To the Graduate Council:

I am submitting herewith a dissertation written by Jing Zhang entitled "Selective Taylor Weak Statement Based, Quadrature-free Finite Element Algorithm for Aerodynamic Flow". I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Engineering Science.

A. J. Baker, Major Professor

We have read this thesis
and recommend its acceptance:

Accepted by the Council:

Associate Vice Chancellor and
Dean of The Graduate School

A SELECTIVE TAYLOR WEAK STATEMENT BASED, QUADRATURE-FREE FINITE ELEMENT ALGORITHM FOR AERODYNAMIC FLOW

A Dissertation Presented for the

Doctor of Philosophy Degree

The University of Tennessee, Knoxville

Jing Zhang

December, 1995

Copyright © Jing Zhang 1995,

ALL RIGHTS RESERVED.

To my parents, Mr. Yan Zhang and Ms. Yuzhen Hou;

张岩先生　侯玉珍女士

to my friend, Dr. Kehui Guo.

郭克辉博士

Jing Zhang of Qingdao, China.

张晶 张家强

ACKNOWLEDGMENT

There are several groups of people who I wish to acknowledge upon the completion of my dissertation. First of all, I want to thank my parents, Mr. Yan Zhang and Ms. Yuzhen Hou, for their love that has never wavered. I also owe tremendous gratitude to my deceased friend, Dr. Guo Kehui, from whom I never ceased learning in our more than fifteen year friendship. If nothing else, I wish to present this dissertation as a gift to them.

Secondly, I am deeply indebted to Dr. Baker, who was my major professor throughout my study. We have had countless discussions, without which this dissertation would be impossible. My acknowledgement to Dr. Baker should also be extended to my fellow members of his Computational Fluid Dynamics (CFD) research group. I have learned a lot from them, including much of the knowledge documented in this dissertation.

I further wish to acknowledge my faculty committee members, Dr. Carley, Dr. Dongarra and Dr. Iannelli, for their advice.

I am also really "in debt" to several departments of The University of Tennessee, Knoxville, who have sponsored my study by offering me teaching or research assistantship. They are: The Department of Engineering Science and Mechanics (ESM),

The Engineering Computing Center (ECC), and the Chromatography Group headed by Dr. Guiochon of The Chemistry Department. What I have gained working with these organizations is not only financial but academic. In this regard, I would like to specifically acknowledge Dr. Guiochon. It is working with him that I learned about the *chemical* shock, as well as the chemists' magic of manipulating it. I have no doubt that I will therefore be better able to deal with the *aerodynamic* shock in the future.

I must not forget Mr. Robert Alexander, who has been a steadfast and loyal friend of mine for eight years now. Always, I would turn to him for advice when my heart/mind were troubled. Several times, it was because of his generosity that I could have overcome financially, and resumed my slow "momentum" in this study. Best to you, my friend.

Finally, I acknowledge Ms. Nancy Lang. She is another tremendous friend, and has spent much time checking and correcting my writing. Thank you, Nancy.

A SELECTIVE TAYLOR WEAK STATEMENT-BASED, QUADRATURE-FREE FINITE ELEMENT ALGORITHM FOR AERODYNAMIC FLOW

Jing Zhang

The University of Tennessee, Knoxville

Abstract

A new finite element computational fluid dynamics (CFD) algorithm has been developed for efficiently solving multi-dimensional compressible flow in general geometry. The algorithm employs a metric data mean (MDM) technique that enables quadrature-free finite element computing, a selective Taylor Weak Statement (S-TWS) procedure that yields diagonalized and positive-definite numerical dissipation for the partial differential equation systems of aerodynamics, and an invariant dissipation length (IDL) scheme that ensures just measurement/bounding of numerical dissipation in mesh. Coded jointly, the two-dimensional compressible potential flow (CPF) and Euler equations have been solved, observing stiff shock/stability. Observed in this process is also the necessity for the "boundary closure" of FE numerical dissipation, in terms of its impact on both the solution accuracy and the matrix-conditioning.

Several lnear algebra solvers have been implementated and tested, including the tensor product (TP) approximate factorization, the line Gauss-Seidel relaxation (LGS), the conjugate gradient square (CGS), and the LGS and TP preconditioned CGS. The tests, which are conducted using rather large Courant numbers (between 100 and 150), demonstrate the LGS preconditioned CGS being the best solver and the "cycled composite LGS" the best line relaxation solver (far superior to TP).

There are also some side-aspects of the documented research, such as the establishment of a single variable convection model of clear yet analyzabe nonlinearity/shock mechanism (serving for numerical verification), the theoretical justification of the positive-definiteness of the numerical dissipation for CPF (being a necessary condition upon the Second Law of Thermodynamics), and the development and application of an incompressible potential flow based procedure for solution initialization in general geometry (by solving the SPD Laplacian).

Contents

1 Introduction **1**

1.1 Overview 1

1.2 Background 6

1.2.1 The Weak Statement Algorithms of CFD 6

1.2.2 Compressible Potential Flow 10

1.2.3 Linear Algebra 12

1.3 Organization and Nomenclature 15

2 Partial Diffential Equations **18**

2.1 Aerodynamic Navier-Stokes Equations 18

2.2 Aerodynamic Euler Equations 24

2.3 Compressible Potential Flow 28

2.4 Convection-Diffusion Problem 30

3 The Metric Data Mean Technique for Finite Element Method **39**

3.1 Problem Statement 39

3.1.1 Galerkin FE-CFD Algorithm 40
3.1.2 The Format of FE Matrices 43
3.1.3 The Problem of Numerical Quadrature 46
3.2 Metric Data Mean Algorithm 48
3.2.1 The Concept Metric Data Mean 48
3.2.2 Termwise Tactics and Metric Polynomial 49
3.2.3 Polynomial Based MDM Algorithm 51
3.2.4 Submeshing Algorithm 53
3.3 Tensor Product Representation 58
3.4 Bilinear Finite Element Matrices 61
3.4.1 Finite Element Definitions 61
3.4.2 Mass and Convection Matrices 62
3.4.3 Diffusion Matrices 63

4 An Invariantly Rescaled, Selective Taylor Weak Statement Algorithm 72

4.1 Taylor Weak Statement Method 73
4.1.1 Perturbation Algorithm 73
4.1.2 Linear Analysis 75
4.1.3 Real Flows 78
4.2 Selective TWS Algorithm 84
4.2.1 Convection Diffusion Problem 84

4.2.2 Compressible Potential Flow 86

4.2.3 Euler Equations 94

4.3 The Invariant Dissipation Length Algorithm 100

4.4 The Invariantly Rescaled Weak Statements 112

5 Finite Element Implementations **115**

5.1 The Nonlinear Convection Problem 115

5.2 Compressible Potential Flow 119

5.3 Euler Equations 126

5.4 Tensor Product Factorization 135

5.4.1 General Description 135

5.4.2 Compressible Potential Flow 137

5.4.3 Euler System 140

5.5 Numerical Linear Algebra Solvers 147

5.5.1 Line Gauss-Seidel Relaxation 148

5.5.2 Gauss-Seidel Iteration 151

5.5.3 Conjugate Gradient Square Algorithm 154

5.5.4 Preconditioned CGS 155

6 Numerical Results **158**

6.1 Introduction 158

6.2 The Nonlinear Convection Problem 159

6.3 Aerodynamic Flows 168

6.3.1 General Statements 168

6.3.2 Numerical Results 174

6.4 Iterative Behavior 190

7 Conclusions **200**

BIBLIOGRAPHY **204**

APPENDICES **221**

A The Derivation of Aerodynamic Navier-Stokes System **222**

A.1 Navier-Stokes Equations 222

A.2 Constitutive Equations 225

B Metric Data Mean Data Base **229**

B.1 Quasi-linear Element 230

B.2 Quasi-quadratic Element 236

B.3 Quasi-cubic Element 251

C The Jacobian Matrices of Aerodynamic Flux-vectors **278**

C.1 General Formulation 278

C.2 Concrete Formulas 282

VITA **284**

List of Figures

2.1 1-D illustration for the nonlinear convection in the case of convex profile: (**a**) the initial symmetric profile; (**b**) the nonlinearly deviated profile a short time later. 32

2.2 1-D illustration for the nonlinear convection in the case of concave profile: (**a**) The initial symmetric profile; (**b**) the nonlinearly deviated profile a short time later. 33

2.3 Mass transfer around a shock for the nonlinear convection problem. The rectangular frame is an infinitesimal stream-pipe control volume. C_1 and C_1 are the components that are the *stationarily* absorbed, and come across at the shock speed v_s; q_1 and q_2 are the components that are linearly convected, and come across at the relative speed $v_s - v_0$. 35

2.4 Illustration for the interaction between shock and characteristics for nonlinear the convection problem. Point **1** — shock ground; point **2** — shock plateau. Shock and characteristics moves respectively at the *secant* and *tangential* speed. Rule of thumb: shock is *overtaken* on the top and *overtaking* at the bottom. 37

3.1 Illustration for the *submeshing* based MDM. Solid quadrilaterals: the *original* mesh cell defined by FE coordinate transform $x^e_{j_\beta}\tilde{N}_\beta(\eta_i)$. Dashed triangles: the *linear submeshes* that approximate $x^e_{j_\beta}\tilde{N}_\beta(\eta_i)$. 55

3.2 Illustration for the optimal *submeshing* (no internal nodes). 56

3.3 Bilinear FE Illustration (**I**). P — the mid-point of diagonal $\overline{13}$; Q — the mid-point of diagonal $\overline{24}$. 65

3.4 Bilinear FE Illustration (**II**): Convex and concave mesh cells. 67

3.5 Submeshes for the *submesh* based MDM alogorithm in bilinear element (two triangles). 71

4.1 Illustration for the supersonic domain of dependence. 88

4.2 Illustration of "incompatibility" between disspation length and mesh size. Mesh: rectangular and sized $\Delta X \times \Delta Y$. Disspation length: $\Delta x_j = \int_{t_1}^{t_2} v_j dt = \bar{v}_j (t_2 - t_1)$, with $t_2 - t_1$ being such that vector is Δx_j is *just* enough to cross the mesh cell. 104

4.3 Illustration for the invariant dissipation length (IDL) in quadrilateral and hexahedron master-elements (citing the rectangle example); mesh size: $\Delta\eta$; IDL: $\Delta\eta_i = |\bar{\tilde{v}}_i| \frac{\Delta t_*}{\beta^*}$, which *just* crosses the mesh cell. IDL algorithm: $\Delta t_* = \beta^* \frac{\Delta\eta}{\max_i |\bar{\tilde{v}}_i|} = \beta^* \frac{\Delta\eta}{|\bar{\tilde{v}}_1|}$. 108

4.4 Illustration for the invariant dissipation length (IDL) in triangle and tetrahedron master-elements (citing the triangle example); mesh size: $\Delta\eta$; IDL: $\Delta\eta_i = |\bar{\tilde{v}}_i| \frac{\Delta t_*}{\beta^*}$, which *just* crosses the mesh cell and defines a point on plane/straight-line $\sum_i \eta_i = \Delta\eta$ if translated to being originated from the origin. 110

5.1 Illustration for the "domain of dependence" of the line Gauss-Seidel algorithm, with • and o being respectively the locations of previous and current. LV mode: relaxation from *left* to right, lined *vertically*; RV mode: relaxation from *right* to left, lined *vertically*; BH mode: relaxation from *bottom* up, lined *horizontally*; TH mode: relaxation from *top* down, lined *horizontally*. 149

5.2 Illustration for the "domain of dependence" of the Gauss-Seidel algorithm, with • and ○ being respectively the locations of previous and current. LBH: solution marching *horizontally* line by line up from the *left/bottom* corner; LBV: solution marching *vertically* line by line to the right from the *left/bottom* corner; RTH: solution marching *horizontally* line by line down from the *right/top* corner; RTV: solution marching *vertically* line by line to the left from the *right/top* corner. 152

6.1 Outlet profiles **I** of the nonlinear convection problem with a square-cap inlet BC. Mesh numbers: A — 1000; B — 900; C — 700; D — 300; E — 5000. 162

6.2 Outlet profiles **II** of the nonlinear convection problem with a square-cap inlet BC. Mesh number: A — 1000; B — 5000. 163

6.3 3-D display for the FE solution of the nonlinear convection problem with a square-cap inlet BC ($N = 300$). 164

6.4 3-D display for the FE solution of the nonlinear convection problem with step function inlet BC ($N = 300$). 167

6.5 Configuration of the arc-bump flow problem: (**a**) geometry; (**b**) general boundary conditions; (**c**) potential boundary conditions. 169

6.6 Mesh for Subsonic flow. 175

6.7 The subsonic solution of the S-TWS FE algorithm for Euler equations. The surface and contour are respectively for pressure and Mach number. 176

6.8 The boundary *unclosed* solution of Euler equations for subsonic flow — an upwinding side-view of the Mach number ("assessed" using the v-form approximate TWS dissipation). 179

6.9 The boundary *closed* solution of Euler equations for subsonic flow — an upwinding side-view of the Mach number ("assessed" using the v-form approximate TWS dissipation). 180

6.10 The convergence history of GS, in boundary unclosed FE solution of Euler equations. (**a**) LBH mode; (**b**) LBV mode; (**c**) RTH mode; (**d**) RTV mode. 181

6.11 The divergence history of GS, in boundary closed FE solution of Euler equations. (**a**) LBH mode; (**b**) LBV mode; (**c**) RTH mode; (**d**) RTV mode. 182

6.12 Refined mesh for Supersonic flow. 184

6.13 3-D display for the Mach number of transonic potential flow, which is the solution using the S-TWS $v_j v_k$-form numerical dissipation (with $M_{in} = 0.675$). 185

6.14 Contour display for the Mach number of the transonic potential flow, which is the solution using the S-TWS $v_j v_k$-form numerical dissipation (with $M_{in} = 0.675$). 186

6.15 3-D display for the transonic solution (Mach number) of Euler equations, which is the solution using the S-TWS $w_j w_k$-form numerical dissipation (with $M_{in} = 0.675$). 188

6.16 Contour display for the transonic solution (Mach number) of Euler euqaiotns, Euler equations, which is the solution using the S-TWS $w_j w_k$-form numerical dissipation (with $M_{in} = 0.675$). 189

6.17 The convergence history of line Gauss-Seidel relaxation. (**a**) LV mode; (**b**) RV mode; (**c**) BH mode; (**d**) TH mode. 191

6.18 The convergence history of the tesor product factorization solver. • $J1 \otimes J2$; ∘ $J2 \otimes J1$. 192

6.19 The convergence history of the "composite" line Gauss-Seidel solvers; □ — the LV-RV pair; ∘ — the BH-TH pair; • — the cycled LV-BH-RV-TH. 194

6.20 The convergence history of the plain conjugate gradient sqaure solver. 195

6.21 The convergence history of the TP preconditioned conjugate gradient square method. Preconditioner: • — $J1 \otimes J2$; ∘ $J2 \otimes J1$ 196

6.22 The convergence history of the "composite" LGS preconditioned conjugate gradient square method. Preconditioners: □ — the LV-RV pair; ∘ — the BH-TH pair; • — the cycled LV-BH-RV-TH. 197

6.23 The convergence history of conjugate gradient square method with the LV-BH-RV-TH cycled LGS as the preconditioner; • — residual; o — preconditioned residual. 199

List of Tables

6.1 The parameters of the nonlinear convection problem. 160

6.2 The convergence of convection in terms of conservation. 165

Chapter 1

Introduction

1.1 Overview

What is documented in this dissertation is a study belonging to the Taylor Weak Statement (TWS) finite element (FE) algorithm research project of the computational fluid dynamics (CFD) group, The University of Tennessee at Knoxville (UTK), which is led by Dr. Baker, and of which this author is a member. Given the founded stability mechanism of TWS-FE algorithm ([14, 15, 17, 18, 21]), the primary goal of this study is to enhance its capability and to improve numerical efficiency. In perspective, the following are the subjects/issues that this author has looked into in promoting such a goal:

1. **Finite Element Quadrature**, an issue in nonlinear FE-CFD. Simply stated, the problem is that there is a necessity of either globally storing or repeatedly quadrating

the FE matrices caused jointly by the nonlinearity of the CFD system and the finite element approximation ([2, 20]). Accordingly, solving the problem means to memorize the FE matrices without consuming a global data base, and in a way that they can be retrieved less costly than being quadrated from scratch ([2, 20]).

2. Approximate TWS Dissipation (A-TWS), which has been a focus topic in real flow TWS research, with the *approximation* being a lever to simplify/reduce the artificial dissipation ([17, 20, 21]). The rationale for A-TWS is simply — why not, since after all the whole TWS procedure is an *approximate* process, and its sole purpose is to stabilize the numerical solution. However, what is interesting to this study in this area is mainly the *difference* between the scalar (v) and tensor ($v_j v_k$) forms of dissipation, which are the two major A-TWS algorithms, developed respectively by Baker and Iannelli ([17, 20]) and Baker and Freels ([21]). Since they are quite *different*, such as the v-form *scalar* scheme is indifferent about *directions* (flow or cross-flow), whereas the *tensor* $v_j v_k$-form correlates the dissipation fluxes by the cross products of velocities, I became convinced that the "gap" in-between should be worth digging into, for better understanding and for the possibility of other options that are numerically meaningful.

3. Compressible Potential Flow. Compressible potential flow (CPF) is a special aerodynamic model which the TWS method was neither invented/designed for nor applied to. That I decided to look into it from a TWS point of view was mainly due to noticing the *fragility* of the empirical dissipation schemes, which are typically

reflected in their *fragile* centralization ([75, 76, 77]). Basically, there is only one serious problem transplanting the TWS algorithm to CPF, i.e., since the CPF equations are second order differential, a syntactically performed TWS procedure leads to a third order "companion" potential system ([17, 20]), and thus requires Hermite C^1 elements in the subsequent Galerkin FE algorithm. This is a problem particularly because the original CPF system permits the Lagrange C^0 elements, which is much less demanding than the Hermite C^1 numerically. Furthermore, the extra computing needed in the Hermite C^1 element is equivalent to including in DOF (degrees of freedom) the first order derivatives of potential, i.e., the velocities, which are all what the CPF model/approximation means to save in the first place.

4. Dissipation Rescaling. By dissipation *rescaling*, I mean the conversion of the TWS dissipation from its original *time* scale to a *length* scale ([17, 20, 21, 24, 25]). The reason for the necessity of doing so is that the time step used is often large, as in most implicit computations, and therefore, there is a need to measure/bound the dissipation on the basis of the spatial mesh size in order to ensure that the numerical solution is not overly dissipated. Symbolically, the general solution for dissipation rescaling is the dimensional condition

$$[\Delta L] \;=\; [V]\,[\Delta t] \tag{1.1}$$

where V is a kind of velocity/speed. The question, however, is how such a symbolic solution should be interpreted — in choosing V, and in associating the subsequently obtained ΔL with *mesh size* in a multi-dimensional flow-field — where even the

definition of *mesh size* is not unambiguous ([17, 20, 21, 24, 25]).

It should be pointed out that dissipation rescaling is not only an issue of TWS. In fact, both Streamline Upwinding Petrov-Galerkin (SUPG, [27, 28, 29, 30, 31, 32, 34]) and Least Squares (LS, [48, 49, 50]) are *temporal* dissipation procedures, although the dissipation in the latter is totally implied and can not be touched in any way other than the adjustment of time step.

□

Regarding the above subjects/issues, this study has resulted in the development of several algorithms:

- A metric data mean (MDM) algorithm;
- A *selective* TWS (S-TWS) dissipation algorithm;
- An invariant dissipation length (IDL) algorithm.

The concept of MDM resembles that of MDO (metric data organization [17, 20], see Section 1.2), i.e., to extract out the metric data in FE integrals, and to admit memory storage of the FE matrices on the master-element basis. The central issue dealt in this developement is the mean value of metric data, which is often unwarranted, and which is why, ultimately, MDM was shaped extracting either *termwise* (polynomial based) or *piecewise* (submesh based) data. Typically speaking, the matrix-computing consumed in an MDM embodied FE-CFD algorithm is about once the metric data per entry, which is less than the consumption at a single quadrature point, being basically comparable with the finite difference (FD) method.

Selective dissipation is a concept that is observed in Navier-Stokes computing — where, for obvious reason, the viscous fluxes are always *selected* out of the numerical dissipation process ([20, 21, 24, 25, 28, 30, 31, 34]). What I call *selective* TWS (S-TWS) algorithm is the same idea, except that the *selection* has been exercised to filter out not only the viscous fluxes but the inviscid terms that are *non-convective*. The concrete results of S-TWS are as follows. For CPF, it turns out to be *just* "scratching off" the aforesaid problem, i.e., the third order TWS artificial potential terms, yielding an positive-definite dissipation scheme. For Euler/Navier-Stokes equations, it is a platform offering insight and choice for approximate TWS (A-TWS) dissipation, which, under certain rules of selection, leads to a "$w_j w_k$-form" dissipation, with w_j being, like the $\bar{v}_k$ of the $v_j v_k$-form A-TWS ([21], see Chapter 4), a variation of the velocity, though distinctively.

As to the invariant *dissipation length* (IDL), theoretically, it is nothing more than such an interpretation of (1.1) that the original of dissipation *time* is a scalar, which, rescaled or not, should neither discriminate nor privilege any direction in the flow field. Mathemtically, the IDL is a length scale defined parallel to the streamlines. It is named as being "invariant" because, in matching it with the mesh size, an invariant transform has been completed to overcome the "ambiguity"and incompatibility between them on the physical domain.

□

In addition, a good portion of this research is devoted to testing linear algebra

solvers, covering tensor product approximate factorization (TP, [21]), conjugate gradient square (CGS, [65, 66]), Gauss-Seidel (GS) and *line* Gauss-Seidel (LGS) relaxation. The initial goal was to compare the factorized TP and the unfactorized CGS, but then extended to including LGS, considering its numerical similarity to TP, and GS, as a conventional solver for behavior reference. What is more, there are also some miscellaneous issues that are of particular interest in the numerical analysis, such as the *boundary closure* of the TWS dissipation ([1, 17, 20]).

1.2 Background

1.2.1 The Weak Statement Algorithms of CFD

Weak statement (WS) based FE methods are equipped with a variety of numerical advantages over FD methods, such as its capability of using arbitrary mesh, its natural Neumann boundary condition, and its usually better accuracy.

However, to properly define WS for fluid dynamics problems is not a staightforward task. The challenge is not only the issue of stability, which requires an *upwind-biased* numerical convection, but that the *upwinding bias* must be implemented in the totally *central* FE approximation. From such a point of view, it is perhaps the foremost important event in CFD that the equivalence between *upwinding* and *dissipation* is understood.

Taylor Weak Statement

The primitive form of TWS algorithm is the Taylor-Galerkin (TG) method, developed by Donea ([14]), who, in the spirit of Lax-Wendroff ([55, 56]), discovered that numerical dissipation can be attained by enforcing the conservation laws to a higher order in the temporal Taylor expansion of the solution. Baker and Kim ([15, 16]) then generalized the idea, by *extremumizing* the semi-discrete Taylor procedure of TG onto the continuous level, which is now known as the TWS "companion conservation laws" ([17, 18, 19, 20]), and by introducing a parametric control mechanism over the dissipation and phase-errors in the system.

Following Baker and Kim's work, TWS FE algorithm research became extended to computing multiple dimensional real flows. Baker and Iannelli studied aerodynamic flows ([17, 18, 19, 20]), presenting a series of techniques including a flow-speed based *approximate* TWS dissipation (A-TWS), which is what I referred to previously as the "v-form" dissipation in Section 1.1. Baker and Freels explored a different approach, the "$v_j v_k$-form" ([21]), using the cross-product of velocities (weighted) as the approximate coefficient of dissipation. Both these algorithms demonstrated to be thoroughly stable under the Tensor Product (TP) factorization. Baker and Noronha ([24, 64]) applied the TWS procedure to the 2-D incompressible Navier-Stokes equations, in which a consistent sparse factorization (CSF) iteration scheme was developed solving the Poisson equation for pressure correction. Lately, Baker and Williams ([25]) implemented TWS FE algorithm for the 3-D incompressible Navier-Stokes, using GMRES ([68]) as the linear algebra solver.

