j,k,n)$. Obviously, they can be considered as the numerical analogues of the values of u_m, $\partial u_m/\partial x$, $\partial u_m/\partial y$, and $\partial u_m/\partial t$ at $(x_{j,k}, y_{j,k}, t^n)$, respectively.

Let $(f_m^x)_{j,k}^n$, $(f_m^y)_{j,k}^n$, $(f_{m,\ell}^x)_{j,k}^n$, and $(f_{m,\ell}^y)_{j,k}^n$ denote the values of f_m^x, f_m^y, $f_{m,\ell}^x$, and $f_{m,\ell}^y$, respectively, when u_m, $m=1,2,3,4$, respectively, assume the values of $(u_m)_{j,k}^n$, $m=1,2,3,4$. For any $m=1,2,3,4$, let

$$(f_{mx}^x)_{j,k}^n \stackrel{def}{=} \sum_{\ell=1}^4 (f_{m,\ell}^x)_{j,k}^n (u_{\ell x})_{j,k}^n \tag{6.16}$$

$$(f_{my}^x)_{j,k}^n \stackrel{def}{=} \sum_{\ell=1}^4 (f_{m,\ell}^x)_{j,k}^n (u_{\ell y})_{j,k}^n \tag{6.17}$$

$$(f_{mt}^x)_{j,k}^n \stackrel{def}{=} \sum_{\ell=1}^4 (f_{m,\ell}^x)_{j,k}^n (u_{\ell t})_{j,k}^n \tag{6.18}$$

$$(f_{mx}^y)_{j,k}^n \stackrel{def}{=} \sum_{\ell=1}^4 (f_{m,\ell}^y)_{j,k}^n (u_{\ell x})_{j,k}^n \tag{6.19}$$

$$(f_{my}^y)_{j,k}^n \stackrel{def}{=} \sum_{\ell=1}^4 (f_{m,\ell}^y)_{j,k}^n (u_{\ell y})_{j,k}^n \tag{6.20}$$

and

$$(f_{mt}^y)_{j,k}^n \stackrel{def}{=} \sum_{\ell=1}^4 (f_{m,\ell}^y)_{j,k}^n (u_{\ell t})_{j,k}^n \tag{6.21}$$

Because (i)

$$\frac{\partial f_m^x}{\partial x} = \sum_{\ell=1}^4 f_{m,\ell}^x \frac{\partial u_\ell}{\partial x}, \qquad m=1,2,3,4 \tag{6.22}$$

and (ii) the expression on the right side of Eq. (6.16) is the numerical analogue of that on the right side of Eq. (6.22) at $(x_{j,k}, y_{j,k}, t^n)$, $(f_{mx}^x)_{j,k}^n$ can be considered as the numerical analogue of the value of $\partial f_m^x/\partial x$ at $(x_{j,k}, y_{j,k}, t^n)$. Similarly, $(f_{my}^x)_{j,k}^n$, $(f_{mt}^x)_{j,k}^n$, $(f_{mx}^y)_{j,k}^n$, $(f_{my}^y)_{j,k}^n$, and

$(f^y_{mt})^n_{j,k}$ can be considered as the numerical analogues of the values of $\partial f^x_m/\partial y$, $\partial f^x_m/\partial t$, $\partial f^y_m/\partial x$, $\partial f^y_m/\partial y$, and $\partial f^y_m/\partial t$ at$(x_{j,k}, y_{j,k}, t^n)$, respectively. As a result, we define

$$f^{x*}_m(x,y,t\,;j,k,n) \stackrel{def}{=} (f^x_m)^n_{j,k} + (f^x_{mx})^n_{j,k}(x - x_{j,k}) + (f^x_{my})^n_{j,k}(y - y_{j,k}) + (f^x_{mt})^n_{j,k}(t - t^n), \qquad m = 1,2,3,4 \tag{6.23}$$

and

$$f^{y*}_m(x,y,t\,;j,k,n) \stackrel{def}{=} (f^y_m)^n_{j,k} + (f^y_{mx})^n_{j,k}(x - x_{j,k}) + (f^y_{my})^n_{j,k}(y - y_{j,k}) + (f^y_{mt})^n_{j,k}(t - t^n), \qquad m = 1,2,3,4 \tag{6.24}$$

Also, as an analogue to Eq. (6.12), we define

$$\vec{h}^*_m(x,y,t\,;j,k,n) \stackrel{def}{=} \Big(f^{x*}_m(x,y,t\,;j,k,n), f^{y*}_m(x,y,t\,;j,k,n), u^*_m(x,y,t\,;j,k,n)\Big), \qquad m = 1,2,3,4 \tag{6.25}$$

Note that, by their definitions: (i) $(f^x_m)^n_{j,k}$, $(f^y_m)^n_{j,k}$, $(f^x_{m,\ell})^n_{j,k}$, and $(f^y_{m,\ell})^n_{j,k}$ are functions of $(u_m)^n_{j,k}$, $m = 1,2,3,4$; (ii) $(f^x_{mx})^n_{j,k}$ and $(f^y_{mx})^n_{j,k}$ are functions of $(u_m)^n_{j,k}$ and $(u_{mx})^n_{j,k}$, $m = 1,2,3,4$; (iii) $(f^x_{my})^n_{j,k}$ and $(f^y_{my})^n_{j,k}$ are functions of $(u_m)^n_{j,k}$ and $(u_{my})^n_{j,k}$, $m = 1,2,3,4$; and (iv) $(f^x_{mt})^n_{j,k}$ and $(f^y_{mt})^n_{j,k}$ are functions of $(u_m)^n_{j,k}$ and $(u_{mt})^n_{j,k}$, $m = 1,2,3,4$.

Moreover, we assume that, for any $(x,y,t) \in \mathrm{SE}(j,k,n)$, and any $m = 1,2,3,4$,

$$\frac{\partial u^*_m(x,y,t\,;j,k,n)}{\partial t} + \frac{\partial f^{x*}_m(x,y,t\,;j,k,n)}{\partial x} + \frac{\partial f^{y*}_m(x,y,t\,;j,k,n)}{\partial y} = 0 \tag{6.26}$$

Note that Eq. (6.26) is the numerical analogue of Eq. (6.10). With the aid of Eqs. (6.15), (6.23), (6.24), (6.16), and (6.20), Eq. (6.26) implies that, for any $m = 1,2,3,4$,

$$(u_{mt})^n_{j,k} = -(f^x_{mx})^n_{j,k} - (f^y_{my})^n_{j,k} = -\sum_{\ell=1}^{4}\left[f^x_{m,\ell}\,u_{\ell x} + f^y_{m,\ell}\,u_{\ell y}\right]^n_{j,k} \tag{6.27}$$

Thus $(u_{mt})^n_{j,k}$ is a function of $(u_m)^n_{j,k}$, $(u_{mx})^n_{j,k}$, and $(u_{my})^n_{j,k}$. From this result and the facts stated following Eq. (6.25), one concludes that *the only independent discrete variables needed to be solved for in the current marching scheme are* $(u_m)^n_{j,k}$, $(u_{mx})^n_{j,k}$, *and* $(u_{my})^n_{j,k}$.

Consider the conservation elements depicted in Figs. 10(a) and 11(a). The Euler counterpart to Eq. (4.11) is

$$\oint_{S(CE_r(j,k,n))} \vec{h}^*_m \cdot d\vec{s} = 0, \qquad r = 1,2,3, \qquad m = 1,2,3,4 \tag{6.28}$$

Next we shall introduce the Euler counterparts of Eqs. (4.22), (4.23), (4.27), and (4.28). For any $(j,k,n) \in \Omega$, let

$$\begin{pmatrix} (f^\zeta_{m,\ell})^n_{j,k} \\ (f^\eta_{m,\ell})^n_{j,k} \end{pmatrix} \stackrel{def}{=} T^{-1} \begin{pmatrix} (f^x_{m,\ell})^n_{j,k} \\ (f^y_{m,\ell})^n_{j,k} \end{pmatrix}, \qquad m,\ell = 1,2,3,4 \tag{6.29}$$

and

$$\begin{pmatrix} (u_{m\zeta})^n_{j,k} \\ \\ (u_{m\eta})^n_{j,k} \end{pmatrix} \stackrel{def}{=} T^t \begin{pmatrix} (u_{mx})^n_{j,k} \\ \\ (u_{my})^n_{j,k} \end{pmatrix}, \qquad m = 1, 2, 3, 4 \tag{6.30}$$

The *normalized* counterparts of those parameters defined in Eqs. (6.29) and (6.30) are

$$(f^{\zeta+}_{m,\ell})^n_{j,k} \stackrel{def}{=} \frac{3\Delta t}{2\Delta\zeta}(f^{\zeta}_{m,\ell})^n_{j,k}, \qquad \text{and} \qquad (f^{\eta+}_{m,\ell})^n_{j,k} \stackrel{def}{=} \frac{3\Delta t}{2\Delta\eta}(f^{\eta}_{m,\ell})^n_{j,k} \tag{6.31}$$

and

$$(u^+_{m\zeta})^n_{j,k} \stackrel{def}{=} \frac{\Delta\zeta}{6}(u_{m\zeta})^n_{j,k}, \qquad \text{and} \qquad (u^+_{m\eta})^n_{j,k} \stackrel{def}{=} \frac{\Delta\eta}{6}(u_{m\eta})^n_{j,k} \tag{6.32}$$

In the following development, for simplicity, we may strip from every variable in an equation its indices j, k, and n if all variables are associated with the same mesh point $(j, k, n) \in \Omega$. Let $F^{\zeta+}$ and $F^{\eta+}$, respectively, denote the matrices formed by $f^{\zeta+}_{m,\ell}$ and $f^{\eta+}_{m,\ell}$, $m, \ell = 1, 2, 3, 4$. Let I be the 4×4 identity matrix. Then the current counterparts to Eqs. (4.29)–(4.46) are

$$\Sigma^{(1)\pm}_{11} \stackrel{def}{=} I - F^{\zeta+} - F^{\eta+} \tag{6.33}$$

$$\Sigma^{(1)\pm}_{12} \stackrel{def}{=} \pm(I - F^{\zeta+} - F^{\eta+})(I + F^{\zeta+}) \tag{6.34}$$

$$\Sigma^{(1)\pm}_{13} \stackrel{def}{=} \pm(I - F^{\zeta+} - F^{\eta+})(I + F^{\eta+}) \tag{6.35}$$

$$\Sigma^{(1)\pm}_{21} \stackrel{def}{=} I + F^{\zeta+} \tag{6.36}$$

$$\Sigma^{(1)\pm}_{22} \stackrel{def}{=} \mp(I + F^{\zeta+})(2I - F^{\zeta+}) \tag{6.37}$$

$$\Sigma^{(1)\pm}_{23} \stackrel{def}{=} \pm(I + F^{\zeta+})(I + F^{\eta+}) \tag{6.38}$$

$$\Sigma^{(1)\pm}_{31} \stackrel{def}{=} I + F^{\eta+} \tag{6.39}$$

$$\Sigma^{(1)\pm}_{32} \stackrel{def}{=} \pm(I + F^{\eta+})(I + F^{\zeta+}) \tag{6.40}$$

$$\Sigma^{(1)\pm}_{33} \stackrel{def}{=} \mp(I + F^{\eta+})(2I - F^{\eta+}) \tag{6.41}$$

$$\Sigma^{(2)\pm}_{11} \stackrel{def}{=} I + F^{\zeta+} + F^{\eta+} \tag{6.42}$$

$$\Sigma^{(2)\pm}_{12} \stackrel{def}{=} \mp(I + F^{\zeta+} + F^{\eta+})(I - F^{\zeta+}) \tag{6.43}$$

$$\Sigma^{(2)\pm}_{13} \stackrel{def}{=} \mp(I + F^{\zeta+} + F^{\eta+})(I - F^{\eta+}) \tag{6.44}$$

$$\Sigma^{(2)\pm}_{21} \stackrel{def}{=} I - F^{\zeta+} \tag{6.45}$$

$$\Sigma^{(2)\pm}_{22} \stackrel{def}{=} \pm(I - F^{\zeta+})(2I + F^{\zeta+}) \tag{6.46}$$

$$\Sigma^{(2)\pm}_{23} \stackrel{def}{=} \mp(I - F^{\zeta+})(I - F^{\eta+}) \tag{6.47}$$

$$\Sigma^{(2)\pm}_{31} \stackrel{def}{=} I - F^{\eta+} \tag{6.48}$$

$$\Sigma^{(2)\pm}_{32} \stackrel{def}{=} \mp(I - F^{\eta+})(I - F^{\zeta+}) \tag{6.49}$$

and

$$\Sigma_{33}^{(2)\pm} \stackrel{def}{=} \pm(I - F^{\eta+})(2I + F^{\eta+}) \tag{6.50}$$

Note that Eqs. (4.29)–(4.46) become Eqs. (6.33)–(6.50), respectively, under the following substitution rules:

§1: 1, ν_ζ, and ν_η, be replaced by I, $F^{\zeta+}$, and $F^{\eta+}$, respectively.

§2: $\sigma_{rs}^{(q)\pm}$ be replaced by $\Sigma_{rs}^{(q)\pm}$, $q = 1,2$ and $r, s = 1,2,3$, respectively.

As will be shown, under the above and other rules of substitution to be given later, many other equations given in Secs. 4 and 5 can be converted to their Euler counterparts given in this section. The latter will be referred to as the Euler *images* of the former.

Equation (6.28) is evaluated in Appendix C. Let $(j,k,n) \in \Omega_q$. Let $\vec{u}$, $\vec{u}_t$, $\vec{u}_\zeta^+$, and $\vec{u}_\eta^+$, respectively, be the 4×1 column matrices formed by u_m, u_{mt}, $u_{m\zeta}^+$, and $u_{m\eta}^+$, $m = 1,2,3,4$. Then, with the aid of Eq. (6.14), for any pair of q and r ($q = 1,2$ and $r = 1,2,3$), the results with $m = 1,2,3,4$ can be combined into the matrix form

$$\left[\Sigma_{r1}^{(q)+}\vec{u} + \Sigma_{r2}^{(q)+}\vec{u}_\zeta^+ + \Sigma_{r3}^{(q)+}\vec{u}_\eta^+\right]_{j,k}^n = \left[\Sigma_{r1}^{(q)-}\vec{u} + \Sigma_{r2}^{(q)-}\vec{u}_\zeta^+ + \Sigma_{r3}^{(q)-}\vec{u}_\eta^+\right]_{(j,k;q,r)}^{n-1/2} \tag{6.51}$$

Eq. (6.51) is the Euler image of Eq. (4.51) under the substitution rules §2 and

§3: u, u_t, u_ζ^+, and u_η^+ be replaced by $\vec{u}$, $\vec{u}_t$, $\vec{u}_\zeta^+$, and $\vec{u}_\eta^+$, respectively.

As a result of Eqs. (6.33)–(6.50), we have

$$\Sigma_{11}^{(q)\pm} + \Sigma_{21}^{(q)\pm} + \Sigma_{31}^{(q)\pm} = 3I, \qquad q = 1,2 \tag{6.52}$$

and

$$\Sigma_{12}^{(q)\pm} + \Sigma_{22}^{(q)\pm} + \Sigma_{32}^{(q)\pm} = \Sigma_{13}^{(q)\pm} + \Sigma_{23}^{(q)\pm} + \Sigma_{33}^{(q)\pm} = 0, \qquad q = 1,2 \tag{6.53}$$

Equations (6.52) and (6.53) are the Euler images of Eqs. (4.47) and (4.48), respectively. For either $q = 1$ or $q = 2$, by summing over the three equations $r = 1,2,3$ given in Eq. (6.51), and using Eqs. (6.52) and (6.53), one concludes that, for any $(j,k,n) \in \Omega_q$,

$$\vec{u}_{j,k}^n = \frac{1}{3}\sum_{r=1}^{3}\left[\Sigma_{r1}^{(q)-}\vec{u} + \Sigma_{r2}^{(q)-}\vec{u}_\zeta^+ + \Sigma_{r3}^{(q)-}\vec{u}_\eta^+\right]_{(j,k;q,r)}^{n-1/2}, \quad q = 1,2 \tag{6.54}$$

As a result, $\vec{u}_{j,k}^n$ can be evaluated in terms of the marching variables at the $(n-1/2)$th time level.

Note that, with the aid of Eqs. (6.33)–(6.50), Eq. (6.54) can be expressed explicitly as

$$\vec{u}_{j,k}^n = \frac{1}{3}\left[\left(I - F^{\zeta+} - F^{\eta+}\right)_{(j,k;1,1)}^{n-1/2}\vec{s}_1^{(1)} + \left(I + F^{\zeta+}\right)_{(j,k;1,2)}^{n-1/2}\vec{s}_2^{(1)} + \left(I + F^{\eta+}\right)_{(j,k;1,3)}^{n-1/2}\vec{s}_3^{(1)}\right] \tag{6.54a}$$

if $(j,k,n) \in \Omega_1$; or

$$\vec{u}_{j,k}^n = \frac{1}{3}\left[\left(I + F^{\zeta+} + F^{\eta+}\right)_{(j,k;2,1)}^{n-1/2}\vec{s}_1^{(2)} + \left(I - F^{\zeta+}\right)_{(j,k;2,2)}^{n-1/2}\vec{s}_2^{(2)} + \left(I - F^{\eta+}\right)_{(j,k;2,3)}^{n-1/2}\vec{s}_3^{(2)}\right] \tag{6.54b}$$

if $(j,k,n) \in \Omega_2$. Here (i)

$$\vec{s}_1^{(1)} \overset{def}{=} \left[\vec{u} - \left(I + F^{\zeta+}\right)\vec{u}_\zeta^+ - \left(I + F^{\eta+}\right)\vec{u}_\eta^+\right]_{(j,k;1,1)}^{n-1/2} \tag{6.55}$$

$$\vec{s}_2^{(1)} \overset{def}{=} \left[\vec{u} + \left(2I - F^{\zeta+}\right)\vec{u}_\zeta^+ - \left(I + F^{\eta+}\right)\vec{u}_\eta^+\right]_{(j,k;1,2)}^{n-1/2} \tag{6.56}$$

and

$$\vec{s}_3^{(1)} \overset{def}{=} \left[\vec{u} - \left(I + F^{\zeta+}\right)\vec{u}_\zeta^+ + \left(2I - F^{\eta+}\right)\vec{u}_\eta^+\right]_{(j,k;1,3)}^{n-1/2} \tag{6.57}$$

with $(j,k,n) \in \Omega_1$; and (ii)

$$\vec{s}_1^{(2)} \overset{def}{=} \left[\vec{u} + \left(I - F^{\zeta+}\right)\vec{u}_\zeta^+ + \left(I - F^{\eta+}\right)\vec{u}_\eta^+\right]_{(j,k;2,1)}^{n-1/2} \tag{6.58}$$

$$\vec{s}_2^{(2)} \overset{def}{=} \left[\vec{u} - \left(2I + F^{\zeta+}\right)\vec{u}_\zeta^+ + \left(I - F^{\eta+}\right)\vec{u}_\eta^+\right]_{(j,k;2,2)}^{n-1/2} \tag{6.59}$$

and

$$\vec{s}_3^{(2)} \overset{def}{=} \left[\vec{u} + \left(I - F^{\zeta+}\right)\vec{u}_\zeta^+ - \left(2I + F^{\eta+}\right)\vec{u}_\eta^+\right]_{(j,k;2,3)}^{n-1/2} \tag{6.60}$$

with $(j,k,n) \in \Omega_2$. Eqs. (6.54a)–(6.60) are the Euler images of Eqs. (4.65), (4.68) and (4.59)–(4.64), respectively, under the substitution rules §1, §3 and

§4: $s_r^{(q)}$ be replaced by $\vec{s}_r^{(q)}$, $q = 1, 2$, and $r = 1, 2, 3$, respectively.

For any $(j,k,n) \in \Omega_q$, the matrices $(\Sigma_{r1}^{(q)+})_{j,k}^n$, $r = 1, 2, 3$, are known functions of $\vec{u}_{j,k}^n$. Thus they can be evaluted after the latter is evaluated using Eq. (6.54). Assuming the existence of the inverse of each of the matrices $(\Sigma_{r1}^{(q)+})_{j,k}^n$ (see Appendix D.3 for an existence theorem), it follows that one can also evaluate $\vec{S}_r^{(q)}$ ($q = 1, 2$ and $r = 1, 2, 3$) where

$$\vec{S}_r^{(q)} \overset{def}{=} \left[\left(\Sigma_{r1}^{(q)+}\right)_{j,k}^n\right]^{-1} \times \left[\Sigma_{r1}^{(q)-}\vec{u} + \Sigma_{r2}^{(q)-}\vec{u}_\zeta^+ + \Sigma_{r3}^{(q)-}\vec{u}_\eta^+\right]_{(j,k;q,r)}^{n-1/2} \tag{6.61}$$

Note that, in this paper, the inverse of a matrix A is denoted by $[A]^{-1}$.

At this juncture, note that $\vec{S}_r^{(q)}$ can be evaluated by a direct application of Eq. (6.61), if one does not mind inverting the 4×4 matrices $\left(\Sigma_{r1}^{(q)+}\right)_{j,k}^n$. Alternatively, for each pair of q and r, one may use the method of Gaussian elimination to obtain *the* 4×1 *column matrix* $\vec{S}_r^{(q)}$ as the solution to the matrix equation

$$\left(\Sigma_{r1}^{(q)+}\right)_{j,k}^n \vec{S}_r^{(q)} = \left[\Sigma_{r1}^{(q)-}\vec{u} + \Sigma_{r2}^{(q)-}\vec{u}_\zeta^+ + \Sigma_{r3}^{(q)-}\vec{u}_\eta^+\right]_{(j,k;q,r)}^{n-1/2} \tag{6.62}$$

Furthermore, by multiplying Eq. (6.51) from the left with

$$\left[\left(\Sigma_{r1}^{(q)+}\right)_{j,k}^n\right]^{-1}$$

repeatedly with all possible pairs of q and r, and using Eqs. (6.33)–(6.50) and (6.61), one has [9] (i)

$$\left[\vec{u} + \left(I + F^{\zeta+}\right) \vec{u}_\zeta^+ + \left(I + F^{\eta+}\right) \vec{u}_\eta^+\right]_{j,k}^n = \vec{S}_1^{(1)} \tag{6.63}$$

$$\left[\vec{u} - \left(2I - F^{\zeta+}\right) \vec{u}_\zeta^+ + \left(I + F^{\eta+}\right) \vec{u}_\eta^+\right]_{j,k}^n = \vec{S}_2^{(1)} \tag{6.64}$$

and

$$\left[\vec{u} + \left(I + F^{\zeta+}\right) \vec{u}_\zeta^+ - \left(2I - F^{\eta+}\right) \vec{u}_\eta^+\right]_{j,k}^n = \vec{S}_3^{(1)} \tag{6.65}$$

where $(j,k,n) \in \Omega_1$; and (ii)

$$\left[\vec{u} - \left(I - F^{\zeta+}\right) \vec{u}_\zeta^+ - \left(I - F^{\eta+}\right) \vec{u}_\eta^+\right]_{j,k}^n = \vec{S}_1^{(2)} \tag{6.66}$$

$$\left[\vec{u} + \left(2I + F^{\zeta+}\right) \vec{u}_\zeta^+ - \left(I - F^{\eta+}\right) \vec{u}_\eta^+\right]_{j,k}^n = \vec{S}_2^{(2)} \tag{6.67}$$

and

$$\left[\vec{u} - \left(I - F^{\zeta+}\right) \vec{u}_\zeta^+ + \left(2I + F^{\eta+}\right) \vec{u}_\eta^+\right]_{j,k}^n = \vec{S}_3^{(2)} \tag{6.68}$$

where $(j,k,n) \in \Omega_2$.