The dissipation *rescaling* of TWS algorithm can be categorized in two types, both connected the *convection* grounded measurement of *dissipation time*. One relates the *dissipation length* to the area or volume of FE mesh cell ([24, 25]). The other takes into account the "components" of the mesh size, thus yielding a symbolically *vectorized* or *tensorized* Courant number ([17, 20, 21]).

Metric data averaging (MDA) is the shared approach tackling the issue of numerical quadrature in all nonlinear TWS FE computations except that of [17, 20], where a metric data optimization (MDO) algorithm was developed. MDA is a common sense idea, and its weakness is the extra error that it induces, which can be severe even for quasi-linear elements, the cases of nonlinear FE where metric data enjoys the least variation (see Chapter 3). MDO seeks to overcome such a weakness by *optimally* evaluating the mean value of metric data.

Streamline Upwind Petrov-Galerkin (SUPG)

The SUPG algorithm was reeported by Hughes et al. ([27, 28, 30, 34]). The basic idea is to upwind displace the center of the test function for the convection terms, which is equivalent to adding a numerical dissipation term. Along with SUPG, Hughes and his colleagues also proposed an entropy-based conservation law approach, which is able to yield a symmetric positive-definite (SPD) matrix.

The SUPG has been widely applied by many practitioners. Habashi et al. implemented a SUPG based full Jacobian implicit algorithm for incompressible Navier-Stokes equations using conjugate gradient (CG) method as the linear algebra solver

([41, 42]). Oden et al. incorporated SUPG with multiple/adaptive mesh methods for compressible flow ([38, 39, 40]).

The *dissipation length* (DL) used in SUPG is the same as that of the "vectorized Courant number" scheme in TWS, except that it is interpreted as the measure of the *dissipation time* (DT) over propagation instead of convection. What reflects this is that the spectrum radius of the flux-Jacobian matrix is taken as the speed/velocity connecting DT and DL (via (1.1), [34]).

Least Squares (LS)

Numerous researchers have been practicing the Least Squares (LS) FE algorithm, an approach which is featured of its symmetric positive definite (SPD) matrix system. Jiang et al. investigated the LS FE algorithms for 2D compressible Euler equations and incompressible Navier-Stokes ([48, 49, 50]). Glowinki, et al., incorporated the LS FE algorithms for both incompressible and compressible Navier-Stokes equations with the domain decomposition strategy, for vector-parallel super-computing ([51, 52]). Bristeau and Periaux ([43, 44, 45, 46, 47]) conducted a comparative research about several FE methods, centralized on LS.

The stability of the LS-FE algorithm relies on its *implied* numerical dissipation, which in fact is identical to that of TG ([50]). Pertinently, there is an interesting dilemma that LS-FE algorithm has to cope with: On the one hand, because of its SPD structure, it is an ideal approach for full-Jacobian implicit algorithm that permits marching in large time step; on the other, since its *implied* dissipation cannot be

"rescaled" onto the spatial mesh, its time step must be small enough in order not to over-dissipate the solution.

Another distinctive feature of LS FE method is that it is a *lesser* weak solution approach than Galerkin, since it permits no reduction of the differentiation order. Consequently, it is inevitable that the C^1 Hermite element has to be used for second order PDE, which, as pointed out in Secttion 1.1, means to suffer extra nodal DOFs. A *trick* minimizing such extra cost for incompressible Navier-Stokes algorithm is to transform the system into first order, by defining vorticity as the chain variable between velocity and shear stress ([48, 49]). And it is quite effective in the case of 2-D flow, adding just one extra DOF — since the vorticity for 2-D flow has only one nonzero component.

1.2.2 Compressible Potential Flow

The earliest stable algorithm for compressible potential flow (CPF) is Murman and Cole's "typed difference" scheme ([81]) for the small perturbation potential equation. The idea was to imitate the *domain of dependence* of flow, by using and successively maintaining a *central* difference scheme in the elliptic subsonic region and an *upwinding* one in the hyperbolic supersonic region. The full potential version of "typed difference" is the so called "rotated difference" scheme accomplished by Jameson ([80]), who considered indefinite curvature of streamlines in full potential flow, and *rotated* the implementation of the concept following the streamlines.

Later on, in order to sustain application for general configurations, the research turned to centralizing the typed difference, by adding to the system a *switched* artificial dissipation (AD), modelling the physical dissipation of the Navier-Stokes equations ([77, 78, 79]). Depending on the emphasis in modelling the Navier-Stokes dissipation, potential AD schemes can be categorized in two groups, artificial *compressiblity* ([77]) and artificial *viscosity* ([78]).

However, it is unclear whether the centralization of CPF algorithm has been vigorously achieved. Per the reported computations to date, there is almost always some sort of *upwinding* allocation needed for the maintenance of stability ([75, 76, 77]). One suggestion for improvement is to go for *full physical consistency* ([73, 74]), which, in brief terms, means to "potentially" *describe* rather than *model* the Navier-Stokes dissipation. Although a *fully physically consistent* dissipation algorithm should have all the reasons to be adequate for ensuring stability, it encounters exactly the same problem as that of a syntatically transplanted TWS, since the potential description of the Navier-Stokes dissipation have to be third order differential.

□

What is also worth noting is the *segregated relaxation* (SR) for compressible potential algorithm. Unanimously, it is applied in the fashion of solving the continuity equation for potential and the energy equation for density ([78, 73, 74, 75]). For steady flow, such a SR should be the optimal choice, since it is able to take advantage of the fact that the continuity equation is a *Laplacian* for potential (with density as

the temporary diffusion coefficient) and the energy equation is *algebraic* for density. Contrastively, to implement it for unsteady CPF, one needs to *twist* the order of physical government in the first place ([73, 74, 75]), hence ends up dealing with a very sophisticated second order (in time) potential system, for which not only must the time-integration scheme be multi-level but only first order accuracy can be handily maintained. However, it is this author's opinion that the whole idea of SR is out of date, given that even the much larger Euler and Navier-Stokes equations are now usually solved simultaneousely as a result of the much improved computing facility and numerical techniques.

1.2.3 Linear Algebra

While a good portion of FE-CFD solution literature was based on explicit time schemes, in which the problem of matrix-inversing can be detoured by lumping the mass matrix, it is the implicit algorithm that possesses the overall better efficiency for allowing larger time step. Therfore, finding good solvers has been another central area of CFD.

Reddy and Reddy ([57]) presented a symmetrical two-step Gauss-Seidel scheme for full matrix inversing, applied with the FE implicit Euler algorithm. The typical feature of this method is that it is implemented on fully unstructured mesh.

Tensor Product (TP) factorization remains as popular an algorithm as the ADI in FD ([88, 89, 90, 92]). Fletcher first implemented it on 2-D rectangular mesh, and

demonstrated it with the implicit Euler method for compressible Navier-Stokes equations ([58]). The algorithm was reported to have better accuracy than the FD ADI algorithm for a sample problem. The solver was then generalized to non-rectangular mesh by Hayes et al. ([59, 60, 61]), who used a "patch" technique for metric data approximation and dummy index contraction. The TWS implementation of Baker and Iannelli can be viewed as a combination of the MDO and the "patch" algorithm, with the former taken as the means for *patching* ([17, 20]).

The appealing character of TP factorization is its temporal "consistency" ([20, 90]), which is why TP algorithms are often the basis of non-iterative CFD solvers. However, the temporal consistency of TP is often not as meaningful as it looks, since the time step used in implicit algorithm is usually not small (e.g., the 100 Courant number of [20]). In comparison, there is another type of factorization of which the consistency is always fully meaningful, the spatially consistent Strongly implicit Procedure (SIP), a method initiated by Stone ([62]) for the five-point FD scheme for the Poisson equation. Schneider et al. ([63]) later modified it to nine diagonal system. The consistent sparse factorization method of Baker, Noronha and Wachspress (CSF, [64]) is a further modification, in which the consistency was pursued to the third order.

Another method comparable to TP is Line Gauss-Seidel (LGS) relaxation, in that it also involves inversing 1D-line-matrices. LGS is disadvantageous to TP for the lack of "consistent" accuracy, and advantageous for not being limited to the TP type

elements.

As all the above are submatrix methods, conjugate gradient (CG) is a full-matrix based solver for SPD system which is favored by numerous authors in LS ([51, 52]) and SUPG ([41, 42]) FE computations. Theoretically, CG is featured being an iterative yet finite-step procedure, which reaches the true solution in at most N steps, where N is the order of the matrix system. As a practical method, the attraction of CG is that it can often be sufficiently converged in far less than N steps. Furthermore, CG has been successfully generalized to non-SPD system in the past few years ([69, 68, 67, 65]). Such generalizations can be grouped into the following two categories:

A. Those that maintained the minimization/orthogonalization property of CG, such as the GCG ([69]) and GMRES ([68]).

B. Those that re-interpreted the CG polynomial — as an error contractor, such as the Bi-CG ([67] and CGS ([65]).

Both of these approaches have gained CFD application (e.g., [25, 66]), but it is unclear which of them performs better. The A-type algorithms are sound but only theoretically, since their successive/orthogonal minimization of residual usually have to be "restarted" ([68, 65]), in order for their increasing expense to be contained. The B-type are advantageous in being *cheap* — consuming comparable per-step work with CG ([67, 65]), but less sound, and depending on whether the condition of matrix is *good* enough for the CG polynomial to be *error contracting*. Seemingly, it is the B-type that is more desirable, considering that the CFD systems, though *bad*, are usually

not difficult to be *preconditioned* close to SPD, which is where the CG polynomial can indeed be safely assumed *error contracting.*

1.3 Organization and Nomenclature

This dissertation is composed of seven chapters. Chapter 1 is an overview about this study, as shown. Chapter 2 introduces the partial differential equations (PDE) that have been encountered. Chapter 3 presents the metric data mean (MDM) algorithm. Chapter 4 presents the *selective* TWS (S-TWS) dissipation algorithm, as well as the invariant dissipation length method. Chapter 5 "checklists" the FE implementation (in detail). Chapter 6 documents the numerical results. The dissertation is concluded in Chapter 7.

There are also three appendices included in this dissertation. Appendix A is a brief yet somewhat specific reference to the aerodynamic Navier-Stokes equations. Appendix B establishes the data base of MDM algorithm for several elements. Appendix C derives the Jacobian matrices of the aerodynamic flux-vectors.

Einstein Convention

Index contraction obeys the Einstein convention in this dissertation, i.e., summation is implied over the repeating indexes of a *term.* Here, by *term*, it is meant to be one that is the product of numbers, variables and/or functions. In association, an

index is *underlined* to be *exempted* from such a convention. For instance,

$$\frac{\partial u_j}{\partial x_j} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \tag{1.2}$$

whereas,

$$\frac{\partial u_j}{\partial x_{\underline{j}}} = \frac{\partial u_{\underline{j}}}{\partial x_j} = \frac{\partial u_1}{\partial x_1} \; or \; \frac{\partial u_2}{\partial x_2} \; or \; \frac{\partial u_3}{\partial x_3} \tag{1.3}$$

FE Matrix Notation

Finite element matrices in this dissertation are denoted according to the nomenclature of Baker for the Galerkin FE-CFD method ([2]). The format of this nomenclature can be summarized as below.

1. Each matrix is represented by a string started with a English letter which is the matrix *identifier*. There are four possible values of a matrix *identifier*: M, A, B and C. An M-matrix is a FE matrix defined on physical domain, which is indifferent about the dimension of system, and could be both global and elemental. The A-B-and-C-matrices are master-element matrices, respectively for 1-2-and-3-D. whereas, .

2. The matrix *identifier* is then followed by a integer, which is the number of FE basis functions in the FE integral, and a few more differentiation *identifiers* for each individual FE function. A *zero* differentiation ID identifies the fact the FE function is *zeroth* order differentiated. A *non-zero* differentiation ID identifies (a) the fact that the FE function is *first* order differentiated, and (b) the coordinate that the differentiation is with respect to.

To illustrate these rules, I cite here the three-index dissipation matrices:

$$A3101_{\alpha\gamma\beta} = \int_{\Omega^*} \frac{dN_\alpha(\xi)}{d\xi} N_\gamma(\xi) \frac{dN_\beta(\xi)}{d\xi} \, d\xi \tag{1.4}$$

$$B3I0J_{\alpha\gamma\beta} = \int_{\Omega^*} \frac{\partial N_\alpha(\eta_1,\eta_2)}{\partial \eta_i} N_\gamma(\eta_1,\eta_2) \frac{\partial N_\beta(\eta_1,\eta_2)}{\partial \eta_j} \, d\eta_1 \, d\eta_2 \tag{1.5}$$

$$C3I0J_{\alpha\gamma\beta} = \int_{\Omega^*} \frac{\partial N_\alpha(\eta_1,\eta_2,\eta_3)}{\partial \eta_i} N_\gamma(\eta_1,\eta_2,\eta_3) \frac{\partial N_\beta(\eta_1,\eta_2,\eta_3)}{\partial \eta_j} \, d\eta_1 \, d\eta_2 \, d\eta_3 \tag{1.6}$$

$$M3I0J^e_{\alpha\gamma\beta} = \int_{\Omega_e} \frac{\partial N_\alpha}{\partial x_i} N_\gamma \frac{\partial N_\beta}{\partial x_j} \, d\Omega \tag{1.7}$$

$$M3I0J_{\alpha\gamma\beta} = \sum_e M3I0J^e_{\alpha\gamma\beta} \tag{1.8}$$

where N_α is the αth FE basis (locally or globally numbered), Ω^e the eth element, and Ω^* the master element. The summation over e in (1.8) is what is usually called the "assembly" of FE matrix.

Note that Einstein summation convention applies to the differentiation ID — in case the latter is an index. For instance,

$$M3J0J_{\alpha\gamma\beta} = M3101_{\alpha\gamma\beta} + M3202_{\alpha\gamma\beta} + \cdots \tag{1.9}$$

Tensor Product Finite Element

A short-hand notation will also be frequently used in referring to the tensor product finite element (FE), the so called $\alpha \Leftrightarrow (i,j,k)$. What it means is nothing more than that there is such a correspondence between α and $(i,\, j,\, k)$ that

$$N_\alpha(\eta_1,\eta_2,\eta_3) = N_i(\eta_1)\, N_j(\eta_2)\, N_k(\eta_3) \tag{1.10}$$

where $N_\alpha(\eta_1,\eta_2,\eta_3)$ is the αth FE basis, and $\{N_l(\varsigma)\}$ are the 1-D shape functions that form it.

Chapter 2

Partial Diffential Equations

This chapter is arranged for describing the partial differential equation (PDE) systems encountered in this dissertation, which include the Navier-Stokes, Euler and potential equations for aerodynamic flow, and a convection/diffusion problem. The aerodynamic PDEs are classic, except being normalized upon a set of somewhat non-standard *reference* states that I found convenient. The nonlinear case of the convection/diffusion system is self-designed, aiming to a clear yet analysable shock/propagation mechanism.

2.1 Aerodynamic Navier-Stokes Equations

The Original Form

As derived in Appendix A, under the assumption that the body force and radiation are neglectable, the Navier-Stokes equations for perfect gas flow with constant heat

capacity are as follows.

$$\frac{\partial U}{\partial t} + \frac{\partial}{\partial x_j}\left(F_j - F_j^{\mu}\right) = 0 \tag{2.1}$$

$$U = \{\rho,\, m_i,\, E\}^T \tag{2.2}$$

$$F_j = \{m_j,\, m_i v_j + p\delta_{ij},\, (E+p)\, v_j\}^T \tag{2.3}$$

$$F_j^{\mu} = \{0,\, \tau_{ji},\, \tau_{ji} v_i + \lambda \frac{\partial T}{\partial x_j}\}^T \tag{2.4}$$

$$E = \frac{1}{2}\rho v_i^2 + \frac{p}{\gamma - 1} \tag{2.5}$$

$$\tau_{ij} = \mu\left[\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) - \frac{2}{3}\frac{\partial v_j}{\partial x_j}\delta_{ij}\right] \tag{2.6}$$

$$p = \rho R T \tag{2.7}$$

where ρ, p, T and R are respectively density, pressure, temperature and gas constant; v_i is the ith component of velocity and $m_i = \rho v_i$ the corresponding momentum; λ and μ are the thermal conductivity and absolute viscosity. Equation (2.1) is the collective form of $\{L(\rho),\, L(m_i), L(E)\}^T$ in (A.15-A.16-A.17) as defined in Appendix A, except that the total energy E has been re-expressed considering (2.7) or (A.22), which is the the equation of state for perfect gas, and (A.23), which is the internal energy formula of perfect gas as heat capacities are constant.

Equations (2.1-2.7) are often referred to as "conservation laws", for that they are *just* in the form revealing the fact that mass, momentum and energy are constantly maintained balancing the nonequilibrium of the corresponding *fluxes*. Identically, they can be written *non-conservatively* as ([9]):

$$L(\rho): \quad \frac{D\rho}{Dt} + \rho\frac{\partial v_j}{\partial x_j} = 0 \tag{2.8}$$

$$L(v_i): \quad \rho\frac{Dv_i}{Dt} + \frac{\partial p}{\partial x_i} - \frac{\partial \tau_{ji}}{\partial x_j} = 0 \tag{2.9}$$

$$L(e): \quad \rho\frac{De}{Dt} + p\frac{\partial v_j}{\partial x_j} - \tau_{ij}\epsilon_{ij} - \frac{\partial}{\partial x_j}\left(\lambda\frac{\partial T}{\partial x_j}\right) = 0 \tag{2.10}$$

in which $\frac{D}{Dt}$ is the same Lagrangian time differentiation as in (A.1-A.2-A.3), except that here it is applied *locally* following the motion of "fluid-particles", in other words,

$$\begin{aligned} \frac{D}{Dt} &= \frac{\partial}{\partial t} + \frac{\partial x_j}{\partial t}\frac{\partial}{\partial x_j} \\ &= \frac{\partial}{\partial t} + v_j\frac{\partial}{\partial x_j} \end{aligned} \tag{2.11}$$

Continuing, $\tau_{ij}\epsilon_{ij}$ and $-p\frac{\partial v_j}{\partial x_j}$ in (2.10) are respectively the *deformative* and *volumetric* work done by stresses. Both are transfered to the *internal* energy of flow, but in different manner: the *deformative* work is "positive definite" ([9, 10]), hence, always a *heating* source in the flow-field; the *volumetric* work, on the orther hand, may both *heat-up* and *cool-down* the flow, depending on the fluid being compressed or dilated. Such a difference can be better observed if one applies the Gibbs equality

$$\begin{aligned} T\,dS &= de + p\,d(\frac{1}{\rho}) \\ &= dh - \frac{1}{\rho}dp \end{aligned} \tag{2.12}$$

which leads to (2.10) in its entropy form:

$$L(S): \ \rho T\frac{DS}{Dt} = \tau_{ij}\epsilon_{ij} + \frac{\partial}{\partial x_j}\lambda\frac{\partial T}{\partial x_j} \tag{2.13}$$

As shown, the *volumetric* work has nothing to do with the entropy increase, and its *heating/cooling*, thus, are nothing more than the different sides of a *reversible* process.

In contrast, the *deformative* work is contributed to the entropy growth in full, hence, forever dissipative.

The Dimensionless Form

Let

t_0 $\equiv$ the reference time,

L_0 $\equiv$ the reference length,

v_0 $\equiv$ the reference speed,

ρ_0 $\equiv$ the reference density,

p_0 $\equiv$ the reference pressure,

T_0 $\equiv$ the reference temperature;

μ_0 $\equiv$ the reference viscosity;

λ_0 $\equiv$ the reference heat conductivity;

and let these reference states be constrained by

$$t_0 = \frac{L_0}{v_0} \tag{2.14}$$

$$T_0 = \frac{p_0}{R\rho_0} \tag{2.15}$$

$$p_0 = \rho_0 v_0^2 \tag{2.16}$$

Then, by taking

$$L^*(\rho^*) \equiv \frac{L_0}{\rho_0 v_0} L(\rho) \tag{2.17}$$

$$L^*(m_i^*) \equiv \frac{L_0}{\rho_0 v_0^2} L(m_i) \tag{2.18}$$

$$L^*(E^*) \equiv \frac{L_0}{\rho_0 v_0^3} L(E) \tag{2.19}$$

one can normalize (2.1-2.7) into the following dimensionless form.

$$\frac{\partial U^*}{\partial t^*} + \frac{\partial}{\partial x_j^*}\left(F_j^* - F_j^{\mu^*}\right) = 0 \tag{2.20}$$

$$U^* = \{\rho^*, m_i^*, E^*\}^T \tag{2.21}$$

$$F_j^* = \{m_j^*, m_i^* v_j^* + p^* \delta_{ij}, (E^* + p^*) v_j^*\}^T \tag{2.22}$$

$$F_j^{\mu^*} = \{0, \tau_{ji}^*, \tau_{ji}^* v_i^* + \frac{\lambda^*}{R_e P_r} \frac{\gamma}{\gamma - 1} \frac{\partial T^*}{\partial x_j^*}\}^T \tag{2.23}$$

$$E^* = \frac{1}{2}\rho^* v_i^{*2} + \frac{p^*}{\gamma - 1} \tag{2.24}$$

$$\tau_{ij}^* = \frac{\mu^*}{R_e}\left[\left(\frac{\partial v_i^*}{\partial x_j^*} + \frac{\partial v_j^*}{\partial x_i^*}\right) - \frac{2}{3}\frac{\partial v_j^*}{\partial x_j^*}\delta_{ij}\right] \tag{2.25}$$

$$p^* = \rho^* T^* \tag{2.26}$$

where $R_e \equiv \frac{\rho_0 v_0 L_0}{\mu_0}$ and $P_r \equiv \frac{\mu_0 C_p}{\lambda_0}$ are respectively the Reynolds and Prandtl numbers, and shoulder $*$ is the dimensioinless marks, which is defined as

$$(\cdot)^* = \frac{(\cdot)}{(\cdot)_0} \tag{2.27}$$

Note that the Strouhal and Euler numbers ([9, 10]) have been *missing* in (2.20), and so does the gas constant R in (2.26), which are being caused as a result of the extra constraints (2.14-2.15-2.16).

□

Corresponding to (2.20), the *non-conservative* version of Navier-Stokes equations (2.8-2.9-2.10) become

$$L^*(\rho^*): \quad \frac{D\rho^*}{Dt^*} + \rho^* \frac{\partial v_j^*}{\partial x_j^*} = 0 \tag{2.28}$$

$$L^*(v_i^*): \quad \rho^* \frac{Dv_i^*}{Dt^*} + \frac{\partial p^*}{\partial x_i^*} - \frac{\partial \tau_{ji}^*}{\partial x_j^*} = 0 \tag{2.29}$$

$$L^*(e^*): \quad \rho^* \frac{De^*}{Dt^*} + p^* \frac{\partial v_j^*}{\partial x_j^*} - \tau_{ij}^* \epsilon_{ij}^* - \frac{\partial}{\partial x_j^*} \left(\frac{\lambda^*}{R_e P_r} \frac{\gamma}{\gamma - 1} \frac{\partial T^*}{\partial x_j^*} \right) = 0 \quad (2.30)$$

which can be obtained by either re-writing (2.20) or directly normalizing (2.8-2.9-2.10).