Note that, with the aid of Eqs. (6.33), (6.36), (6.39), (6.42), (6.45), (6.48) and (6.61), Eq. (6.54) can also be expressed as

$$\vec{u}_{j,k}^n = \frac{1}{3}\left[\left(I - F^{\zeta+} - F^{\eta+}\right)_{j,k}^n \vec{S}_1^{(1)} + \left(I + F^{\zeta+}\right)_{j,k}^n \vec{S}_2^{(1)} + \left(I + F^{\eta+}\right)_{j,k}^n \vec{S}_3^{(1)}\right] \tag{6.69}$$

if $(j,k,n) \in \Omega_1$; or

$$\vec{u}_{j,k}^n = \frac{1}{3}\left[\left(I + F^{\zeta+} + F^{\eta+}\right)_{j,k}^n \vec{S}_1^{(2)} + \left(I - F^{\zeta+}\right)_{j,k}^n \vec{S}_2^{(2)} + \left(I - F^{\eta+}\right)_{j,k}^n \vec{S}_3^{(2)}\right] \tag{6.70}$$

if $(j,k,n) \in \Omega_2$. Furthermore, by subtracting Eqs. (6.64) and (6.65), respectively, from Eq. (6.63), one obtains

$$\left(\vec{u}_\zeta^+\right)_{j,k}^n = (\vec{u}_\zeta^{a+})_{j,k}^n \stackrel{def}{=} \frac{1}{3}\left(\vec{S}_1^{(1)} - \vec{S}_2^{(1)}\right) \tag{6.71}$$

and

$$\left(\vec{u}_\eta^+\right)_{j,k}^n = (\vec{u}_\eta^{a+})_{j,k}^n \stackrel{def}{=} \frac{1}{3}\left(\vec{S}_1^{(1)} - \vec{S}_3^{(1)}\right) \tag{6.72}$$

respectively, where $(j,k,n) \in \Omega_1$. Next, by subtracting Eq. (6.66) from Eqs. (6.67) and (6.68), respectively, one obtains

$$\left(\vec{u}_\zeta^+\right)_{j,k}^n = (\vec{u}_\zeta^{a+})_{j,k}^n \stackrel{def}{=} \frac{1}{3}\left(\vec{S}_2^{(2)} - \vec{S}_1^{(2)}\right) \tag{6.73}$$

and

$$\left(\vec{u}_\eta^+\right)_{j,k}^n = (\vec{u}_\eta^{a+})_{j,k}^n \stackrel{def}{=} \frac{1}{3}\left(\vec{S}_3^{(2)} - \vec{S}_1^{(2)}\right) \tag{6.74}$$

respectively, where $(j,k,n) \in \Omega_2$.

Note that, under the substitution rules §1, §3,

§5: u_ζ^{a+} and u_η^{a+} be replaced by $\vec{u}_\zeta^{a+}$ and $\vec{u}_\eta^{a+}$, respectively.

§6: $s_r^{(q)}$ be replaced by $\vec{S}_r^{(q)}$, $q = 1, 2$, and $r = 1, 2, 3$, respectively.

Eqs. (6.63)–(6.74) are the Euler images of Eqs. (4.53)–(4.58), (4.65), (4.68), (4.66), (4.67), (4.69) and (4.70), repectively.

The 2D Euler a scheme is formed by repeatedly applying the two marching steps defined, respectively, by (i) Eqs. (6.54a), (6.71) and (6.72); and (ii) Eqs. (6.54b), (6.73) and (6.74). Note that: (i) because $\vec{S}_r^{(q)}$ can not be evaluated without $\vec{u}_{j,k}^n$ being known first, one cannot evaluate $\vec{u}_{j,k}^n$ using Eqs. (6.69) and (6.70); and (ii) the 2D Euler a scheme is a two-way marching scheme in the sense that the conservation conditions Eq. (6.28) can also be used to construct its backward time marching version.

At this juncture, note that the 2D Euler a scheme is greatly simplified by the fact that $\vec{u}_{j,k}^n$ can be evaluated explicitly in terms of the marching variables at the $(n - 1/2)$th time levels using Eq. (6.54). As a result, the matrices $(\Sigma_{rs}^{(q)+})_{j,k}^n$, which are *nonlinear* functions of $\vec{u}_{j,k}^n$, can be evaluated easily. In other words, nonlinearity of the above matrix functions does not pose a difficult problem for the 2D Euler a scheme.

To explain how Eq. (6.54) arises, note that, because of Eq. (5.1),

$$\oint_{S(CE(j,k,n))} \vec{h}_m^* \cdot d\vec{s} = 0, \qquad (j,k,n) \in \Omega \tag{6.75}$$

is the direct result of Eq. (6.28), the basic assumptions of the 2D Euler a scheme. According to Eq. (5.1), $CE(j,k,n)$ is the hexagonal cylinder $A'B'C'D'E'F'ABCDEF$ depicted in Figs. 10(a) and 11(a). Except for the top face $A'B'C'D'E'F'$, the other boundaries of this cylinder are the subsets of three solution elements at the $(n - 1/2)$th time level. Thus, for any $m = 1, 2, 3, 4$, the flux of $\vec{h}_m^*$ leaving $CE(j,k,n)$ through all the boundaries except the top face can be evaluated in terms of the marching variables at the $(n - 1/2)$th time level. On the other hand, because the top face is a subset of SE(j,k,n), the flux leaving there is a function of the marching variables associated with the mesh point (j,k,n). Furthermore, because the outward normal to the top face has no spatial component, the total flux of $\vec{h}_m^*$ leaving $CE(j,k,n)$ through the top face is the surface integral of u_m^* over the top face. *Because the center of SE(j,k,n) coincides with the center of the top face*, it is easy to see that the first-order terms in Eqs. (6.15) do not contribute to the *total* flux leaving the top face. It follows that the total flux leaving the top face is a function of $(u_m)_{j,k}^n$ only. As a result of the above considerations, $\vec{u}_{j,k}^n$ can be determined in terms of the marching variables at the $(n - 1/2)$th time level by using Eq. (6.75) only. Equation (6.54) is the direct results of Eq. (6.75).

Because implementation of the 2D Euler a scheme requires, at each mesh point $(j,k,n) \in \Omega$, the solution of the three matrix equations (corresponding to $r = 1, 2, 3$) given in Eq. (6.62), the scheme is referred to as *locally implicit* [1, p.22]. A simplified and completely explicit version of it will be described immediately.

6.2. The Simplified 2D Euler a Scheme

Eq. (6.75) is assumed in the 2D Euler a scheme. As a result, Eq. (6.54) is also applicable to the new scheme.

To construct the rest of the simplified scheme, note that, with the aid of Eqs. (6.33)–(6.50), a substitution of the approximations

$$\left(\Sigma_{r1}^{(q)+}\right)_{j,k}^{n} \approx \left(\Sigma_{r1}^{(q)+}\right)_{(j,k;q,r)}^{n-1/2} \tag{6.76}$$

into Eq. (6.61) reveals that

$$\vec{S}_r^{(q)} \approx \vec{s}_r^{(q)}, \qquad q = 1,2; \quad r = 1,2,3 \tag{6.77}$$

where $\vec{s}_r^{(q)}$ are defined in Eqs. (6.55)–(6.60).

As a result of Eq. (6.77), Eqs. (6.71) and (6.72) can be approximated by

$$\left(\vec{u}_\zeta^{+}\right)_{j,k}^{n} = (\vec{u}_\zeta^{a'+})_{j,k}^{n} \stackrel{def}{=} \frac{1}{3}\left(\vec{s}_1^{(1)} - \vec{s}_2^{(1)}\right) \tag{6.78}$$

and

$$\left(\vec{u}_\eta^{+}\right)_{j,k}^{n} = (\vec{u}_\eta^{a'+})_{j,k}^{n} \stackrel{def}{=} \frac{1}{3}\left(\vec{s}_1^{(1)} - \vec{s}_3^{(1)}\right) \tag{6.79}$$

respectively, where $(j,k,n) \in \Omega_1$. Similarly, Eqs. (6.73) and (6.74) can be approximated by

$$\left(\vec{u}_\zeta^{+}\right)_{j,k}^{n} = (\vec{u}_\zeta^{a'+})_{j,k}^{n} \stackrel{def}{=} \frac{1}{3}\left(\vec{s}_2^{(2)} - \vec{s}_1^{(2)}\right) \tag{6.80}$$

and

$$\left(\vec{u}_\eta^{+}\right)_{j,k}^{n} = (\vec{u}_\eta^{a'+})_{j,k}^{n} \stackrel{def}{=} \frac{1}{3}\left(\vec{s}_3^{(2)} - \vec{s}_1^{(2)}\right) \tag{6.81}$$

respectively, where $(j,k,n) \in \Omega_2$.

Note that Eqs. (6.78)–(6.81) are the Euler images of Eqs. (4.66), (4.67), (4.69) and (4.70) under the substitution rules §3, §4 and

§7: u_ζ^{a+} and u_η^{a+} be replaced by $\vec{u}_\zeta^{a'+}$ and $\vec{u}_\eta^{a'+}$, respectively.

The first marching step of the simplified 2D Euler a scheme is formed by Eqs. (6.54a), (6.78) and (6.79). The second marching step is formed by Eqs. (6.54b), (6.80) and (6.81). Moreover, because every $\vec{s}_r^{(q)}$ (and thus every $(\vec{u}_\zeta^{a'+})_{j,k}^{n}$ and $(\vec{u}_\eta^{a'+})_{j,k}^{n}$ with $(j,k,n) \in \Omega$) can be evaluated without solving a system of equations, the simplified version is computationally more efficient than the original scheme.

6.3. The 2D Euler a-ϵ Scheme

Eq. (6.75) is assumed in the 2D Euler a-ϵ scheme. As a result, Eq. (6.54) is also applicable to the new scheme. As will be shown shortly, by considering their component equations separately, the vector equations that form the rest of the 2D Euler a-ϵ can be developed in a fashion similar to that which was used to develop the 2D a-ϵ scheme.

Let $(j,k,n) \in \Omega_q$ and consider any $m = 1,2,3,4$. Let $(u'_m)^n_{(j,k;q,r)}$, $(u^c_m)^n_{j,k}$, $(u^c_{m\zeta})^n_{j,k}$, and $(u^c_{m\eta})^n_{j,k}$ be defined by a set of equations identical to Eqs. (5.3) and (5.6)–(5.8) except that the symbols u', u, u_t, u^c, u^c_ζ and u^c_η in the latter equations are replaced, respectively, by the symbols (u'_m), u_m, u_{mt}, u^c_m, $u^c_{m\zeta}$ and $u^c_{m\eta}$ in the former equations. Let P_m, Q_m and R_m (see Figs. 16(a) and 16(b)) be the three points in the ζ-η-u space with (i) their ζ- and η-coordinates being those of the mesh points $((j,k;q,r), n-1/2)$, $r = 1,2,3$, respectively, and (ii) their u-coordinates being $(u'_m)^n_{(j,k;q,r)}$, $r = 1,2,3$, respectively. It can be shown that the plane in the ζ-η-u space that intersects the above three points is represented by an equation that is identical to Eq. (5.5) except that the symbols u^c, u^c_ζ and u^c_η in Eq. (5.5) are now replaced by u^c_m, $u^c_{m\zeta}$ and $u^c_{m\eta}$, respectively. As a result, for every point on the plane referred to above, we have two relations that are identical to those given in Eq. (5.9) except that the symbols u^c_ζ and u^c_η in Eq. (5.9) are now replaced by $u^c_{m\zeta}$ and $u^c_{m\eta}$, respectively. Furthermore, let $(u^{c+}_{m\zeta})^n_{j,k}$ and $(u^{c+}_{m\eta})^n_{j,k}$ be defined using an equation that is identical to Eq. (5.10) except that the symbols u^{c+}_ζ, u^c_ζ, u^{c+}_η and u^c_η in the latter equation are replaced, respectively, by the symbols $u^{c+}_{m\zeta}$, $u^c_{m\zeta}$, $u^{c+}_{m\eta}$ and $u^c_{m\eta}$ in the former equation.

Moreover, let $\vec{u}'$, $\vec{u}^c$ $\vec{u}^c_\zeta$, $\vec{u}^c_\eta$, $\vec{u}^{c+}_\zeta$ and $\vec{u}^{c+}_\eta$, respectively, denote the 4×1 column matrices formed by u'_m, u^c_m, $u^c_{m\zeta}$, $u^c_{m\eta}$, $u^{c+}_{m\zeta}$ and $u^{c+}_{m\eta}$, $m = 1,2,3,4$. Then, with the aid of the relation

$$\vec{u}_t = -\frac{4}{\Delta t}\left(F^{\zeta+}\vec{u}^+_\zeta + F^{\eta+}\vec{u}^+_\eta\right) \tag{6.82}$$

which follows from Eqs. (6.27), (6.29), and (6.30), it becomes evident that we can obtain a set of equations that are the Euler images of Eqs. (5.3), (5.4), (5.6)–(5.8), and (5.10)–(5.12) under the substitution rules §1, §3, §5 and

§8: u', u^c, u^c_ζ, u^c_η, u^{c+}_ζ and u^{c+}_η be replaced by $\vec{u}'$, $\vec{u}^c$, $\vec{u}^c_\zeta$, $\vec{u}^c_\eta$, $\vec{u}^{c+}_\zeta$ and $\vec{u}^{c+}_\eta$, respectively.

Note that the Euler images of Eqs. (5.13)–(5.16) under the substitution rules §3, §5 and §8 are not valid for the current scheme because (i) $(\vec{u}^{a+}_\zeta)^n_{j,k}$ and $(\vec{u}^{a+}_\eta)^n_{j,k}$ are defined in terms of $\vec{S}^{(q)}_r$, $q = 1,2$, $r = 1,2,3$ (see Eqs. (6.71)–(6.74)), while $(u^{a+}_\zeta)^n_{j,k}$ and $(u^{a+}_\eta)^n_{j,k}$ are defined in terms of $s^{(q)}_r$, $q = 1,2$, $r = 1,2,3$ (see Eqs. (4.66), (4.67), (4.69), and (4.70)); and (ii) $\vec{S}^{(q)}_r$, which were defined by Eq. (6.61), are structually different from $s^{(q)}_r$, which were defined by Eqs. (4.59)–(4.64). However, as will be shown shortly, the Euler images of Eqs. (5.13)–(5.16) under the substitution rules §3, §7 and §8 do exist.

For future reference, several key equations associated with the 2D Euler a-ϵ scheme will be given explicitly. They are:

$$\vec{u}'^n_{(j,k;q,r)} \overset{def}{=} \left(\vec{u} + \frac{\Delta t}{2}\vec{u}_t\right)^{n-1/2}_{(j,k;q,r)} = \left[\vec{u} - 2\left(F^{\zeta+}\vec{u}^+_\zeta + F^{\eta+}\vec{u}^+_\eta\right)\right]^{n-1/2}_{(j,k;q,r)} \tag{6.83}$$

$$(\vec{u}^{c+}_\zeta)^n_{j,k} \overset{def}{=} \frac{(-1)^q}{6}\left(\vec{u}'^n_{(j,k;q,2)} - \vec{u}'^n_{(j,k;q,1)}\right) \tag{6.84}$$

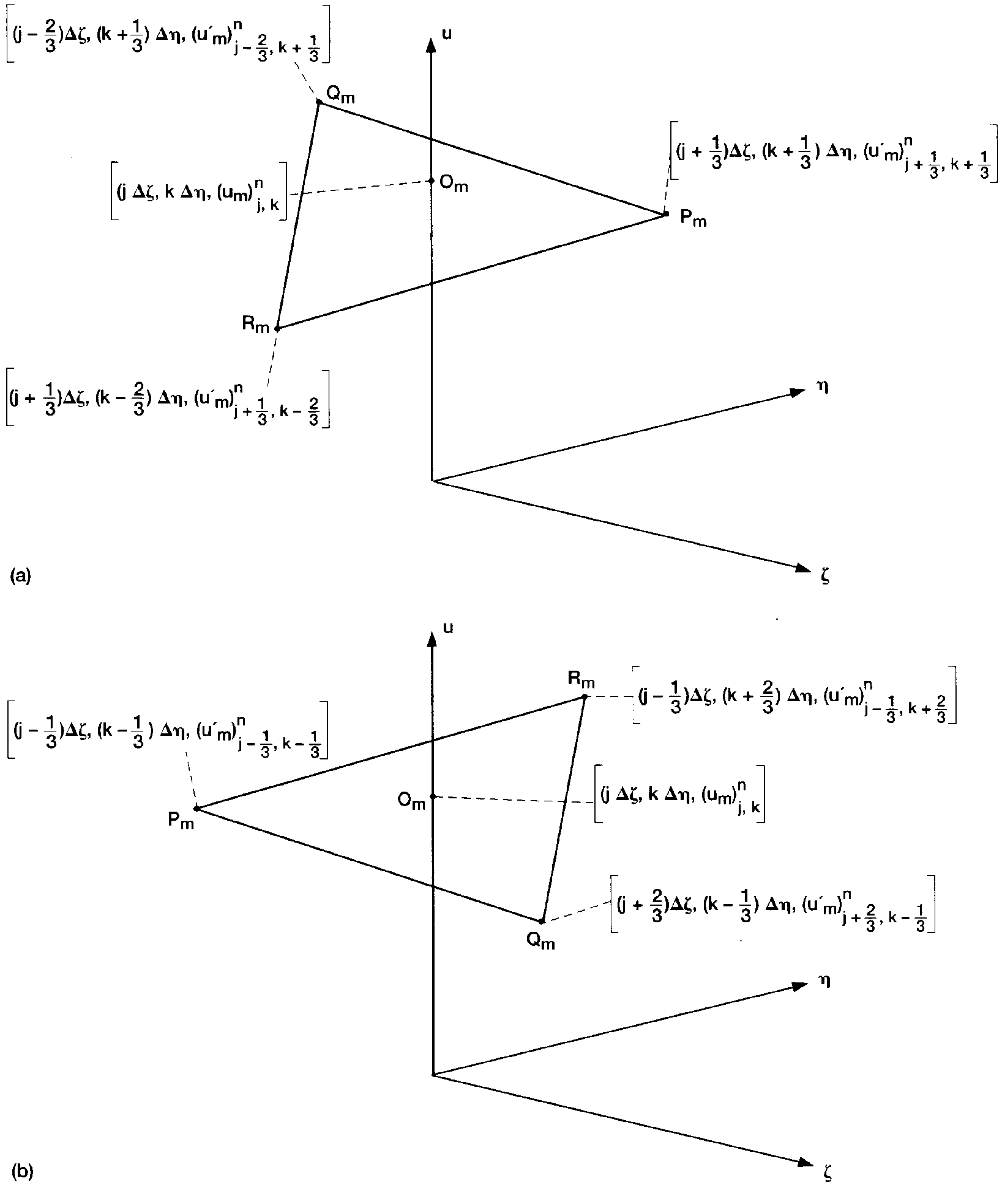


Figure 16: Construction of the 2D Euler a-ϵ and a-ϵ-α-β schemes ($m = 1, 2, 3, 4$). (a) $(j, k, n) \in \Omega_1$. (b) $(j, k, n) \in \Omega_2$.

$$(\vec{u}_\eta^{c+})_{j,k}^n \stackrel{def}{=} \frac{(-1)^q}{6}\left(\vec{u}'^{\,n}_{(j,k;q,3)} - \vec{u}'^{\,n}_{(j,k;q,1)}\right) \tag{6.85}$$

$$\left(\vec{u}_\zeta^+\right)_{j,k}^n = (\vec{u}_\zeta^{a+})_{j,k}^n + 2\epsilon(\vec{u}_\zeta^{c+} - \vec{u}_\zeta^{a+})_{j,k}^n \tag{6.86}$$

and

$$\left(\vec{u}_\eta^+\right)_{j,k}^n = (\vec{u}_\eta^{a+})_{j,k}^n + 2\epsilon(\vec{u}_\eta^{c+} - \vec{u}_\eta^{a+})_{j,k}^n \tag{6.87}$$

where $(j,k,n) \in \Omega_q$, $q = 1,2$. The 2D Euler a-ϵ scheme is formed by Eqs. (6.54), (6.86) and (6.87) for any $(j,k,n) \in \Omega_q$.

6.4. The Simplified 2D Euler a-ϵ Scheme

The defining equations of the simplified 2D Euler a-ϵ scheme are identical to those of the 2D Euler a-ϵ scheme except that Eqs. (6.86) and (6.87) should be replaced by

$$\left(\vec{u}_\zeta^+\right)_{j,k}^n = (\vec{u}_\zeta^{a'+})_{j,k}^n + 2\epsilon(\vec{u}_\zeta^{c+} - \vec{u}_\zeta^{a'+})_{j,k}^n \tag{6.88}$$

and

$$(u_\eta^+)_{j,k}^n = (\vec{u}_\eta^{a'+})_{j,k}^n + 2\epsilon(\vec{u}_\eta^{c+} - \vec{u}_\eta^{a'+})_{j,k}^n \tag{6.89}$$

respectively.