A Supplementary Note

Entropy and heat capacities have not been included in the above normalization. However, it will be shown in this segment that they can all be *supplemented* by requiring that the form of the Gibbs equality remain, i.e.,

$$\begin{aligned} T^* \, dS^* &= C_v^* \, dT^* + p^* \, d(\frac{1}{\rho^*}) \\ &= C_p^* \, dT^* - \frac{dp^*}{\rho^*} \end{aligned} \quad (2.31)$$

In fact, according to (2.27), to the original Gibbs equality (2.12) reads

$$\begin{aligned} T_0 \, S_0 \, T^* \, dS^* &= T_0 \, C_{p_0} \, C_p^* \, dT^* - \frac{p_0}{\rho_0} \frac{dp^*}{\rho^*} \\ &= T_0 \, C_{v_0} \, C_v^* \, dT^* + \frac{p_0}{\rho_0} p^* \, d(\frac{1}{\rho^*}) \end{aligned} \quad (2.32)$$

where S_0, C_{p_0} and C_{v_0} denote respectively the *scales* for entropy and heat capacities to be determined. Dividing both sides by $T_0 S_0$ and considering (2.15), one can reform (2.32) into

$$\begin{aligned} T^* \, dS^* &= \frac{C_{p_0}}{S_0} C_p^* \, dT^* - \frac{R}{S_0} \frac{dp^*}{\rho^*} \\ &= \frac{C_{v_0}}{S_0} C_v^* \, dT^* + \frac{R}{S_0} p^* \, d(\frac{1}{\rho^*}) \end{aligned} \quad (2.33)$$

Consequently, (2.31) is true if and only if

$$S_0 = C_{p_0} = C_{v_0} = R \quad (2.34)$$

□

Subsequent to (2.34),

$$C_v^* = \frac{1}{\gamma - 1} \tag{2.35}$$

$$C_p^* = \frac{\gamma}{\gamma - 1} \tag{2.36}$$

in other words, like the equation of state (2.26), a (2.31) supplemented normalization of heat capacities abide by the "missing R" rule.

Since the form of Gibbs equality remains, one can immediately draw from (2.30) that

$$L^*(S^*): \ \rho^* T^* \frac{DS^*}{Dt^*} = \tau_{ij}^* \epsilon_{ij}^* + \frac{\partial}{\partial x_j^*}\left(\frac{\lambda^*}{R_e P_r}\frac{\gamma}{\gamma - 1}\frac{\partial T^*}{\partial x_j^*}\right) \tag{2.37}$$

which is the dimensionless form of (2.13).

□

The dimensionless mark "$*$"will be omitted in the rest of this dissertation, though all flow variables will be assumed dimensionless.

2.2 Aerodynamic Euler Equations

The Euler equations are the model governing *inviscid* flow, the case of Navier-Stokes equations as as μ and λ vanish, i.e.,

$$\frac{\partial U}{\partial t} + \frac{\partial F_j}{\partial x_j} = 0 \tag{2.38}$$

$$U = \{\rho, m_i, E\}^T$$

$$F_j = \{m_j, m_i v_j + p\delta_{ij}, (E + p)v_j\}^T$$

$$E = \frac{1}{2}\rho v_i^2 + \frac{p}{\gamma - 1}$$

or in *non-conservative* form,

$$L(\rho): \quad \frac{D\rho}{Dt} + \rho\frac{\partial v_j}{\partial x_j} = 0 \tag{2.39}$$

$$L(v_i): \quad \rho\frac{Dv_i}{Dt} + \frac{\partial p}{\partial x_i} = 0 \tag{2.40}$$

$$L(e): \quad \rho\frac{De}{Dt} + p\frac{\partial v_j}{\partial x_j} = 0 \tag{2.41}$$

Physically, the Euler equations, as shown in (2.20-2.26), represent a situation where the viscous flux-vector F_j^μ is negligible compared with the inviscid fluxt-vector F_j, which is the asymptotic state of real flow as Reynolds number tends to infinity.

□

In association with the absence of viscosity, the Euler equations are entitled to many properties that do not exist in the Navier-Stokes system. A typical example is the *shock*. Shock is possible in inviscid flow because, short of viscosity, there is no mechanism damping *discontinuity*, as is discussed in Appendix A. Of course, a shock is not totally equivalent to the *discontinuity* in mathematic sense. In fact, a *discontinuity* may qualify to being a shock only if it is *compressive* ([9, 10]).

Furthermore, shock is the sole entropy source in inviscid flow, since, unless shock is encountered, inviscid flow-field *is* differentiable, and/hence according to (2.13),

$$\frac{DS}{Dt} = 0 \tag{2.42}$$

□

Using formula ([9, 10])

$$(\mathbf{v}\cdot\nabla)\mathbf{v} \;=\; \nabla\left(\frac{v^2}{2}\right) - \mathbf{v}\times\boldsymbol{\Omega} \tag{2.43}$$

where $\boldsymbol{\Omega}$ is the curl of the velocity, one may re-write momentum equation (2.40) as

$$\frac{\partial\mathbf{v}}{\partial t} + \nabla\left(\frac{v^2}{2}\right) - \mathbf{v}\times\boldsymbol{\Omega} + \frac{\nabla p}{\rho} \;=\; 0 \tag{2.44}$$

Numerous properties of inviscid flow can be derived subsequently ([9, 10]), including the following related to compressible potential flow (CPF).

- **Coroco Theorem**: For steady inviscid flow,

$$\mathbf{v}\times\boldsymbol{\Omega} \;=\; \nabla h_0 - T\nabla S \tag{2.45}$$

  where h_0 is the total enthalpy.

- **Cauchy-Lagrange integration**: For potential inviscid flow,

$$\int_*(2.44)_j\,dx_j: \quad \frac{\partial\phi}{\partial t} + \frac{1}{2}\frac{\partial\phi}{\partial x_j}\frac{\partial\phi}{\partial x_j} + \int_*\frac{dp}{\rho} \;=\; \left(\frac{\partial\phi}{\partial t} + \frac{1}{2}v_j^2\right)_* \tag{2.46}$$

  where $(\cdot)_*$ denotes the states at a reference point which could be anywhere in the flow-field, and ϕ is the velocity potential, defined as

$$v_i \;=\; \frac{\partial\phi}{\partial x_i} \tag{2.47}$$

□

Coroco theorem *locally* specifies the relation between the vorticity and the entropy production in steady inviscid flow, — considering that the distribution of h_0, as

is warranted by the Benooli theorem ([9, 10]), is inherited from the upstream far-field. In comparison, under the condition that $\frac{dp}{\rho}$ is a total differential, there is the Kelvin theorem ([8, 9]) that governs the unsteady transportation of vorticity from a *Lagrangian* perspective:

$$\begin{aligned} \frac{D}{Dt}\oint_\Gamma v_i\,dx_i &\equiv \frac{D}{Dt}\int_A \Omega_i\,n_i\,dA \\ &\equiv -\oint_\Gamma \frac{1}{\rho}\,dp \\ &= 0 \end{aligned} \tag{2.48}$$

where Γ is a closed path of "fluid particles", A a surface bordered by Γ, and n_i the normal unit vector on A, which is outward on the side that is "kept left" to circulation $\oint_\Gamma v_i\,dx_i$ ([9]).

By definition and according to the Gibbs equality, the total differentiality of $\frac{dp}{\rho}$ condition is equivalent to that

$$\oint_\Gamma T\,dS \equiv 0 \tag{2.49}$$

is true and independent of the path Γ. As a result, Kelvin theorem is not granted across a shock, since usually $\oint_\Gamma T\,dS \neq 0$ as soon as there is a shock cutting through Γ. However, this is not a profound weakness, and it can be simply released by allowing the remains of $\oint_\Gamma \frac{1}{\rho}\,dp$:

$$\frac{D}{Dt}\oint_\Gamma v_i\,dx_i \equiv \frac{D}{Dt}\oint_A \Omega_i\,n_i\,dA = \oint_\Gamma T\,dS \tag{2.50}$$

2.3 Compressible Potential Flow

If the compressible inviscid flow is *potential*, i.e., isentropic and irrotational, the conservation laws can then be further simplified to ([9]):

$$L(\rho): \quad \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho \frac{\partial \phi}{\partial x_j}\right) = 0 \tag{2.51}$$

$$L(\phi): \quad \frac{\partial \phi}{\partial t} + \frac{1}{2}\frac{\partial \phi}{\partial x_j}\frac{\partial \phi}{\partial x_j} + \frac{\gamma}{\gamma - 1}\frac{T_*}{\rho_*^{\gamma-1}}\rho^{\gamma-1} = \left(\frac{\partial \phi}{\partial t} + \frac{1}{2}v_j^2 + \frac{\gamma}{\gamma-1}T\right)_* \tag{2.52}$$

Here, equation (2.51) is the same continuity condition as in the Euler/Navier-Stokes system, though *potentially* described via (2.47). Further, equation (2.52) is identical to the Cauchy-Lagrange integration (2.46), except that $\int_* \frac{dp}{\rho}$ has been isentropically specified, i.e., letting

$$\begin{aligned} \int_* \frac{dp}{\rho} &= \frac{\gamma}{\gamma - 1}(T - T_*) \\ &= \frac{\gamma}{\gamma - 1}\left(\frac{T_*}{\rho_*^{\gamma-1}}\rho^{\gamma-1} - T_*\right) \end{aligned} \tag{2.53}$$

□

Equations (2.51-2.52) are the general model for compressible potential flow (CPF). In this dissertation, consideration is limited to the CPF with a steady far-field (at least in the upstream). As a result, by setting the reference point $(\cdot)_*$ of the Cauchy-Lagrange integration in the far-field, denoted hereafter as $(\cdot)_\infty$, one can discard the unsteady potential term on the right hand side (RHS) of (2.52). Consequently, the

eventual form of the CPF equations I will be using is

$$L(\rho): \quad \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho \frac{\partial \phi}{\partial x_j}\right) = 0 \tag{2.54}$$

$$L(\phi): \quad \frac{\partial \phi}{\partial t} + \frac{1}{2}\frac{\partial \phi}{\partial x_j}\frac{\partial \phi}{\partial x_j} + \frac{\gamma}{\gamma - 1}\frac{T_\infty}{\rho_\infty^{\gamma-1}}\rho^{\gamma-1} = \frac{\gamma}{\gamma - 1}T_{0\infty} \tag{2.55}$$

where T_0 is the total temperature.

□

A fundamental question about CPF is why/how it occurs and can be maintained. And the answer can be summarized as follows, as being provided by the Kelvin and Coroco theorems, and the entropy convection equation (2.42).

i. The necessary condition for inviscid flow to be *potential* is that it is *potential* in the upstream far-field.

ii. Under the condition i, CPF remains until a shock is encountered.

iii. CPF is theoretically impossible downstream of a shock. A CPF model based shock is isentropic, and it is acceptable only if the simulated real shock is *weak* enough to permit an isentropic approximation.

□

Remarks:

Numerically, the whole point of the CPF model is that it reduces the DOFs, as a result of substituting the velocity *vector* with its *scalar* potential. That is why I insisted in Chapter 1 that the approach that demands C^1 Hermite elements should

not be appropriate for CPF, because of their re-committing/re-visiting of the DOFs that have been "potentially" saved.

2.4 Convection-Diffusion Problem

In addition to the aerodynamic equations that have been described, I will also frequently refer to the scalar convection/diffusion problem, i.e.,

$$\frac{\partial q}{\partial t} + \frac{\partial q v_j}{\partial x_j} - \frac{\partial}{\partial x_j}\mu\frac{\partial q}{\partial x_j} = 0 \tag{2.56}$$

or

$$\frac{\partial q}{\partial t} + v_j\frac{\partial q}{\partial x_j} - \frac{\partial}{\partial x_j}\mu\frac{\partial q}{\partial x_j} = 0 \tag{2.57}$$

where q is the transported scalar, v_j is the velocity, and μ is the diffusitivity.

□

The application of this model in most cases is symbolic, assuming linearity or nonlinearity without specification. The only exception is in the eventual numerical analysis, where the following nonlinear version of (2.57) is considered:

$$\frac{\partial q}{\partial t} + v_j\frac{\partial q}{\partial x_j} = 0 \tag{2.58}$$

$$v_j = \frac{v_{0_j}}{1 + af(q)} \tag{2.59}$$

wherein $a \geq 0$ is a parameter controlling the nonlinearity; v_{0_j} is the *linear* velocity (the v_j as $a = 0$), which is assumed to be continuous, incompressible (divergence free)

and time-independent; $f(q) > 0$ is a monotonely decreasing function, bounded by

$$f(q) = \begin{cases} 1, & as\ q = 0 \\ 0, & as\ q = \infty \end{cases} \tag{2.60}$$

The nonlinear regulation (2.59-2.60) results in several properties of (2.58). First of all, since $f(q)$ monotonely decreases, $v_j(q)$ monotonely increases. Consequently, a q-profile convected by (2.58) gets constantly *deviated*, in the way that, while its "up-sides" (defined as where $\frac{\partial q}{\partial x_j} > 0$) become gradually *stretched* apart, its "down-sides" (defined as where $\frac{\partial q}{\partial x_j} < 0$) are increasingly *tightened* up. Ultimately, the "down-sides" evolve into being shocks. Illustrated for this is in Figure 2.1-2.2.

Secondly, the *deviation* of profile and the *formation* of shocks described above occurs at a *bounded* rate, since according to (2.60),

$$v_{j_{min}} = v_j(q)|_{q=0} = \frac{v_{0_j}}{1+a} \tag{2.61}$$

$$v_{j_{max}} = v_j(q)|_{q=\infty} = v_{0_j} \tag{2.62}$$

Thirdly, one can also formulate the velocity of shock for (2.58-2.59-2.60):

$$v_{s_j} = \frac{v_{0_j}}{1 + \frac{\Delta C}{\Delta q}} \tag{2.63}$$

where v_{s_j} is the shock velocity, $\Delta q = q_2 - q_1$ is the shock amplititude and ΔC is the incremental corresponding to Δq of function

$$C(q) = a \int_0^q f(q)\, dq \tag{2.64}$$

i.e.,

$$\Delta C = C(q_2) - C(q_1) = a \int_{q_1}^{q_2} f(q)\, dq \tag{2.65}$$

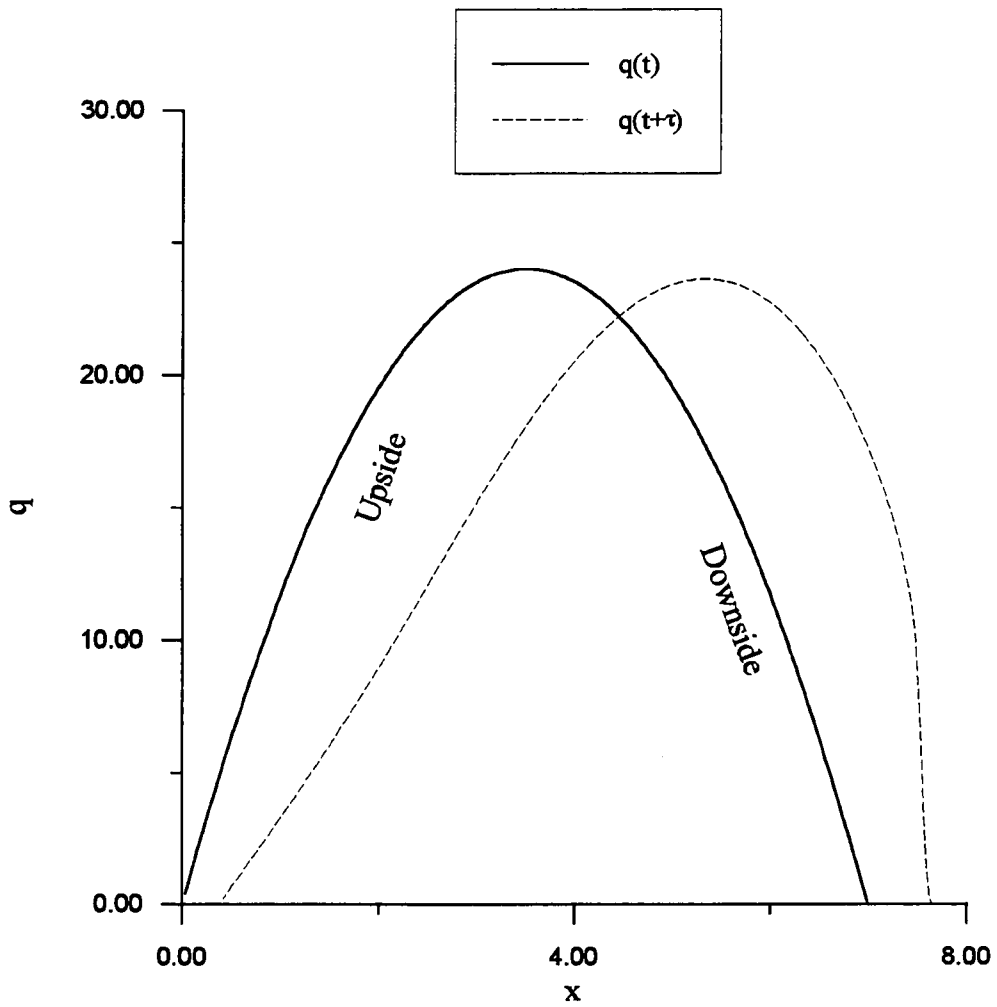


Figure 2.1: 1-D illustration for the nonlinear convection in the case of convex profile: (**a**) the initial symmetric profile; (**b**) the nonlinearly deviated profile a short time later.

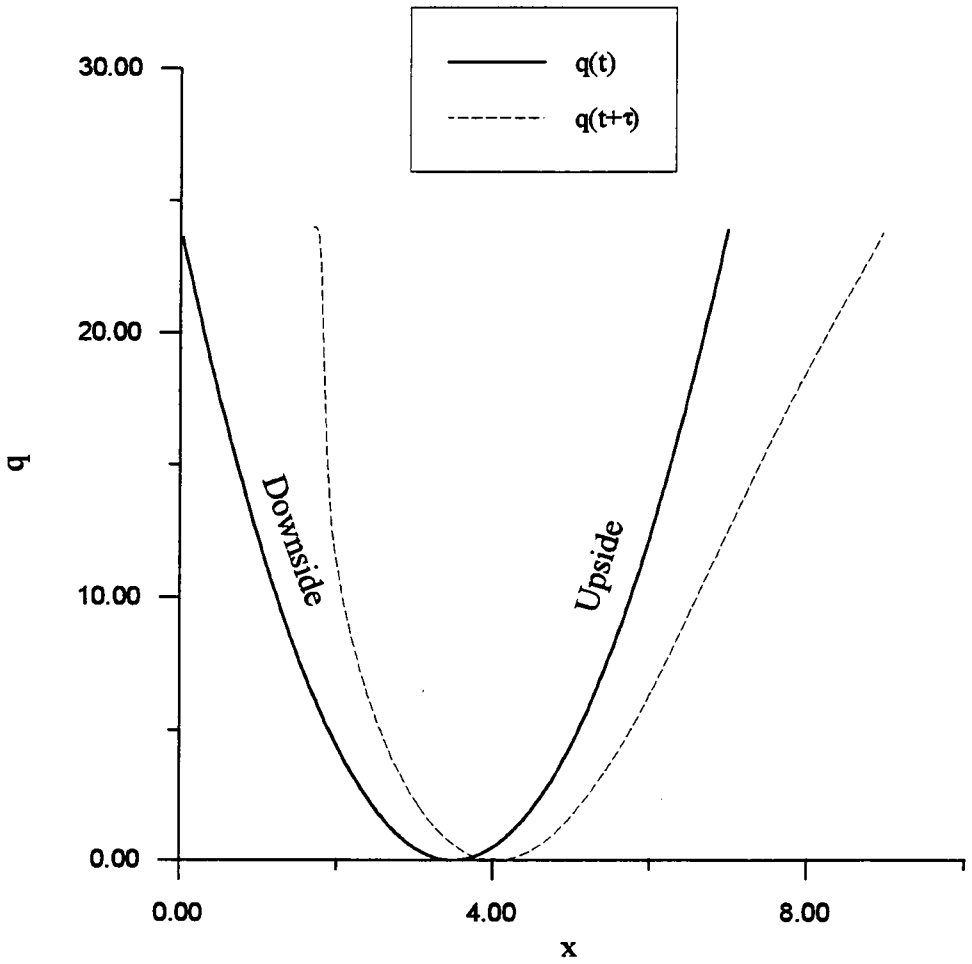


Figure 2.2: 1-D illustration for the nonlinear convection in the case of concave profile: (**a**) The initial symmetric profile; (**b**) the nonlinearly deviated profile a short time later.

To prove (2.63), the following needs to be noticed:

i. (2.58-2.59) is a *pure* convection system, which has neither the mechanism transfering q across streamlines (such as *diffusion*), nor the one correlating the convection on different streamlines (such as sonic *propagation* in flow). As a result, the convection on different streamlines are independent of one another, as if each stands alone.

ii. By multiplying both sides of (2.58) with $[1 + af(q)]$, and considering the chain rule of differentiation via (2.64), i.e.,

$$\frac{\partial C}{\partial t} = \frac{dC}{dq}\frac{\partial q}{\partial t} = a f(q)\frac{\partial q}{\partial t} \tag{2.66}$$

one can transform (2.58-2.59) into

$$\frac{\partial q}{\partial t} + \frac{\partial C}{\partial t} + v_{0_j}\frac{\partial q}{\partial x_j} = 0 \tag{2.67}$$

In other words, the nonlinear convection governed by (2.58-2.59) is equivalent to the convection in the linear velocity-field $\mathbf{v}_0$ embedded with absorption at rate $\frac{\partial C}{\partial t}$.

With **i** in view, it is immediate that $\mathbf{v}_s \parallel \mathbf{v}_0$ (or $\mathbf{v}$), wherever it is on the shock.

With **ii** in view, and if the shock is isolated by *fixing* to it a micro-streampipe, as illustrated in Figure 2.3, the convection around the micro-streampipe then should be such a picture. While there is no flux on the "side-wall", the *stationarily* absorbed C-components and the *linearly* convected q-components come across respectively at

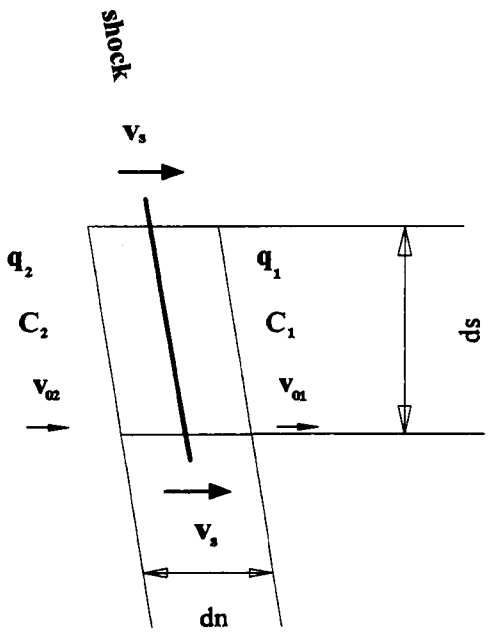


Figure 2.3: Mass transfer around a shock for the nonlinear convection problem. The rectangular frame is an infinitesimal stream-pipe control volume. C_1 and C_1 are the components that are the *stationarily* absorbed, and come across at the shock speed v_s; q_1 and q_2 are the components that are linearly convected, and come across at the relative speed $v_s - v_0$.

the *shock* speed v_s and relative speed $v_s - v_0$. Consequently, the net "mass balance" through the micro-streampipe is as follows.

$$[C_1 v_s + q_1 (v_s - v_{0_1})]\, ds - [C_2 v_s + q_2 (v_s - v_{0_2})]\, ds = (\cdot)\, ds\, dn \tag{2.68}$$

where the RHS volumetric term accounts for the "mass" that remains in the micro-streampipe. Taking the limit of (2.68) as $ds \to 0$ and $dn \to 0$, and considering the continuity of v_{0_j}, there is

$$v_s = \frac{v_0}{1 + \frac{\Delta C}{\Delta q}} \tag{2.69}$$

and this, together with the fact that $\mathbf{v}_s \parallel \mathbf{v}_0$ as pointed out earlier, leads to (2.63).

□

It is interesting that, if (2.64) is substituted into (2.59), v_j becomes

$$v_j = \frac{v_0}{1 + \frac{dC}{dq}} \tag{2.70}$$

which provided a verification of the shock velocity formula that has been established, in that, therefore, (2.63) is correct at least when shock is infinitesimal. Furthermore, by combining (2.70) and (2.63), one can geometrically characterize the interaction between shock and characteristics upon function $C(q)$, as shown in Figure 2.4. The key information in Figure 2.4 is the monotone increasingness and convexity of $C(q)$, i.e.,

$$\left(\frac{dC}{dq}\right)_{q_1} > \left(\frac{\Delta C}{\Delta q}\right)_{q_1}^{q_2} > \left(\frac{dC}{dq}\right)_{q_2} > 0 \tag{2.71}$$

for any $0 < q_1 < q_2$, which is as a result of the given the $f(q)$. That is why it is the

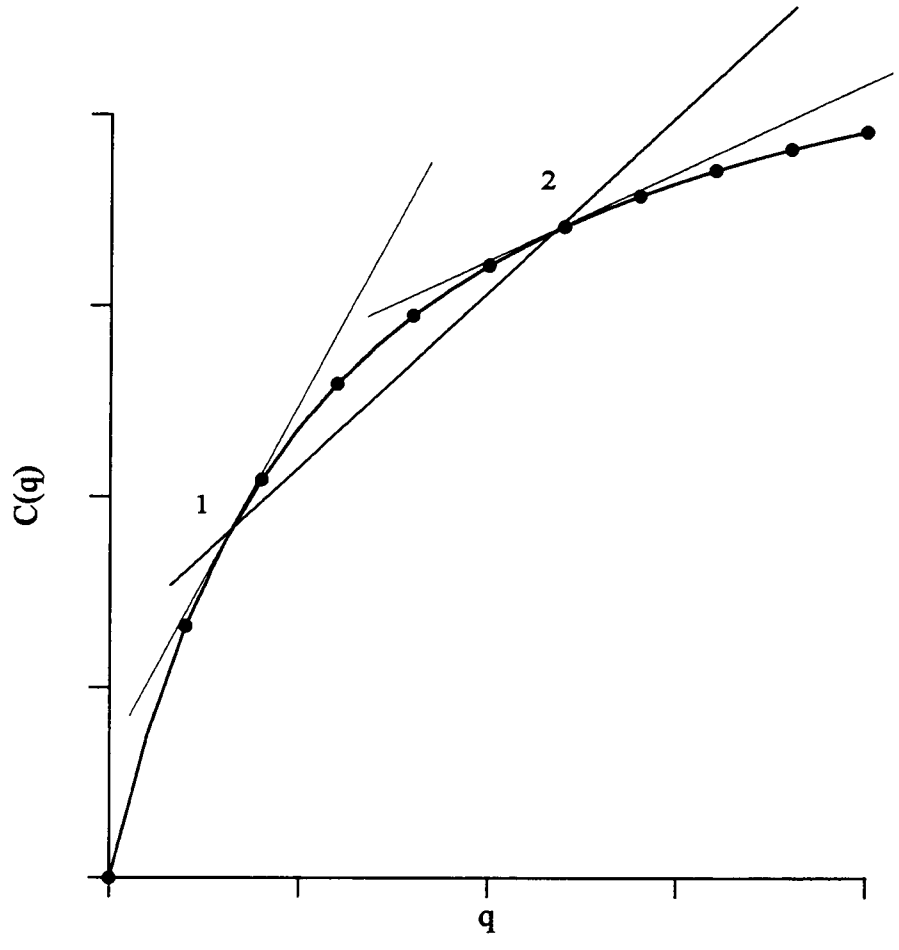


Figure 2.4: Illustration for the interaction between shock and characteristics for non-linear the convection problem. Point **1** — shock ground; point **2** — shock plateau. Shock and characteristics moves respectively at the *secant* and *tangential* speed. Rule of thumb: shock is *overtaken* on the top and *overtaking* at the bottom.

"rule of thumb" that "shock is *overtaken* on the top and *overtaking* at the bottom" (quote the caption of Figure 2.4).

Remarks:

The net "mass balance" equation (2.68) can be analytically proven. In fact, if one denotes the moving coordnates that are fixed with the shock as (t', x'_j), then, by definition,

$$\begin{aligned} t' &= t \\ x'_j &= x_j - v_{s_j} t \end{aligned} \tag{2.72}$$

and/thus,

$$\begin{aligned} \partial_t &= \frac{\partial t'}{\partial t}\partial_{t'} + \frac{\partial x'_j}{\partial t}\partial_{x'_j} \\ &= \partial_{t'} - v_{s_j}\partial_{x'_j} \end{aligned} \tag{2.73}$$

$$\begin{aligned} \partial_{x_j} &= \frac{\partial t'}{\partial x_j}\partial_{t'} + \frac{\partial x'_i}{\partial x_j}\partial_{x'_i} \\ &= \frac{1}{v_{s_j}}\partial_{t'} + \partial_{x'_j} \end{aligned} \tag{2.74}$$

Consequently, (2.67) can be transformed as follows.

$$\begin{aligned} &\frac{\partial q}{\partial t'} + \frac{\partial C}{\partial t'} + \frac{1}{v_{s_j}}\frac{\partial q v_{0_j}}{\partial t'} + \\ &\frac{\partial q v_{0_j}}{\partial x'_j} - v_{s_j}\left(\frac{\partial q}{\partial x'_j} + \frac{\partial C}{\partial x'_j}\right) = 0 \end{aligned} \tag{2.75}$$

The integration (2.75) over the said micro-streampipe is (2.68), with the unsteady terms contributing to the volumetric terms on the RHS.

Note that the incompressibility of v_{0_j} has been used in getting (2.75).

Chapter 3

The Metric Data Mean Technique for Finite Element Method

Metric data mean (MDM) is an algorithm designed to extract out FE metric data, hence, *unload* the *burden* of numerical quadrature in nonlinear FE-CFD ([20]). The major issue dealt with is the mean value of metric data, which is often unwarranted, and which is why, ultimately, MDM was shaped extracting either *termwise* (polynomial based) or *piecewise* (submesh based) data.

3.1 Problem Statement

Simply stated, the "quadrature problem" in nonlinear FE-CFD is that, caused jointly by the nonlinearity of the CFD system and the finite element, there is a necessity of either globally storing or repeatedly quadrating the FE matrices ([2, 20]). In

this section, I will first observe the context of such a problem, sampling the convection/diffusion system (2.57).

3.1.1 Galerkin FE-CFD Algorithm

Galerkin Weak Statement

The theoretical origin of Galerkin FE algorithm is the Galerkin weak statement, which, for (2.57), is defined as for $q \in H^1(\Omega)$ that satifies

$$\begin{aligned} & \int_\Omega w L(q)\, d\Omega \\ = & \int_\Omega \left(w \frac{\partial q}{\partial t} + w v_j \frac{\partial q}{\partial x_j} + \mu \frac{\partial w}{\partial x_j} \frac{\partial q}{\partial x_j} \right) d\Omega - \oint_{\partial\Omega} \mu w \frac{\partial q}{\partial x_j} n_j d\sigma \\ = & \ 0 \end{aligned} \tag{3.1}$$

for all $w \in H^1(\Omega)$. Function q is often referred to as the *weak* solution, because

- $\int_\Omega w L(q)\, d\Omega = 0$ is a weaker condition than $L(q) = 0$, which, in fact, permits $L(q) \neq 0$ to take place on a *zero-measured* subset of Ω, and
- $q \in H^1(\Omega)$, is a weaker requirement on the "smoothness" of q than that it has to be second order differentiable as in the original PDE (2.57).

Finite Element Subspace

Let $S^h(\Omega)$ be the FE subspace of $H^1(\Omega)$, and $\{N_\alpha(\eta_j)\}$ the shape functions of its basis. In S^h, then,

$$q(t, x_j) \simeq q^h(t, x_j) = \bigcup_e q_\beta(t)\, N^e_\beta(\eta_j) \tag{3.2}$$

$$v_j(t,x_j) \simeq v_j^h(t,x_j) = \bigcup_e v_{j_\beta}(t)\, N_\beta^e(\eta_j) \tag{3.3}$$

$$x_j^e(\eta_i) = x_{j\beta}^e \tilde{N}_\beta(\eta_i) \tag{3.4}$$

where $x_j^e(\eta_i)$ is the FE coordinate transform, which is defined by a set of independent shape functions $\{\tilde{N}_\beta(\eta_i)\}$. The choice of $x_j^e(\eta_i)$ is rather free; as long as being non-singular, it just has to be such that

$$\begin{aligned} \partial\Omega^h &= \bigcup_e \partial\Omega^e \\ &= \bigcup_e \partial\{x_{j\beta}^e \tilde{N}_\beta(\eta_i) \mid \eta_i \in \Omega^*\} \end{aligned} \tag{3.5}$$

is a good enough approximation to $\partial\Omega$ ([2]).

FE Semi-discretization

With (3.2-3.3), (3.1) becomes *semi-discrete* (spatially):

$$\begin{aligned} &\dot{q}_\beta \int_{\Omega^h} N_\alpha N_\beta \, d\Omega + v_{j_\theta} q_\beta \int_{\Omega^h} N_\alpha N_\theta \frac{\partial N_\beta}{\partial x_j} d\Omega + \\ &\mu\, q_\beta \int_{\Omega^h} \frac{\partial N_\alpha}{\partial x_j} \frac{\partial N_\beta}{\partial x_j} d\Omega - \mu \oint_{\partial\Omega} N_\alpha \frac{\partial q}{\partial x_j} n_j d\sigma = 0 \end{aligned} \tag{3.6}$$

which is an temporal ODE system, and can be re-expressed in the following mass-residual form

$$M200_{\alpha\beta} \cdot \dot{q}_\beta + R_\alpha = 0 \tag{3.7}$$

where $M200_{\alpha\beta}$ and R_α are the "assembly" of terms

$$\begin{aligned} R_\alpha^e &= v_{J_\theta} q_\beta M3J00_{\beta\theta\alpha}^e + \mu\, q_\beta M2JJ_{\alpha\beta}^e - \\ &\oint_{\partial\Omega^e \cap \Omega} N_\alpha \mu \frac{\partial q}{\partial x_j} n_j d\sigma \end{aligned} \tag{3.8}$$

$$M200^e_{\alpha\beta} = \int_{\Omega^e} N_\alpha N_\beta \, d\Omega = \int_{\Omega^*} N_\alpha N_\beta \, D \, d\Omega^* \tag{3.9}$$

and

$$M3J00^e_{\beta\theta\alpha} = \int_{\Omega^e} \frac{\partial N_\beta}{\partial x_J} N_\theta N_\alpha \, d\Omega = \int_{\Omega^*} E_{MJ} \frac{\partial N_\beta}{\partial \eta_M} N_\theta N_\alpha \, d\Omega^* \tag{3.10}$$

$$M2JJ^e_{\alpha\beta} = \int_{\Omega^e} \frac{\partial N_\alpha}{\partial x_J} \frac{\partial N_\beta}{\partial x_J} \, d\Omega = \int_{\Omega^*} \frac{E_{MJ} E_{NJ}}{D} \frac{\partial N_\alpha}{\partial \eta_M} \frac{\partial N_\beta}{\partial \eta_N} \, d\Omega^* \tag{3.11}$$

The chain rules ([2])

$$d\Omega = D \, d\Omega^* \tag{3.12}$$

$$\frac{\partial}{\partial x_j} = \frac{E_{mj}}{D} \frac{\partial}{\partial \eta_m} \tag{3.13}$$

have been used in (3.9-3.10-3.11), where D and E_{ij} are what is called FE metric data ([20, 21]), which are respectively the determinent and cofactors of the Jacobian matrix of the FE coordinate transform (3.4):

$$D = | D_{ij} | = | \frac{\partial x_j}{\partial \eta_i} | \tag{3.14}$$

$$[D_{ij}]^{-1} = \frac{1}{D} [E_{ji}] \tag{3.15}$$

□

Θ Scheme

Temporally integrating (3.7) leads us to the eventual FE numerical algorithm. If a single step Θ-implicit scheme is conducted then

$$M200_{\alpha\beta} \left(q^{n+1}_\beta - q^n_\beta \right) + \Theta \, \Delta t \, R^{n+1}_\alpha + (1 - \Theta) \, \Delta t \, R^n_\alpha = 0 \tag{3.16}$$

where the shoulder indexes n and $n+1$ denote time levels.

If (2.57) is nonlinear, then so is (3.16), and its solution has to be through linearization/relaxation such as the Newton iteration algorithm ([2]):

$$\begin{aligned} & \left[M200_{\alpha\beta} + \Theta\Delta t J_{\alpha\beta}^{p,n+1}\right] \delta q_{\beta}^{p,n+1} \\ = & -M200_{\alpha\beta}\left(q_{\beta}^{p,n+1} - q_{\beta}^{n}\right) - \\ & \Theta\Delta t R_{\alpha}^{p,n+1} - (1-\Theta)\Delta t R_{\alpha}^{n} \end{aligned} \tag{3.17}$$

where $(\cdot)^{p,n+1}$ is the pth approximation of $(\cdot)^{n+1}$,

$$q_{\beta}^{p+1,n+1} = q_{\beta}^{p,n+1} + \delta q_{\beta}^{p,n+1} \tag{3.18}$$

and

$$\begin{aligned} J_{\alpha\beta} = & \mu\, M2II_{\alpha\beta} - v_{I_\theta}\, M3I00_{\beta\theta\alpha} - \\ & q_{\theta}\frac{\partial v_{I_{\underline{\beta}}}}{\partial q_{\underline{\beta}}} M3I00_{\theta\beta\alpha} - \frac{\partial}{\partial q_{\beta}} \oint_{\partial\Omega} N_{\alpha}\mu \frac{\partial q}{\partial x_j} n_j d\sigma \end{aligned} \tag{3.19}$$

which is the Jacoban matrix of residual $\{R_\alpha\}$, i.e., $\left[\frac{\partial R_\alpha}{\partial q_\beta}\right]$.

Note that what causes "nonlinearity" here is the dependence of v_j on q, i.e., a non-vanishing $\frac{\partial v_j}{\partial q}$.

3.1.2 The Format of FE Matrices

Several features of the Galerkin FE-CFD have been reflected in the above sample algorithm ([2]): First of all, it is that the Galerkin weak statement is first order differential — since both w and $q \in H^1(\Omega)$. Subsequently, one has the following

general representation of the Galerkin FE matrices.

$$M^e_{\alpha_1\alpha_2\cdots} = \int_{\Omega^e} \prod_{k=1}^{\tilde{k}} \left(\frac{\partial^{r_k} N_{\alpha_{\underline{k}}}}{\partial x_{i_{\underline{k}}}^{r_{\underline{k}}}} \right) d\Omega$$
$$= \int_{\Omega^*} \prod_{k=1}^{\tilde{k}} \left(\frac{E^{r_k}_{n_{\underline{k}} i_{\underline{k}}}}{D^{r_{\underline{k}}}} \frac{\partial^{r_{\underline{k}}} N_{\alpha_{\underline{k}}}}{\partial \eta^{r_{\underline{k}}}_{n_{\underline{k}}}} \right) D \, d\Omega^* \tag{3.20}$$

where $\prod$ is the multiplicator; $\tilde{k}$ the number of FE shape functions involved; r_k is the differentiation order for the kth FE shape function, which is either 0 or 1, and x_{i_k} is the coordinate that the differentiation is with respect to as $r_k \neq 0$. The second step in (3.20) employs the same chain rules as in (3.9-3.10-3.11), i.e., (3.12-3.13), except that r_k has been included, so that $\frac{E^{r_k}_{n_{\underline{k}} i_{\underline{k}}}}{D^{r_{\underline{k}}}} \frac{\partial^{r_{\underline{k}}} N_{\alpha_{\underline{k}}}}{\partial \eta^{r_{\underline{k}}}_{n_{\underline{k}}}}$ can be a general expression for both N_{α_k} and $\frac{\partial N_{\alpha_k}}{\partial x_{i_{\underline{k}}}}$.

Secondly, using (3.20), one can generally categorize the Galerkin FE matrices as:

- Mass matrix, which is the case of (3.20) with $r_k \equiv 0$,

$$M^e_{\alpha_1\alpha_2\cdots} = \int_{\Omega^*} D \left(\prod_{k=1}^{\tilde{k}} N_{\alpha_k} \right) d\Omega^* \tag{3.21}$$

- Convection matrix, which is the case of (3.20) with $r_k \neq 0$ at only one k,

$$M^e_{\alpha_1\alpha_2\cdots} = \int_{\Omega^*} \left(E_{n_1 I} \frac{\partial N_{\alpha_1}}{\partial \eta_{n_1}} \right) \left(\prod_{2}^{\tilde{k}} N_{\alpha_k} \right) d\Omega^* \tag{3.22}$$

- Diffusion matrix, which is the case of (3.20) when there are *two* $r_k \neq 0$,

$$M^e_{\alpha_1\alpha_2\cdots} = \int_{\Omega^*} \left(\frac{E_{m_1 I} E_{n_2 J}}{D} \frac{\partial N_{\alpha_1}}{\partial \eta_{m_1}} \frac{\partial N_{\alpha_2}}{\partial \eta_{n_2}} \right) \left(\prod_{k=3}^{\tilde{k}} N_{\alpha_k} \right) d\Omega^* \tag{3.23}$$

Note that the differentiated FE basis have been permuted to the front in both (3.22) and (3.23). Note also that, although the above categorization is generally

valid, the matrices may not be in their "right" names if the considered system is not the convection/diffusion problem (2.57). For example, it is the convection terms that generate the diffusion matrices in compressible potential flow (see Chapter 4).

Thirdly, $M^e_{\alpha_1\alpha_2\cdots}$ in the above formulates a *physical* FE matrix, which is the combination of a group of matrices after the FE coordinate transform — as shown in the implied Einstein summation over n (the only dummy index). If one wants to further empower it to represent each "submatrix", one can do so by further exempting n from the Einstein convention:

$$M^e_{\alpha_1\alpha_2\cdots} = \int_{\Omega^*} p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right] F\left(N_{\alpha_1}, N_{\alpha_2}, \cdots\right) d\Omega^* \tag{3.24}$$

$$p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right] = \prod_{k=1}^{\tilde{k}} \left(\frac{E^{\underline{r}_k}_{n_k i_k}}{D^{\underline{r}_k}}\right) \tag{3.25}$$

$$F\left(N_{\alpha_1}, N_{\alpha_2}, \cdots\right) = \prod_{k=1}^{\tilde{k}} \left(\frac{\partial^{\underline{r}_k} N_{\alpha_k}}{\partial \eta^{\underline{r}_k}_{\underline{n}_k}}\right) \tag{3.26}$$

Let me illustrate this taking the example of $M3J00^e$: Using (3.20) or (3.22), the result is (3.10), i.e., the following after the summation is expanded.

$$\begin{aligned} M^e_{\alpha_1\alpha_2\alpha_3} = & \int_{\Omega^*} E_{1I} \frac{\partial N_{\alpha_1}}{\partial \eta_1} N_{\alpha_2} N_{\alpha_3}\, d\Omega^* + \\ & \int_{\Omega^*} E_{2I} \frac{\partial N_{\alpha_1}}{\partial \eta_2} N_{\alpha_2} N_{\alpha_3}\, d\Omega^* + \\ & \int_{\Omega^*} E_{3I} \frac{\partial N_{\alpha_1}}{\partial \eta_3} N_{\alpha_2} N_{\alpha_3}\, d\Omega^* \end{aligned} \tag{3.27}$$

Using (3.24-3.25-3.26), on the other hand, leads to

$$M^e_{\alpha_1\alpha_2\alpha_3} = \int_{\Omega^*} E_{n_1 I} \frac{\partial N_{\alpha_1}}{\partial \eta_{\underline{n}_1}} N_{\alpha_2} N_{\alpha_3}\, d\Omega^* \tag{3.28}$$

$$= \begin{cases} \int_{\Omega^*} E_{1I} \frac{\partial N_{\alpha_1}}{\partial \eta_1} N_{\alpha_2} N_{\alpha_3} \, d\Omega^* & or \\ \int_{\Omega^*} E_{2I} \frac{\partial N_{\alpha_1}}{\partial \eta_2} N_{\alpha_2} N_{\alpha_3} \, d\Omega^* & or \\ \int_{\Omega^*} E_{3I} \frac{\partial N_{\alpha_1}}{\partial \eta_3} N_{\alpha_2} N_{\alpha_3} \, d\Omega^* \end{cases}$$

□

It is imperative to notice that the element dependence of FE matrices is caused solely due to the presence of FE metric data. An immediate consequense of this is that, while in general it is necessary to *globally* quadrate an FE matrix, it is not if the FE coordinate transform $x_i^e(\eta_j)$ is linear, since by definition, both E_{ij} and D are *constant* (though still element dependent) in the case of a linear $x_i^e(\eta_j)$, and/hence,

$$M^e_{\alpha_1\alpha_2\cdots} = p^e(E_{ij}, D) \int_{\Omega^*} F(N_{\alpha_1}, N_{\alpha_2}, \cdots) \, d\Omega^* \tag{3.29}$$

3.1.3 The Problem of Numerical Quadrature

As described, due to nonlinearity, the solution of the FE algorithm (3.16) can only be approached via iteration that requires computing of $\{R_\alpha\}$ and $[J_{\alpha\beta}]$ at every iterative step. Consequently, FE matrices have to be either stored or repeatedly computed at the same frequency. At the same time, because of element dependence of FE matrices as a result of the presence of the FE metric data, to compute an FE matrix means to quadrate it globlally except in the case of a linear FE coordinate transform.

If the FE coordinate transform *is* nonlinear, and as long as one is not willing or can not afford to store FE matrices, it is then a forceful choice to repeatedly perform numerical quadrature throughout the solution process — not only on each

time station but at every step within. This is what I referred to as the "quadrature problem" in nonlinear FE-CFD.

The "to store FE matrices" alternative should not be confused with the storage of the coefficient matrix of the eventual linear algebra system, i.e., $[M200_{\alpha\beta} + \Theta \Delta t J_{\alpha\beta}]$, as is often imposed by linear algebra solvers. Here, it is the *global* storage all FE matrices that constitute $[M200_{\alpha\beta} + \Theta \Delta t J_{\alpha\beta}]$, i.e., $M200_{\alpha\beta}$, $M3100_{\alpha\theta\beta}$, $M3200_{\alpha\theta\beta}$, $M3300_{\alpha\theta\beta}$, $M211_{\alpha\beta}$, $M221_{\alpha\beta}$, $M231_{\alpha\beta}$, $M212_{\alpha\beta}$, $M222_{\alpha\beta}$, $M232_{\alpha\beta}$, $M213_{\alpha\beta}$, $M223_{\alpha\beta}$ and $M233_{\alpha\beta}$, — even for a system as simple as (2.57).

The "quadrature problem" is *minor* in terms of the proportion that it consumes in the FE-CFD algorithm, considering that, after all, it is only about calculating the residuals and coefficients ($\{R_\alpha\}$ and $[J_{\alpha\beta}]$), which is usually much less demanding than actually solving the system. However, it is this "minor" problem that is primarily responsible for the efficiency gap between FE and FD methods (since, in terms of solving the system, FE and FD are really comparable).

Remarks:

It is not entirely fair to *blame* the "quadrature problem " on the nonlinearity of PDE. In fact, the "problem" exists even if (2.57) is linear, as long as its coefficients, i.e., the velocities v_j, are time dependent. Of course, a "quadrature problem " caused by the time-dependence of v_j would appear less frequently (once per *time station*, instead of once per *iterative step* as in the case of nonlinearity).

3.2 Metric Data Mean Algorithm

3.2.1 The Concept Metric Data Mean

Accordingly, solving the "quadrature problem" means to *memorize* the FE matrices without consuming a global data base, and in the way that they can be retrieved not as costly as being quadrated from scratch. But how ?

Well, there is only one way out, and that is to extract out the metric data from the FE integrals ([2, 20]). And the reason is simply that one can only put into memory storage the element-independent part of the FE matrices without consuming a global data base, i.e., the part other than metric data as discussed earlier.