Moreover, with the aid of Eqs. (6.78)–(6.81) and (6.83)–(6.85), it can be shown that (i)

$$(\vec{u}_\zeta^{c+} - \vec{u}_\zeta^{a'+})_{j,k}^n = \frac{1}{6}\left[\left(\vec{u} + 4\vec{u}_\zeta^+ - 2\vec{u}_\eta^+\right)_{(j,k;1,2)}^{n-1/2} - \left(\vec{u} - 2\vec{u}_\zeta^+ - 2\vec{u}_\eta^+\right)_{(j,k;1,1)}^{n-1/2}\right] \tag{6.90}$$

and

$$(\vec{u}_\eta^{c+} - \vec{u}_\eta^{a'+})_{j,k}^n = \frac{1}{6}\left[\left(\vec{u} - 2\vec{u}_\zeta^+ + 4\vec{u}_\eta^+\right)_{(j,k;1,3)}^{n-1/2} - \left(\vec{u} - 2\vec{u}_\zeta^+ - 2\vec{u}_\eta^+\right)_{(j,k;1,1)}^{n-1/2}\right] \tag{6.91}$$

if $(j,k,n) \in \Omega_1$; and (ii)

$$(\vec{u}_\zeta^{c+} - \vec{u}_\zeta^{a'+})_{j,k}^n = \frac{1}{6}\left[\left(\vec{u} + 2\vec{u}_\zeta^+ + 2\vec{u}_\eta^+\right)_{(j,k;2,1)}^{n-1/2} - \left(\vec{u} - 4\vec{u}_\zeta^+ + 2\vec{u}_\eta^+\right)_{(j,k;2,2)}^{n-1/2}\right] \tag{6.92}$$

and

$$(\vec{u}_\eta^{c+} - \vec{u}_\eta^{a'+})_{j,k}^n = \frac{1}{6}\left[\left(\vec{u} + 2\vec{u}_\zeta^+ + 2\vec{u}_\eta^+\right)_{(j,k;2,1)}^{n-1/2} - \left(\vec{u} + 2\vec{u}_\zeta^+ - 4\vec{u}_\eta^+\right)_{(j,k;2,3)}^{n-1/2}\right] \tag{6.93}$$

if $(j,k,n) \in \Omega_2$.

Note that, under the substitution rules §3, §7 and §8, Eqs. (6.90)–(6.93) are the Euler images of Eqs. (5.13)–(5.16), respectively. Also note that $(\vec{u}_\zeta^{c+})_{j,k}^n$, $(\vec{u}_\zeta^{a'+})_{j,k}^n$, $(\vec{u}_\eta^{c+})_{j,k}^n$ and $(\vec{u}_\eta^{a'+})_{j,k}^n$ are explicitly dependent on $F^{\zeta+}$ and $F^{\eta+}$ (and, as a result of Eq. (6.31), also explicitly dependent on Δt). However, according to Eqs. (6.90)–(6.93), $(\vec{u}_\zeta^{c+} - \vec{u}_\zeta^{a'+})_{j,k}^n$ and $(\vec{u}_\eta^{c+} - \vec{u}_\eta^{a'+})_{j,k}^n$ are free from this depenency.

6.5. The 2D Euler a-ϵ-α-β Scheme

In this subsection, the techniques used in constructing the 1D Euler a-ϵ-α-β scheme and the 2D a-ϵ-α-β scheme will be combined and used to construct the 2D Euler a-ϵ-α-β scheme.

To proceed, for any $(j, k, n) \in \Omega_q$, any $m = 1, 2, 3, 4$, and any $r = 1, 2, 3$, let

$$x_{m,r} \stackrel{def}{=} (-1)^q \left[(u_m)^n_{j,k} - (u'_m)^n_{(j,k;q,r)} \right] \tag{6.94}$$

$$(u^{(r)}_{m\zeta})^n_{j,k} \stackrel{def}{=} f^{(r)}_\zeta(x_{m,1}, x_{m,2}, x_{m,3}), \qquad (u^{(r)}_{m\eta})^n_{j,k} \stackrel{def}{=} f^{(r)}_\eta(x_{m,1}, x_{m,2}, x_{m,3}) \tag{6.95}$$

$$(u^{(r)}_{mx})^n_{j,k} \stackrel{def}{=} f^{(r)}_x(x_{m,1}, x_{m,2}, x_{m,3}), \qquad (u^{(r)}_{my})^n_{j,k} \stackrel{def}{=} f^{(r)}_y(x_{m,1}, x_{m,2}, x_{m,3}) \tag{6.96}$$

where $f^{(r)}_\zeta$, $f^{(r)}_\eta$, $f^{(r)}_x$, and $f^{(r)}_y$ are the functions defined in Eqs. (5.24)–(5.29). Note that Eqs. (6.94)–(6.96) are the Euler counterparts of Eqs. (5.30)–(5.32), respectively.

To proceed further, for either $(j, k, n) \in \Omega_1$ or $(j, k, n) \in \Omega_2$, consider any *fixed* value of $m = 1, 2, 3, 4$. Let P_m, Q_m and R_m be the three points defined in Sec. 6.3. Let O_m (see Figs. 16(a) and 16(b)) denote the point in the ζ-η-u space with the coordinates $(j\Delta\zeta, k\Delta\eta, (u_m)^n_{j,k})$. Let planes #1, #2, and #3, respectively, be the planes containing the following trios of points: (i) points O_m, Q_m, and R_m; (ii) points O_m, R_m, and P_m; and (iii) points O_m, P_m, and Q_m. Then it can be shown that, for each $r = 1, 2, 3$, plane $\# r$ is represented by an equation that is identical to Eq. (5.33) except that the symbols $u^{(r)}_\zeta$, $u^{(r)}_\eta$, and u on the *right* side of Eq. (5.33) are now replaced by $u^{(r)}_{m\zeta}$, $u^{(r)}_{m\eta}$, and (u_m), respectively. Alternatively, the plane $\# r$ can be represented by another equation that is identical to Eq. (5.34) except that the symbols $u^{(r)}_x$, $u^{(r)}_y$, and u on the *right* side of Eq. (5.34) are now replaced by $u^{(r)}_{mx}$, $u^{(r)}_{my}$, and (u_m), respectively. As a result, for every point on the plane $\# r$, we have a set of relations that are identical to those given in Eqs. (5.35) and (5.36) except that the symbols $u^{(r)}_\zeta$, $u^{(r)}_\eta$, $u^{(r)}_x$, and $u^{(r)}_y$ in the latter equations are now replaced by $u^{(r)}_{m\zeta}$, $u^{(r)}_{m\eta}$, $u^{(r)}_{mx}$, and $u^{(r)}_{my}$, respectively. It follows that, at any point on plane $\# r$, we have

$$|\nabla u| = (\theta_{mr})^n_{j,k} \stackrel{def}{=} \left[\sqrt{(u^{(r)}_{mx})^2 + (u^{(r)}_{my})^2} \right]^n_{j,k} \tag{6.97}$$

Furthermore, let

$$(u^{(r)+}_{m\zeta})^n_{j,k} \stackrel{def}{=} \frac{\Delta\zeta}{6} (u^{(r)}_{m\zeta})^n_{j,k}, \qquad (u^{(r)+}_{m\eta})^n_{j,k} \stackrel{def}{=} \frac{\Delta\eta}{6} (u^{(r)}_{m\eta})^n_{j,k} \tag{6.98}$$

Then Eqs. (6.84), (6.85), (5.24)–(5.26) and (6.94)–(6.96) imply that

$$(u^{c+}_{m\zeta})^n_{j,k} = \frac{1}{3} \left[u^{(1)+}_{m\zeta} + u^{(2)+}_{m\zeta} + u^{(3)+}_{m\zeta} \right]^n_{j,k} \tag{6.99}$$

and

$$(u^{c+}_{m\eta})^n_{j,k} = \frac{1}{3} \left[u^{(1)+}_{m\eta} + u^{(2)+}_{m\eta} + u^{(3)+}_{m\eta} \right]^n_{j,k} \tag{6.100}$$

i.e., (i) $u^{c+}_{m\zeta}$ is the simple average of $u^{(r)+}_{m\zeta}$, $r = 1, 2, 3$; and (ii) $u^{c+}_{m\eta}$ is the simple average of $u^{(r)+}_{m\eta}$, $r = 1, 2, 3$. Equations (6.97)–(6.100) are the Euler counterparts of Eqs. (5.37)–(5.40), respectively.

With the above preliminaries, it becomes obvious that $u_{m\zeta}^{w+}$ and $u_{m\eta}^{w+}$, respectively the present counterparts of the weighted averages u_{ζ}^{w+} and u_{η}^{w+} defined in Eqs. (5.41) and (5.42), should be defined by

$$u_{m\zeta}^{w+} \stackrel{def}{=} \begin{cases} 0, & \text{if } \theta_{m1}=\theta_{m2}=\theta_{m3}=0 \\ \dfrac{(\theta_{m2}\theta_{m3})^{\alpha}\, u_{m\zeta}^{(1)+} + (\theta_{m3}\theta_{m1})^{\alpha}\, u_{m\zeta}^{(2)+} + (\theta_{m1}\theta_{m2})^{\alpha}\, u_{m\zeta}^{(3)+}}{(\theta_{m1}\theta_{m2})^{\alpha} + (\theta_{m2}\theta_{m3})^{\alpha} + (\theta_{m3}\theta_{m1})^{\alpha}}, & \text{otherwise} \end{cases} \tag{6.101}$$

and

$$u_{m\eta}^{w+} \stackrel{def}{=} \begin{cases} 0, & \text{if } \theta_{m1}=\theta_{m2}=\theta_{m3}=0 \\ \dfrac{(\theta_{m2}\theta_{m3})^{\alpha}\, u_{m\eta}^{(1)+} + (\theta_{m3}\theta_{m1})^{\alpha}\, u_{m\eta}^{(2)+} + (\theta_{m1}\theta_{m2})^{\alpha}\, u_{m\eta}^{(3)+}}{(\theta_{m1}\theta_{m2})^{\alpha} + (\theta_{m2}\theta_{m3})^{\alpha} + (\theta_{m3}\theta_{m1})^{\alpha}}, & \text{otherwise} \end{cases} \tag{6.102}$$

respectively. Note that, to avoid dividing by zero, in practice a small positive number such as 10^{-60} is added to the denominators in Eqs. (6.101) and (6.102).

Let $\vec{u}_{\zeta}^{w+}$ ($\vec{u}_{\eta}^{w+}$) be the column matrix formed by $u_{m\zeta}^{w+}$ ($u_{m\eta}^{w+}$), $m = 1, 2, 3, 4$. Then, for any $(j, k, n) \in \Omega$, the 2D Euler a-ϵ-α-β scheme is defined by Eq. (6.54) and

$$\left(\vec{u}_{\zeta}^{+}\right)_{j,k}^{n} = (\vec{u}_{\zeta}^{a+})_{j,k}^{n} + 2\epsilon(\vec{u}_{\zeta}^{c+} - \vec{u}_{\zeta}^{a+})_{j,k}^{n} + \beta(\vec{u}_{\zeta}^{w+} - \vec{u}_{\zeta}^{c+})_{j,k}^{n} \tag{6.103}$$

and

$$\left(\vec{u}_{\eta}^{+}\right)_{j,k}^{n} = (\vec{u}_{\eta}^{a+})_{j,k}^{n} + 2\epsilon(\vec{u}_{\eta}^{c+} - \vec{u}_{\eta}^{a+})_{j,k}^{n} + \beta(\vec{u}_{\eta}^{w+} - \vec{u}_{\eta}^{c+})_{j,k}^{n} \tag{6.104}$$

where ϵ and β are adjustable parameters.

6.6. The Simplified 2D Euler a-ϵ-α-β Scheme

For any $(j, k, n) \in \Omega$, the simplified 2D Euler a-ϵ-α-β scheme is formed by Eq. (6.54) and

$$\left(\vec{u}_{\zeta}^{+}\right)_{j,k}^{n} = (\vec{u}_{\zeta}^{a'+})_{j,k}^{n} + 2\epsilon(\vec{u}_{\zeta}^{c+} - \vec{u}_{\zeta}^{a'+})_{j,k}^{n} + \beta(\vec{u}_{\zeta}^{w+} - \vec{u}_{\zeta}^{c+})_{j,k}^{n} \tag{6.105}$$

and

$$\left(\vec{u}_{\eta}^{+}\right)_{j,k}^{n} = (\vec{u}_{\eta}^{a'+})_{j,k}^{n} + 2\epsilon(\vec{u}_{\eta}^{c+} - \vec{u}_{\eta}^{a'+})_{j,k}^{n} + \beta(\vec{u}_{\eta}^{w+} - \vec{u}_{\eta}^{c+})_{j,k}^{n} \tag{6.106}$$

where ϵ and β are adjustable parameters.

6.7. The 2D CE/SE Shock-Capturing Scheme

Let $\epsilon = 1/2$ and $\beta = 1$. Then the 2D Euler a-ϵ-α-β scheme and the simplified 2D Euler a-ϵ-α-β scheme reduce to the same scheme. For any $(j, k, n) \in \Omega$, the reduced scheme is formed by Eq. (6.54) and

$$\left(\vec{u}_{\zeta}^{+}\right)_{j,k}^{n} = (\vec{u}_{\zeta}^{w+})_{j,k}^{n} \tag{6.107}$$

and

$$\left(\vec{u}_{\eta}^{+}\right)_{j,k}^{n} = (\vec{u}_{\eta}^{w+})_{j,k}^{n} \tag{6.108}$$

The above scheme is one of the simplest among the 2D Euler solvers known to the authors. *The value of α is the only adustable parameter allowed in this scheme.* Because this scheme is the 2D counterpart of the 1D CE/SE shock-capturing scheme and shares with the latter all the distinctive features described in Sec. 2.8, it will be referred to as the 2D CE/SE shock-capturing scheme.

7. Stability

In this section, stability of the 2D a and a-ϵ schemes will be studied using the von Neumann analysis. Note that Eqs. (4.73) and (4.74) are valid for these two schemes if the matrices $Q_r^{(q)}$ ($q = 1, 2$ and $r = 1, 2, 3$) are defined using Eqs. (5.18)–(5.23) with the understanding that $\epsilon = 0$ should be assumed for the 2D a scheme.

To proceed, let

$$M^{(1)}(\theta_\zeta, \theta_\eta) \stackrel{def}{=} Q_1^{(1)} e^{(i/3)(\theta_\zeta + \theta_\eta)} + Q_2^{(1)} e^{(i/3)(-2\theta_\zeta + \theta_\eta)} + Q_3^{(1)} e^{(i/3)(\theta_\zeta - 2\theta_\eta)} \tag{7.1}$$

and

$$M^{(2)}(\theta_\zeta, \theta_\eta) \stackrel{def}{=} Q_1^{(2)} e^{-(i/3)(\theta_\zeta + \theta_\eta)} + Q_2^{(2)} e^{-(i/3)(-2\theta_\zeta + \theta_\eta)} + Q_3^{(2)} e^{-(i/3)(\theta_\zeta - 2\theta_\eta)} \tag{7.2}$$

Furthermore, for all $(j, k, n) \in \Omega$, let

$$\vec{q}(j, k, n) = \vec{q}^*(n, \theta_\zeta, \theta_\eta) e^{i(j\theta_\zeta + k\theta_\eta)}, \qquad (i \stackrel{def}{=} \sqrt{-1}, \quad -\pi < \theta_\zeta, \theta_\eta \le \pi) \tag{7.3}$$

where $\vec{q}^*(n, \theta_\zeta, \theta_\eta)$ is a 3×1 column matrix (see Sec. 4 in [1]). Substituting Eq. (7.3) into Eqs. (4.73) and (4.74), one concludes that: (i)

$$\vec{q}^*(n + m, \theta_\zeta, \theta_\eta) = \left[M^{(1)}(\theta_\zeta, \theta_\eta) M^{(2)}(\theta_\zeta, \theta_\eta)\right]^m \vec{q}^*(n, \theta_\zeta, \theta_\eta) \tag{7.4}$$

where $n = \pm 1/2, \pm 3/2, \pm 5/2, \ldots$, and $m = 0, 1, 2, \ldots$; and (ii)

$$\vec{q}^*(n + m, \theta_\zeta, \theta_\eta) = \left[M^{(2)}(\theta_\zeta, \theta_\eta) M^{(1)}(\theta_\zeta, \theta_\eta)\right]^m \vec{q}^*(n, \theta_\zeta, \theta_\eta) \tag{7.5}$$

where $n = 0, \pm 1, \pm 2, \ldots$, and $m = 0, 1, 2, \ldots$. Equation (7.4) implies that the amplification matrix among the half-integer time levels is $M^{(1)}(\theta_\zeta, \theta_\eta) M^{(2)}(\theta_\zeta, \theta_\eta)$; while Eq. (7.5) implies that the amplification matrix among the whole-integer time levels is $M^{(2)}(\theta_\zeta, \theta_\eta) M^{(1)}(\theta_\zeta, \theta_\eta)$.

Let A and B be two arbitrary $n \times n$ matrices. Then AB and BA have the same eigenvalues, counting multiplicity [54, p.53]. Thus the 3×3 amplification matrix among the half-integer time levels and that among the whole-integer time levels have the same eigenvalues. These eigenvalues may be referred to as the amplification factors. The amplification factors are functions of phase angles θ_ζ and θ_η. In addition, they are functions of a set of coefficients that are dependent on the physical properties and the mesh parameters. These coefficients are (i) ν_ζ and ν_η for the 2D a scheme; and (ii) ν_ζ, ν_η, and ϵ for the 2D a-ϵ scheme. Let λ_1, λ_2, and λ_3 denote the amplification factors. In the current stability analysis, *a scheme is said to be stable in a domain of the above coefficients if, for all values of the coefficients belonging to this domain, and all θ_ζ and θ_η with $-\pi < \theta_\zeta, \theta_\eta \le \pi$,*

$$|\lambda_1| \le 1, \qquad |\lambda_2| \le 1, \qquad \text{and} \qquad |\lambda_3| \le 1 \tag{7.6}$$

Consider the 2D a scheme. By using its two-way marching nature and the fact that its stencil is invariant under space-time inversion, it is shown in [9] that, for any given ν_ζ, ν_η, θ_ζ, and θ_η,

$$|\lambda_1 \lambda_2 \lambda_3| = 1 \tag{7.7}$$

It follows from Eqs. (7.6) and (7.7) that the 2D a scheme must be neutrally stable, i.e.,

$$|\lambda_1| = |\lambda_2| = |\lambda_3| = 1, \qquad -\pi < \theta_\zeta, \theta_\eta \leq \pi \tag{7.8}$$

if it is stable. In other words, *the 2D a scheme is non-dissipative if it is stable.* Moreover, *a systematic numerical evaluation of λ_1, λ_2, and λ_3, for different values of ν_ζ, ν_η, θ_ζ, and θ_η, has confirmed that the 2D a scheme is indeed neutrally stable in the stability domain defined by Eq. (4.75).* In the following, we shall discuss the meaning of this stability domain.

Let $(j,k,n) \in \Omega$. According to Eqs. (4.73) and (4.74), the marching variables at the mesh point (j,k,n) are completely determined by those of seven mesh points at the $(n-1)$th time level (i.e., the mesh point $(j,k,n-1)$, and points A, B, C, D, E and F shown in Figs. 17(a) and 17(b)). As a result, in this paper, *the interior and boundary of the hexagon $ABCDEF$ shall be referred to as the numerical domain of dependence of the mesh point (j,k,n) at the $(n-1)$th time level.* Note that the dashed lines depicted in Figs. 17(a) and 17(b) are the spatial projections of boundaries of CEs.

The 2D a scheme is designed to solve Eq. (4.1). For Eq. (4.1), the value of u is a constant along a characteristic line. The characteristic line passing through the mesh point (j,k,n) will intersect a point on the plane $t = t^{n-1}$. The point of intersection, *referred to as the backward characteristic projection of the mesh point (j,k,n) at the $(n-1)$th time level*, is the "domain" of dependence at the $(n-1)$th time level for the value of u at the mesh point (j,k,n). It is shown in Appendix D.1 that the backward characteristic projection is in the interior of the numerical domain of dependence if and only if Eq. (4.75) is satisfied.

At this juncture, note that the concept of characteristics was never used in the design of the 2D a scheme. Nevertheless, its stability condition is completely consistent with the general stability requirement of an explicit solver of a hyperbolic equation, i.e., the analytic domain of dependence be a subset of the numerical domain of dependence.

Next we consider the stability of the 2D a-ϵ scheme. Recall that the 1-D a-ϵ scheme is not stable for any Courant number ν if $\epsilon < 0$, or $\epsilon > 1$ [2]. Similarly, the results of numerical experiments indicate that the 2D a-ϵ scheme is not stable in any domain on the ν_ζ-ν_η plane if $\epsilon < 0$ or $\epsilon > 1$. For any ϵ with $0 \leq \epsilon \leq 1$, the 2D a-ϵ scheme has a stability domain on the ν_ζ-ν_η plane. The stability domains for several values of ϵ were obtained numerically. As shown in Figs. 18(a)-(c), these domains (shaded areas) vary only slightly in shape and size from that depicted in Fig. 14. They become smaller in size as ϵ increases.