I call the method that enables such extraction the *metric data mean* (MDM). But then, of course, what does it mean to "metric data mean", whether it is permissible, and/or what kind ? A good way spotting the answers is to look into the "metric data mean" in its literal sense — the mean value theorem of integration (MVTI):

$$\begin{aligned} M^e_{\alpha_1\alpha_2\cdots} &= \int_{\Omega^*} p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right] F\left(N_{\alpha_1}, N_{\alpha_2}, \cdots\right) d\Omega^* \\ &= \overline{p\left(E_{ij}, D\right)}^e_{\underline{\alpha_1\alpha_2}\cdots} \int_{\Omega^*} F\left(N_{\alpha_1}, N_{\alpha_2}, \cdots\right) d\Omega^* \end{aligned} \tag{3.30}$$

where $\overline{p\left(E_{ij}, D\right)}^e_{\alpha_1\alpha_2\cdots}$ is the mean value of $p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right]$ relative to entry $F\left(N_{\alpha_1}, N_{\alpha_2}, \cdots\right)$.

Immediately, it is obvious that it is actually a dilemmatic task to seek the said MDM algorithm, since the mean value $\overline{p\left(E_{ij}, D\right)}^e_{\alpha_1\alpha_2\cdots}$ may not exist. In fact, the existence of $\overline{p\left(E_{ij}, D\right)}^e_{\alpha_1\alpha_2\cdots}$ is warranted only if $F\left(N_{\alpha_1}, N_{\alpha_2}, \cdots\right)$, i.e., N_α and $\frac{\partial N_\alpha}{\partial \eta_i}$ as according to (3.26), is non-trivial and does not change sign ([20]), which is a

condition met only in the case of quasi-linear elements. In other words, if one seeks to go beyond quasi-linear elements, $p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right]$ then has to be extracted from $M^e_{\alpha_1\alpha_2\cdots}$ in such a way that (a) $p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right]$ is not *all out*, so that its mean value can be avoided, and (b) yet the extraction is still good enough to put *all out* the element dependent information, so that the remaining part of $M^e_{\alpha_1\alpha_2\cdots}$ becomes element independent.

3.2.2 Termwise Tactics and Metric Polynomial

Because one is confined by that the mean value $\overline{p\left(E_{ij}, D\right)}^e_{\alpha_1\alpha_2\cdots}$ rarely exists, a natural idea beating it is to go *termwise*. Symbolically, a termwise based algorithm to extract-out metric data metric data can be described as follows.

$$p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right] = a^e_i\, g_i\left(\eta_j\right) \tag{3.31}$$

$$M^e_{\alpha_1\alpha_2\cdots} = a^e_i \int_{\Omega^*} g_i\left(\eta_j\right)\, F\left(N_{\alpha_1\alpha_2\cdots}\right)\, d\Omega^* \tag{3.32}$$

where (3.31) is an expansion of some sort for $p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right]$, in which, as noted, only the coefficients are element dependent; (3.32) is the MDM algorithm. The data base for (3.32) is $\int_{\Omega^*} g_i\left(\eta_j\right) F\left(N_{\alpha_1\alpha_2\cdots}\right)\, d\Omega^*$, which constitues an array of *master matrices*.

□

The question, thus, comes down to whether there is the structure of $p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right]$ fitting (3.31). And the answer *depends*:

For mass and convection matrices, as shown (3.21-3.22),

$$p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right] = \begin{cases} D\left(\eta_l\right) \\ E_{ij}\left(\eta_l\right) \end{cases} \tag{3.33}$$

which are polynomials by definition, and/hence, can both be "extracted" via (3.31-3.32) straightforwardly, — by taking

$$\begin{aligned} p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right] = \; & a_0^e + a_1^e\eta_1 + a_2^e\eta_2 + a_3^e\eta_3 + a_4^e\eta_1^2 + \\ & a_5^e\eta_2^2 + a_6^e\eta_3^2 + a_7^e\eta_1\eta_2 + a_8^e\eta_2\eta_3 + \\ & a_9^e\eta_3\eta_1 + \cdots \end{aligned} \tag{3.34}$$

For diffusion matrices,

$$p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right] = \frac{E_{mi}E_{nj}}{D} \tag{3.35}$$

which is a rational function. Subsequently, the termise MDM applies to diffusion matrices only if the Taylor expansion of $\frac{E_{mi}E_{nj}}{D}$ converges, and *fast* enough (so that its polynomial approximation can be truncated *short* enough).

Remarks:

If one separates the average and varying part of $D\left(\eta_l\right)$, i.e., to let

$$D\left(\eta_l\right) = D^* + \tilde{D}\left(\eta_l\right) \tag{3.36}$$

$$D^* = \frac{1}{\Omega^*}\int_{\Omega^*} D\left(\eta_l\right)\, d\Omega^* \tag{3.37}$$

then

$$\frac{E_{mi}E_{nj}}{D} = \frac{E_{mi}E_{nj}}{D^* + \tilde{D}\left(\eta_l\right)} \tag{3.38}$$

It is consequent that the Taylor expansion of $\frac{E_{mi}E_{nj}}{D}$ converges if and only if $\frac{\tilde{D}}{D^*} < 1$. And if it does converge,

$$\frac{E_{mi}E_{nj}}{D} = \frac{E_{mi}E_{nj}}{D^*}\left[1 + \left(\frac{\tilde{D}}{D^*}\right) + \left(\frac{\tilde{D}}{D^*}\right)^2 + \cdots\right] \tag{3.39}$$

3.2.3 Polynomial Based MDM Algorithm

General Algorithm

Described in programmable fashion, the polynomial (3.34) based implementation of (3.31-3.32) is as follows.

- Data base:

$$C(\eta_1^r\eta_2^s\eta_3^t)_{\alpha_1\alpha_2\cdots} = \int_{\Omega^*} \eta_1^r\eta_2^s\eta_3^t\, F(N_{\alpha_1}, N_{\alpha_2}, \cdots)\, d\eta \tag{3.40}$$

  where r, s and t run through all terms in (3.34).

- Algorithm:

$$\begin{aligned} M^e_{\alpha_1\alpha_2\cdots} &= \int_{\Omega^*} p^e\left(E_{ij}(\eta), D(\eta)\right)\, F(N_{\alpha_1}, N_{\alpha_2}, \cdots)\, d\Omega^* \\ &= a^e_0 C_{\alpha_1\alpha_2\cdots} + a^e_1 C(\eta_1)_{\alpha_1\alpha_2\cdots} + a^e_2 C(\eta_2)_{\alpha_1\alpha_2\cdots} + \\ &\quad a^e_3 C(\eta_3)_{\alpha_1\alpha_2\cdots} + a^e_4 C(\eta_1^2)_{\alpha_1\alpha_2\cdots} + a^e_5 C(\eta_2^2)_{\alpha_1\alpha_2\cdots} + \\ &\quad a^e_6 C(\eta_3^2)_{\alpha_1\alpha_2\cdots} + a^e_7 C(\eta_1\eta_2)_{\alpha_1\alpha_2\cdots} + a^e_8 C(\eta_2\eta_3)_{\alpha_1\alpha_2\cdots} + \\ &\quad a^e_9 C(\eta_3\eta_1)_{\alpha_1\alpha_2\cdots} + \cdots \end{aligned} \tag{3.41}$$

□

Numerical Efficiency

The consumption of the above algorithm is computing the metric data once per entry, — less than that of quadrating a single point (since, in quadrature, you have to calculate not only $p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right]$ but also $F\left(N_{\alpha_1}, N_{\alpha_2}, \cdots\right)$). Of course, such an efficiency is only "namely" meaningful for diffusion matrices, unless its metric polynomial converges *fast* enough as discussed earlier.

Mean Value Form Algorithm

Algorithm (3.40-3.41) acquire its following mean value form, if $C_{\alpha_1\alpha_2\cdots} \neq 0$, where $C_{\alpha_1\alpha_2\cdots}$ is the case of $C(\eta_1^r\eta_2^s\eta_3^t)_{\alpha\beta\cdots}$ as $r = s = t = 0$, i.e.,

$$C_{\alpha_1\alpha_2\cdots} = \int_{\Omega^*} F\left(N_{\alpha_1}, N_{\alpha_2}, \cdots\right) d\Omega^* \tag{3.42}$$

- Data base:

$$\left(\overline{\eta_1^r\eta_2^s\eta_3^t}\right)_{\alpha_1\alpha_2\cdots} = \frac{C(\eta_1^r\eta_2^s\eta_3^t)_{\alpha_1\alpha_2\cdots}}{C_{\underline{\alpha}_1\underline{\alpha}_2\cdots}} \tag{3.43}$$

- Algorithm:

$$\begin{aligned} M^e_{\alpha_1\alpha_2\cdots} &= \overline{p\left(E_{ij}, D\right)}^e_{\underline{\alpha}_1\underline{\alpha}_2\cdots} \int_{\Omega^*} F(N_{\alpha_1}, N_{\alpha_2}, \cdots)\, d\Omega^* \\ &= \overline{p\left(E_{ij}, D\right)}^e_{\underline{\alpha}_1\underline{\alpha}_2\cdots} C_{\alpha_1\alpha_2\cdots} \end{aligned} \tag{3.44}$$

$$\begin{aligned} \overline{p\left(E_{ij}, D\right)}^e_{\alpha_1\alpha_2\cdots} &= a_0^e + a_1^e\overline{\eta_1}_{\alpha_1\alpha_2\cdots} + a_2^e\overline{\eta_2}_{\alpha_1\alpha_2\cdots} + a_3^e\overline{\eta_3}_{\alpha_1\alpha_2\cdots} + \\ & a_4^e\overline{\eta_1^2}_{\alpha_1\alpha_2\cdots} + a_5^e\overline{\eta_2^2}_{\alpha_1\alpha_2\cdots} + a_6^e\overline{\eta_3^2}_{\alpha_1\alpha_2\cdots} + a_7^e\overline{\eta_1\eta_2}_{\alpha_1\alpha_2\cdots} + \\ & a_8^e\overline{\eta_2\eta_3}_{\alpha_1\alpha_2\cdots} + a_9^e\overline{\eta_3\eta_1}_{\alpha_1\alpha_2\cdots} + \cdots \end{aligned} \tag{3.45}$$

The advantage of the "mean value" form MDM algorithm is its reduced data base — which involves only master-matrix $C_{\alpha_1\alpha_2\cdots}$. And its disadvantage is that it

consumes an extra multiplication compared with the "general" MDM algorithm: the one in (3.44) to combine $\overline{p\left(E_{ij}, D\right)}^{e}_{\alpha_1\alpha_2\ldots}$ and $C_{\alpha_1\alpha_2\ldots}$, which is the price paid for the reduced data base.

Note that, although the $\overline{p\left(E_{ij}, D\right)}^{e}_{\underline{\alpha\beta}\ldots}$ in (3.45) is identical with the mean value of $p^e\left[E_{ij}\left(\eta_l\right), D\left(\eta_m\right)\right]$ if the latter exists, it does not mean that the validity of (3.44-3.45) is thus conditioned. The validity of (3.44-3.45) requires only that $C_{\alpha\beta\ldots} \neq 0$, which is a weaker condition, and in fact satisfied by many matrices of many elements that do not "honor" MVTI (see the appendices of [2]).

3.2.4 Submeshing Algorithm

So far, I have completed constructing the polynomial based *termwise* MDM (P-MDM) algorithm. As said, the algorithm is fast and accurate in analytical sense for mass and convection matrices, however, meaningful for diffusion matrix only if the Taylor expansion of the metric data converges, and *fast* enough.

Now what if the Taylor expansion of the diffusive metric data does not converge or not fast enough? In this section, I will discuss an alternative approach to deal with these situations, one that is *piecewise* built in *submeshes*, which I call the submesh based MDM (S-MDM).

Conceptually, S-MDM is a simple thought, i.e., to utilize our freedom in choosing the FE coordinate transform as discussed in Section 3.1.1, and submesh the FE cell to the extent that the FE coordinate transform can be piecewise-linearly constructed

(see Figure 3.1 for illustration):

$$\tilde{N}_\beta(\eta_i) \simeq \bigcup_s \tilde{N}^s_\beta(\eta_i) \tag{3.46}$$

$$\begin{aligned} x^e_j(\eta_i) &= x^e_{j\beta}\tilde{N}_\beta(\eta_i) \\ &\simeq x^{es}_{j\beta}\tilde{N}^s_\beta(\eta_i) \end{aligned} \tag{3.47}$$

where s denotes a submesh cell in element Ω^e, and $\tilde{N}^s_\beta(\eta_i)$ are the linear FE shape functions spanning Ω^{es}. Since the metric data defined by (3.47) are piecewise-constant, they can be directly extracted out:

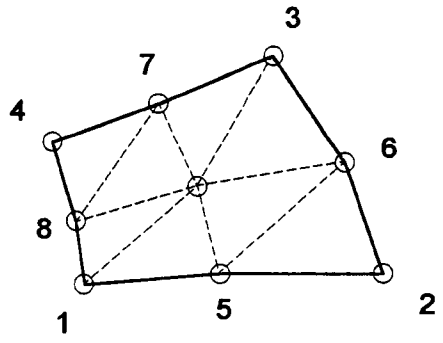
$$\begin{aligned} M^e_{\alpha_1\alpha_2\cdots} &= p^{es}(E_{ij}, D)\int_{\Omega^{*s}} F(N_{\alpha_1}, N_{\alpha_2}, \cdots)\, d\Omega^* \\ & p^{es}(E_{ij}, D)\, C^s_{\alpha_1\alpha_2\cdots} \end{aligned} \tag{3.48}$$

where $p^{es}(E_{ij}, D)$ is the metric data on Ω^{es}.

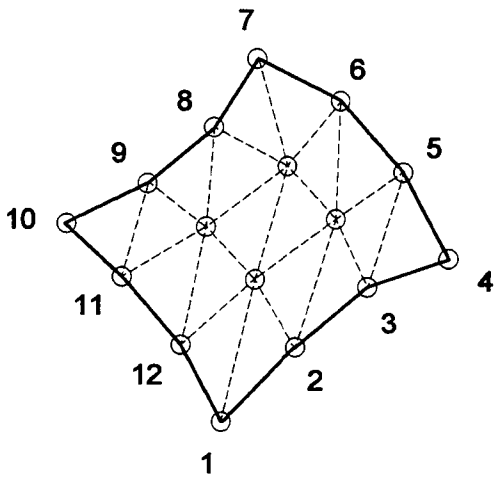
□

S-MDM is an approximate method, unlike the P-MDM in the case of mass and convection matrices. The error of S-MDM, as illustrated in Figure 3.1, occurs as a result of the "short-cut" between $\partial\Omega^e$ and $\partial(\cup_s\Omega^{es})$, which, however, is usually insignificant (being the second order of the size of submesh on $\partial\Omega^e$).

Pertinently, submeshes, say, $\cup_s\Omega^{es}_1$ and $\cup_s\Omega^{es}_2$, are indifferent if their borders are identical, i.e., $\partial(\cup_s\Omega^{es}_1) = \partial(\cup_s\Omega^{es}_2)$. The importance of noticing this is that it provides one with a lever to minimize the computing. For example, the bi-quadratic and bi-cubic elements can be submeshed as in either Figure 3.1 or Figure 3.2. The $M^e_{\alpha_1\alpha_2\cdots}$ that they generate via (3.48) are exactly the same, and the only difference

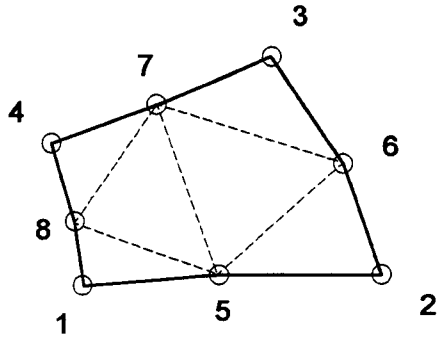


Quadratic Element

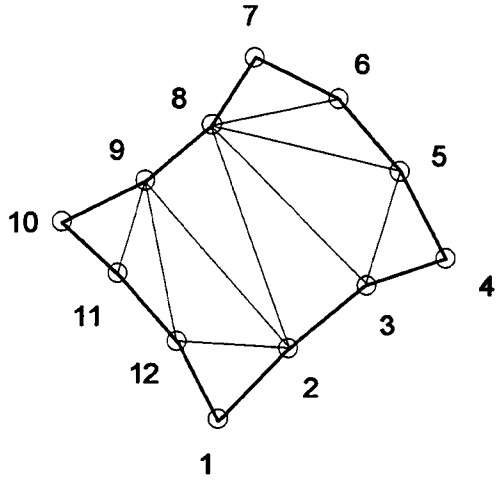


Cubic Element

Figure 3.1: Illustration for the *submeshing* based MDM. Solid quadrilaterals: the *original* mesh cell defined by FE coordinate transform $x^e_{j\beta}\tilde{N}_\beta(\eta_i)$. Dashed triangles: the *linear submeshes* that approximate $x^e_{j\beta}\tilde{N}_\beta(\eta_i)$.



Quadratic Element



Cubic Element

Figure 3.2: Illustration for the optimal *submeshing* (no internal nodes).

between them is that the Figure 3.2 approach costs less.

□

It is worthnoting that (3.48) also has its *mean value* form, if $C_{\alpha_1\alpha_2\cdots} \neq 0$:

$$\begin{aligned} M^e_{\alpha_1\alpha_2\cdots} &= p^{es}(E_{ij}, D)\, C^s_{\alpha_1\alpha_2\cdots} \\ &= \overline{p(E_{ij}, D)}^e_{\underline{\alpha_1\alpha_2}\cdots} C_{\alpha_1\alpha_2\cdots} \qquad (3.49) \\ \overline{p(E_{ij}, D)}^e_{\alpha_1\alpha_2\cdots} &= p^{es}(E_{ij}, D)\, \frac{C^s_{\underline{\alpha_1\alpha_2}\cdots}}{C_{\alpha_1\alpha_2\cdots}} \qquad (3.50) \end{aligned}$$

3.3 Tensor Product Representation

For tensor-product (TP) finite element, the data base of MDM can be correspondingly factorized, hence, represented onto the 1-D basis. — Let us pursue it.

Suppose $\alpha \Leftrightarrow (i,j,k)$, i.e.,

$$\{N_\alpha(\eta_1,\eta_2,\eta_3)\} \;=\; \{N_i(\eta_1)\cdot N_j(\eta_2)\cdot N_k(\eta_3)\} \tag{3.51}$$

Let us first notice that, because of (3.51), the lth term of $F(N_{\alpha_1},N_{\alpha_2},\cdots)$, as defined in (3.26), can be expanded as

$$\begin{aligned} &\frac{\partial^{r_{\underline{l}}}N_{\alpha_{\underline{l}}}}{\partial\eta_{n_{\underline{l}}}^{r_{\underline{l}}}} \\ =\;& \delta_{1n_{\underline{l}}}\frac{d^{r_{\underline{l}}}N_{i_{\underline{l}}}(\eta_1)}{d\eta_1^{r_{\underline{l}}}}N_{j_{\underline{l}}}(\eta_2)N_{k_{\underline{l}}}(\eta_3)+\delta_{2n_{\underline{l}}}N_{i_{\underline{l}}}(\eta_1)\frac{d^{r_{\underline{l}}}N_{j_{\underline{l}}}(\eta_2)}{d\eta_2^{r_{\underline{l}}}}N_{k_{\underline{l}}}(\eta_3)+ \\ &\delta_{3n_{\underline{l}}}N_{i_{\underline{l}}}(\eta_1)N_{j_{\underline{l}}}(\eta_2)\frac{d^{r_{\underline{l}}}N_{j_{\underline{l}}}(\eta_3)}{d\eta_3^{r_{\underline{l}}}} \end{aligned} \tag{3.52}$$

where

$$\delta_{jn_{\underline{l}}} \;=\; \frac{\partial^{r_{\underline{l}}}\eta_j}{\partial\eta_{n_{\underline{l}}}^{r_{\underline{l}}}} \tag{3.53}$$

Since (3.52) has only one non-zero term on the right-hand-side, it can thus be re-expressed without losing generality:

$$\frac{\partial^{r_{\underline{l}}}N_{\alpha_{\underline{l}}}}{\partial\eta_{n_{\underline{l}}}^{r_{\underline{l}}}} \;=\; \frac{d^{r1_{\underline{l}}}N_{i_{\underline{l}}}(\eta_1)}{d\eta_1^{r1_{\underline{l}}}}\,\frac{d^{r2_{\underline{l}}}N_{j_{\underline{l}}}(\eta_2)}{d\eta_2^{r2_{\underline{l}}}}\,\frac{d^{r3_{\underline{l}}}N_{j_{\underline{l}}}(\eta_2)}{d\eta_3^{r3_{\underline{l}}}} \tag{3.54}$$

where $r1_l$, $r2_l$ and $r3_l$ are the 1-D differentiation indicator, and still being either 0 or 1.

Consequently,

$$\begin{aligned} F\left(N_{\alpha_1}, N_{\alpha_2}, \cdots\right) &= F_1\left(N_{i_1}, N_{i_2}, \cdots\right) \\ & \quad F_2\left(N_{j_1}, N_{j_2}, \cdots\right) \\ & \quad F_3\left(N_{k_1}, N_{k_2}, \cdots\right) \end{aligned} \tag{3.55}$$

where

$$F_1(N_{i_1}, N_{i_2}, \cdots) = \frac{d^{r_1} N_{i_1}}{d\eta_1^{r_1}} \frac{d^{r_2} N_{i_2}}{d\eta_1^{r_2}} \cdots \tag{3.56}$$

$$F_2(N_{j_1}, N_{j_2}, \cdots) = \frac{d^{r_1} N_{j_1}}{d\eta_2^{r_1}} \frac{d^{r_2} N_{j_2}}{d\eta_2^{r_2}} \cdots \tag{3.57}$$

$$F_3(N_{k_1}, N_{k_2}, \cdots) = \frac{d^{r_1} N_{k_1}}{d\eta_3^{r_1}} \frac{d^{r_2} N_{k_2}}{d\eta_3^{r_2}} \cdots \tag{3.58}$$

(3.55) is the TP factorization of $F\left(N_{\alpha_1}, N_{\alpha_2}, \cdots\right)$.