Given any pair of ν_ζ and ν_η, λ_1, λ_2 and λ_3 are functions of θ_ζ and θ_η. Let (i)

$$|\lambda_3| \leq |\lambda_2| \leq |\lambda_1| \leq 0, \qquad -\pi < \theta_\zeta, \theta_\eta \leq \pi \tag{7.9}$$

and (ii) $\lambda_1 = 1$ when $\theta_\zeta = \theta_\eta = 0$. Then λ_1 can be referred to as the principal amplification factor; while λ_2 and λ_3 are referred to as the spurious amplification factors [1]. In general, the principal amplification factor is the deciding factor in determining the accuracy of computations [1]. Specifically, numerical solutions may suffer annihilations of sharply different degrees at different locations and different frequencies if numerical diffusion associated with λ_1 varies greatly with respect to θ_ζ, θ_η, ν_ζ, and ν_η [7, p.20]. Moreover, note that $(1 - |\lambda_r|)$ is

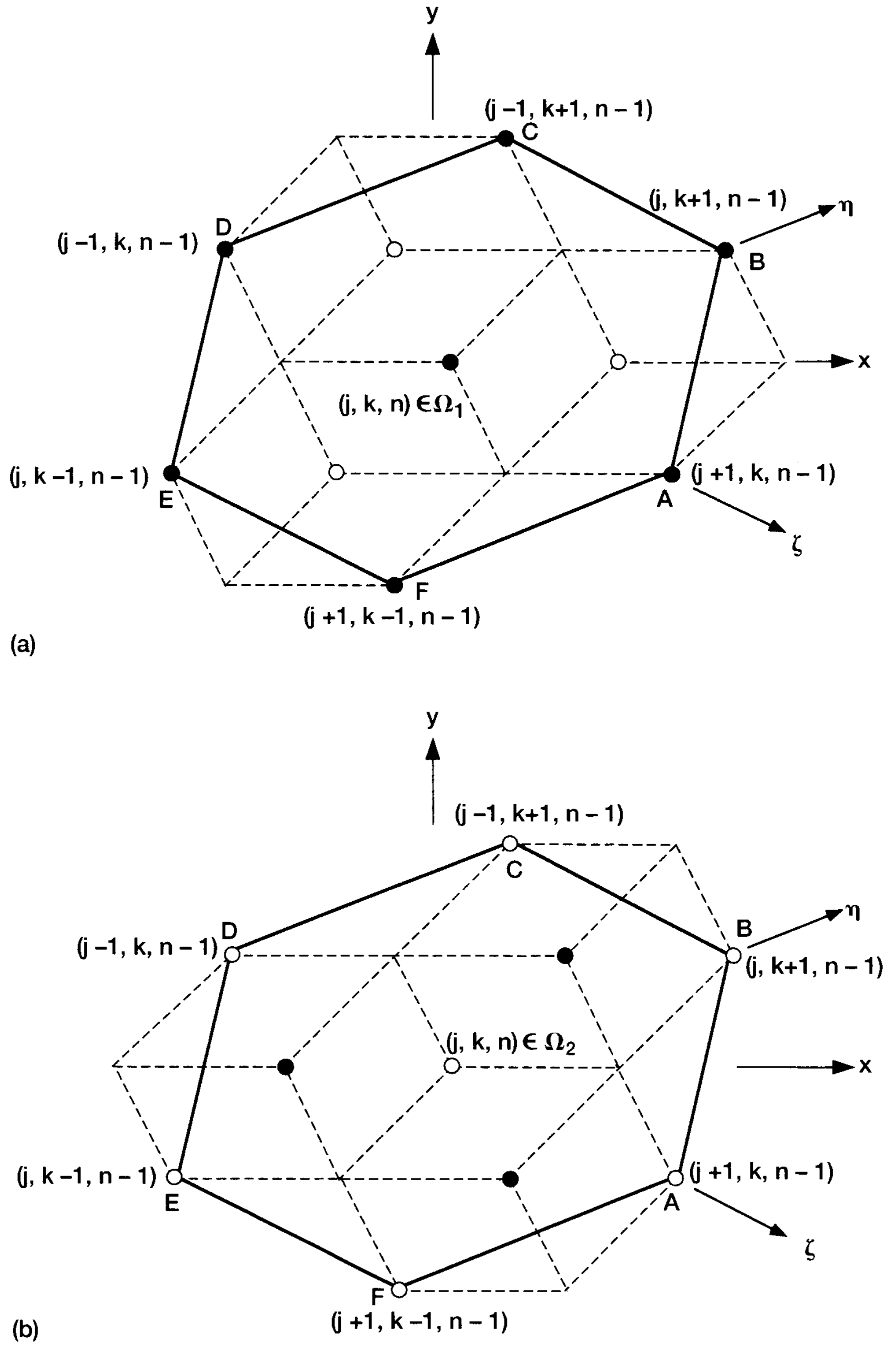


Figure 17: The numerical domains of dependence associated with the 2D CE/SE solvers. (a) $(j,k,n) \in \Omega_1$. (b) $(j,k,n) \in \Omega_2$.

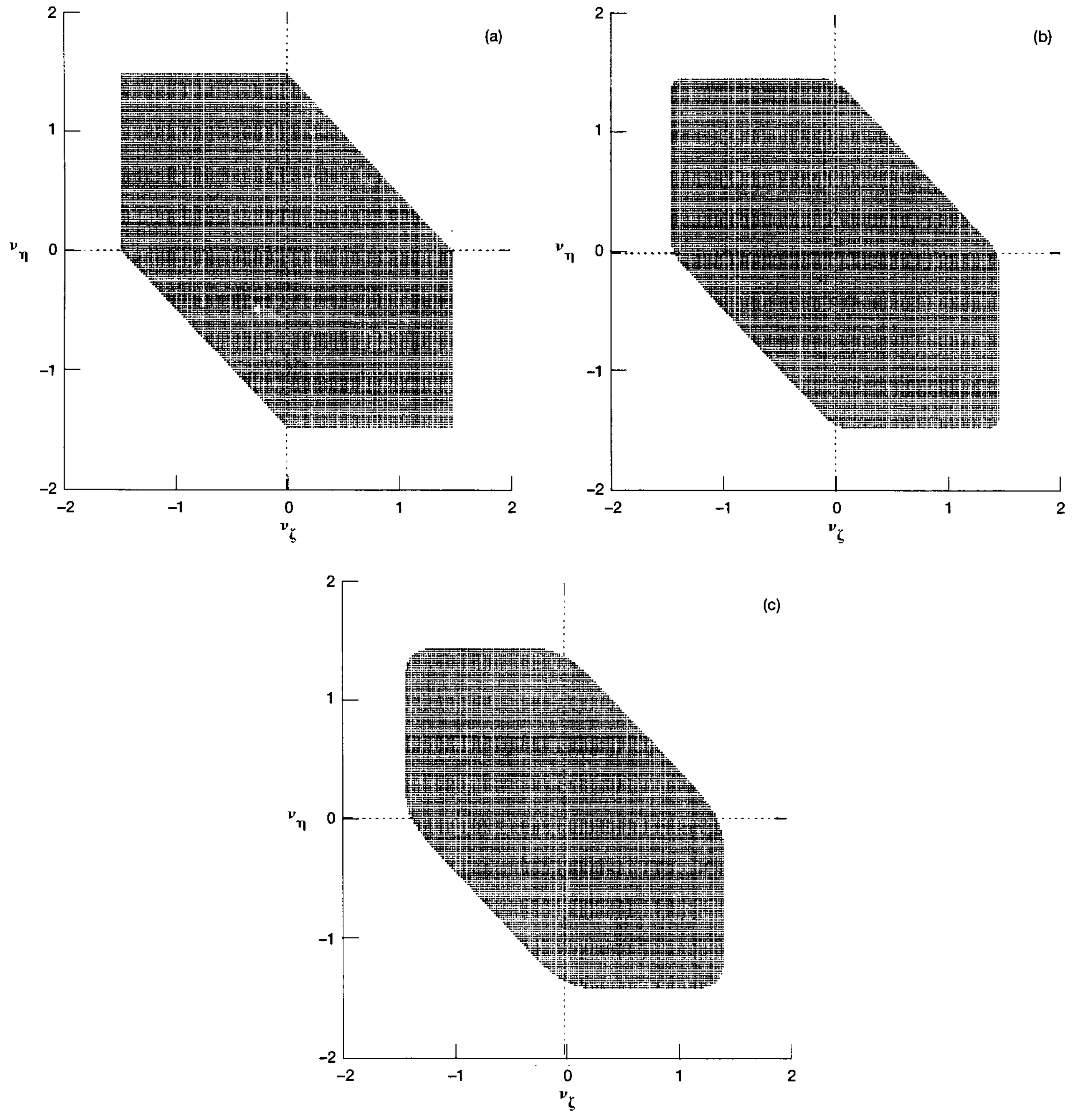


Figure 18: The stability domain of the 2D a-ϵ scheme. (a) ϵ=0.1. (b) ϵ=0.5. (c) ϵ=0.8.

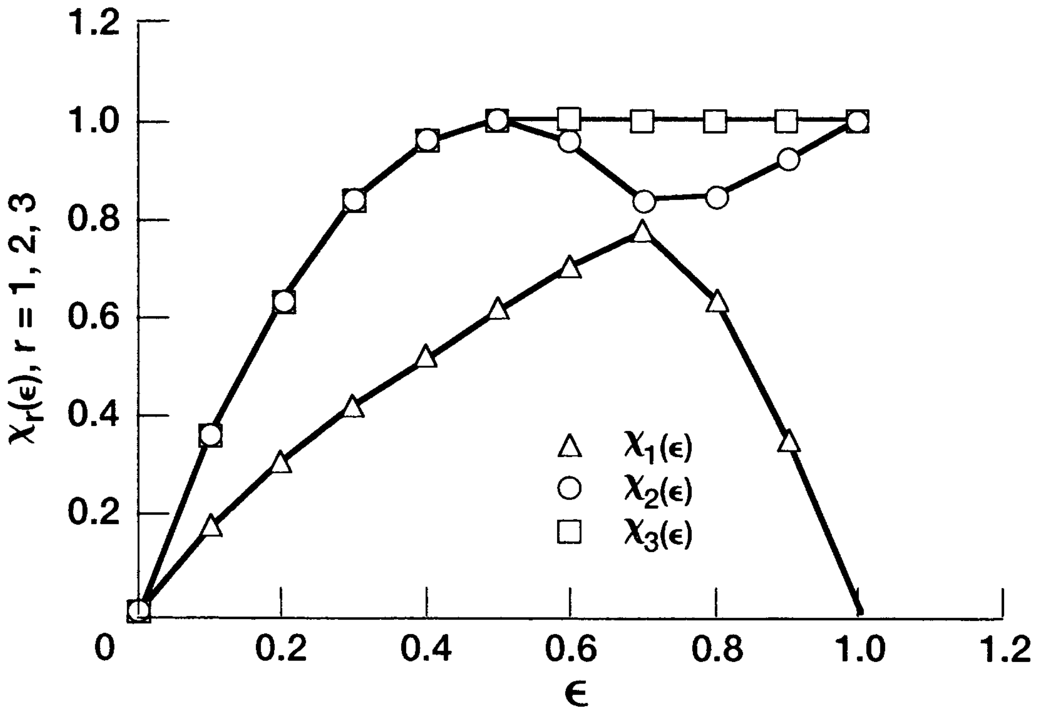


Figure 19: The functions $\chi_r(\epsilon)$, $r = 1, 2, 3$.

a measure of the numerical diffusion associated with λ_r, $r = 1, 2, 3$. For a given ϵ, let $D(\epsilon)$ denote the stability domain of the 2D a-ϵ scheme on the ν_ζ-ν_η plane. Let

$$\chi_r(\epsilon) \stackrel{def}{=} \max_{-\pi<\theta_\zeta,\theta_\eta\leq\pi;\ (\nu_\zeta,\nu_\eta)\in D(\epsilon)} (1 - |\lambda_r|), \qquad r = 1, 2, 3; \quad 0 \leq \epsilon \leq 1 \tag{7.10}$$

Then, for a given ϵ and each r, $(1 - |\lambda_r|)$ is bounded *uniformly* from above by $\chi_r(\epsilon)$. The numerically estimated values of $\chi_r(\epsilon)$ are plotted in Fig. 19. From this figure, one concludes that the numerical diffusion, particularly that associated with λ_1, can be bounded *uniformly* from above by an arbitrary small number by choosing an ϵ small enough. Note that this property is also shared by the 1-D a-ϵ scheme (see Eq. (3.19) in [2]). Moreover, the results shown in Fig. 19 indicate that $\chi_2(\epsilon)$ and $\chi_3(\epsilon)$ are much larger than $\chi_1(\epsilon)$ in the range of $0 \leq \epsilon \leq 0.5$. Thus, in this range, the spurious part of a numerical solution is annihilated much faster than the principal part. Also it is seen that the numerical diffusion associated with the principal solution, measured by $\chi_1(\epsilon)$, increases with ϵ in the range of $0 \leq \epsilon \leq 0.7$.

Because of the appearance of *nonlinear* weighted-average terms in its defining equations, stability of the 2D a-ϵ-α-β scheme is difficult to study analytically. However, results from numerical experiments indicate that the stability domain of this scheme is slightly larger than that of the 2D a-ϵ scheme when $\alpha > 0$ and $\beta > 0$.

Before we proceed further, several concepts related to stability need to be clarified. First note that, to define a numerical problem, one must specify (i) the main scheme (such as any solver described in Secs. 4–6) used in the updating of the marching variables at the interior mesh points, and (ii) the auxiliary discrete initial/boundary conditions. Thus, generally stability is not a concept involving only the main scheme.

Next note that use of the von Neumann stability analysis can be rigorously justified only if the numerical problem under consideration satisfies a set of strict conditions [1]. They include (i) the mesh used should be uniform in both spatial and temporal directions, (ii) the main scheme used should be linear in the discrete variables, and (iii) the boundary conditions used should be periodic in nature. Because (i) the stability conditions generated using the von Neumann analysis are expressed in terms of the coefficients of the discrete variables and the mesh parameters only, and (ii) the above coefficients and mesh parameters are constant and independent of the initial/boundary conditions, the stability of a numerical problem that satisfies the above strict conditions (i)–(iii) is completely independent of the initial/boundary conditions. For this special numerical problem, stability can be considered as a concept involving only the main scheme.

For a uniform-mesh linear problem with non-periodic boundary conditions, the stability conditions generated from the von Neumann analysis generally are necessary but not sufficient conditions for stability. For such a problem, the initial/boundary conditions may have an impact on stability and numerical diffusion. Note that the results given earlier in this section are obtained without considering this impact.

Generally, stability of a nonlinear problem is highly dependent on the initial/boundary conditions, and therefore highly problem-dependent. As a result, a discussion of the stability of nonlinear solvers without specifying the exact initial/boundary conditions, such as that to be given immediately, is inherently imprecise in nature.

To proceed, for each mesh point $(j,k,n) \in \Omega$, a local Euler CFL number $\nu_e \geq 0$ is introduced in Appendix D.2 (see Eqs. (D.32)–(D.35)). This number has the following property: For the flow variables at the mesh point (j,k,n), its analytical domain of dependence at the $(n-1)$th time level lies within the corresponding numerical domain of dependence if and only if $\nu_e < 1$. According to the results of numerical experiments, both the 2D Euler a scheme and the simplified 2D Euler a scheme are generally unstable. However the former is only marginally unstable when $\nu_{e,max} < 1$ where $\nu_{e,max}$ is the maximum value of ν_e ever reached in a numerical experiment. As a matter of fact, in simulating smooth flows, its round-off error often never reaches an appreciable level before the end of the simulation run. As for the other solvers described in Sec. 6, stability generally can be realized if $\nu_{e,max} < 1$ and $0.05 < \epsilon < 1$. However, for a nonsmooth flow problem, stricter stability conditions such as $\nu_{e,max} < 2/3$, $0.1 < \epsilon < 1$ and $\alpha \geq 1$ may apply.

8. Conclusions and Discussions

The space-time CE/SE method was conceived from a global CFD perspective and designed to avoid the limitations of the traditional methods. It was built from ground zero with a foundation which is solid in physics and yet mathematically simple enough that one can build from it a coherent, robust, efficient and accurate CFD numerical framework. A clear and thorough discussion of these basic motivating ideas was given in Sec. 1.

The 1D CE/SE schemes [2] were reformulated in Sec. 2 such that the reader can see more clearly the structural similarity between the solvers of the 1D convection equation Eq. (1.1) and those of the 1D Euler equations. In addition, this reformulation also paves the way for the construction of the 2D CE/SE schemes and makes it easier for the reader to appreciate the consistency between the construction of the 1D CE/SE schemes and that of the 2D schemes.

It was shown in Sec. 3 that the basic building blocks of the spatial meshes used in the 2D CE/SE schemes are triangles. As a result, these schemes are compatible with the simplest unstructured meshes, and therefore are applicable to 2D problems with complex geometries. Furthermore, because they are constructed without using the dimensional-splitting approach, these schemes are genuinely multidimensional.

The 2D a scheme, a nondissipative solver for the 2D convection equation Eq. (4.1), was constructed in Sec. 4. It is a natural extension of the 1D a scheme and shares with the latter several nontraditional features which are listed following Eq. (4.74).

Because a nonlinear extension of a nondissipative linear solver generally is unstable or highly dispersive, the 2D a scheme was modified in Sec. 5 to become the dissipative 2D a-ϵ and a-ϵ-α-β schemes before it was extended to model the 2D Euler equations. It was clearly explained in Sec. 5 that these 2D dissipative schemes are the natural extensions of the 1D a-ϵ and a-ϵ-α-β schemes, respectively. Moreover, as in the case of the latter schemes, numerical dissipation introduced in the former schemes is controlled by the parameters ϵ, α and β.

A family of solvers for the 2D Euler equations were constructed in Sec. 6. Not only are these solvers the natural extensions of the 1D CE/SE Euler solvers, but their algebraic structures are strikingly similar to those of the 2D a, a-ϵ and a-ϵ-α-β schemes.

Next, stability of the 2D solvers described in Sec. 4–6 was discussed in Sec. 7. It was shown that the 2D a scheme is nondissipative in the stability domain defined by Eq. (4.75). It was also shown that the necessary stability conditions for the 2D solvers include: (i) the local CFL number < 1 at every mesh point, and (ii) $1 \geq \epsilon \geq 0$, $\alpha \geq 0$ and $\beta \geq 0$ if applicable. Note that these conditions are also necessary stability conditions for the 1D CE/SE solvers.

A summary of the key results of the present paper has been given. It is seen that each of the present 2D schemes is constructed in a very simple and consistent manner as the natural extension of its 1D counterpart. This is made possible because of the present development's strict adherence to its two basic beliefs which were stated in Sec. 1.

To evaluate the accuracy and robustness of the CE/SE schemes, the two simplest schemes among them, i.e., the 1D and 2D CE/SE shock-capturing schemes, will be used in Part II [3] to simulate flows involving phenomena such as shock waves, contact discontinuities, expansion waves and their interactions. The numerical results, when compared with experimental data, exact solutions or numerical solutions by other methods, indicate that these schemes can consistently resolve shock and contact discontinuities with high accuracy. Note that other CE/SE schemes described in this paper have also been shown to be accurate solvers for other applications [11,13–17,20,24,26-28]. Furthermore, using the present method, Yu *et al.* have successfully constructed several accurate solvers for 1D and 2D problems with stiff source terms [21,22,32].

Note that the 1D CE/SE schemes have been extended to become accurate 2D and 3D solvers by others without using the current approach. After constructing their 1D CE/SE solver for the Saint Venant equations, Molls *et al.* [29] construct the 2D version using the Strang's splitting technique [56]. Furthermore, several 2D and 3D non-splitting Euler solvers have also been constructed by Zhang *et al.* [57–61] without using triangular or tetrahedral meshes.

The triangles depicted in Fig. 5 are obtained by sectioning each parallelogram depicted in the same figure into two triangles. The 2D CE/SE solvers can also be constructed using the triangles that are obtained by sectioning each parallelogram into four triangles. These solvers along with other CE/SE solvers with nonuniform spatial meshes [4] will be described in future papers.

This paper is concluded with a discussion of several other extensions.

8.1. A sketch of a 3D Euler solver

The CE/SE method can be extended to three spatial dimensions using the same procedure that was used in extending the method from one spatial dimension to two spatial dimensions. In the 3D case, at each mesh point, the mesh values of any physical variable and its three spatial gradient components are considered as independent variables. Because there are four independent discrete variables per physical variable (or per conservation law to be solved), construction of the 3D a scheme and the 3D Euler a scheme demands that four CEs be defined at each mesh point. As will be shown immediately, this requirement can be met by using tetrahedrons as the basic building blocks of the 3D spatial mesh.

To pave the way, consider the 2D case and Figs. 5 and 6(a). The quadrilaterals $GFAB$, $GBCD$ and $GDEF$ are the spatial projections of the CEs associated with the point G'. The CEs in the 3D case can be constructed in a similar fashion. Consider the tetrahedrons $ABCD$ and $ABCP$ depicted in Fig. 20. Points G and H are the centroids of $ABCD$ and $ABCP$, respectively. The two tetrahedrons share the face ABC. The polyhedron $GABCH$ is then defined as the spatial projection of a CE associated with a space-time mesh point G'. The CE is thus a right cylinder in space-time, with $GABCH$ as its spatial base. The point G is the spatial projection of point G'.

In a similar fashion, three additional CEs associated with the mesh point G' can be

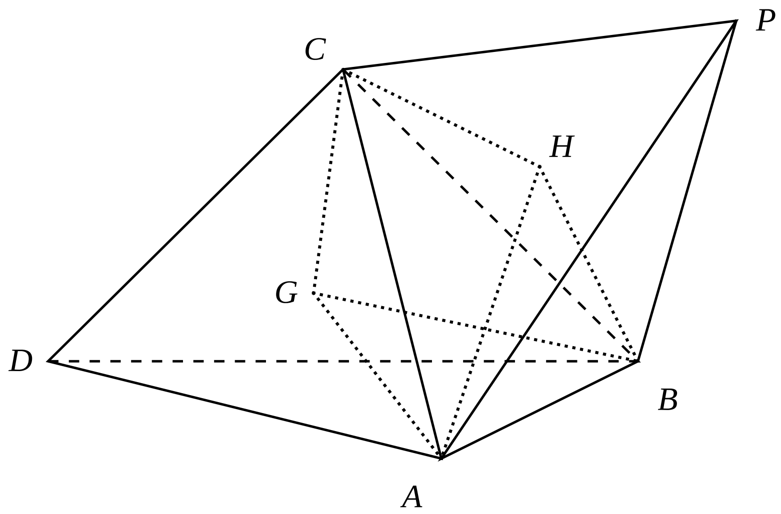


Figure 20: Spatial projection of part of a 3D space-time mesh, showing the construction of a CE.

constructed by considering in turn the three tetrahedrons that share with $ABCD$ one of its other three faces.

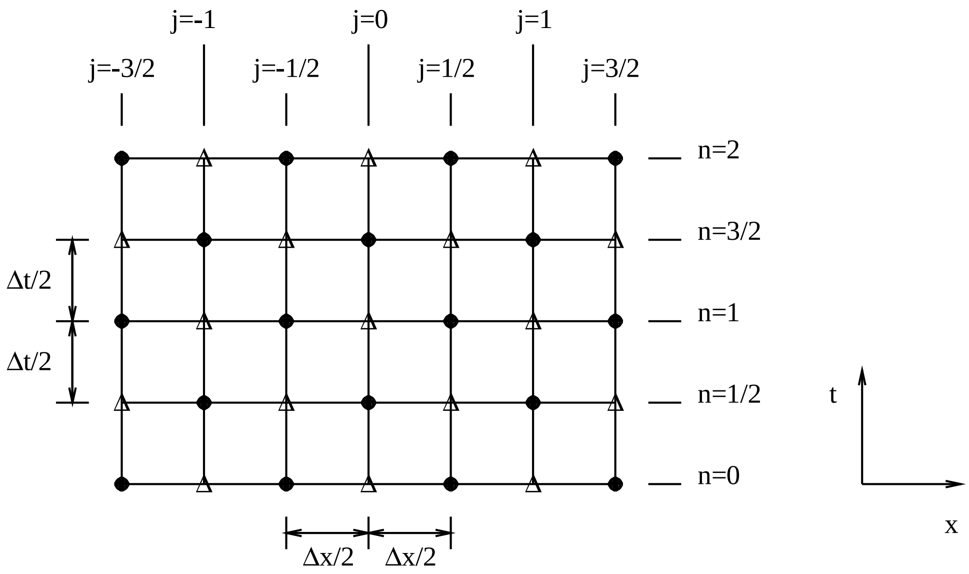
Note that a structured 3D spatial mesh can be constructed from the tetrahedrons that are obtained by sectioning the parallelepipeds occupying a spatial region. The details will be given in a separate paper.

8.2. Concept of Dual Space-Time Meshes and Its Applications

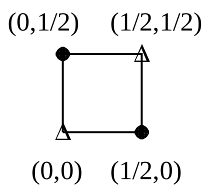
The mesh depicted in Fig. 4(a) is staggered in time, i.e., the mesh points that have the same spatial locations appear only at alternating time levels. In Fig. 21(a), the mesh depicted in Fig. 4(a) (referred to as the mesh 1) is superimposed on another staggered mesh (referred to as the mesh 2), with the mesh points of the latter being marked by solid triangular symbols. The combination of the meshes 1 and 2 shall be referred to as the dual mesh. As shown in Fig. 21(b), a CE of a mesh point marked by a triangle may coincide with a CE of another mesh point marked by a dot.

Obviously the 1D a scheme can also be constructed using mesh 2. As a matter of fact, one can even combine two independent 1D a schemes, one constructed on the mesh 1, and the other on the mesh 2, into a "single" scheme referred to as the 1D dual a scheme. Similarly one can also construct the dual 1D a-ϵ and a-ϵ-α-β schemes. Each of the new schemes has two completely decoupled solutions. Without considering this decoupled nature in the von Neumann analysis, it can be shown that the resulting amplification factors of the dual 1D a scheme are identical to those of the Leapfrog scheme as given in [52, p.100]. Note that the deficiency of the standard practice that the amplification factors of the Leapfrog scheme are obtained without taking into account the decoupled nature of its solutions was addressed in Sec. 1.

Let (j, n) be a mesh point of mesh 1 (mesh 2). Then $(j \pm 1/2, n)$ are mesh points of



(a). — The dual space-time mesh



(b). — A rectangular space-time region shared by $CE_-(1/2,1/2)$ and $CE_+(0,1/2)$

Figure 21: Concept of dual space-time meshes. (a) The dual space-time mesh. (b) A rectangular space-time region shared by $CE_-(1/2,1/2)$ and $CE_+(0,1/2)$.

mesh 2 (mesh 1). Recall that ${u'}^n_{j\pm1/2}$ (see Eq. (2.10)) are defined in terms of the marching variables at $(j \pm 1/2, n - 1/2)$, which are on the same mesh with (j, n). Thus the two solutions on meshes 1 and 2 of either the dual 1D a-ϵ scheme or the dual 1D a-ϵ-α-β scheme are decoupled. However, by replacing ${u'}^n_{j\pm1/2}$ with $u^n_{j\pm1/2}$ (which are evaluated using Eq. (2.8) with the understanding that j be replaced by $j \pm 1/2$) in their construction, each of the above two schemes will turn into a new scheme in which the solutions on meshes 1 and 2 become coupled. The coupling results from the fact that u^n_j and $u^n_{j\pm1/2}$ are not associated with the same mesh. Note that the solutions of the new schemes generally are indistinguishable from (or only slightly more diffusive than) those of the original schemes.

In [12,25], two implicit schemes for solving the convection-diffusion equation Eq. (1.2) were constructed using a dual space-time mesh. In the case that $\mu = 0$, both the above implicit schemes reduce to the explicit non-dissipative dual a scheme. As a result, the amplification factors of these schemes reduce to those of the Leapfrog scheme if $\mu = 0$. Furthermore, these two implicit schemes have the property that their numerical dissipation approaches zero as the physical dissipation approaches zero. The significance of this property was discussed in Sec. 1.

In case that $\mu > 0$, both the above implicit schemes are truly implicit. This implicit nature is consistent with the fact that, for $\mu > 0$, the value of a solution to Eq. (1.2) at any point (x, t) depends on the initial data and all the boundary data up to the time t. In other words, generally an implicit scheme should be used to solve an initial/boundary-value problem, such as one involving Eq. (1.2) with $\mu > 0$. This requirement becomes more important as the diffusion term in Eq. (1.2) becomes more dominant.

In addition, for both the above implicit schemes, the solution at the mesh points marked by dots, through the diffusion term in Eq. (1.2), is coupled with that at the mesh points marked by triangles if $\mu > 0$. Also it was shown in [12,25] that, in the pure diffusion case (i.e., when $a = 0$), the principal amplification factors of both the above implicit schemes reduce to the amplification factor of the Crank-Nicolson scheme [52]. Note that the latter has only one amplification factor.

The concept of dual space-time meshes also is applicable to the 2D and 3D cases. As an example, consider a 2D mesh (the mesh 1) with the mesh points marked by circles in Fig. 6(a)–(c). For this case, the mesh points of the mesh 2 are points G, C', E', G'', I'' and K''. In general, if (j, k, n) represents a mesh point of the mesh 1, then (j, k, n') represents a mesh point of the mesh 2 if and only if $(n - n')$ is a half-integer. Note that a more complete discussion of the concept of dual meshes will be given in Part II [3].

Note that not only can the concept of dual meshes be used to construct implicit schemes, but it can also be used to implement reflecting boundary conditions (see the following paper [3]). In addition, this concept is indispensable in the development of a 2D triangular unstructured-mesh CE/SE scheme [31].

8.3. A discussion on locally adjustable numerical dissipation

Consider the 1D a-ϵ-α-β scheme, i.e., the scheme defined by Eqs. (2.7) and (2.60). With

ϵ, α and β being held constant, generally numerical dissipation associated with this scheme increases as the Courant number ν decreases. To compensate for this effect, Eq. (2.60) may be replaced by

$$(u_x^+)_j^n = (u_x^{a+})_j^n + 2\epsilon(\nu)(u_x^{c+} - u_x^{a+})_j^n + \beta(\nu)(u_x^{w+} - u_x^{c+})_j^n \tag{8.1}$$

where $\epsilon(\nu)$ and $\beta(\nu)$ are monotonically decreasing functions of ν with $\epsilon(0) = \beta(0) = 0$. The optimal forms of these functions generally are problem-dependent. The scheme defined by Eqs. (2.7) and (8.1) has the property that

$$(u_x^+)_j^n \to (u_x^{a+})_j^n \quad \text{as} \quad \Delta t \to 0 \tag{8.2}$$

With the aid of Eq. (8.2), it is easy to see that the new scheme shares with the a scheme the same property Eq. (2.19) in [2], i.e.,

$$u_j^{n+1} \to u_j^n \quad \text{and} \quad (u_x^+)_j^{n+1} \to (u_x^+)_j^n \quad \text{as} \quad \Delta t \to 0 \tag{8.3}$$

In the new scheme introduced above, numerical dissipation is controlled by the parameters $\epsilon(\nu)$, $\beta(\nu)$ and α with the first two being the functions of the convection speed a, the mesh interval Δx and the time-step size Δt. In similar extensions involving solvers of more complicated nonlinear equations, the values of these parameters may vary with space and time, and their local values generally will be functions of local values of dynamic variables, mesh intervals and time-step size.

Appendix A. A CE/SE Solver for the Sod's Shock Tube Problem with Non-Reflecting Boundary Conditions

```
      implicit real*8(a-h,o-z)
      dimension  q(3,999), qn(3,999), qx(3,999), qt(3,999),
     *           s(3,999), vxl(3), vxr(3), xx(999)
c
c     nx must be an odd integer.
      nx = 101
      it = 100
      dt = 0.4d-2
      dx = 0.1d-1
      ga = 1.4d0
      rhol = 1.d0
      ul = 0.d0
      pl = 1.d0
      rhor = 0.125d0
      ur = 0.d0
      pr = 0.1d0
      ia = 1
c
      nx1 = nx + 1
      nx2 = nx1/2
      hdt = dt/2.d0
      tt = hdt*dfloat(it)
      qdt = dt/4.d0
      hdx = dx/2.d0
      qdx = dx/4.d0
      dtx = dt/dx
      a1 = ga - 1.d0
      a2 = 3.d0 - ga
      a3 = a2/2.d0
      a4 = 1.5d0*a1
      u2l = rhol*ul
      u3l = pl/a1 + 0.5d0*rhol*ul**2
      u2r = rhor*ur
      u3r = pr/a1 + 0.5d0*rhor*ur**2
      do 5 j = 1,nx2
      q(1,j) = rhol
      q(2,j) = u2l
      q(3,j) = u3l
      q(1,nx2+j) = rhor
      q(2,nx2+j) = u2r
      q(3,nx2+j) = u3r
      do 5 i = 1,3
```

```
      qx(i,j) = 0.d0
      qx(i,nx2+j) = 0.d0
5     continue
c
      open (unit=8,file='for008')
      write (8,10) tt,it,ia,nx
      write (8,20) dt,dx,ga
      write (8,30) rhol,ul,pl
      write (8,40) rhor,ur,pr
c
      do 400 i = 1,it
      m = nx + i - (i/2)*2
      do 100 j = 1,m
      w2 = q(2,j)/q(1,j)
      w3 = q(3,j)/q(1,j)
      f21 = -a3*w2**2
      f22 = a2*w2
      f31 = a1*w2**3 - ga*w2*w3
      f32 = ga*w3 - a4*w2**2
      f33 = ga*w2
      qt(1,j) = -qx(2,j)
      qt(2,j) = -(f21*qx(1,j) + f22*qx(2,j) + a1*qx(3,j))
      qt(3,j) = -(f31*qx(1,j) + f32*qx(2,j) + f33*qx(3,j))
      s(1,j) = qdx*qx(1,j) + dtx*(q(2,j) + qdt*qt(2,j))
      s(2,j) = qdx*qx(2,j) + dtx*(f21*(q(1,j) + qdt*qt(1,j)) +
     *         f22*(q(2,j) + qdt*qt(2,j)) + a1*(q(3,j) + qdt*qt(3,j)))
      s(3,j) = qdx*qx(3,j) + dtx*(f31*(q(1,j) + qdt*qt(1,j)) +
     *         f32*(q(2,j) + qdt*qt(2,j)) + f33*(q(3,j) + qdt*qt(3,j)))
100   continue
      if (i.ne.(i/2)*2) goto 150
      do 120 k = 1,3
      qx(k,nx1) = qx(k,nx)
      qn(k,1) = q(k,1)
      qn(k,nx1) = q(k,nx)
120   continue
150   j1 = 1 - i + (i/2)*2
      mm = m - 1
      do 200 j = 1,mm
      do 200 k = 1,3
      qn(k,j+j1) = 0.5d0*(q(k,j) + q(k,j+1) + s(k,j) - s(k,j+1))
      vxl(k) = (qn(k,j+j1) - q(k,j) - hdt*qt(k,j))/hdx
      vxr(k) = (q(k,j+1) + hdt*qt(k,j+1) - qn(k,j+j1))/hdx
      qx(k,j+j1) = (vxl(k)*(dabs(vxr(k)))**ia + vxr(k)*(dabs(vxl(k)))
     *    **ia)/((dabs(vxl(k)))**ia + (dabs(vxr(k)))**ia + 1.d-60)
200   continue
```

```
      m = nx1 - i + (i/2)*2
      do 300 j = 1,m
      do 300 k = 1,3
      q(k,j) = qn(k,j)
300   continue
400   continue
c
      m = nx1 -it + (it/2)*2
      mm = m - 1
      xx(1) = -0.5d0*dx*dfloat(mm)
      do 500 j = 1,mm
      xx(j+1) = xx(j) + dx
500   continue
      do 600 j = 1,m
      x = q(2,j)/q(1,j)
      y = a1*(q(3,j) - 0.5d0*x**2*q(1,j))
      z = x/dsqrt(ga*y/q(1,j))
      write (8,50) xx(j),q(1,j),x,y,z
600   continue
c
      close (unit=8)
10    format(' t = ',g14.7,' it = ',i4,' ia = ',i4,' nx = ',i4)
20    format(' dt = ',g14.7,' dx = ',g14.7,' gamma = ',g14.7)
30    format(' rhol = ',g14.7,' ul = ',g14.7,' pl = ',g14.7)
40    format(' rhor = ',g14.7,' ur = ',g14.7,' pr = ',g14.7)
50    format(' x =',f8.4,' rho =',f8.4,' u =',f8.4,' p =',f8.4,
     *       ' M =',f8.4)
      stop
      end
```

Appendix B. Proof for Eq. (4.51)

To proceed, first we shall evaluate the flux leaving each of the six quadrilaterals that form the boundary of a CE (see Figs. 10(a) and 11(a)). As a preliminary, note that, in Fig. 10(a),

$$\text{area of } ABGF = \text{area of } CDGB = \text{area of } EFGD = \frac{2wh}{3} \tag{B.1}$$

In Fig. 11(a), we have

$$\text{area of } BCGA = \text{area of } DEGC = \text{area of } FAGE = \frac{2wh}{3} \tag{B.2}$$

Equations (B.1) and (B.2) can be proved easily using the information provided in Fig. 12(a). Moreover, because $u^*(x, y, t; j, k, n)$ is linear in x, y, and t (see Eq. (4.10)), *its average value over any quadrilateral is equal to its value at the geometric center (centroid) of the quadrilateral.* With the above preparations, flux evaluation can be carried out easily using Eqs. (4.6a)–(4.6c), (4.8), (4.10), (B.1), and (B.2).

For each quadrilateral, the result of flux evaluation is a formula involving a_x, a_y, $u^n_{j,k}$, $(u_x)^n_{j,k}$, and $(u_y)^n_{j,k}$. It can be converted to another formula involving ν_ζ, ν_η, $u^n_{j,k}$, $(u^+_\zeta)^n_{j,k}$, and $(u^+_\eta)^n_{j,k}$. To carry out the above conversion, note that Eqs. (4.19), (4.20), (4.22), (4.23), (4.27), and (4.28) imply that

$$\begin{pmatrix} a_x \\ \\ a_y \end{pmatrix} = \frac{2}{3\Delta t} \begin{pmatrix} w-b & w+b \\ \\ -h & h \end{pmatrix} \begin{pmatrix} \nu_\zeta \\ \\ \nu_\eta \end{pmatrix} \tag{B.3}$$

and, for any $(j, k, n) \in \Omega$,

$$\begin{pmatrix} (u_x)^n_{j,k} \\ \\ (u_y)^n_{j,k} \end{pmatrix} = \frac{3}{w} \begin{pmatrix} 1 & 1 \\ \\ -\dfrac{w+b}{h} & \dfrac{w-b}{h} \end{pmatrix} \begin{pmatrix} (u^+_\zeta)^n_{j,k} \\ \\ (u^+_\eta)^n_{j,k} \end{pmatrix} \tag{B.4}$$

Let $(u_x)^n_{j,k}$, $(u_y)^n_{j,k}, \ldots$, be abbreviated as u_x, $u_y, \ldots$, respectively. Then Eqs. (B.3) and (B.4) imply that

$$a_y = \frac{2h}{3\Delta t}\left(\nu_\eta - \nu_\zeta\right) \tag{B.5}$$

$$ha_x + \left(\frac{w}{3} - b\right) a_y = \frac{4wh}{9\Delta t}\left(\nu_\zeta + 2\nu_\eta\right) \tag{B.6}$$

$$ha_x - \left(\frac{w}{3} + b\right) a_y = \frac{4wh}{9\Delta t}\left(2\nu_\zeta + \nu_\eta\right) \tag{B.7}$$

$$u_x = \frac{3}{w}\left(u^+_\zeta + u^+_\eta\right) \tag{B.8}$$

$$u_y = \frac{3}{wh}\left[(w-b)u_\eta^+ - (w+b)u_\zeta^+\right] \tag{B.9}$$

$$\frac{\Delta t}{4}(a_x u_x + a_y u_y) = \nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+ \tag{B.10}$$

$$\left(\frac{b}{2}+\frac{w}{6}\right) u_x + \frac{h}{2} u_y = 2u_\eta^+ - u_\zeta^+ \tag{B.11}$$

and

$$\left(\frac{b}{2}-\frac{w}{6}\right) u_x + \frac{h}{2} u_y = u_\eta^+ - 2u_\zeta^+ \tag{B.12}$$

The conversion referred to above can be carried out using Eqs. (B.5)–(B.12).

Consider Fig. 10(a). The results of flux evaluation involving the quadrilaterals that form the boundaries of $\mathrm{CE}_r(j,k,n)$, $r=1,2,3$, and $(j,k,n)\in\Omega_1$ are given here:

(1) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $G'F'A'B'$ is

$$\frac{2wh}{3}\left(u + u_\zeta^+ + u_\eta^+\right)_{j,k}^n$$

(2) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $G'GFF'$ is

$$-\frac{2wh}{9}\Big(\nu_\zeta + 2\nu_\eta\Big)\left[u + 2u_\zeta^+ - u_\eta^+ + \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\right]_{j,k}^n$$

(3) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $G'B'BG$ is

$$-\frac{2wh}{9}\Big(2\nu_\zeta + \nu_\eta\Big)\left[u - u_\zeta^+ + 2u_\eta^+ + \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\right]_{j,k}^n$$

(4) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $AFGB$ is

$$-\frac{2wh}{3}\left(u - u_\zeta^+ - u_\eta^+\right)_{j+1/3,k+1/3}^{n-1/2}$$

(5) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $ABB'A'$ is

$$\frac{2wh}{9}\Big(\nu_\zeta + 2\nu_\eta\Big)\left[u - 2u_\zeta^+ + u_\eta^+ - \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\right]_{j+1/3,k+1/3}^{n-1/2}$$

(6) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $AA'F'F$ is

$$\frac{2wh}{9}\Big(2\nu_\zeta + \nu_\eta\Big)\left[u + u_\zeta^+ - 2u_\eta^+ - \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\right]_{j+1/3,k+1/3}^{n-1/2}$$

(7) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $G'B'C'D'$ is

$$\frac{2wh}{3}\left(u-2u_\zeta^+ + u_\eta^+\right)_{j,k}^n$$

(8) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $G'GBB'$ is

$$\frac{2wh}{9}\Big(2\nu_\zeta+\nu_\eta\Big)\Big[u-u_\zeta^+ + 2u_\eta^+ + \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j,k}^n$$

(9) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $G'D'DG$ is

$$\frac{2wh}{9}\Big(\nu_\zeta-\nu_\eta\Big)\Big[u-u_\zeta^+ - u_\eta^+ + \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j,k}^n$$

(10) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $CBGD$ is

$$-\frac{2wh}{3}\left(u+2u_\zeta^+ - u_\eta^+\right)_{j-2/3,k+1/3}^{n-1/2}$$

(11) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $CDD'C'$ is

$$-\frac{2wh}{9}\Big(2\nu_\zeta+\nu_\eta\Big)\Big[u+u_\zeta^+ - 2u_\eta^+ - \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j-2/3,k+1/3}^{n-1/2}$$

(12) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $CC'B'B$ is

$$\frac{2wh}{9}\Big(\nu_\eta-\nu_\zeta\Big)\Big[u+u_\zeta^+ + u_\eta^+ - \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j-2/3,k+1/3}^{n-1/2}$$

(13) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $G'D'E'F'$ is

$$\frac{2wh}{3}\left(u+u_\zeta^+ - 2u_\eta^+\right)_{j,k}^n$$

(14) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $G'GDD'$ is

$$\frac{2wh}{9}\Big(\nu_\eta-\nu_\zeta\Big)\Big[u-u_\zeta^+ - u_\eta^+ + \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j,k}^n$$

(15) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $G'F'FG$ is

$$\frac{2wh}{9}\Big(\nu_\zeta+2\nu_\eta\Big)\Big[u+2u_\zeta^+ - u_\eta^+ + \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j,k}^n$$

(16) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $EDGF$ is

$$-\frac{2wh}{3}\left(u-u_\zeta^+ + 2u_\eta^+\right)_{j+1/3,k-2/3}^{n-1/2}$$

(17) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $EFF'E'$ is

$$\frac{2wh}{9}\Big(\nu_\zeta - \nu_\eta\Big)\Big[u + u_\zeta^+ + u_\eta^+ - \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j+1/3,k-2/3}^{n-1/2}$$

(18) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $EE'D'D$ is

$$-\frac{2wh}{9}\Big(\nu_\zeta + 2\nu_\eta\Big)\Big[u - 2u_\zeta^+ + u_\eta^+ - \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j+1/3,k-2/3}^{n-1/2}$$

Consider Fig. 11(a). The results of flux evaluation involving the quadrilaterals that form the boundaries of $\mathrm{CE}_r(j,k,n)$, $r = 1,2,3$, and $(j,k,n) \in \Omega_2$, are given here:

(19) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $G'C'D'E'$ is

$$\frac{2wh}{3}\left(u - u_\zeta^+ - u_\eta^+\right)_{j,k}^{n}$$

(20) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $G'GCC'$ is

$$\frac{2wh}{9}\Big(\nu_\zeta + 2\nu_\eta\Big)\Big[u - 2u_\zeta^+ + u_\eta^+ + \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j,k}^{n}$$

(21) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $G'E'EG$ is

$$\frac{2wh}{9}\Big(2\nu_\zeta + \nu_\eta\Big)\Big[u + u_\zeta^+ - 2u_\eta^+ + \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j,k}^{n}$$

(22) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $DCGE$ is

$$-\frac{2wh}{3}\left(u + u_\zeta^+ + u_\eta^+\right)_{j-1/3,k-1/3}^{n-1/2}$$

(23) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $DEE'D'$ is

$$-\frac{2wh}{9}\Big(\nu_\zeta + 2\nu_\eta\Big)\Big[u + 2u_\zeta^+ - u_\eta^+ - \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j-1/3,k-1/3}^{n-1/2}$$