General Algorithm:

Accordingly,

$$\begin{aligned} C(\eta_1^r \eta_2^s \eta_3^t)_{\alpha_1 \alpha_2 \cdots} &= \int_{\Omega^*} \eta_1^r \eta_2^s \eta_3^t \, F(N_\alpha(\eta_1, \eta_2, \eta_3), N_\beta(\eta_1, \eta_2, \eta_3), \cdots) \, d\Omega^* \\ &= \int_{-1}^{1} \eta_1^r \, F_1(N_{i_1}(\eta_1), N_{i_2}(\eta_1), \cdots) \, d\eta_1 \cdot \\ & \quad \int_{-1}^{1} \eta_2^s \, F_2(N_{j_1}(\eta_2), N_{j_2}(\eta_2), \cdots) \, d\eta_2 \cdot \\ & \quad \int_{-1}^{1} \eta_3^t \, F_3(N_{k_1}(\eta_3), N_{k_2}(\eta_3), \cdots) \, d\eta_3 \end{aligned} \tag{3.59}$$

In other words,

$$C(\eta_1^r \eta_2^s \eta_3^t)_{\alpha_1 \alpha_2 \cdots} = A1(\varsigma^r)_{i_1 i_2 \cdots} \cdot A2(\varsigma^s)_{j_1 j_2 \cdots} \cdot A3(\varsigma^t)_{k_1 k_2 \cdots} \tag{3.60}$$

where

$$\begin{cases} A1(\varsigma^r)_{i_1 i_2 \cdots} = \int_{-1}^{1} \varsigma^r \, F_1(N_{i_1}(\varsigma), N_{i_2}(\varsigma), \cdots) \, d\varsigma \\ A2(\varsigma^s)_{j_1 j_2 \cdots} = \int_{-1}^{1} \varsigma^s \, F_2(N_{j_1}(\varsigma), N_{j_2}(\varsigma), \cdots) \, d\varsigma \\ A3(\varsigma^t)_{k_1 k_2 \cdots} = \int_{-1}^{1} \varsigma^t \, F_3(N_{k_1}(\varsigma), N_{k_2}(\varsigma), \cdots) \, d\varsigma \end{cases} \tag{3.61}$$

Mean Value Form Algorithm:

Similarly, inserting (3.55) into (3.43), we have

$$\begin{aligned}(\overline{\eta_1^r \eta_2^s \eta_3^t})_{\alpha_1 \alpha_2 \cdots} &= \frac{C(\eta_1^r \eta_2^s \eta_3^t)_{\alpha_1 \alpha_2 \cdots}}{C_{\underline{\alpha}_1 \underline{\alpha}_2 \cdots}} \\ &= (\overline{\varsigma^r})_{i_1 i_2 \cdots} \cdot (\overline{\varsigma^s})_{j_1 j_2 \cdots} \cdot (\overline{\varsigma^t})_{k_1 k_2 \cdots} \end{aligned} \tag{3.62}$$

where

$$(\overline{\varsigma^r})_{i_1 i_2 \cdots} = \frac{\int_{-1}^{1} \varsigma^r F_1(N_{i_1}(\varsigma), N_{i_2}(\varsigma), \cdots) \, d\varsigma}{\int_{-1}^{1} F_1(N_{\underline{i}_1}(\varsigma), N_{\underline{i}_2}(\varsigma), \cdots) \, d\varsigma} = \frac{A1(\varsigma^r)_{i_1 i_2 \cdots}}{A1_{\underline{i}_1 \underline{i}_2 \cdots}} \tag{3.63}$$

$$(\overline{\varsigma^s})_{j_1 j_2 \cdots} = \frac{\int_{-1}^{1} \varsigma^s F_2(N_{j_1}(\varsigma), N_{j_2}(\varsigma), \cdots) \, d\varsigma}{\int_{-1}^{1} F_2(N_{\underline{j}_1}(\varsigma), N_{\underline{j}_2}(\varsigma), \cdots) \, d\varsigma} = \frac{A2(\varsigma^s)_{j_1 j_2 \cdots}}{A2_{\underline{j}_1 \underline{j}_2 \cdots}} \tag{3.64}$$

$$(\overline{\varsigma^t})_{k_1 k_2 \cdots} = \frac{\int_{-1}^{1} \varsigma^t F_3(N_{k_1}(\varsigma), N_{k_2}(\varsigma), \cdots) \, d\varsigma}{\int_{-1}^{1} F_3(N_{\underline{k}_1}(\varsigma), N_{\underline{n}_2}(\varsigma), \cdots) \, d\varsigma} = \frac{A3(\varsigma^t)_{k_1 k_2 \cdots}}{A3_{\underline{k}_1 \underline{k}_2 \cdots}} \tag{3.65}$$

□

Appendix B enlists such TP represented MDM data base for several the Lagrange elements.

3.4 Bilinear Finite Element Matrices

3.4.1 Finite Element Definitions

$$\{N_\alpha(\eta_j)\} = \frac{1}{4}\begin{Bmatrix} (1-\eta_1)(1-\eta_2) \\ (1+\eta_1)(1-\eta_2) \\ (1+\eta_1)(1+\eta_2) \\ (1-\eta_1)(1+\eta_2) \end{Bmatrix} \tag{3.66}$$

$$E_{ij} = \overline{E}^*_{ij} + (-1)^{i+j}\overline{E}^0_j\eta_i \tag{3.67}$$

$$\begin{aligned} \overline{E}^*_{ij} &= \frac{\int_{\Omega^*} E_{ij}\, d\Omega^*}{\int_{\Omega^*} d\Omega^*} \\ &= \frac{1}{4}\int_{\Omega^*} E_{ij}\, d\Omega^* \end{aligned} \tag{3.68}$$

$$\{\overline{E}^0_j\} = \frac{1}{2}\begin{Bmatrix} \frac{y_1+y_3}{2} - \frac{y_2-y_4}{2} \\ \frac{x_1+x_3}{2} - \frac{x_2-x_4}{2} \end{Bmatrix} \tag{3.69}$$

$$D = \overline{D}^*(1 + A\eta_1 + B\eta_2) \tag{3.70}$$

$$\begin{aligned} \overline{D}^* &= \frac{\int_{\Omega^*} D\, d\Omega^*}{\int_{\Omega^*} d\Omega^*} \\ &= \frac{1}{4}\int_{\Omega^*} D\, d\Omega^* \\ &= \overline{E}^*_{11}\overline{E}^*_{22} - \overline{E}^*_{21}\overline{E}^*_{12} \end{aligned} \tag{3.71}$$

$$\begin{cases} A = \frac{1}{\overline{D}^*}(\overline{E}^*_{22}\overline{E}^0_1 + \overline{E}^*_{21}\overline{E}^0_2) \\ B = \frac{1}{\overline{D}^*}(\overline{E}^*_{11}\overline{E}^0_2 + \overline{E}^*_{12}\overline{E}^0_1) \end{cases} \tag{3.72}$$

Note that the varying part of both $E_{ij}(\eta_k)$ and $D(\eta_k)$ are proportional to $\{\overline{E}^0_j\}$, and $\{\overline{E}^0_j\}$, as indicated in (3.69), vanishes as the FE cell becomes a *paralellgram.*

3.4.2 Mass and Convection Matrices

Let $\alpha \Leftrightarrow (i,j)$, $\theta \Leftrightarrow (k,l)$, and $\beta \Leftrightarrow (m,n)$.

Mass Matrix $\left[M200^e_{\alpha\beta}\right]$:

$$M200^e_{\alpha\beta} = \left(\overline{D}\right)^e_{\underline{\alpha\beta}} B200_{\alpha\beta} \tag{3.73}$$

$$B200_{\alpha\beta} = A200_{im}\, A200_{jn} \tag{3.74}$$

Therefore, according to (3.70),

$$\left(\overline{D}\right)^e_{\alpha\beta} = \overline{D}^* \left(1 + A\overline{\varsigma 200}_{im} + B\overline{\varsigma 200}_{jn}\right) \tag{3.75}$$

where $\overline{\varsigma 200}_{im}$ can be found in Appendix B.

Convection Matrix $\left[M3J00^e_{\alpha\theta\beta}\right]$:

$$\begin{aligned} M3J00^e_{\alpha\theta\beta} &= (\overline{E_{MJ}})^e_{\underline{\alpha\theta\beta}}\, B3M00_{\alpha\theta\beta} \\ &= (\overline{E_{1J}})^e_{\underline{\alpha\theta\beta}}\, B3100_{\alpha\theta\beta} + (\overline{E_{2J}})^e_{\underline{\alpha\theta\beta}}\, B3200_{\alpha\theta\beta} \end{aligned} \tag{3.76}$$

$$B3100_{\alpha\theta\beta} = A3100_{ikm}\, A3000_{jln} \tag{3.77}$$

$$B3200_{\alpha\theta\beta} = A3000_{ikm}\, A3100_{jln} \tag{3.78}$$

Therefore, according to (3.67),

$$\begin{cases} (\overline{E_{1J}})^e_{\alpha\theta\beta} = \overline{E}^*_{1J} + (-1)^{1+J}\overline{E}^0_J\overline{\varsigma 3100}_{ikm} \\ (\overline{E_{2J}})^e_{\alpha\theta\beta} = \overline{E}^*_{2J} + (-1)^{2+J}\overline{E}^0_J\overline{\varsigma 3100}_{jln} \end{cases} \tag{3.79}$$

where $\overline{\varsigma 3100}_{ikm}$ should be found in Appendix B.

3.4.3 Diffusion Matrices

The Convergence of Metric Polynomial

As said, for the diffusion matrix, one has to at least establish that its metric data has indeed a convergent Taylor series prior to applying the polynomial based MDM algorithm to it. In the following, such establishment will be presented, under the condition that the bilinear FE mesh cell is *convex* and *non-degenerated* (to a triangle).

Proof:

As discussed in Section 3.2.2, what needs to be shown is that

$$| A\eta_1 + B\eta_2 | < 1 \tag{3.80}$$

Let

$$\begin{cases} \tilde{A} = \overline{D}^* A = \overline{E}^*_{22}\overline{E}^0_1 + \overline{E}^*_{21}\overline{E}^0_2 \\ \tilde{B} = \overline{D}^* B = \overline{E}^*_{11}\overline{E}^0_2 + \overline{E}^*_{12}\overline{E}^0_1 \end{cases} \tag{3.81}$$

As a result, (3.80) is equivalent to

$$| \tilde{A}\eta_1 + \tilde{B}\eta_2 | < \overline{D}^* \tag{3.82}$$

Since the left hand side of (3.82) is linear, it has maximum and minimum values, and reaches them at the corners of element. In turn, (3.82) is equivalent to

$$\begin{cases} | \tilde{A} + \tilde{B} | < \overline{D}^* \\ | \tilde{A} - \tilde{B} | < \overline{D}^* \end{cases} \tag{3.83}$$

i.e.,

$$\begin{cases} \tilde{A} + \tilde{B} + \overline{D^*} > 0 \\ \tilde{A} + \tilde{B} - \overline{D^*} < 0 \end{cases} \tag{3.84}$$

and

$$\begin{cases} \tilde{A} - \tilde{B} + \overline{D^*} > 0 \\ \tilde{A} - \tilde{B} - \overline{D^*} < 0 \end{cases} \tag{3.85}$$

Now notice (3.71), (3.72), (3.68) and (3.69), there are

$$\begin{aligned} \tilde{A} + \tilde{B} + \overline{D^*} &= \tfrac{1}{4}\left[(x_3 - \tfrac{x_2+x_4}{2})(y_4 - y_2) - (y_3 - \tfrac{y_2+y_4}{2})(x_4 - x_2)\right] \\ &= \tfrac{1}{4}(\vec{Q3} \times \vec{24}) \cdot \vec{e_3} \end{aligned} \tag{3.86}$$

$$\begin{aligned} \tilde{A} + \tilde{B} - \overline{D^*} &= \tfrac{1}{4}\left[(x_1 - \tfrac{x_2+x_4}{2})(y_4 - y_2) - (y_1 - \tfrac{y_2+y_4}{2})(x_4 - x_2)\right] \\ &= \tfrac{1}{4}(\vec{Q1} \times \vec{24}) \cdot \vec{e_3} \end{aligned} \tag{3.87}$$

$$\begin{aligned} \tilde{A} - \tilde{B} + \overline{D^*} &= \tfrac{1}{4}\left[(x_2 - \tfrac{x_1+x_3}{2})(y_3 - y_1) - (y_2 - \tfrac{y_1+y_3}{2})(x_3 - x_1)\right] \\ &= \tfrac{1}{4}(\vec{P2} \times \vec{13}) \cdot \vec{e_3} \end{aligned} \tag{3.88}$$

$$\begin{aligned} \tilde{A} - \tilde{B} - \overline{D^*} &= \tfrac{1}{4}\left[(x_4 - \tfrac{x_1+x_3}{2})(y_3 - y_1) - (y_4 - \tfrac{y_1+y_3}{2})(x_3 - x_1)\right] \\ &= \tfrac{1}{4}(\vec{P4} \times \vec{13}) \cdot \vec{e_3} \end{aligned} \tag{3.89}$$

where P and Q are respectively the mid-point of the FE-cell diagonal $\overline{13}$ and $\overline{24}$, as illustrated in Figure 3.3; $\vec{e_3}$ is the unit vector along the z axis that is perpendicularly outward from the plane of paper.

With (3.86), (3.87), (3.88) and (3.89) in view, inequalities (3.84) and (3.85) are simply the results of the *convexity* and *non-degeneratedness* of the FE mesh as assumed. For example, suppose that $\tilde{A} + \tilde{B} + \overline{D^*} > 0$ is untrue, i.e., $\tilde{A} + \tilde{B} + \overline{D^*} \leq 0$.

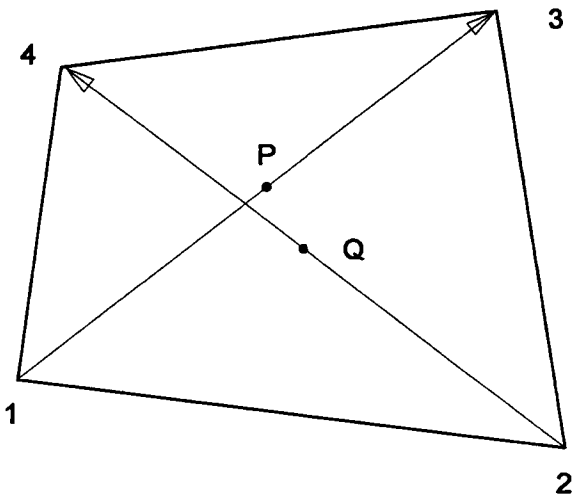


Figure 3.3: Bilinear FE Illustration (**I**). P — the mid-point of diagonal $\overline{13}$; Q — the mid-point of diagonal $\overline{24}$.

According to (3.86), this means that there must be either $\left(\vec{Q3} \times \vec{24}\right) \cdot \vec{e_3} < 0$ or $\vec{Q3} \times \vec{24} = 0$, which, however, are both impossible on a "convex" and "non-degenerated" quadrilateral:

- $\left(\vec{Q3} \times \vec{24}\right) \cdot \vec{e_3} < 0$ is impossible because it requires that $\vec{Q3}$, $\vec{24}$ and $\vec{e_3}$ form a left-hand-rotational frame, which occurs only if node 3 comes below diagonal $\overline{24}$ (see Figure 3.4), i.e., when the mesh-cell is *concave*.

- $\vec{Q3} \times \vec{24} = 0$ is impossible because it happens only if node 3 is on diagonal $\overline{24}$, i.e., when the FE mesh cell is *degenerated* to a triangle.

□

MDM Algorithm

- Metric Data Formula:

$$\begin{aligned} \frac{E_{mi}E_{nj}}{D} &= \frac{E_{mi}E_{nj}}{\overline{D^*}(1 + A\eta_1 + B\eta_2)} \\ &= \frac{E_{mi}E_{nj}}{\overline{D^*}} \cdot (1 - A\eta_1 - B\eta_2 + A^2\eta_1^2 + B^2\eta_2^2 + 2AB\eta_1\eta_2 + \cdots \end{aligned} \tag{3.90}$$

Since

$$\eta_i = \delta_{1i}\eta_1 + \delta_{2i}\eta_2 \tag{3.91}$$

(3.67) can be re-written as

$$\begin{aligned} E_{mi}E_{nj} = {} & A0_{minj} + A1_{minj}\eta_1 + A2_{minj}\eta_2 + \\ & A3_{minj}\eta_1^2 + A4_{minj}\eta_2^2 + A5_{minj}\eta_1\eta_2 \end{aligned} \tag{3.92}$$

where

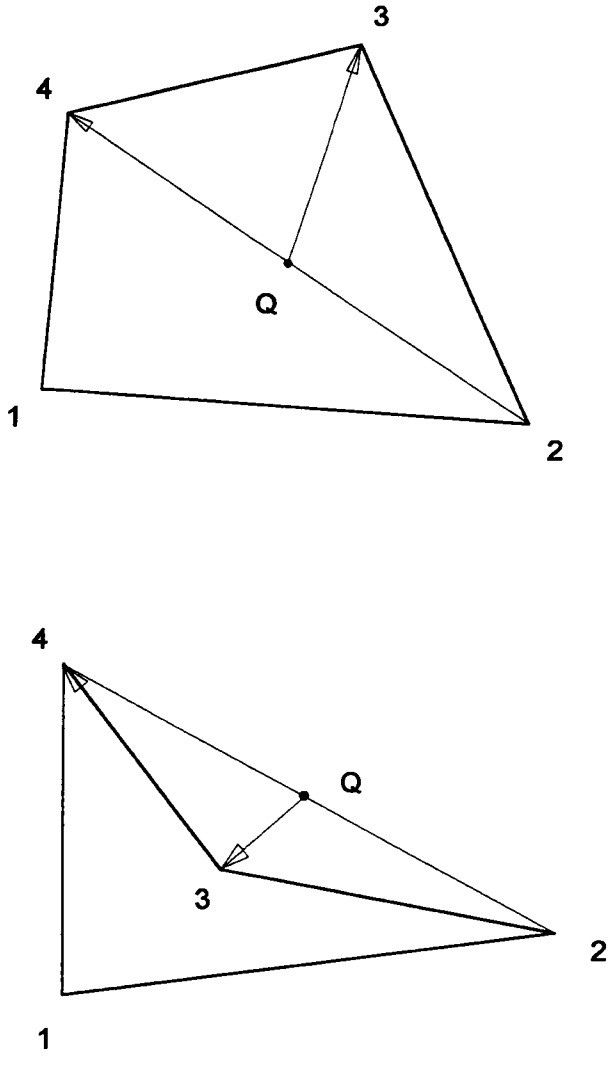


Figure 3.4: Bilinear FE Illustration (**II**): Convex and concave mesh cells.

$$\begin{cases} A0_{minj} &= \overline{E}^{*}_{mi}\overline{E}^{0}_{nj} \\ A1_{minj} &= (-1)^{n+j}\delta_{1n}\overline{E}^{0}_{j}\overline{E}^{*}_{mi} + (-1)^{m+i}\delta_{1m}\overline{E}^{0}_{i}\overline{E}^{*}_{nj} \\ A2_{minj} &= (-1)^{n+j}\delta_{2n}\overline{E}^{0}_{j}\overline{E}^{*}_{mi} + (-1)^{m+i}\delta_{2M}\overline{E}^{0}_{i}\overline{E}^{*}_{nj} \\ A3_{minj} &= (-1)^{m+i+n+j}\delta_{1m}\delta_{1n}\overline{E}^{0}_{i}\overline{E}^{0}_{j} \\ A4_{minj} &= (-1)^{m+i+n+j}\delta_{2m}\delta_{2n}\overline{E}^{0}_{i}\overline{E}^{0}_{j} \\ A5_{minj} &= (-1)^{m+i+n+j}(\delta_{1m}\delta_{2n} + \delta_{2m}\delta_{1n})\overline{E}^{0}_{i}\overline{E}^{0}_{j} \end{cases} \tag{3.93}$$

Subsequently,

$$\begin{aligned} \frac{E_{mi}E_{nj}}{|D|} &= \frac{1}{D^{*}}(P0_{minj} + P1_{minj}\eta_1 + P2_{minj}\eta_2 + \\ &\quad P3_{minj}\eta_1^2 + P4_{minj}\eta_2^2 + P5_{minj}\eta_1\eta_2 + \cdots) \end{aligned} \tag{3.94}$$

in which,

$$\begin{cases} P0_{minj} &= A0_{minj} \\ P1_{minj} &= A1_{minj} - A \cdot A0_{minj} \\ P2_{minj} &= A2_{minj} - B \cdot A0_{minj} \\ P3_{minj} &= A^2 \cdot A0_{minj} - A \cdot A1_{minj} + A3_{minj} \\ P4_{minj} &= B^2 \cdot A0_{minj} - B \cdot A1_{minj} + A4_{minj} \\ P5_{minj} &= 2 \cdot A \cdot B \cdot A0_{minj} - A \cdot A2_{minj} - B \cdot A1_{minj} + A5_{minj} \end{cases} \tag{3.95}$$

- Matrix Formula:

$$\begin{aligned} M2IJ^{e}_{\alpha\beta} &= (\overline{\frac{E_{MI}E_{NJ}}{D}})^{e}_{\underline{\alpha\beta}} \cdot B2MN_{\alpha\beta} \\ &= (\overline{\frac{E_{1I}E_{1J}}{D}})^{e}_{\underline{\alpha\beta}} \cdot B211_{\alpha\beta} + (\overline{\frac{E_{2I}E_{1J}}{D}})^{e}_{\underline{\alpha\beta}} \cdot B221_{\alpha\beta} + \\ &\quad (\overline{\frac{E_{1I}E_{2J}}{D}})^{e}_{\underline{\alpha\beta}} \cdot B212_{\alpha\beta} + (\overline{\frac{E_{2I}E_{2J}}{D}})^{e}_{\underline{\alpha\beta}} \cdot B222_{\alpha\beta} \end{aligned} \tag{3.96}$$

Considering

$$\begin{cases} B211_{\alpha\beta} = A211_{im}A200_{jn} \\ B221_{\alpha\beta} = A201_{im}A210_{jn} \\ B212_{\alpha\beta} = A210_{im}A201_{jn} \\ B222_{\alpha\beta} = A200_{im}A211_{jn} \end{cases} \tag{3.97}$$

where $\alpha \Leftrightarrow (i,j)$ and $\beta \Leftrightarrow (m,n)$, we have the following TP based MDM algorithm for the $(\overline{\frac{E_{MI}E_{NJ}}{D}})^e_{\underline{\alpha\beta}}$ in (3.96).

$$\begin{cases} (\overline{\frac{E_{1I}E_{1J}}{D}})^e_{\alpha\beta} = \frac{1}{D^*}(P0_{1I1J} + P1_{1I1J}\overline{\varsigma 211}_{im} + P2_{1I1J}\overline{\varsigma 200}_{jn} + \\ \qquad P3_{1I1J}\overline{\varsigma 211}^2_{im} + P4_{1I1J}\overline{\varsigma 200}^2_{jn} + P5_{1I1J}\overline{\varsigma 211}_{im}\overline{\varsigma 200}_{jn}) \\ (\overline{\frac{E_{2I}E_{1J}}{|D|}})^e_{\alpha\beta} = \frac{1}{|D^*|}(P0_{2I1J} + P1_{2I1J}\overline{\varsigma 201}_{im} + P2_{2I1J}\overline{\varsigma 210}_{jn} + \\ \qquad P3_{2I1J}\overline{\varsigma 201}^2_{im} + P4_{2I1J}\overline{\varsigma 210}^2_{jn} + P5_{2I1J}\overline{\varsigma 201}_{im}\overline{\varsigma 210}_{jn}) \\ (\overline{\frac{E_{1I}E_{2J}}{|D|}})^e_{\alpha\beta} = \frac{1}{|D^*|}(P0_{1I2J} + P1_{1I2J}\overline{\varsigma 210}_{im} + P2_{1I2J}\overline{\varsigma 201}_{jn} + \\ \qquad P3_{1I2J}\overline{\varsigma 210}^2_{im} + P4_{1I2J}\overline{\varsigma 201}^2_{jn} + P5_{1I2J}\overline{\varsigma 210}_{im}\overline{\varsigma 201}_{jn}) \\ (\overline{\frac{E_{2I}E_{2J}}{|D|}})^e_{\alpha\beta} = \frac{1}{|D^*|}(P0_{2I2J} + P1_{2I2J}\overline{\varsigma 200}_{im} + P2_{2I2J}\overline{\varsigma 211}_{jn} + \\ \qquad P3_{2I2J}\overline{\varsigma 200}^2_{im} + P4_{2I2J}\overline{\varsigma 211}^2_{jn} + P5_{2I2J}\overline{\varsigma 200}_{im}\overline{\varsigma 211}_{jn}) \end{cases} \tag{3.98}$$

Apparently, algorithm (3.98) is effective only if the bilinear FE mesh cell is *parallellgram* enough.

Submeshing Algorithm

Primarily, the S-MDM algorithm is featured as follows for bilinear element.

- It is *fast*, since the FE cell needs only to be submeshed into two triangles as

shown Figure 3.5:

$$M2IJ^e_{\alpha\beta} = \left(\frac{E_{MI}E_{NJ}}{D}\right)^e_1 B2MN^1_{\alpha\beta} + \left(\frac{E_{MI}E_{NJ}}{D}\right)^e_2 B2MN^2_{\alpha\beta} \qquad (3.99)$$

- It is *accurate* — in analytical sense, since the bilinear FE basis itself is linear on the edge of the mesh cell.

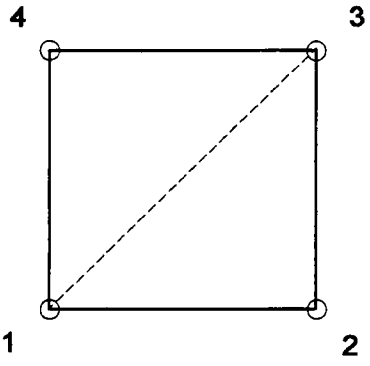
- Since $B2MN\alpha\beta \neq 0$ for sure, (3.99) is warranted of its "mean value" form:

$$M2IJ^e_{\alpha\beta} = \left(\frac{E_{MI}E_{NJ}}{D}\right)^e_{\underline{\alpha\beta}} B2MN_{\alpha\beta} \qquad (3.100)$$

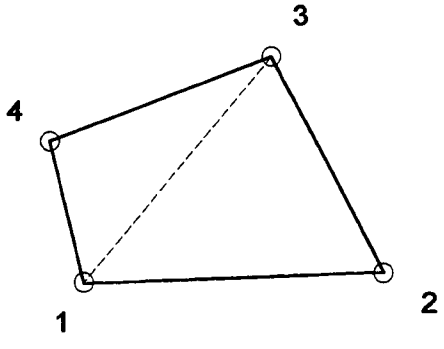
where

$$\left(\frac{E_{MI}E_{NJ}}{D}\right)^e_{\alpha\beta} = \left(\frac{E_{MI}E_{NJ}}{D}\right)^e_1 \frac{B2\underline{MN}^1_{\underline{\alpha\beta}}}{B2MN_{\alpha\beta}} + \left(\frac{E_{MI}E_{NJ}}{D}\right)^e_2 \frac{B2\underline{MN}^2_{\underline{\alpha\beta}}}{B2MN_{\alpha\beta}} \qquad (3.101)$$

It is obvious that, subsequent to these features, S-MDM should be the most appropriate scheme for bilinear diffusion matrices.



Master Element



Physical Element

Figure 3.5: Submeshes for the *submesh* based MDM alogorithm in bilinear element (two triangles).

Chapter 4

An Invariantly Rescaled, Selective Taylor Weak Statement Algorithm

This chapter is devoted to presenting the *selective* TWS (S-TWS) dissipation algorithm and its *invariant* rescaling. The presentation is organized in four sections. Section 4.1 is a general discussion about TWS; Section 4.2 characterizes S-TWS; Section 4.3 defines the "invariant dissipation length". The eventual "invariantly rescaled weak statements" are summarized in Section 4.4.

4.1 Taylor Weak Statement Method

4.1.1 Perturbation Algorithm

Essentially, the TWS algorithm is a procedure to *reenforce* the to-be-solved PDE to a *higher* order, and to retain the information that would be otherwise lost. For the convection problem (2.58-2.59), i.e.,

$$\frac{\partial q}{\partial t} + \frac{\partial C}{\partial t} + v_{0_j}\frac{\partial q}{\partial x_j} = 0 \tag{4.1}$$

TWS comprises the following aspects ([14, 15, 17, 18, 21]).

- **Taylor Expansion:**

$$\begin{aligned} & \frac{q^{n+1}-q^n}{\Delta t} + \frac{C^{n+1}-C^n}{\Delta t} \\ = \; & q_t^n + C_t^n + \frac{\Delta t}{2}\left(q_{tt}^n + C_{tt}^n\right) + 0\left(\Delta t^2\right) \\ = \; & -v_{0_j}\frac{\partial q}{\partial x_j} + \frac{\Delta t}{2}\left(q_{tt}^n + C_{tt}^n\right) + 0\left(\Delta t^2\right) \end{aligned} \tag{4.2}$$

- **Reenforcing the PDE:**

  Because of (4.1),

$$\begin{aligned} q_{tt} + C_{tt} &= -\frac{\partial}{\partial x_j}v_{0_j}\frac{\partial q}{\partial t} \\ &= \frac{\partial}{\partial x_j}v_{0_j}v_{0_k}\frac{\partial q}{\partial x_k} \end{aligned} \tag{4.3}$$

  and/thus,

$$q_{tt} + C_{tt} = \beta\frac{\partial}{\partial x_j}v_{0_j}v_{0_k}\frac{\partial q}{\partial x_k} - \alpha\frac{\partial}{\partial x_j}v_{0_j}\frac{\partial q}{\partial t} \tag{4.4}$$

where α and β are real numbers satisfying $\alpha + \beta = 1$. Consequently, (4.2) becomes

$$\frac{q^{n+1} - q^n}{\Delta t} + \frac{C^{n+1} - C^n}{\Delta t} + v_{0_j}\left(\frac{\partial q}{\partial x_j}\right)^n + \frac{\Delta t}{2}\frac{\partial}{\partial x_j}\left(\alpha v_{0_j}\frac{\partial q}{\partial t} - \beta v_{0_j}v_{0_k}\frac{\partial q}{\partial x_k}\right)^n = 0\left(\Delta t^2\right) \tag{4.5}$$

- **Companion Conservation Law:**

  Symbolically extremizing (4.5), by taking the limit of (4.5) for $\Delta t \to 0$, and yet maintaining the Δt terms, yields ([17, 18, 19, 20])

$$\frac{\partial q}{\partial t} + \frac{\partial C}{\partial t} + v_{0_j}\frac{\partial q}{\partial x_j} + \frac{\Delta t_*}{2}\frac{\partial}{\partial x_j}\left(\alpha v_{0_j}\frac{\partial q}{\partial t} - \beta v_{0_j}v_{0_k}\frac{\partial q}{\partial x_k}\right) = 0 \tag{4.6}$$

  which is the "TWS companion conservation law" ([20]), and in which Δt_* is what I call the *dissipation time.*

□

Note that, in (4.3-4.4-4.5), the "reenforcement" of the to-be-solved PDE has taken place twice, and they are combined convexly to initiate a parametric control mechanism. The "reenforced" Taylor expansion (4.5) as $\alpha = 0$ (or equivalently $\beta = 1$) is the second order case of the Taylor Galerkin algorithm (TG [14]). The symbolic extremization that brings about the "companion conservation law" is of profound importance, in that the temporal Taylor expansion that carries the above *perturbation* thus becomes irrelevant to how the *perturbed* system is to be time-integrated. As a result, the TWS algorithm supports both explicit and implicit computing — unlike the explicit TG. Noticing this, one should also notice that there is a *draw-back* in

association: The numerical solution now in general endures error of order $0(\Delta t_*)$, which in turn forces that the *dissipation time* Δt_* must be very small.

Note also that the above has not yet been a complete description of the TWS algorithm. To complete it, the Galerkin weak statement of (4.6) has to be included (see § 4.4).

4.1.2 Linear Analysis

Amplification Factor

Consider the linear/constant-coefficient case of (4.6), i.e.,

$$\frac{\partial q}{\partial t} + v_{0_j}\frac{\partial q}{\partial x_j} + \frac{\Delta t_*}{2}\frac{\partial}{\partial x_j}\left(\alpha v_{0_j}\frac{\partial q}{\partial t} - \beta v_{0_j}v_{0_k}\frac{\partial q}{\partial x_k}\right) = 0 \tag{4.7}$$

and its Fourier single wave:

$$q(t, x_j) = Q(t)\, e^{\mathbf{i}\sigma_j x_j} \tag{4.8}$$

where Q_t is the magnitude and σ_j the *wave number.*

Substituting (4.8) into (4.7), we have

$$\frac{Q'}{Q} = -\frac{\frac{\beta}{2}v_\sigma^2\Delta t_* + \mathbf{i}v_\sigma}{1 + \mathbf{i}\frac{\alpha}{2}v_\sigma\Delta t_*} \tag{4.9}$$

$$Q(t) = Q(0)\, e^{-\frac{\frac{\beta}{2}v_\sigma^2\Delta t_* + \mathbf{i}v_\sigma}{1+\mathbf{i}\frac{\alpha}{2}v_\sigma\Delta t_*}t} \tag{4.10}$$

where the denominator and numerator of the coefficient $\frac{\frac{\beta}{2}v_\sigma^2\Delta t_* + \mathbf{i}v_\sigma}{1+\mathbf{i}\frac{\alpha}{2}v_\sigma\Delta t_*}$, with $v_\sigma = v_{0_j}\sigma_j$ which is the *phase-speed,* account respectively for the time and space differential terms in (4.7).

Rationalizing the coefficient $\frac{\frac{\beta}{2}v_\sigma^2\Delta t_* + \mathrm{i}v_\sigma}{1+\mathrm{i}\frac{\alpha}{2}v_\sigma\Delta t_*}$ in (4.10) yields

$$\begin{aligned} Q(t) &= Q(0)\, e^{-\frac{\left(\frac{\beta}{2}v_\sigma^2\Delta t_* + \mathrm{i}v_\sigma\right)\left(1-\mathrm{i}\frac{\alpha}{2}v_\sigma\Delta t_*\right)}{1+\frac{\alpha^2}{4}v_\sigma^2\Delta t_*^2}t} \\ &= Q(0)\, e^{-\frac{\frac{1}{2}(\alpha+\beta)v_\sigma^2\Delta t_* + \mathrm{i}\left(v_\sigma - \frac{\alpha\beta}{4}v_\sigma^3\Delta t_*^2\right)}{1+\frac{\alpha^2}{4}v_\sigma^2\Delta t_*^2}t} \end{aligned} \tag{4.11}$$

Since $\alpha+\beta > 0$, $Q(t)$ hence $q(t, x_j)$ is *decaying* — as a result of the TWS dissipation.

The ratio

$$g(t) = \left| \frac{Q(t)}{Q(0)} \right| = e^{-\frac{(\alpha+\beta)v_\sigma^2\Delta t_*}{2+\frac{\alpha^2}{2}v_\sigma^2\Delta t_*^2}t} \tag{4.12}$$

is often called the *amplification factor*, which, under the condition that $\alpha+\beta \equiv 1$ is maintained, is a monotonely increasing function of α, with

$$g_{min} = g_{\alpha=0} = e^{-\frac{v_\sigma^2\Delta t_*}{2}t} \tag{4.13}$$

$$g_{max} = g_{\alpha=1} = e^{-\frac{v_\sigma^2\Delta t_*}{2+\frac{1}{2}v_\sigma^2\Delta t_*^2}t} \tag{4.14}$$

3rd Order TWS Perturbation

Procedure (4.2-4.6) may be pursued to even higher order, such as the following third order algorithm ([15, 21]).

$$\frac{\partial q}{\partial t} + v_{0_j}\frac{\partial q}{\partial x_j} + \frac{\Delta t_*}{2}\frac{\partial}{\partial x_j}\left(\alpha v_{0_j}\frac{\partial q}{\partial t} - \beta v_{0_j}v_{0_k}\frac{\partial q}{\partial x_k}\right) = \frac{\Delta t_*^2}{6}q_{ttt} \tag{4.15}$$

where q_{ttt}, similar to the q_{tt} in (4.2), can be *reenforced* parametrically:

$$q_{ttt} = -\kappa\frac{\partial}{\partial x_j}v_{0_j}\frac{\partial^2 q}{\partial t^2} + \gamma\frac{\partial}{\partial x_j}v_{0_j}v_{0_k}\frac{\partial q_t}{\partial x_k} - \vartheta\frac{\partial}{\partial x_j}v_{0_j}v_{0_k}v_{0_l}\frac{\partial^2 q}{\partial x_k\partial x_l} \tag{4.16}$$

Both the κ-and-ϑ-terms complicate the system *fundamentally*, in that they are either second order in time and third order in space differential, and/thus, force in

demand for either the multi-level time scheme or the Hermite C^1 elements in the subesequent FE algorithm. The viable 3rd order TWS dissipation algorithm is its γ-option (i.e., the case as $\kappa = \vartheta = 0$ and $\gamma = 1$):

$$\frac{\partial q}{\partial t} + \frac{\partial C}{\partial t} + v_{0_j}\frac{\partial q}{\partial x_j} + \frac{\Delta t_*}{2}\frac{\partial}{\partial x_j}\left(\alpha v_{0_j}\frac{\partial q}{\partial t} - \beta v_{0_j}v_{0_k}\frac{\partial q}{\partial x_k} - \frac{\Delta t_*}{3}v_{0_j}v_{0_k}\frac{\partial q_t}{\partial x_k}\right) = 0 \qquad (4.17)$$

The Fourier single wave for (4.17) obey

$$\frac{Q'}{Q} = -\frac{\frac{\beta}{2}v_\sigma^2\Delta t_* + \mathrm{i}v_\sigma}{1 + \frac{1}{6}v_\sigma^2\Delta t_*^2 + \mathrm{i}\frac{\alpha}{2}v_\sigma\Delta t_*} \qquad (4.18)$$

where $\frac{1}{4}v_\sigma^2\Delta t_*^2$ in the denominator on the RHS is contributing to the γ-term. Accordingly,

$$\begin{aligned} Q(t) &= Q(0)\,e^{-\frac{\frac{\beta}{2}v_\sigma^2\Delta t_*+\mathrm{i}v_\sigma}{1+\frac{1}{6}v_\sigma^2\Delta t_*^2+\mathrm{i}\frac{\alpha}{2}v_\sigma\Delta t_*}t} \\ &= Q(0)\,e^{-\frac{\left(\frac{\beta}{2}v_\sigma^2\Delta t_*+\mathrm{i}v_\sigma\right)\left(1+\frac{1}{6}v_\sigma^2\Delta t_*^2-\mathrm{i}\frac{\alpha}{2}v_\sigma\Delta t_*\right)}{\left(1+\frac{1}{6}v_\sigma^2\Delta t_*^2\right)^2+\frac{\alpha^2}{4}v_\sigma^2\Delta t_*^2}t} \\ &= Q(0)\,e^{-\frac{\frac{1}{2}(\alpha+\beta)v_\sigma^2\Delta t_*+\frac{\beta}{12}v_\sigma^4\Delta t_*^3+\mathrm{i}\left(v_\sigma-\frac{\alpha\beta}{4}v_\sigma^3\Delta t_*^2+\frac{1}{6}v_\sigma^3\Delta t_*^2\right)}{\left(1+\frac{1}{6}v_\sigma^2\Delta t_*^2\right)^2+\frac{\alpha^2}{4}v_\sigma^2\Delta t_*^2}t} \end{aligned} \qquad (4.19)$$

From (4.19), one sees that although the 3rd order γ-term participates in altering the amplification factor, it does so only in the presense of the 2nd order α-and-β-terms. If the latter is omitted, then

$$Q(t) = Q(0)\,e^{-\frac{\mathrm{i}v_\sigma}{1+\frac{1}{6}v_\sigma^2\Delta t_*^2}t} \qquad (4.20)$$

Consequently, the only efffect of a stand-alone γ-term is to *slow down* the phase-speed by factor $\frac{1}{1+\frac{1}{4}v_\sigma^2\Delta t_*^2}$. Such a slow-down can be very severe, especially for the

high order Fourier components. For instance, consider in a rectangular mesh $\Omega^h = L_1 \otimes L_2 \otimes L_3 = n_1\Delta x_1 \otimes n_2\Delta x_2 \otimes n_3\Delta x_3$ the Fourier component which is (a) *flat* in all but the jth direction, i.e., $\sigma_i \equiv 0$ except if $i = j$, and (b) the most frequent *oscillation* that the mesh can resolve, i.e., the one with $\sigma_j = n_j\pi$ or Δx_j being half its wave length:

Since $v_\sigma = v_{0_i}\sigma_i = \sigma_j v_{0_j}$,

$$\begin{aligned} Q(t) &= Q(0)\, e^{-\frac{\mathbf{i}v_\sigma}{1+\frac{1}{6}\left(\sigma_j v_{0_j}\Delta t_*\right)^2}t} \\ &= Q(0)\, e^{-\frac{\mathbf{i}v_\sigma}{1+\frac{1}{6}\left(\pi n_j v_{0_j}\Delta t_*\right)^2}t} \end{aligned} \tag{4.21}$$

Now suppose that Δt_* corresponds to the Courant number C_j^* in the jth direction, i.e., $v_{0_j}\Delta t_* = C_j^*\Delta x_j$. Then,

$$\begin{aligned} Q(t) &= Q(0)\, e^{-\frac{\mathbf{i}v_\sigma}{1+\frac{1}{6}\left(\pi C_j^* n_j \Delta x_j\right)^2}t} \\ &= Q(0)\, e^{-\frac{\mathbf{i}v_\sigma}{1+\frac{1}{6}\left(\pi C_j^* L_j\right)^2}t} \end{aligned} \tag{4.22}$$

Therefore, if you are computing using only the γ-term, and do not see *oscillations* as anticipated, you'd better take it as *a result of no surprise* and prolong your observation for some time, instead of jumping to the conclusion about stability in a hurry.

4.1.3 Real Flows

Euler/Navier-Stokes Equations

TWS algorithm has gained applications to both Euler and Navier-Stokes equations. The following are the "companion conservation laws" used ([15, 17, 18, 20, 21]).

$$\frac{\partial U}{\partial t} + \frac{\partial}{\partial x_j}\left(F_j - F_j^{\nu}\right) + \frac{\Delta t_*}{2}\frac{\partial}{\partial x_j}\left(\alpha A_j \frac{\partial U}{\partial t} - \beta A_j A_k \frac{\partial U}{\partial x_k}\right) = 0 \quad (4.23)$$

$$\frac{\partial U}{\partial t} + \frac{\partial}{\partial x_j}\left(F_j - F_j^{\nu}\right) - \frac{\Delta t_*}{2}\frac{\partial}{\partial x_j} v^2 \frac{\partial U}{\partial x_j} = 0 \quad (4.24)$$

$$\frac{\partial U}{\partial t} + \frac{\partial}{\partial x_j}\left(F_j - F_j^{\nu}\right) - \frac{\Delta t_*}{2}\frac{\partial}{\partial x_j} \bar{v}_j \bar{v}_k \frac{\partial U}{\partial x_k} = 0 \quad (4.25)$$

where A_j is the Jacobian matrix of flux-vector F_j, i.e., $A_j = \frac{\partial F_j}{\partial U}$, which is derived in full in Appendix C; v is the velocity norm ([20]); $\bar{v}_j$ is a *weighted* v_j as follows ([21]),

$$\bar{v}_j = diag\left\{1,\, 1+\delta_{\underline{j}i},\, 1\right\} v_j \quad (4.26)$$

$$= \begin{bmatrix} v_j & 0 & 0 & 0 & 0 \\ 0 & \left(1+\delta_{\underline{j}1}\right) v_j & 0 & 0 & 0 \\ 0 & 0 & \left(1+\delta_{\underline{j}2}\right) v_j & 0 & 0 \\ 0 & 0 & 0 & \left(1+\delta_{\underline{j}3}\right) v_j & 0 \\ 0 & 0 & 0 & 0 & v_j \end{bmatrix} \quad (4.27)$$

or

$$\bar{v}_1 = \begin{bmatrix} v_1 & 0 & 0 & 0 & 0 \\ 0 & 2v_1 & 0 & 0 & 0 \\ 0 & 0 & v_1 & 0 & 0 \\ 0 & 0 & 0 & v_1 & 0 \\ 0 & 0 & 0 & 0 & v_1 \end{bmatrix} \quad (4.28)$$

$$\bar{v}_2 = \begin{bmatrix} v_2 & 0 & 0 & 0 & 0 \\ 0 & v_2 & 0 & 0 & 0 \\ 0 & 0 & 2v_2 & 0 & 0 \\ 0 & 0 & 0 & v_2 & 0 \\ 0 & 0 & 0 & 0 & v_2 \end{bmatrix} \tag{4.29}$$

$$\bar{v}_3 = \begin{bmatrix} v_3 & 0 & 0 & 0 & 0 \\ 0 & v_3 & 0 & 0 & 0 \\ 0 & 0 & v_3 & 0 & 0 \\ 0 & 0 & 0 & 2v_3 & 0 \\ 0 & 0 & 0 & 0 & v_3 \end{bmatrix} \tag{4.30}$$

Equation (4.23) is the *standard* TWS algorithm, which results from the same procedure as that of (4.1) defined in § 4.1.1, except that, in the case of viscous flow, *selection* has been conducted to filter out the viscous fluxes ([20, 21]):

$$\begin{aligned} U_{tt} &= -\frac{\partial}{\partial t}\frac{\partial}{\partial x_j}\left(F_j - F_j^{\nu}\right) \\ &\simeq -\frac{\partial}{\partial x_j} A_j \frac{\partial U}{\partial t} \\ &= \frac{\partial}{\partial x_j} A_j \frac{\partial}{\partial x_k}\left(F_k - F_k^{\nu}\right) \\ &\simeq \frac{\partial}{\partial x_j} A_j A_k \frac{\partial U}{\partial x_k} \end{aligned} \tag{4.31}$$

Equation (4.24) and (4.25), on the other hand, formulates the v-and-$v_j v_k$-form *approximate* TWS (A-TWS) dissipation ([20, 21]), with the *approximations* being

respectively

$$A_j A_k \simeq v^2 \delta_{jk} \tag{4.32}$$

$$A_j A_k \simeq \bar{v}_j \bar{v}_k \tag{4.33}$$

which both are *diagonalized* coefficients that minimize the artificial TWS computing effort. However, these two approximations are not only "quite different", as pointed out in Chpater 1, but also very *distant* — in terms of the approaches that let them. In fact, they are derived via totally unconnective "approximations" which permit "bifurcation" almost everywhere, and that is another reason (in addition to the issue of CPF) why this author became convinced that it it is necessary to "dig" back into the TWS process, which ultimately leads to the development of the *selective* TWS algorithm.

Compressible Potential Flow

Finally, let me conclude this section "looking into" the compressible potential flow (CPF). Apparently, although the CPF model does not conform to the first order convection problem which the TWS method was intended/designed for, there is no reason why we can not "syntactically transplant" the idea — to "reenforce" it "to a *higher* order" and "to retain the information that would be otherwise lost".