(24) The flux leaving $\mathrm{CE}_1(j,k,n)$ through $DD'C'C$ is

$$-\frac{2wh}{9}\Big(2\nu_\zeta + \nu_\eta\Big)\Big[u - u_\zeta^+ + 2u_\eta^+ - \Big(\nu_\zeta u_\zeta^+ + \nu_\eta u_\eta^+\Big)\Big]_{j-1/3,k-1/3}^{n-1/2}$$

(25) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $G'E'F'A'$ is

$$\frac{2wh}{3}\left(u + 2u_\zeta^+ - u_\eta^+\right)_{j,k}^{n}$$

(26) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $G'GEE'$ is

$$-\frac{2wh}{9}\Big(2\nu_\zeta+\nu_\eta\Big)\Big[u+u_\zeta^+-2u_\eta^++\Big(\nu_\zeta u_\zeta^++\nu_\eta u_\eta^+\Big)\Big]_{j,k}^{n}$$

(27) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $G'A'AG$ is

$$\frac{2wh}{9}\Big(\nu_\eta-\nu_\zeta\Big)\Big[u+u_\zeta^++u_\eta^++\Big(\nu_\zeta u_\zeta^++\nu_\eta u_\eta^+\Big)\Big]_{j,k}^{n}$$

(28) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $FEGA$ is

$$-\frac{2wh}{3}\left(u-2u_\zeta^++u_\eta^+\right)_{j+2/3,k-1/3}^{n-1/2}$$

(29) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $FAA'F'$ is

$$\frac{2wh}{9}\Big(2\nu_\zeta+\nu_\eta\Big)\Big[u-u_\zeta^++2u_\eta^+-\Big(\nu_\zeta u_\zeta^++\nu_\eta u_\eta^+\Big)\Big]_{j+2/3,k-1/3}^{n-1/2}$$

(30) The flux leaving $\mathrm{CE}_2(j,k,n)$ through $FF'E'E$ is

$$\frac{2wh}{9}\Big(\nu_\zeta-\nu_\eta\Big)\Big[u-u_\zeta^+-u_\eta^+-\Big(\nu_\zeta u_\zeta^++\nu_\eta u_\eta^+\Big)\Big]_{j+2/3,k-1/3}^{n-1/2}$$

(31) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $G'A'B'C'$ is

$$\frac{2wh}{3}\left(u-u_\zeta^++2u_\eta^+\right)_{j,k}^{n}$$

(32) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $G'GAA'$ is

$$\frac{2wh}{9}\Big(\nu_\zeta-\nu_\eta\Big)\Big[u+u_\zeta^++u_\eta^++\Big(\nu_\zeta u_\zeta^++\nu_\eta u_\eta^+\Big)\Big]_{j,k}^{n}$$

(33) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $G'C'CG$ is

$$-\frac{2wh}{9}\Big(\nu_\zeta+2\nu_\eta\Big)\Big[u-2u_\zeta^++u_\eta^++\Big(\nu_\zeta u_\zeta^++\nu_\eta u_\eta^+\Big)\Big]_{j,k}^{n}$$

(34) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $BAGC$ is

$$-\frac{2wh}{3}\left(u+u_\zeta^+-2u_\eta^+\right)_{j-1/3,k+2/3}^{n-1/2}$$

(35) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $BCC'B'$ is

$$\frac{2wh}{9}\Big(\nu_\eta-\nu_\zeta\Big)\Big[u-u_\zeta^+-u_\eta^+-\Big(\nu_\zeta u_\zeta^++\nu_\eta u_\eta^+\Big)\Big]_{j-1/3,k+2/3}^{n-1/2}$$

(36) The flux leaving $\mathrm{CE}_3(j,k,n)$ through $BB'A'A$ is

$$\frac{2wh}{9}\Big(\nu_\zeta+2\nu_\eta\Big)\Big[u+2u_\zeta^+-u_\eta^+-\Big(\nu_\zeta u_\zeta^++\nu_\eta u_\eta^+\Big)\Big]_{j-1/3,k+2/3}^{n-1/2}$$

With the aid of Eqs. (4.29)–(4.46) and (4.49a)–(4.50c), Eq. (4.51) is the result of (1)–(36) and Eq. (4.11). QED.

Appendix C. Proof for Eq. (6.51)

As a preliminary, note that Eqs. (6.18), (6.21), and (6.27) can be used to obtain

$$f^x_{mt} = -\sum_{\ell,q=1}^{4} f^x_{m,\ell}\left(f^x_{\ell,q}u_{qx} + f^y_{\ell,q}u_{qy}\right) \tag{C.1}$$

and

$$f^y_{mt} = -\sum_{\ell,q=1}^{4} f^y_{m,\ell}\left(f^x_{\ell,q}u_{qx} + f^y_{\ell,q}u_{qy}\right) \tag{C.2}$$

In this appendix, we adopt the same convention stated following Eq. (6.32). It follows from Eqs. (6.29)–(6.32) that

$$\begin{pmatrix} f^x_{m,\ell} \\ f^y_{m,\ell} \end{pmatrix} = \frac{2}{3\Delta t}\begin{pmatrix} w-b & w+b \\ -h & h \end{pmatrix}\begin{pmatrix} f^{\zeta+}_{m,\ell} \\ f^{\eta+}_{m,\ell} \end{pmatrix}, \qquad m=1,2,3,4 \tag{C.3}$$

and

$$\begin{pmatrix} u_{mx} \\ u_{my} \end{pmatrix} = \frac{3}{w}\begin{pmatrix} 1 & 1 \\ -\dfrac{w+b}{h} & \dfrac{w-b}{h} \end{pmatrix}\begin{pmatrix} u^+_{m\zeta} \\ u^+_{m\eta} \end{pmatrix}, \qquad m=1,2,3,4 \tag{C.4}$$

An immediate result of Eqs. (C.3) and (C.4) is

$$\sum_{\ell=1}^{4}\left(f^x_{m,\ell}\,u_{\ell x} + f^y_{m,\ell}\,u_{\ell y}\right) = \frac{4}{\Delta t}\sum_{\ell=1}^{4}\left(f^{\zeta+}_{m,\ell}\,u^+_{\ell\zeta} + f^{\eta+}_{m,\ell}\,u^+_{\ell\eta}\right), \qquad m=1,2,3,4 \tag{C.5}$$

By using Eqs. (6.14), (6.16)–(6.21), and (C.1)–(C.5), it can be shown that

$$u_{mx} = \frac{3}{w}\left(u^+_{m\zeta} + u^+_{m\eta}\right) \tag{C.6}$$

$$\left(\frac{b}{2} + \frac{w}{6}\right)u_{mx} + \frac{h}{2}\,u_{my} = 2u^+_{m\eta} - u^+_{m\zeta} \tag{C.7}$$

$$\left(\frac{b}{2} - \frac{w}{6}\right)u_{mx} + \frac{h}{2}\,u_{my} = u^+_{m\eta} - 2u^+_{m\zeta} \tag{C.8}$$

$$hf^x_m + \left(\frac{w}{3} - b\right)f^y_m = \frac{4wh}{9\Delta t}\sum_{\ell=1}^{4}\left(f^{\zeta+}_{m,\ell} + 2f^{\eta+}_{m,\ell}\right)u_\ell \tag{C.9}$$

$$hf^x_m - \left(\frac{w}{3} + b\right)f^y_m = \frac{4wh}{9\Delta t}\sum_{\ell=1}^{4}\left(2f^{\zeta+}_{m,\ell} + f^{\eta+}_{m,\ell}\right)u_\ell \tag{C.10}$$

$$
\begin{aligned}
& h\Big[\Big(\frac{b}{2}+\frac{w}{6}\Big)f_{mx}^{x}+\frac{h}{2}f_{my}^{x}\Big]-\Big(\frac{w}{3}+b\Big)\Big[\Big(\frac{b}{2}+\frac{w}{6}\Big)f_{mx}^{y}+\frac{h}{2}f_{my}^{y}\Big] \\
= \; & \frac{4wh}{9\Delta t}\sum_{\ell=1}^{4}\Big(2f_{m,\ell}^{\zeta+}+f_{m,\ell}^{\eta+}\Big)\Big(2u_{\ell\eta}^{+}-u_{\ell\zeta}^{+}\Big)
\end{aligned}
\qquad \text{(C.11)}
$$

$$
\begin{aligned}
& h\Big[\Big(\frac{b}{2}-\frac{w}{6}\Big)f_{mx}^{x}+\frac{h}{2}f_{my}^{x}\Big]+\Big(\frac{w}{3}-b\Big)\Big[\Big(\frac{b}{2}-\frac{w}{6}\Big)f_{mx}^{y}+\frac{h}{2}f_{my}^{y}\Big] \\
= \; & \frac{4wh}{9\Delta t}\sum_{\ell=1}^{4}\Big(f_{m,\ell}^{\zeta+}+2f_{m,\ell}^{\eta+}\Big)\Big(u_{\ell\eta}^{+}-2u_{\ell\zeta}^{+}\Big)
\end{aligned}
\qquad \text{(C.12)}
$$

$$
hf_{mt}^{x}+\left(\frac{w}{3}-b\right)f_{mt}^{y}=-\frac{16wh}{9(\Delta t)^{2}}\sum_{\ell,q=1}^{4}\left(f_{m,\ell}^{\zeta+}+2f_{m,\ell}^{\eta+}\right)\left(f_{\ell,q}^{\zeta+}u_{q\zeta}^{+}+f_{\ell,q}^{\eta+}u_{q\eta}^{+}\right) \qquad (C.13)
$$

$$
-hf_{mt}^{x}+\left(\frac{w}{3}+b\right)f_{mt}^{y}=\frac{16wh}{9(\Delta t)^{2}}\sum_{\ell,q=1}^{4}\left(2f_{m,\ell}^{\zeta+}+f_{m,\ell}^{\eta+}\right)\left(f_{\ell,q}^{\zeta+}u_{q\zeta}^{+}+f_{\ell,q}^{\eta+}u_{q\eta}^{+}\right) \qquad (C.14)
$$

$$
\begin{aligned}
& f_{m}^{y}\pm\frac{w}{3}f_{mx}^{y}\pm\frac{\Delta t}{4}f_{mt}^{y} \\
= \; & \frac{2h}{3\Delta t}\sum_{\ell=1}^{4}\Big(f_{m,\ell}^{\eta+}-f_{m,\ell}^{\zeta+}\Big)\Big[u_{\ell}\pm u_{\ell\zeta}^{+}\pm u_{\ell\eta}^{+}\mp\sum_{q=1}^{4}\Big(f_{\ell,q}^{\zeta+}u_{q\zeta}^{+}+f_{\ell,q}^{\eta+}u_{q\eta}^{+}\Big)\Big]
\end{aligned}
\qquad \text{(C.15)}
$$

and

$$
\begin{aligned}
& f_{m}^{y}\pm\frac{w}{3}f_{mx}^{y}\mp\frac{\Delta t}{4}f_{mt}^{y} \\
= \; & \frac{2h}{3\Delta t}\sum_{\ell=1}^{4}\Big(f_{m,\ell}^{\eta+}-f_{m,\ell}^{\zeta+}\Big)\Big[u_{\ell}\pm u_{\ell\zeta}^{+}\pm u_{\ell\eta}^{+}\pm\sum_{q=1}^{4}\Big(f_{\ell,q}^{\zeta+}u_{q\zeta}^{+}+f_{\ell,q}^{\eta+}u_{q\eta}^{+}\Big)\Big]
\end{aligned}
\qquad \text{(C.16)}
$$

Note that each of Eqs. (C.15) and (C.16) represents two equations. One corresponds to the upper signs; while the other, to the lower signs.

Next we shall evaluate the flux of $\vec{h}_m^*$ leaving each of the six quadrilaterals that form the boundary of a CE (see Figs. 10(a) and 11(a)). The evaluation procedure is similar to that described in Appendix B. For the current case, the key equations used are Eqs. (4.6a)–(4.6c), (6.15), (6.23)–(6.25), and (C.6)–(C.16). Futhermore, as will be shown shortly, *the structures of the results obtained here are very similar to those given in Appendix B.*

Consider Fig. 10(a). The results of flux evaluation involving the quadrilaterals that form the boundaries of $\mathrm{CE}_r(j,k,n)$, $r=1,2,3$, and $(j,k,n)\in\Omega_1$, are given here:

(1) The flux of $\vec{h}^*_m$ leaving $\text{CE}_1(j,k,n)$ through $G'F'A'B'$ is

$$\frac{2wh}{3}\left(u_m + u^+_{m\zeta} + u^+_{m\eta}\right)^n_{j,k}$$

(2) The flux of $\vec{h}^*_m$ leaving $\text{CE}_1(j,k,n)$ through $G'GFF'$ is

$$-\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f^{\zeta+}_{m,\ell} + 2f^{\eta+}_{m,\ell}\right)\left[u_\ell + 2u^+_{\ell\zeta} - u^+_{\ell\eta} + \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^n_{j,k}$$

(3) The flux of $\vec{h}^*_m$ leaving $\text{CE}_1(j,k,n)$ through $G'B'BG$ is

$$-\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(2f^{\zeta+}_{m,\ell} + f^{\eta+}_{m,\ell}\right)\left[u_\ell - u^+_{\ell\zeta} + 2u^+_{\ell\eta} + \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^n_{j,k}$$

(4) The flux of $\vec{h}^*_m$ leaving $\text{CE}_1(j,k,n)$ through $AFGB$ is

$$-\frac{2wh}{3}\left(u_m - u^+_{m\zeta} - u^+_{m\eta}\right)^{n-1/2}_{j+1/3,k+1/3}$$

(5) The flux of $\vec{h}^*_m$ leaving $\text{CE}_1(j,k,n)$ through $ABB'A'$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f^{\zeta+}_{m,\ell} + 2f^{\eta+}_{m,\ell}\right)\left[u_\ell - 2u^+_{\ell\zeta} + u^+_{\ell\eta} - \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n-1/2}_{j+1/3,k+1/3}$$

(6) The flux of $\vec{h}^*_m$ leaving $\text{CE}_1(j,k,n)$ through $AA'F'F$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(2f^{\zeta+}_{m,\ell} + f^{\eta+}_{m,\ell}\right)\left[u_\ell + u^+_{\ell\zeta} - 2u^+_{\ell\eta} - \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n-1/2}_{j+1/3,k+1/3}$$

(7) The flux of $\vec{h}^*_m$ leaving $\text{CE}_2(j,k,n)$ through $G'B'C'D'$ is

$$\frac{2wh}{3}\left(u_m - 2u^+_{m\zeta} + u^+_{m\eta}\right)^n_{j,k}$$

(8) The flux of $\vec{h}^*_m$ leaving $\text{CE}_2(j,k,n)$ through $G'GBB'$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(2f^{\zeta+}_{m,\ell} + f^{\eta+}_{m,\ell}\right)\left[u_\ell - u^+_{\ell\zeta} + 2u^+_{\ell\eta} + \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^n_{j,k}$$

(9) The flux of $\vec{h}^*_m$ leaving $\text{CE}_2(j,k,n)$ through $G'D'DG$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f^{\zeta+}_{m,\ell} - f^{\eta+}_{m,\ell}\right)\left[u_\ell - u^+_{\ell\zeta} - u^+_{\ell\eta} + \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^n_{j,k}$$

(10) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_2(j,k,n)$ through $CBGD$ is

$$-\frac{2wh}{3}\left(u_m + 2u^+_{m\zeta} - u^+_{m\eta}\right)^{n-1/2}_{j-2/3,k+1/3}$$

(11) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_2(j,k,n)$ through $CDD'C'$ is

$$-\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(2f^{\zeta+}_{m,\ell} + f^{\eta+}_{m,\ell}\right)\left[u_\ell + u^+_{\ell\zeta} - 2u^+_{\ell\eta} - \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n-1/2}_{j-2/3,k+1/3}$$

(12) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_2(j,k,n)$ through $CC'B'B$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f^{\eta+}_{m,\ell} - f^{\zeta+}_{m,\ell}\right)\left[u_\ell + u^+_{\ell\zeta} + u^+_{\ell\eta} - \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n-1/2}_{j-2/3,k+1/3}$$

(13) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_3(j,k,n)$ through $G'D'E'F'$ is

$$\frac{2wh}{3}\left(u_m + u^+_{m\zeta} - 2u^+_{m\eta}\right)^{n}_{j,k}$$

(14) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_3(j,k,n)$ through $G'GDD'$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f^{\eta+}_{m,\ell} - f^{\zeta+}_{m,\ell}\right)\left[u_\ell - u^+_{\ell\zeta} - u^+_{\ell\eta} + \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n}_{j,k}$$

(15) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_3(j,k,n)$ through $G'F'FG$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f^{\zeta+}_{m,\ell} + 2f^{\eta+}_{m,\ell}\right)\left[u_\ell + 2u^+_{\ell\zeta} - u^+_{\ell\eta} + \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n}_{j,k}$$

(16) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_3(j,k,n)$ through $EDGF$ is

$$-\frac{2wh}{3}\left(u_m - u^+_{m\zeta} + 2u^+_{m\eta}\right)^{n-1/2}_{j+1/3,k-2/3}$$

(17) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_3(j,k,n)$ through $EFF'E'$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f^{\zeta+}_{m,\ell} - f^{\eta+}_{m,\ell}\right)\left[u_\ell + u^+_{\ell\zeta} + u^+_{\ell\eta} - \sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta} + f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n-1/2}_{j+1/3,k-2/3}$$

(18) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_3(j,k,n)$ through $EE'D'D$ is

$$-\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f^{\zeta+}_{m,\ell}+2f^{\eta+}_{m,\ell}\right)\left[u_\ell-2u^+_{\ell\zeta}+u^+_{\ell\eta}-\sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta}+f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n-1/2}_{j+1/3,k-2/3}$$

Consider Fig. 11(a). The results of flux evaluation involving the quadrilaterals that form the boundaries of $\mathrm{CE}_r(j,k,n)$, $r=1,2,3$, and $(j,k,n)\in\Omega_2$, are given here:

(19) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_1(j,k,n)$ through $G'C'D'E'$ is

$$\frac{2wh}{3}\left(u_m-u^+_{m\zeta}-u^+_{m\eta}\right)^n_{j,k}$$

(20) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_1(j,k,n)$ through $G'GCC'$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f^{\zeta+}_{m,\ell}+2f^{\eta+}_{m,\ell}\right)\left[u_\ell-2u^+_{\ell\zeta}+u^+_{\ell\eta}+\sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta}+f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n}_{j,k}$$

(21) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_1(j,k,n)$ through $G'E'EG$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(2f^{\zeta+}_{m,\ell}+f^{\eta+}_{m,\ell}\right)\left[u_\ell+u^+_{\ell\zeta}-2u^+_{\ell\eta}+\sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta}+f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n}_{j,k}$$

(22) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_1(j,k,n)$ through $DCGE$ is

$$-\frac{2wh}{3}\left(u_m+u^+_{m\zeta}+u^+_{m\eta}\right)^{n-1/2}_{j-1/3,k-1/3}$$

(23) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_1(j,k,n)$ through $DEE'D'$ is

$$-\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f^{\zeta+}_{m,\ell}+2f^{\eta+}_{m,\ell}\right)\left[u_\ell+2u^+_{\ell\zeta}-u^+_{\ell\eta}-\sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta}+f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n-1/2}_{j-1/3,k-1/3}$$

(24) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_1(j,k,n)$ through $DD'C'C$ is

$$-\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(2f^{\zeta+}_{m,\ell}+f^{\eta+}_{m,\ell}\right)\left[u_\ell-u^+_{\ell\zeta}+2u^+_{\ell\eta}-\sum_{q=1}^{4}\left(f^{\zeta+}_{\ell,q}u^+_{q\zeta}+f^{\eta+}_{\ell,q}u^+_{q\eta}\right)\right]\right\}^{n-1/2}_{j-1/3,k-1/3}$$

(25) The flux of $\vec{h}^*_m$ leaving $\mathrm{CE}_2(j,k,n)$ through $G'E'F'A'$ is

$$\frac{2wh}{3}\left(u_m+2u^+_{m\zeta}-u^+_{m\eta}\right)^n_{j,k}$$

(26) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_2(j,k,n)$ through $G'GEE'$ is

$$-\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(2f_{m,\ell}^{\zeta+}+f_{m,\ell}^{\eta+}\right)\left[u_\ell+u_{\ell\zeta}^{+}-2u_{\ell\eta}^{+}+\sum_{q=1}^{4}\left(f_{\ell,q}^{\zeta+}u_{q\zeta}^{+}+f_{\ell,q}^{\eta+}u_{q\eta}^{+}\right)\right]\right\}_{j,k}^{n}$$

(27) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_2(j,k,n)$ through $G'A'AG$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f_{m,\ell}^{\eta+}-f_{m,\ell}^{\zeta+}\right)\left[u_\ell+u_{\ell\zeta}^{+}+u_{\ell\eta}^{+}+\sum_{q=1}^{4}\left(f_{\ell,q}^{\zeta+}u_{q\zeta}^{+}+f_{\ell,q}^{\eta+}u_{q\eta}^{+}\right)\right]\right\}_{j,k}^{n}$$

(28) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_2(j,k,n)$ through $FEGA$ is

$$-\frac{2wh}{3}\left(u_m-2u_{m\zeta}^{+}+u_{m\eta}^{+}\right)_{j+2/3,k-1/3}^{n-1/2}$$

(29) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_2(j,k,n)$ through $FAA'F'$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(2f_{m,\ell}^{\zeta+}+f_{m,\ell}^{\eta+}\right)\left[u_\ell-u_{\ell\zeta}^{+}+2u_{\ell\eta}^{+}-\sum_{q=1}^{4}\left(f_{\ell,q}^{\zeta+}u_{q\zeta}^{+}+f_{\ell,q}^{\eta+}u_{q\eta}^{+}\right)\right]\right\}_{j+2/3,k-1/3}^{n-1/2}$$

(30) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_2(j,k,n)$ through $FF'E'E$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f_{m,\ell}^{\zeta+}-f_{m,\ell}^{\eta+}\right)\left[u_\ell-u_{\ell\zeta}^{+}-u_{\ell\eta}^{+}-\sum_{q=1}^{4}\left(f_{\ell,q}^{\zeta+}u_{q\zeta}^{+}+f_{\ell,q}^{\eta+}u_{q\eta}^{+}\right)\right]\right\}_{j+2/3,k-1/3}^{n-1/2}$$

(31) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_3(j,k,n)$ through $G'A'B'C'$ is

$$\frac{2wh}{3}\left(u_m-u_{m\zeta}^{+}+2u_{m\eta}^{+}\right)_{j,k}^{n}$$

(32) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_3(j,k,n)$ through $G'GAA'$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f_{m,\ell}^{\zeta+}-f_{m,\ell}^{\eta+}\right)\left[u_\ell+u_{\ell\zeta}^{+}+u_{\ell\eta}^{+}+\sum_{q=1}^{4}\left(f_{\ell,q}^{\zeta+}u_{q\zeta}^{+}+f_{\ell,q}^{\eta+}u_{q\eta}^{+}\right)\right]\right\}_{j,k}^{n}$$

(33) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_3(j,k,n)$ through $G'C'CG$ is

$$-\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f_{m,\ell}^{\zeta+}+2f_{m,\ell}^{\eta+}\right)\left[u_\ell-2u_{\ell\zeta}^{+}+u_{\ell\eta}^{+}+\sum_{q=1}^{4}\left(f_{\ell,q}^{\zeta+}u_{q\zeta}^{+}+f_{\ell,q}^{\eta+}u_{q\eta}^{+}\right)\right]\right\}_{j,k}^{n}$$

(34) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_3(j,k,n)$ through $BAGC$ is

$$-\frac{2wh}{3}\left(u_m + u_{m\zeta}^+ - 2u_{m\eta}^+\right)_{j-1/3,k+2/3}^{n-1/2}$$

(35) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_3(j,k,n)$ through $BCC'B'$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f_{m,\ell}^{\eta+} - f_{m,\ell}^{\zeta+}\right)\left[u_\ell - u_{\ell\zeta}^+ - u_{\ell\eta}^+ - \sum_{q=1}^{4}\left(f_{\ell,q}^{\zeta+}u_{q\zeta}^+ + f_{\ell,q}^{\eta+}u_{q\eta}^+\right)\right]\right\}_{j-1/3,k+2/3}^{n-1/2}$$

(36) The flux of $\vec{h}_m^*$ leaving $\mathrm{CE}_3(j,k,n)$ through $BB'A'A$ is

$$\frac{2wh}{9}\left\{\sum_{\ell=1}^{4}\left(f_{m,\ell}^{\zeta+} + 2f_{m,\ell}^{\eta+}\right)\left[u_\ell + 2u_{\ell\zeta}^+ - u_{\ell\eta}^+ - \sum_{q=1}^{4}\left(f_{\ell,q}^{\zeta+}u_{q\zeta}^+ + f_{\ell,q}^{\eta+}u_{q\eta}^+\right)\right]\right\}_{j-1/3,k+2/3}^{n-1/2}$$

With the aid of Eqs. (6.33)–(6.50) and (4.49a)–(4.50c), Eq (6.51) is the result of (1)–(36) and Eq. (6.28). QED.

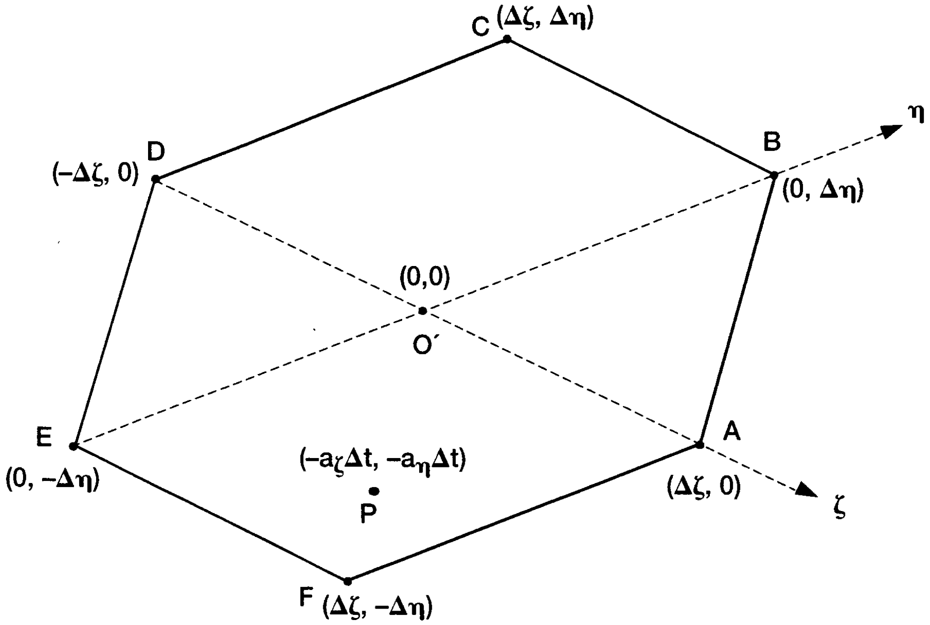


Figure 22: The numerical and analytical domains of dependence associated with the 2D a-scheme.

Appendix D. Supplementary Notes

D.1. A Discussion of Eq. (4.75)

Here we shall prove an assertion made in Sec. 7 about the 2D a scheme, i.e., the backward characteristic projection of a mesh point $(j,k,n) \in \Omega$ at the $(n-1)$th time level is in the interior of the numerical domain of dependence of the same mesh point if and only if Eq. (4.75) is satisfied (see Fig. 22). For simplicity, hereafter the above mesh point will be referred to as point O (not shown). In Fig. 22, *the spatial projection of point O at the $(n-1)$th time level* is represented by point O'; while the backward characteristic projection of point O at the $(n-1)$th time level is represented by point P. Without any loss of generality, we shall assume that $j = k = 0$. Thus (i)

$$\zeta = \eta = 0, \qquad \text{and} \qquad t = n\Delta t \tag{D.1}$$

for point O, and (ii)

$$\zeta = \eta = 0, \qquad \text{and} \qquad t = (n-1)\Delta t \tag{D.2}$$

for point O'.

To simplify the discussion, Eq. (4.1) is converted to an equivalent form in which ζ, η, and t are the independent variables, i.e.,

$$\frac{\partial u}{\partial t} + a_\zeta \frac{\partial u}{\partial \zeta} + a_\eta \frac{\partial u}{\partial \eta} = 0 \tag{D.3}$$

Here a_ζ and a_η are defined in Eq. (4.22). The characteristics of Eq. (D.3) are the family of straight lines defined by

$$\zeta = a_\zeta t + c_1, \qquad \text{and} \qquad \eta = a_\eta t + c_2 \tag{D.4}$$

where c_1 and c_2 are constant along a characteristic, and vary from one characteristic to another. Because points O and P share the same characteristic line, Eqs. (D.1) and (D.4) imply that

$$\zeta = -a_\zeta \Delta t, \qquad \eta = -a_\eta \Delta t, \qquad \text{and} \qquad t = (n-1)\Delta t \tag{D.5}$$

for point P. Note that the temporal coordinate, i.e., $t = (n-1)\Delta t$, of points O' and P are suppressed in Fig. 22.

According to the definition given in Sec. 7, the numerical domain of dependence of point O at the $(n-1)$th time level is the hexagon depicted in Fig. 22. Here the term 'hexagon' refers to both the boundary and the interior. The coordinates (ζ, η) of the vertices A, B, C, D, E, and F are given in the same figure. The six edges of the hexagon and their equations on the ζ-η plane are

$$\begin{array}{ll} \overline{AB}: & \zeta^+ + \eta^+ = 1 \\ \overline{DE}: & \zeta^+ + \eta^+ = -1 \\ \overline{BC}: & \eta^+ = 1 \\ \overline{EF}: & \eta^+ = -1 \\ \overline{CD}: & \zeta^+ = -1 \\ \overline{FA}: & \zeta^+ = 1 \end{array} \tag{D.6}$$

Here the normalized coordinates ζ^+ and η^+ are defined by

$$\zeta^+ \stackrel{def}{=} \zeta/\Delta\zeta, \qquad \text{and} \qquad \eta^+ \stackrel{def}{=} \eta/\Delta\eta \tag{D.7}$$

As a result of Eq. (D.6), a point (ζ, η) is in the interior of the hexagon $ABCDEF$ if and only if

$$|\zeta^+ + \eta^+| < 1, \qquad |\eta^+| < 1, \qquad \text{and} \qquad |\zeta^+| < 1 \tag{D.8}$$

Equations (D.5), (D.7) and (D.8) coupled with Eqs. (4.27) imply point P is in the interior of the hexagon $ABCDEF$ if and only if Eq. (4.75) is satisfied. QED.

D.2. The Local Euler CFL Number

The definition of the local Euler CFL number at the point O (the same point defined in Sec. D.1) is given here.

To proceed, consider Fig. 23. In this figure, point O' and the hexagon $ABCDEF$ are also those defined in Sec. D.1. Let u, v and c be the x-velocity, the y-velocity and the sonic speed at point O, respectively. Let $\vec{e}_x$ and $\vec{e}_y$ be the unit vectors in the x- and the y- directions, respectively. Let $\vec{q}$ denote the velocity vector at point O, i.e.,

$$\vec{q} \stackrel{def}{=} u\vec{e}_x + v\vec{e}_y \tag{D.9}$$

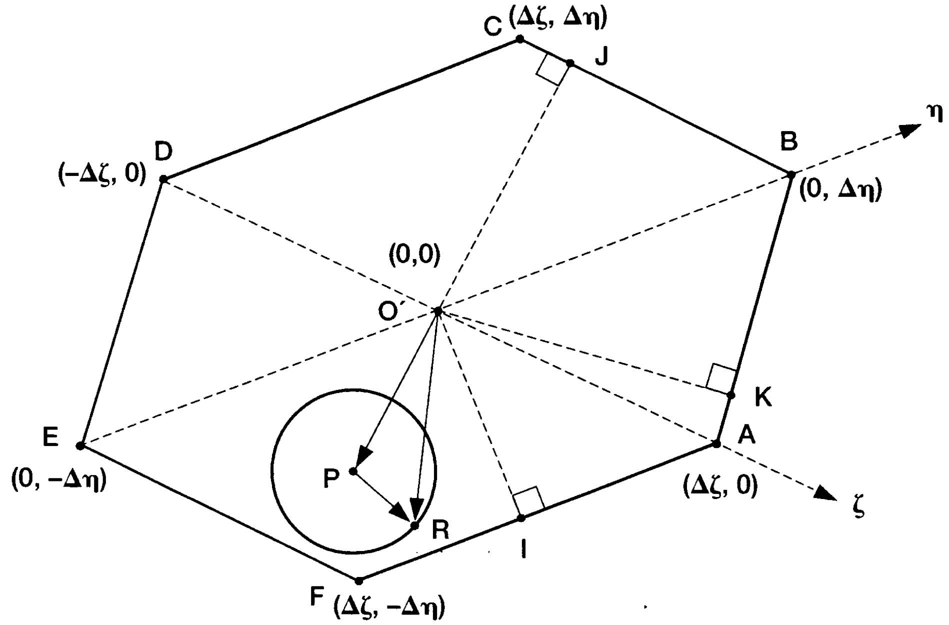


Figure 23: The numerical and analytical domains of dependence associated with the 2D CE/SE Euler solvers.

Let the point P depicted in Fig. 23 be at the $(n-1)$th time level with its spatial position defined by

$$\overrightarrow{O'P} = -\vec{q}\Delta t \tag{D.10}$$

Point P is the center of the circle depicted in Fig. 23. This circle lies at the $(n-1)$th time level and has a radius of $c\Delta t$. Furthermore, it is the intersection of (i) the Mach cone [62, p.425] with point O being its vertex, and (ii) the plane with $t = (n-1)\Delta t$. For the Euler equations Eq. (6.10), and in the limit of $\Delta t \to 0$, this circle is *the domain of dependence of point O at the $(n-1)$th time level.* Here a circle refers to both its circumference and interior. The local Euler CFL number ν_e at point O will be defined such that $\nu_e < 1$ if and only if the domain of dependence of the Euler equations (i.e., the circle) lies in the interior of the numerical domain of dependence (i.e., the hexagon $ABCDEF$). In other words, $\nu_e < 1$ if and only if the normalized coordinates (ζ^+, η^+) of every point on the circumference of the circle satisfy Eq. (D.8).

As a preliminary, let (i) ∂C denote the set of the points on the circumference of the circle defined above, and (ii) S_e denote the set of the unit vectors on the x-y plane. Then, for any point $R \in \partial C$ (see Fig. 23), there exists an $\vec{e} \in S_e$ such that

$$\overrightarrow{PR} = c\Delta t\,\vec{e} \tag{D.11}$$

Combining Eqs. (D.10) and (D.11), one has

$$\overrightarrow{O'R} = (c\vec{e} - \vec{q})\Delta t \tag{D.12}$$

To proceed further, note that Eqs. (4.18), (4.20) and (D.7) imply that

$$\nabla\zeta^{+} = \frac{1}{2w}\,(\vec{e}_x - \frac{w+b}{h}\,\vec{e}_y) \tag{D.13}$$

and

$$\nabla\eta^{+} = \frac{1}{2w}\,(\vec{e}_x + \frac{w-b}{h}\,\vec{e}_y) \tag{D.14}$$

Let (i) $\zeta^+(O')$, $\zeta^+(P)$ and $\zeta^+(R)$ denote the values of ζ^+ at points O', P and R, respectively, and (ii) $\eta^+(O')$, $\eta^+(P)$ and $\eta^+(R)$ denote the values of η^+ at points O', P and R, respectively. Then, because $\zeta^+(O') = \eta^+(O') = 0$ and the gradient vectors given in Eqs. (D.13) and (D.14) are constant, Eqs. (D.10) and (D.12)–(D.14) imply that

$$\zeta^+(P) = -\Delta t\,\vec{q}\cdot\nabla\zeta^+ = -\frac{\Delta t}{2w}\,(u - \frac{w+b}{h}\,v) \tag{D.15}$$

$$\eta^+(P) = -\Delta t\,\vec{q}\cdot\nabla\eta^+ = -\frac{\Delta t}{2w}\,(u + \frac{w-b}{h}\,v) \tag{D.16}$$

$$\zeta^+(R) = \Delta t\,(c\vec{e} - \vec{q})\cdot\nabla\zeta^+ = \zeta^+(P) + c\Delta t\,\vec{e}\cdot\nabla\zeta^+ \tag{D.17}$$

$$\eta^+(R) = \Delta t\,(c\vec{e} - \vec{q})\cdot\nabla\eta^+ = \eta^+(P) + c\Delta t\,\vec{e}\cdot\nabla\eta^+ \tag{D.18}$$

and

$$\zeta^+(R) + \eta^+(R) = \zeta^+(P) + \eta^+(P) + c\Delta t\,\vec{e}\cdot\nabla(\zeta^+ + \eta^+) \tag{D.19}$$

Note that point R is a function of $\vec{e} \in S_e$. In the following, we shall evaluate the maxima and minima of $\zeta^+(R)$, $\eta^+(R)$ and $(\zeta^+(R) + \eta^+(R))$ over the range S_e. To proceed, let

$$\nu_{\pm}^{(1)} \stackrel{def}{=} \left(-\vec{q}\cdot\nabla\zeta^+ \pm c|\nabla\zeta^+|\right)\Delta t \tag{D.20}$$

$$\nu_{\pm}^{(2)} \stackrel{def}{=} \left(-\vec{q}\cdot\nabla\eta^+ \pm c|\nabla\eta^+|\right)\Delta t \tag{D.21}$$

$$\nu_{\pm}^{(3)} \stackrel{def}{=} \left[-\vec{q}\cdot\nabla(\zeta^+ + \eta^+) \pm c|\nabla(\zeta^+ + \eta^+)|\right]\Delta t \tag{D.22}$$

and

$$\vec{e}_1 \stackrel{def}{=} \frac{\nabla\zeta^+}{|\nabla\zeta^+|}, \qquad \vec{e}_2 \stackrel{def}{=} \frac{\nabla\eta^+}{|\nabla\eta^+|}, \qquad \text{and} \qquad \vec{e}_3 \stackrel{def}{=} \frac{\nabla(\zeta^+ + \eta^+)}{|\nabla(\zeta^+ + \eta^+)|} \tag{D.23}$$

With the aid of Eqs. (D.13)–(D.16), (4.14) and (4.15), Eqs. (D.20)–(D.22) imply that

$$\nu_{\pm}^{(1)} = -\frac{\Delta t}{2wh}\,[hu - (w+b)v \mp c\Delta\eta] = \zeta^+(P) \pm \frac{c\Delta t\Delta\eta}{2wh} \tag{D.24}$$

$$\nu_{\pm}^{(2)} = -\frac{\Delta t}{2wh}\,[hu + (w-b)v \mp c\Delta\zeta] = \eta^+(P) \pm \frac{c\Delta t\Delta\zeta}{2wh} \tag{D.25}$$

and

$$\nu_{\pm}^{(3)} = -\frac{\Delta t}{2wh}\,[2hu - 2bv \mp c\Delta\tau] = \zeta^+(P) + \eta^+(P) \pm \frac{c\Delta t\Delta\tau}{2wh} \tag{D.26}$$

where $\Delta\zeta$, $\Delta\eta$ and

$$\Delta\tau \stackrel{def}{=} 2\sqrt{b^2 + h^2} \tag{D.27}$$

respectively, are the lengths of the three sides $\overline{DF}$, $\overline{BD}$, and $\overline{FB}$ of the triangle $\triangle BDF$ depicted in Figs. 12(a)–(c). Furthermore, as a result of Eq. (D.23), (i) $\vec{e}_1$ is normal to any straight line along which ζ^+ is a constant, (ii) $\vec{e}_2$ is normal to any straight line along which η^+ is a constant, and (iii) $\vec{e}_3$ is normal to any straight line along which $\zeta^+ + \eta^+$ is a constant. It follows from the above observations and Eq. (D.6) that $\vec{e}_1$, $\vec{e}_2$ and $\vec{e}_3$, respectively, point in the directions of $\overrightarrow{O'I}$, $\overrightarrow{O'J}$ and $\overrightarrow{O'K}$ (see Fig. 23).

With the aid of Eqs. (D.20)–(D.23), it is easy to conclude from Eqs. (D.17)–(D.19) that:

(a) For all $\vec{e} \in S_e$,

$$\nu_+^{(1)} \geq \zeta^+(R) \geq \nu_-^{(1)} \tag{D.28}$$

with the understanding that the upper bound $\nu_+^{(1)}$ and the lower bound $\nu_-^{(1)}$, respectively, are attained when $\vec{e} = \vec{e}_1$ and $\vec{e} = -\vec{e}_1$.

(b) For all $\vec{e} \in S_e$,

$$\nu_+^{(2)} \geq \eta^+(R) \geq \nu_-^{(2)} \tag{D.29}$$

with the understanding that the upper bound $\nu_+^{(2)}$ and the lower bound $\nu_-^{(2)}$, respectively, are attained when $\vec{e} = \vec{e}_2$ and $\vec{e} = -\vec{e}_2$.

(c) For all $\vec{e} \in S_e$,

$$\nu_+^{(3)} \geq \zeta^+(R) + \eta^+(R) \geq \nu_-^{(3)} \tag{D.30}$$

with the understanding that the upper bound $\nu_+^{(3)}$ and the lower bound $\nu_-^{(3)}$, respectively, are attained when $\vec{e} = \vec{e}_3$ and $\vec{e} = -\vec{e}_3$.

Let

$$\nu^{(\ell)} \stackrel{def}{=} \max\{|\nu_+^{(\ell)}|, |\nu_-^{(\ell)}|\}, \qquad \ell = 1, 2, 3 \tag{D.31}$$

Then Eqs. (D.24)–(D.26) imply that

$$\nu^{(1)} = \frac{\Delta t}{2wh}\left[|hu - (w+b)v| + c\Delta\eta\right] \tag{D.32}$$

$$\nu^{(2)} = \frac{\Delta t}{2wh}\left[|hu + (w-b)v| + c\Delta\zeta\right] \tag{D.33}$$

and

$$\nu^{(3)} = \frac{\Delta t}{2wh}\left[2|hu - bv| + c\Delta\tau\right] \tag{D.34}$$

Let ν_e, the local Euler CFL number at point O, be defined by

$$\nu_e \stackrel{def}{=} \max\{\nu^{(1)}, \nu^{(2)}, \nu^{(3)}\} \tag{D.35}$$

Then the conclusions given in (a)–(c) coupled with Eq. (D.8) imply that the circle depicted in Fig. 23 lies entirely in the interior of the hexagon $ABCDEF$ (i.e., the analytical domain of dependence of point O lies within its numerical domain of dependence) if and only if

$$\nu_e < 1 \tag{D.36}$$

The mesh with $b = 0$ is used in [3]. For this special case, we have

$$\Delta\zeta = \Delta\eta = \sqrt{w^2 + h^2}, \quad \text{and} \quad \Delta\tau = 2h \qquad \text{if} \qquad b = 0 \tag{D.37}$$

As a result, Eqs. (D.32)–(D.35) imply that

$$\nu_e = \max\left\{\frac{(c + |u|)\Delta t}{w}, \frac{\Delta t}{2wh}[h\,|u| + w\,|v| + \sqrt{w^2 + h^2}\,c]\right\} \quad \text{if} \quad b = 0 \tag{D.38}$$

Note that the second component within the parentheses in Eq. (D.38) is a simplified form of the expression given on the extreme right side of Eq. (D.8) in [9]. As a result, ν_e given in Eq. (D.38) is identical to that given in Eq. (D.9) in [9].

D.3. An Existence Theorem

Here we shall prove the following theorem.

Theorem. At any mesh point $(j, k, n) \in \Omega$, existence of

$$\left[\Sigma_{\ell 1}^{(1)+}\right]^{-1} \quad \text{and} \quad \left[\Sigma_{\ell 1}^{(2)+}\right]^{-1}, \qquad \ell = 1, 2, 3$$

is assured if the local CFL number

$$\nu_e < 2/3 \tag{D.39}$$

Proof: As a preliminary, we shall discuss the eigenvalues of the matrix

$$M(k_x, k_y) \stackrel{def}{=} k_x\,F^x + k_y\,F^y \tag{D.40}$$

Here (i) k_x and k_y are arbitrary real numbers, and (ii) F^x and F^y are the matrices formed by $f^x_{m,\ell}$ and $f^y_{m,\ell}$ (see Eq. (6.13)), $m, \ell = 1, 2, 3, 4$, respectively. By using (i) Eqs. (1.1), (1.2), (2.1) and (4.1)–(4.3) in [63], and (ii) the fact that two similar matrices have the same eigenvalues, counting multiplicity [54, p.45], one concludes that the eigenvalues of $M(k_x, k_y)$ are λ_0, λ_0, λ_+ and λ_- with

$$\lambda_0 \stackrel{def}{=} k_x\,u + k_y\,v \tag{D.41}$$

and

$$\lambda_\pm \stackrel{def}{=} \lambda_0 \pm c\sqrt{k_x^2 + k_y^2} \tag{D.42}$$

Note that it is assumed here that the flow variables are evaluated at the mesh point (j, k, n) (i.e., the point O referred to earlier in this appendix).

Because $F^{\zeta+}$ and $F^{\eta+}$, respectively, are the matrices formed by $f^{\zeta+}_{m,\ell}$ and $f^{\eta+}_{m,\ell}$, $m, \ell = 1, 2, 3, 4$, Eqs. (6.29), (6.31) and (4.20) imply that

$$F^{\zeta+} = \frac{3\Delta t}{4w}\left(F^x - \frac{w + b}{h}F^y\right) \tag{D.43}$$

$$F^{\eta+} = \frac{3\Delta t}{4w}\left(F^x + \frac{w - b}{h}F^y\right) \tag{D.44}$$

and

$$F^{\zeta+} + F^{\eta+} = \frac{3\Delta t}{2w}\left(F^x - \frac{b}{h}F^y\right) \tag{D.45}$$

With the aid of Eqs. (4.14), (4.15), (D.27), (D.40)–(D.45), one arrives at the following conclusions:

(a) The eigenvalues of $F^{\zeta+}$ are $\lambda_0^{(1)}$, $\lambda_0^{(1)}$, $\lambda_+^{(1)}$ and $\lambda_-^{(1)}$ where

$$\lambda_0^{(1)} \stackrel{def}{=} \frac{3\Delta t}{4w}\left(u - \frac{w+b}{h}v\right) \tag{D.46}$$

and

$$\lambda_\pm^{(1)} \stackrel{def}{=} \lambda_0^{(1)} \mp \frac{3c\Delta t\Delta\eta}{4wh} \tag{D.47}$$

(b) The eigenvalues of $F^{\eta+}$ are $\lambda_0^{(2)}$, $\lambda_0^{(2)}$, $\lambda_+^{(2)}$ and $\lambda_-^{(2)}$ where

$$\lambda_0^{(2)} \stackrel{def}{=} \frac{3\Delta t}{4w}\left(u + \frac{w-b}{h}v\right) \tag{D.48}$$

and

$$\lambda_\pm^{(2)} \stackrel{def}{=} \lambda_0^{(2)} \mp \frac{3c\Delta t\Delta\zeta}{4wh} \tag{D.49}$$

(c) The eigenvalues of $(F^{\zeta+} + F^{\eta+})$ are $\lambda_0^{(1)} + \lambda_0^{(2)}$, $\lambda_0^{(1)} + \lambda_0^{(2)}$, $\lambda_+^{(3)}$ and $\lambda_-^{(3)}$ where

$$\lambda_\pm^{(3)} \stackrel{def}{=} \lambda_0^{(1)} + \lambda_0^{(2)} \mp \frac{3c\Delta t\Delta\tau}{4wh} \tag{D.50}$$

Let (i) $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of any $n \times n$ matrix A, and (ii) I be the $n \times n$ identity matrix. Then the eigenvalues of the matrix $I - A$ are $1-\lambda_1, 1-\lambda_2, \ldots, 1-\lambda_n$. As a result, Eqs. (6.33), (6.36), (6.39), (6.42), (6.45) and (6.48) coupled with the above results (a)–(c) imply that:

(d) The eigenvalues of $\Sigma_{11}^{(1)+}$ are $1 - \lambda_0^{(1)} - \lambda_0^{(2)}$, $1 - \lambda_0^{(1)} - \lambda_0^{(2)}$, $1 - \lambda_+^{(3)}$ and $1 - \lambda_-^{(3)}$ while the eigenvalues of $\Sigma_{11}^{(2)+}$ are $1 + \lambda_0^{(1)} + \lambda_0^{(2)}$, $1 + \lambda_0^{(1)} + \lambda_0^{(2)}$, $1 + \lambda_+^{(3)}$ and $1 + \lambda_-^{(3)}$.

(e) The eigenvalues of $\Sigma_{21}^{(1)+}$ are $1 + \lambda_0^{(1)}$, $1 + \lambda_0^{(1)}$, $1 + \lambda_+^{(1)}$ and $1 + \lambda_-^{(1)}$, while the eigenvalues of $\Sigma_{21}^{(2)+}$ are $1 - \lambda_0^{(1)}$, $1 - \lambda_0^{(1)}$, $1 - \lambda_+^{(1)}$ and $1 - \lambda_-^{(1)}$.

(f) The eigenvalues of $\Sigma_{31}^{(1)+}$ are $1 + \lambda_0^{(2)}$, $1 + \lambda_0^{(2)}$, $1 + \lambda_+^{(2)}$ and $1 + \lambda_-^{(2)}$, while the eigenvalues of $\Sigma_{31}^{(2)+}$ are $1 - \lambda_0^{(2)}$, $1 - \lambda_0^{(2)}$, $1 - \lambda_+^{(2)}$ and $1 - \lambda_-^{(2)}$.

Note that the matrices referred to in (d)–(f) are nonsingular, and therefore their inverses exist, if the eigenvalues of these matrices are nonzero [54, p.14]. To complete the proof, we need only to show that these eigenvalues are nonzero if $\nu_e < 2/3$.

To proceed, note that, because $c > 0$, it follows from Eqs. (D.24)–(D.26) that

$$\nu_+^{(1)} > \zeta^+(P) > \nu_-^{(1)}, \qquad \text{and} \qquad \nu_+^{(2)} > \eta^+(P) > \nu_-^{(2)} \tag{D.51}$$

and

$$\nu_+^{(3)} > \zeta^+(P) + \eta^+(P) > \nu_-^{(3)} \tag{D.52}$$

With the aid of Eqs. (D.31), (D.35), (D.51) and (D.52), Eq. (D.39), which is equivalent to $(3/2)\nu_e < 1$, implies that

$$\frac{3}{2}\,|\nu_\pm^{(\ell)}| < 1, \qquad \ell = 1,2,3 \tag{D.53}$$

and

$$\frac{3}{2}\,|\zeta^+(P)| < 1, \quad \frac{3}{2}\,|\eta^+(P)| < 1, \quad \text{and} \quad \frac{3}{2}\,|\zeta^+(P) + \eta^+(P)| < 1 \tag{D.54}$$

Next note that Eqs. (D.15), (D.16), (D.24)–(D.26) and Eqs. (D.46)–(D.50) imply that

$$\lambda_\pm^{(\ell)} = -\frac{3}{2}\,\nu_\pm^{(\ell)}, \qquad \ell = 1,2,3 \tag{D.55}$$

and

$$\lambda_0^{(1)} = -\frac{3}{2}\,\zeta^+(P), \quad \lambda_0^{(2)} = -\frac{3}{2}\,\eta^+(P), \quad \text{and} \quad \lambda_0^{(1)} + \lambda_0^{(2)} = -\frac{3}{2}(\zeta^+(P) + \eta^+(P)) \tag{D.56}$$

It now follows from Eqs. (D.53)–(D.56) that each one of the eigenvalues listed in (d)–(f) has the form of $1 \pm x$ with $|x| < 1$ if $\nu_e < 2/3$. Thus these eigenvalues are all positive if $\nu_e < 2/3$. QED.

References

[**1**] S.C. Chang and W.M. To, "A New Numerical Framework for Solving Conservation Laws – The Method of Space-Time Conservation Element and Solution Element," NASA TM 104495, August 1991.

[**2**] S.C. Chang, "The Method of Space-Time Conservation Element and Solution Element – A New Approach for Solving the Navier-Stokes and Euler Equations," *J. Comput. Phys.*, **119**, 1995, pp. 295–324.

[**3**] X.Y. Wang, C.Y. Chow and S.C. Chang, "The Space-Time Conservation Element and Solution Element Method—A New High-Resolution and Genuinely Multidimensional Paradigm for Solving Conservation Laws. II Numerical Simulation of Shock Waves and Contact Discontinuities," NASA TM 208844, December 1998.

[**4**] X.Y. Wang, C.Y. Chow and S.C. Chang, "A Two-Dimensional Euler Solver for Irregular Domain Based on the Space-Time Conservation Element and Solution Element Method," to be published as a NASA TM.

[**5**] S.C. Chang, "On an Origin of Numerical Diffusion: Violation of Invariance Under Space-Time Inversion," *Proceedings of the 23rd Modeling and Simulation Conference*, April 30 - May 1, 1992, Pittsburgh, PA, William G. Vogt and Marlin H. Mickle eds., Part 5, pp. 2727–2738. Also published as NASA TM 105776.

[**6**] S.C. Chang and W.M. To, "A Brief Description of a New Numerical Framework for Solving Conservation Laws – The Method of Space-Time Conservation Element and Solution Element," *Proceedings of the 13th International Conference on Numerical Methods in Fluid Dynamics*, July 6-10, 1992, Rome, Italy, M. Napolitano and F. Sabetta, eds. Also published as NASA TM 105757.

[**7**] S.C. Chang, "New Developments in the Method of Space-Time Conservation Element and Solution Element – Applications to the Euler and Navier-Stokes Equations," Presented at the Second U.S. National Congress on Computational Mechanics, August 16-18, 1993, Washington D.C. Published as NASA TM 106226.

[**8**] X.Y. Wang, C.Y. Chow and S.C. Chang, "Application of the Space-Time Conservation Element and Solution Element Method to Shock-Tube Problem," NASA TM 106806, December 1994.

[**9**] S.C. Chang, X.Y. Wang and C.Y. Chow, "New Developments in the Method of Space-Time Conservation Element and Solution Element – Applications to Two-Dimensional Time-Marching Problems," NASA TM 106758, December 1994.

[**10**] S.C. Chang, X.Y. Wang and C.Y. Chow, "The Method of Space-Time Conservation Element and Solution Element – Applications to One-Dimensional and Two-Dimensional Time-Marching Flow Problems," AIAA Paper 95-1754, in *A Collection of Technical Papers, 12th AIAA CFD Conference*, June 19-22, 1995, San Diego, CA, pp. 1258–1291. Also published as NASA TM 106915.

[11] X.Y. Wang, C.Y. Chow and S.C. Chang, "Application of the Space-Time Conservation Element and Solution Element Method to Two-Dimensional Advection-Diffusion Problems," NASA TM 106946, June 1995.

[12] S.C. Chang, X.Y. Wang, C.Y. Chow and A. Himansu, "The Method of Space-Time Conservation Element and Solution Element – Development of a New Implicit Solver," *Proceedings of the Ninth International Conference on Numerical Methods in Laminar and Turbulent Flow*, C. Taylor and P. Durbetaki eds., Vol. 9, Part 1, pp. 82–93, July 10-14, 1995, Atlanta, GA. Also published as NASA TM 106897.

[13] C.Y. Loh, S.C. Chang, J.R. Scott and S.T. Yu, "Application of the Method of Space-Time Conservation Element and Solution Element to Aeroacoustics Problems," *Proceedings of the 6th International Symposium of CFD*, September 1995, Lake Tahoe, NV.

[14] X.Y. Wang, "Computational Fluid Dynamics Based on the Method of Space-Time Conservation Element and Solution Element," Ph.D. Dissertation, 1995, Department of Aerospace Engineering, University of Colorado, Boulder, CO.

[15] X.Y. Wang, C.Y. Chow and S.C. Chang, "High Resolution Euler Solvers Based on the Space-Time Conservation Element and Solution Element Method," AIAA Paper 96-0764, presented at the 34th AIAA Aerospace Sciences Meeting, January 15-18, 1996, Reno, NV.

[16] C.Y. Loh, S.C. Chang, J.R. Scott and S.T. Yu, "The Space-Time Conservation Element Method – A New Numerical Scheme for Computational Aeroacoustics," AIAA Paper 96-0276, presented at the 34th AIAA Aerospace Sciences Meeting, January 15-18, 1996, Reno, NV.

[17] C.Y. Loh, S.C. Chang and J.R. Scott, "Computational Aeroacoustics via the Space-Time Conservation Element / Solution Element Method," AIAA Paper 96-1687, presented at the 2nd AIAA/CEAS Aeroacoustics Conference, May 6-8, 1996, State College, PA.

[18] X.Y. Wang, C.Y. Chow and S.C. Chang, "Numerical Simulation of Flows Caused by Shock-Body Interaction," AIAA Paper 96-2004, presented at the 27th AIAA Fluid Dynamics Conference, June 17-20, 1996, New Orleans, LA.

[19] X.Y. Wang, C.Y. Chow and S.C. Chang, "An Euler Solver Based on the Method of Space-Time Conservation Element and Solution Element," *Proceedings of the 15th International Conference on Numerical Methods in Fluid Dynamics*, P. Kutler, J. Flores and J.-J. Chattot eds., June 24-28, 1996, Monterey, CA.

[20] S.C. Chang, C.Y. Loh and S.T. Yu, "Computational Aeroacoustics via a New Global Conservation Scheme," *Proceedings of the 15th International Conference on Numerical Methods in Fluid Dynamics*, P. Kutler, J. Flores and J.-J. Chattot eds., June 24-28, 1996, Monterey, CA.

[21] S.T. Yu and S.C. Chang, "Treatments of Stiff Source Terms in Conservation Laws by the Method of Space-Time Conservation Element and Solution Element," AIAA Paper 97-0435, presented at the 35th AIAA Aerospace Sciences Meeting, January 6-10, 1997, Reno, NV.

[22] S.T. Yu and S.C. Chang, "Applications of the Space-Time Conservation Element / Solution Element Method to Unsteady Chemically Reactive Flows," AIAA Paper 97-2099, in *A Collection of Technical Papers, 13th AIAA CFD Conference*, June 29-July 2, 1997, Snowmass, CO.

[23] S.C. Chang, A. Himamsu, C.Y. Loh, X.Y. Wang, S.T. Yu and P. Jorgenson, "Robust and Simple Non-Reflecting Boundary Conditions for the Space-Time Conservation Element and Solution Element Method," AIAA Paper 97-2077, in *A Collection of Technical Papers, 13th AIAA CFD Conference*, June 29-July 2, 1997, Snowmass, CO.

[24] S.C. Chang, S.T. Yu, A. Himansu, X.Y. Wang, C.Y. Chow and C.Y. Loh, "The Method of Space-Time Conservation Element and Solution Element–A New Paradigm for Numerical Solution of Conservation Laws," to appear in *Computational Fluid Dynamics Review 1997*, M.M. Hafez and K. Oshima, eds., John Wiley and Sons, West Sussex, UK.

[25] S.C. Chang and A. Himansu, "The Implicit and Explicit a-μ Schemes," NASA TM 97–206307, November 1997.

[26] X.Y. Wang, C.Y. Chow and S.C. Chang, "Numerical Simulation of Gust Generated Aeroacoustics in a Cascade Using the Space-Time Conservation Element and Solution Element Method," AIAA Paper 98-0178, presented at the 36th AIAA Aerospace Sciences Meeting, January 12-15, 1998, Reno, NV.

[27] X.Y. Wang, C.Y. Chow and S.C. Chang, "Numerical Simulation of shock Reflection over a Dust Layer Model Using the Space-Time Conservation Element and Solution Element Method," AIAA Paper 98-0443, presented at the 36th AIAA Aerospace Sciences Meeting, January 12-15, 1998, Reno, NV.

[28] C.Y. Loh, L.S. Hultgren and S.C. Chang, "Computing Waves in Compressible Flow Using the Space-Time Conservation Element and Solution Element method," AIAA Paper 98-0369, presented at the 36th AIAA Aerospace Sciences Meeting, January 12-15, 1998, Reno, NV.

[29] T. Molls and F. Molls, "Space-Time Conservation Method Applied to Saint Venant Equation," *Journal of Hydraulic Engineering*, **124**, p. 501 (1998).

[30] X.Y. Wang, C.Y. Chow and S.C. Chang, "Non-reflecting Boundary Conditions Based on the Space-Time CE/SE Method for Free Shear Flows," AIAA Paper 98-3020, presented at the 29th AIAA Fluid Dynamics Conference, Albuquerque, New Mexico, June 15-18, 1998.

[**31**] X.Y. Wang, S.C. Chang, K.H. Kao and P. Jorgenson, "A Non-Splitting Unstructured-Triangular-Mesh Euler Solver Based on the Method of Space-Time Conservation Element and Solution Element," To appear in the *Proceedings of the 16th International Conference on Numerical Method in Fluid Dynamics*, Arcachon, France, July 6-July 10, 1998.

[**32**] S.T. Yu, S.C. Chang, P.C.E. Jorgenson, S.J. Park and M.C. Lai, "Treating Stiff Source Terms in Conservation Laws by the Space-Time Conservation Element and Solution Element Method," To appear in the *Proceedings of the 16th International Conference on Numerical Method in Fluid Dynamics*, Arcachon, France, July 6-July 10, 1998.

[**33**] D. Sidilkover, "A genuinely multidimensional upwind scheme and efficient multigrid solver for the compressible Euler equations," ICASE Report No. 94-84, ICASE, 1994.

[**34**] P. Colella and H.M. Glaz, "Efficient Solution Algorithms for the Riemann problem for Real Gases," *J. Comput. Phys.*, **59**, 1985, pp. 264–289.

[**35**] P. Glaister, "An Approximate Linearised Riemann Solver for the Euler Equations for Real Gases," *J. Comput. Phys.*, **74**, 1988, pp. 382–408.

[**36**] Y. Liu and M. Vinokur, "Nonequilibrium Flow Computations. I. An Analysis of Numerical Formulations of Conservation Laws," *J. Comput. Phys.*, **83**, 1989, pp. 373–397.

[**37**] B. Grossman and P. Cinnella, "Flux-Split Algorithms for Flows with Non-equilibrium Chemistry and Vibration Relaxation," *J. Comput. Phys.*, **88**, 1990, pp. 131–168.

[**38**] G.A. Sod, "A Survey of Several Finite Difference Methods for Systems of Nonlinear Hyperbolic Conservation Laws," *J. Comput. Phys.*, **27**, 1 (1978).

[**39**] K.W. Thompson, "Time dependent boundary conditions for hyperbolic systems," *J. Comput. Phys.*, **68**, 1987, pp. 1–24.

[**40**] K.W. Thompson, "Time dependent boundary conditions for hyperbolic systems, II," *J. Comput. Phys.*, **89**, 1990, pp. 439–461.

[**41**] C.K.W. Tam and T.Z. Dong, "Radiation and outflow boundary conditions for direct computation of acoustics and flow disturbance in a nonuniform mean flow," *J. Comput. Acoustics*, **4(2)**, 1996.

[**42**] T.Z. Dong, "On boundary conditions for acoustic computations in non-uniform mean flows," to appear in *J. Comput. Acoustics*.

[**43**] S. Ta'asan and D.M. Nark, "An absorbing buffer zone technique for acoustic wave propagation," AIAA Paper 95-0164, 1995.

[**44**] B. Engquist and A. Majda, "Absorbing boundary conditions for the numerical simulation of waves," *Math. Computation*, **31**, July 1977, pp. 629–651.

[**45**] B.P. Leonard, "Universal Limiter for Transient Interpolation Modeling of the Advective Transport Equations: The ULTIMATE Conservative Difference Scheme," NASA TM 100916, September 1988.

[**46**] R.J. LeVeque, *Numerical Methods for Conservation Laws*, Birkhäuser, Basel, 1990.

[**47**] H.T. Huynh, "Accurate Upwind Methods for the Euler Equations," SIAM J. Numer. Anal., **32**, 5 (1995), pp.1565–1619.

[**48**] K.Y. Choe and K.A. Holsapple, "The discontinuous finite element method with the Taylor-Galerkin aproach for nonlinear hyperbolic conservation laws," *Computer Meth. in Appl. Mech. and Engr.*, **95**, 1992, pp. 141–167.

[**49**] H. Nessyahu and Eitan Tadmor, "Non-Oscillatory Central Differencing for Hyperbolic Conservation Laws," *J. Comput. Phys.*, **87**, 1990, pp. 408–463.

[**50**] R. Sanders and A. Weiser, "A High Resolution Staggered Mesh Approach for Nonlinear Hyperbolic Systems of Conservation Laws," *J. Comput. Phys.*, **101**, 1992, pp. 314–329.

[**51**] H.T. Huynh, "A Piecewise-Parabolic Dual-Mesh Method for the Euler Equations," AIAA paper 95-1739, in *A Collection of Technical Papers, 12th AIAA CFD Conference*, June 19-22, 1995, San Diego, CA, pp. 1054–1066.

[**52**] D.A. Anderson, J.C. Tannehill and R.H. Pletcher, *Computational Fluid Mechanics and Heat Transfer*, Hemisphere Publ. Corp., New York, 1984.

[**53**] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. II, Interscience, 1962.

[**54**] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[**55**] Jon Mathews and R.L. Walker, *Mathematical Methods of Physics*, W.A.Benjamin, Inc., New York, 1965.

[**56**] G. Strang, "On the Construction and Comparison of Difference Schemes," SIAM J. Numer. Anal., **5(3)**, p. 506 (1968).

[**57**] Z. Zhang and M. Shen, "New Approach to Obtain Space-Time Conservation Schemes," *Chinese J. of Aeronautics*, **10(3)**, p. 87 (1997).

[**58**] Z. Zhang and M. Shen, "Improved Scheme of Space-Time Conservation Element and Solution Element," *J. of Tsinghua University (Sci & Tech)*, **37(8)**, p. 65 (1997).

[**59**] Z. Zhang, M. Shen and H. Li, "Modified Space-Time Conservation Schemes for 2D Euler Equations," presented at the 7th International Symposium on Computational Fluid Dynamics, September 1997, Beijing, China.

[**60**] Z. Zhang, "A New General Space-Time Conservation Scheme for 2D Euler Equations," *Chinese J. of Computational Mechanics*, **14(4)**, p. 377 (1997).

[61] Z. Zhang, H. Li and M. Shen, "Space-Time Conservation Scheme for 3D Euler Equations," presented at the 7th International Symposium on Computational Fluid Dynamics, September 1997, Beijing, China.

[62] M.J. Zucrow and J.D. Hoffman, *Gas Dynamics*, Vol. II. John Wiley and Sons, 1977.

[63] R.F. Warming, R.M. Beam and B.J. Hyett, "Diagonalization and Simultaneous Symmetrization of the Gas-Dynamics Matrices," *Mathematics of Computation*, Vol. 29, No. 132, pp. 1037–1045.