As described in § 4.1.1, the process starts from Taylor-expanding the CPF solution:

$$\rho^{n+1} = \rho^n + \Delta t \rho_t + \frac{1}{2}\Delta t^2 \rho_{tt} + \cdots \tag{4.34}$$

$$\phi^{n+1} = \phi^n + \Delta t \phi_t + \frac{1}{2}\Delta t^2 \phi_{tt} + \cdots \tag{4.35}$$

which, after being "symbolically extremized", gives

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho \frac{\partial \phi}{\partial x_j}\right) = \frac{\Delta t_*}{2}\rho_{tt} \tag{4.36}$$

$$\frac{\partial \phi}{\partial t} + \frac{1}{2}\frac{\partial \phi}{\partial x_j}\frac{\partial \phi}{\partial x_j} + \frac{\gamma}{\gamma - 1}\left(\frac{T_\infty}{\rho_\infty^{\gamma-1}}\rho^{\gamma-1} - T_{0\infty}\right) = \frac{\Delta t_*}{2}\phi_{tt} \tag{4.37}$$

"Reenforcing" ρ_{tt} and ϕ_{tt} via (4.36-4.37), i.e., letting

$$\begin{aligned}
\rho_{tt} &= -\frac{\partial}{\partial x_j}\left(\frac{\partial \rho}{\partial t}\frac{\partial \phi}{\partial x_j} + \rho\frac{\partial}{\partial x_j}\frac{\partial \phi}{\partial t}\right) \\
&= \frac{\partial}{\partial x_j}\left(\frac{\partial \phi}{\partial x_j}\frac{\partial}{\partial x_k}\rho\frac{\partial \phi}{\partial x_k}\right) + \frac{\partial}{\partial x_j}\left(\rho\frac{\partial \phi}{\partial x_k}\frac{\partial^2 \phi}{\partial x_j \partial x_k} + \gamma\frac{T_\infty}{\rho_\infty^{\gamma-1}}\frac{\partial}{\partial x_j}\rho^{\gamma-1}\frac{\partial \rho}{\partial x_j}\right) \\
&= \frac{\partial}{\partial x_j}\left(\frac{\partial \phi}{\partial x_j}\frac{\partial}{\partial x_k}\rho\frac{\partial \phi}{\partial x_k} + \rho\frac{\partial \phi}{\partial x_k}\frac{\partial^2 \phi}{\partial x_j \partial x_k} + \frac{T_\infty}{\rho_\infty^{\gamma-1}}\frac{\partial \rho^\gamma}{\partial x_j}\right)
\end{aligned} \tag{4.38}$$

$$\begin{aligned}
\phi_{tt} &= -\frac{\partial \phi}{\partial x_j}\frac{\partial}{\partial x_j}\frac{\partial \phi}{\partial t} - \frac{\gamma}{\gamma - 1}\frac{T_\infty}{\rho_\infty^{\gamma-1}}\frac{\partial \rho^{\gamma-1}}{\partial t} \\
&= \frac{\partial \phi}{\partial x_j}\frac{\partial \phi}{\partial x_k}\frac{\partial^2 \phi}{\partial x_j \partial x_k} + \frac{\partial \phi}{\partial x_j}\gamma\frac{T_\infty}{\rho_\infty^{\gamma-1}}\rho^{\gamma-2}\frac{\partial \rho}{\partial x_j} + \gamma\frac{T_\infty}{\rho_\infty^{\gamma-1}}\rho^{\gamma-2}\frac{\partial}{\partial x_j}\rho\frac{\partial \phi}{\partial x_j} \\
&= \frac{\partial \phi}{\partial x_j}\frac{\partial \phi}{\partial x_k}\frac{\partial^2 \phi}{\partial x_j \partial x_k} + \frac{2\gamma}{\gamma - 1}\frac{T_\infty}{\rho_\infty^{\gamma-1}}\frac{\partial \phi}{\partial x_j}\frac{\partial \rho^{\gamma-1}}{\partial x_j} + \gamma\frac{T_\infty}{\rho_\infty^{\gamma-1}}\rho^{\gamma-1}\frac{\partial^2 \phi}{\partial x_j^2}
\end{aligned} \tag{4.39}$$

we have the TWS "companion conservation laws" as follows,

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho\frac{\partial \phi}{\partial x_j}\right) + \frac{\alpha\Delta t_*}{2}\frac{\partial}{\partial x_j}\left(\frac{\partial \rho}{\partial t}\frac{\partial \phi}{\partial x_j} + \rho\frac{\partial}{\partial x_j}\frac{\partial \phi}{\partial t}\right) - \\
&\frac{\beta\Delta t_*}{2}\frac{\partial}{\partial x_j}\left(\frac{\partial \phi}{\partial x_j}\frac{\partial}{\partial x_k}\rho\frac{\partial \phi}{\partial x_k} + \rho\frac{\partial \phi}{\partial x_k}\frac{\partial^2 \phi}{\partial x_j \partial x_k} + \frac{T_\infty}{\rho_\infty^{\gamma-1}}\frac{\partial \rho^\gamma}{\partial x_j}\right) = 0
\end{aligned} \tag{4.40}$$

$$\begin{aligned}
&\frac{\partial \phi}{\partial t} + \frac{1}{2}\frac{\partial \phi}{\partial x_j}\frac{\partial \phi}{\partial x_j} + \frac{\gamma}{\gamma - 1}\left(\frac{T_\infty}{\rho_\infty^{\gamma-1}}\rho^{\gamma-1} - T_{0\infty}\right) + \\
&\frac{\alpha\Delta t_*}{2}\left(\frac{\partial \phi}{\partial x_j}\frac{\partial}{\partial x_j}\frac{\partial \phi}{\partial t} + \frac{\gamma}{\gamma - 1}\frac{T_\infty}{\rho_\infty^{\gamma-1}}\frac{\partial \rho^{\gamma-1}}{\partial t}\right) - \\
&\frac{\beta\Delta t_*}{2}\left(\frac{\partial \phi}{\partial x_j}\frac{\partial \phi}{\partial x_k}\frac{\partial^2 \phi}{\partial x_j \partial x_k} + \frac{2\gamma}{\gamma - 1}\frac{T_\infty}{\rho_\infty^{\gamma-1}}\frac{\partial \phi}{\partial x_j}\frac{\partial \rho^{\gamma-1}}{\partial x_j} + \gamma\frac{T_\infty}{\rho_\infty^{\gamma-1}}\rho^{\gamma-1}\frac{\partial^2 \phi}{\partial x_j^2}\right) = 0
\end{aligned} \tag{4.41}$$

where α and β are the same parameters as in (4.3-4.6).

□

It is obvious that (4.40-4.41) may be numerically meaningful only if thay can be significantly simplified. Particularly, the simplification must go so far as eliminating those third order TWS *potential* terms, for otherwise, as in the case of the ϑ-option of (4.15-4.16), the Hermite C^1 element would have to be used.

4.2 Selective TWS Algorithm

Selective dissipation is a concept that has in Navier-Stokes computing — where, in both TWS and SUPG and for obvious reason, the viscous fluxes are always *selected* out of the process ([20, 21, 24, 25, 28, 30, 31, 34]). What is called *selective* TWS (S-TWS) algorithm in this dissertation is the same idea, except that the *selection* has been exercised to "filter out" not only the viscous terms but all that which are *non-convective*. In this section, I will define this method, and demonstrate that it can serve as a common platform resolving the issues of both A-TWS and CPF.

4.2.1 Convection Diffusion Problem

The best way illustrating the idea of S-TWS is to do it upon the *conservative* form convection/diffusion problem (2.56), which can be summarized as follows.

$$\frac{\partial q}{\partial t} + v_j\frac{\partial q}{\partial x_j} + q\frac{\partial v_j}{\partial x_j} - \frac{\partial}{\partial x_j}\mu\frac{\partial q}{\partial x_j} = 0 \tag{4.42}$$

$$\frac{\partial q}{\partial t} + \frac{\partial q v_j}{\partial x_j} - \frac{\partial}{\partial x_j}\mu\frac{\partial q}{\partial x_j} = \frac{\Delta t_*}{2}q_{tt} \tag{4.43}$$

$$\begin{aligned} q_{tt} &= -\tfrac{\partial}{\partial x_j}\left(v_j\tfrac{\partial q}{\partial t} + q\tfrac{\partial v_j}{\partial t} - \mu\tfrac{\partial q_t}{\partial x_j}\right) \\ &\simeq -\tfrac{\partial}{\partial x_j}v_j\tfrac{\partial q}{\partial t} \\ &= \tfrac{\partial}{\partial x_j}\left[v_j\left(v_k\tfrac{\partial q}{\partial x_k} + q\tfrac{\partial v_k}{\partial x_k} - \tfrac{\partial}{\partial x_k}\mu\tfrac{\partial q}{\partial x_k}\right)\right] \\ &\simeq \tfrac{\partial}{\partial x_j}v_j v_k\tfrac{\partial q}{\partial x_k} \end{aligned} \tag{4.44}$$

Hereby, (4.42) is the rewritten of (2.56), with the net flux $\frac{\partial q v_j}{\partial x_j}$ being decomposed into term $v_j\frac{\partial q}{\partial x_j}$ and $q\frac{\partial v_j}{\partial x_j}$, which reflect the convection and compression respectively;

(4.43) is the "companion conservation law" prior to "reenforcement"; (4.44) is the actual "reenforcement" of (2.56) or (4.42), — where *selection* has been exercised to filter out all but the convective term $v_j \frac{\partial q}{\partial x_j}$.

Note that I have only considered the $\alpha = 0$ and $\beta = 1$ case of the "reenforcement", which is the scope of this study, though apparently not a limit of S-TWS.

The combination of (4.44) and (4.43) is the eventual S-TWS dissipation scheme:

$$\frac{\partial q}{\partial t} + \frac{\partial q v_j}{\partial x_j} - \frac{\partial}{\partial x_j}\mu\frac{\partial q}{\partial x_j} - \frac{\Delta t_*}{2}\frac{\partial}{\partial x_j}v_j v_k \frac{\partial q}{\partial x_k} = 0 \tag{4.45}$$

Equation (4.45) may be reexpressed:

$$\frac{\partial q}{\partial t} + \frac{\partial \tilde{q} v_j}{\partial x_j} - \frac{\partial}{\partial x_j}\mu\frac{\partial q}{\partial x_j} = 0 \tag{4.46}$$

where

$$\begin{aligned} \tilde{q} &= q(t, x_k) - \frac{1}{2}v_k \Delta t_* \frac{\partial q}{\partial x_k} \\ &\simeq q\left(t, x_k - \frac{1}{2}v_k \Delta t_*\right) + 0\left(\Delta t_*^2\right) \end{aligned} \tag{4.47}$$

Consequently, one knows that the net effect of the *seclected* dissipation (4.44) is that *flux* has been *upwindingly* dislocated, — with the dislocation being for half the distence that the velocity travels in *dissipation time.*

Remarks:

1. Equation (4.45) is identical to (4.6) in the case of "linear and constant coefficient". Accordingly, the "linear analysis" for (4.6) as in § 4.1.2) applies identically to (4.45).

2. Part of the "selections" in (4.44) is the omission of $\frac{\partial v_j}{\partial t}$. In other words, what has been "selected" is not only *out* the viscosity and compressiblity, but *in* a temporarily frozen velocity field.

□

4.2.2 Compressible Potential Flow

Dissipation Scheme

To "potentially"apply the S-TWS dissipation procedure, one needs only to notice these:

- The continuity equation of CPF, i.e., (2.54), is identical to (2.56), — except that the convected variable is density ρ, and the velocity v_j has potential ϕ.
- The energy equation (2.54), at the same time, is *entirely* unconvective, — since, as discussed in Chapter 2, the energy convection in CPF has been Cauchy-Lagrange integrated out.

Consequently, by "selecting out"the energy equation entirely, and by *selectively* dissipating the continuity equation as (2.56) in § 4.2.1, we have the S-TWS "companion conservation laws" for CPF as below.

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho \frac{\partial \phi}{\partial x_j}\right) - \frac{\partial}{\partial x_j}\left(\frac{1}{2}\Delta t_* v_j v_k \frac{\partial \rho}{\partial x_k}\right) = 0 \tag{4.48}$$

$$\frac{\partial \phi}{\partial t} + \frac{1}{2}\frac{\partial \phi}{\partial x_j}\frac{\partial \phi}{\partial x_j} + \frac{\gamma}{\gamma - 1}\left(\frac{T_\infty}{\rho_\infty^{\gamma-1}}\rho^{\gamma-1} - T_{0\infty}\right) = 0 \tag{4.49}$$

Like (4.45), equation (4.48) can be rewritten:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\tilde{\rho}\frac{\partial \phi}{\partial x_j}\right) = 0 \tag{4.50}$$

where, as the $\tilde{q}$ in (4.46),

$$\begin{aligned} \tilde{\rho} &= \rho(t, x_k) - \frac{1}{2} v_k \Delta t_* \frac{\partial \rho}{\partial x_k} \\ &= \rho\left(t,\, x_k - \frac{1}{2} v_k \Delta t_*\right) + 0\left(\Delta t_*^2\right) \end{aligned} \tag{4.51}$$

□

The above "selected"TWS dissipation is apparently *simple* enough, — compared to the "syntatically transplanted" one as in (4.40-4.41). The question, however, is whether it is *good* enough to sustain stability. In the next few paragraphs, this will be examined in the perspection of "supersonic characteristics" and "positivity analysis".

Supersonic Characteristics

Consider the temporal integration of the continuity equation over the dissipation time Δt_* in a supersonic flow-field:

$$\frac{(\rho^{n+1} - \rho^n)_A}{\Delta t_*} + \left(\frac{\partial \rho v_j}{\partial x_j}\right)_A^{n+\theta} = 0 \tag{4.52}$$

where $0 \leq \theta \leq 1$, and A denotes the spatial location of the integration which could be anywhere not in the vincinity of a shock. The "domain of dependence" of A is illustrated in Figure 4.1, about which the *characteristics* are

$$\sin\alpha = \frac{1}{M} \tag{4.53}$$

$$\rho_A = \rho(R_I, R_{II}) \tag{4.54}$$

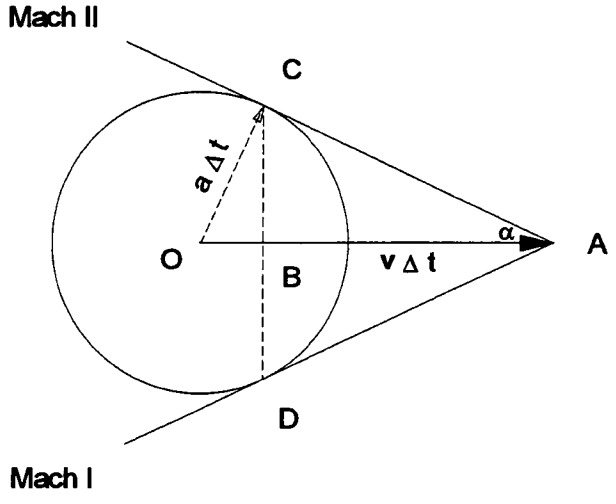


Figure 4.1: Illustration for the supersonic domain of dependence.

where M, R_I and R_{II} are respectively the Mach number and the Rieman invariants along Mach wave I and II.

Since Δt_* is small, the following approximation holds.

$$\begin{aligned}
\rho_A &\simeq \frac{1}{2}(\rho_C + \rho_D) \\
&\simeq \rho_B \\
&\simeq \rho_A + (\rho_B - \rho_A) \\
&\simeq \rho_A - (\nabla\rho)_A \cdot \left[\mathbf{v}\Delta t_* - \frac{\mathbf{v}}{v}(a\Delta t_* \sin\alpha)\right] \\
&= \rho_A - \left(1 - \frac{1}{M^2}\right) v_k \Delta t_* \left(\frac{\partial\rho}{\partial x_k}\right)
\end{aligned} \tag{4.55}$$

Subsequently, (4.52) becomes

$$\frac{(\rho^{n+1} - \rho^n)_A}{\Delta t_*} + \left(\frac{\partial \rho v_j}{\partial x_j}\right)_A^{n+\theta} - \frac{\partial}{\partial x_j}\left[\left(1 - \frac{1}{M^2}\right) v_j v_k \Delta t_* \frac{\partial\rho}{\partial x_k}\right]_A^{n+\theta} \simeq 0 \tag{4.56}$$

Now assuming that what we have got here is *characteristically* correct, so then must be the S-STWS, reflected in that (4.48) and (4.56) are fully consistent in terms of *directionality*. The "error" of S-TWS is that it *upwindingly* off-sets ρ to a different location on the streamline — the mid-point of $\overline{OA}$ (see (4.50-4.51) and Figure 4.1) — instead of point B as in (4.55-4.56). Such an "error", while vanishing at $M = \sqrt{2}$, i.e., when B rides on the mid-point of $\overline{OA}$, reflects

- an over-dissipation as $M < \sqrt{2}$, since that is when B is *downwinding* to the mid-$\overline{OA}$, and
- an under-dissipation as $M > \sqrt{2}$, since that is when B is *upwinding* to the mid-$\overline{OA}$.

However, because of the *correctness* of S-TWS in terms of *directionality*, its "error" may be simply corrected adapting to the "supersonic characteristics":

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho \frac{\partial \phi}{\partial x_j}\right) - \frac{\partial}{\partial x_j}\left[\max\left(0, 1 - \frac{1}{M^2}\right) \Delta t_* v_j v_k \frac{\partial \rho}{\partial x_k}\right] = 0 \qquad (4.57)$$

Equation (4.57), coupled with (4.49), are the actual S-TWS "companion conservation laws" for CPF that are used in this study. Note that we have not only adapted (4.48) to (4.56), but modified the factor into a switcher on the sonic line — to take into consideration the subsonic flow ([76, 77, 78, 79]).

Positivity Analysis

1. Justfication. In "priminitive" terms, the continuity embedded dissipation algorithm for CPF can be rewritten as ([76, 77, 78, 79])

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho v_j}{\partial x_j} - \frac{\partial}{\partial x_j}\left(\mu_{jk} \frac{\partial \rho}{\partial x_k}\right) = 0 \qquad (4.58)$$

$$\rho \frac{D}{Dt}\left(\frac{1}{2} v_i^2\right) + v_i \frac{\partial p}{\partial x_i} = 0 \qquad (4.59)$$

where μ_{jk} is "artificial viscosity", and (4.59) the kinetic energy equation, which is theoretically equivalent to the CPF (4.49) though not yet being Cauchy-Lagrange integrated. Equation (4.58), the *dissipated* continuity, may also be equivalently expressed as

$$\frac{D\rho}{Dt} + \rho \frac{\partial v_j}{\partial x_j} - \frac{\partial}{\partial x_j}\left(\mu_{jk} \frac{\partial \rho}{\partial x_k}\right) = 0 \qquad (4.60)$$

In coping with (4.58-4.59), the following process holds:

$$L(E): \quad \frac{\partial}{\partial t}\left[\rho\left(e + \frac{1}{2} v_i^2\right)\right] + \frac{\partial}{\partial x_j}\left[\rho v_j \left(e + \frac{1}{2} v_i^2 + \frac{p}{\rho}\right)\right] = 0 \qquad (4.61)$$

$$\Longrightarrow \quad \rho\frac{D}{Dt}\left(e+\frac{1}{2}v_i^2\right)+\left(e+\frac{1}{2}v_i^2\right)\left(\frac{\partial\rho}{\partial t}+\frac{\partial\rho v_j}{\partial x_j}\right)+\frac{\partial pv_j}{\partial x_j} = 0 \qquad (4.62)$$

$$\Longrightarrow \quad \rho\frac{D}{Dt}\left(e+\frac{1}{2}v_i^2\right)+\frac{\partial pv_j}{\partial x_j}+\left(e+\frac{1}{2}v_i^2\right)\frac{\partial}{\partial x_j}\left(\mu_{jk}\frac{\partial\rho}{\partial x_k}\right) = 0 \qquad (4.63)$$

$$\Longrightarrow \quad \rho\frac{De}{Dt}+p\frac{\partial v_j}{\partial x_j}+\left(e+\frac{1}{2}v_i^2\right)\frac{\partial}{\partial x_j}\left(\mu_{jk}\frac{\partial\rho}{\partial x_k}\right) = 0 \qquad (4.64)$$

where, (4.61) is the *total* energy equation of the Euler system (2.38); (4.62) is the expansion of (4.61); (4.63) discounts (4.58) from (4.62); equation (4.64) is the internal energy equation, which results from further discounting the kinetic energy equation (4.59).

Consequently, according to the Gibbs equality,

$$\begin{aligned}
\rho T\frac{DS}{Dt} &\equiv \rho\left[\frac{De}{Dt}+p\frac{D}{Dt}\left(\frac{1}{\rho}\right)\right] \\
&\equiv \rho\frac{De}{Dt}-\frac{p}{\rho}\frac{D\rho}{Dt} \\
&= -\left(e+\frac{p}{\rho}+\frac{1}{2}v_i^2\right)\frac{\partial}{\partial x_j}\left(\mu_{jk}\frac{\partial\rho}{\partial x_k}\right) \\
&= -h_0\frac{\partial}{\partial x_j}\left(\mu_{jk}\frac{\partial\rho}{\partial x_k}\right) \qquad (4.65)
\end{aligned}$$

where $h_0 = \left(e+\frac{p}{\rho}+\frac{1}{2}v_i^2\right)$ is the total enthalpy.

Now consider the following integration of (4.65),

$$\begin{aligned}
\int_\Omega \frac{\rho^2 T}{h_0}\frac{DS}{Dt}\,d\Omega &= -\int_\Omega \rho\frac{\partial}{\partial x_j}\left(\mu_{jk}\frac{\partial\rho}{\partial x_k}\right)d\Omega \\
&= -\oint_{\partial\Omega}\rho\mu_{jk}\frac{\partial\rho}{\partial x_k}n_j d\sigma+\int_\Omega\frac{\partial\rho}{\partial x_j}\mu_{jk}\frac{\partial\rho}{\partial x_k}\,d\Omega \\
&= \int_\Omega\frac{\partial\rho}{\partial x_j}\mu_{jk}\frac{\partial\rho}{\partial x_k}\,d\Omega \qquad (4.66)
\end{aligned}$$

where Ω, the domain of the integration, is an *isolated system*, which is why the boundary integral vanishes: Since the left-hand-side of (4.66) is positive (ruled by

the Second Law of Thermodynamics), so is the right-hand-side — regardless to the values of $\frac{\partial \rho}{\partial x_j}$, in other words, μ_{jk} *is* PD.

□

2. Examination. Having said the above, let us examine PD of the S-TWS dissipation, comparatively with that of the Jameson and Habashi scheme which are the two *mainstream* CPF algorithms ([76, 77, 78, 79]), i.e.,

$$\mu_{jk} = \max\left(0, 1 - \frac{1}{M^2}\right) \Delta t_* \, v_j \, v_k \tag{4.67}$$

$$\mu^J_{jk} = \max\left(0, 1 - \frac{1}{M^2}\right) \Delta x_{\underline{j}} \, v_{\underline{j}} \, \delta_{jk} \tag{4.68}$$

$$\mu^H_{jk} = \max\left(0, 1 - \frac{1}{M^2}\right) \Delta x_{\underline{j}} \, \frac{v_j v_k}{v} \tag{4.69}$$

By definition, we only need to check if $Y_i \mu_{jk} Y_j$ is PD — positive regardless to what Y_i is as long as nonzero. The results are as follows.

- The S-TWS Algorithm:

$$\begin{aligned} Y_j \, \mu_{jk} \, Y_k &= \max\left(0, 1 - \frac{1}{M^2}\right) \Delta t_* \, Y_j \, v_j \, v_k Y_k \\ &= \max\left(0, 1 - \frac{1}{M^2}\right) \Delta t_* \, (Y_j \, v_j)^2 \end{aligned} \tag{4.70}$$

  which is a just squared-sum, and/thus PD.

- The Jameson algorithm: ([78, 79]):

$$\begin{aligned} Y_j \, \mu^J_{jk} \, Y_k &= \max\left(0, 1 - \frac{1}{M^2}\right) Y_j \, \Delta x_{\underline{j}} \, v_{\underline{j}} \, \delta_{jk} \, Y_k \\ &= \max\left(0, 1 - \frac{1}{M^2}\right) Y_j \, v_j \, \Delta x_j Y_j \\ &= \max\left(0, 1 - \frac{1}{M^2}\right) \left(Y_1^2 v_1 \Delta x_1 + Y_2^2 v_2 \Delta x_2 + Y_3^2 v_3 \Delta x_3\right) \end{aligned} \tag{4.71}$$

which is not PD. For instance, if v_i are not all positive, such as $v_1 > 0$, $v_3 > 0$ and $v_2 < 0$, a $Y_i = \{0, 1, 0\}^T$ leads to

$$\begin{aligned} Y_j \, \mu^J_{jk} \, Y_k &= \max\left(0, 1 - \frac{1}{M^2}\right) Y_2^2 \, v_2 \, \Delta x_2 \\ &< 0 \end{aligned} \tag{4.72}$$

- The Habashi algorithm:

$$\begin{aligned} Y_j \, \mu^H_{jk} \, Y_k &= \max\left(0, 1 - \frac{1}{M^2}\right) Y_j \, \Delta x_j \, \frac{v_j v_k}{v} \, Y_k \\ &= \max\left(0, 1 - \frac{1}{M^2}\right) (Y_j v_j \Delta x_j) \left(\frac{v_k}{v} Y_k\right) \end{aligned} \tag{4.73}$$

which is PD neither. For example, considering such a rectangular but non-uniform mesh as $\Delta x_1 = \Delta x_3 = \Delta x$ and $\Delta x_2 = 3\Delta x$, and such a velocity field $v_1 = v_3 = \frac{v}{\sqrt{3}}$ and $v_2 = -\frac{v}{\sqrt{3}}$, we then have

$$\begin{aligned} & Y_j \, \mu^H_{jk} \, Y_k \\ = \quad & \max\left(0, 1 - \frac{1}{M^2}\right) (Y_j \, v_j \, \Delta x_j) \left(\frac{v_k}{v} Y_k\right) \\ = \quad & \max\left(0, 1 - \frac{1}{M^2}\right) \frac{v \Delta x}{3} (Y_1 - 3Y_2 + Y_3) (Y_1 - Y_2 + Y_3) \end{aligned} \tag{4.74}$$

which is *negative* for, say, $Y_i = \{1, 1, 1\}^T$.

4.2.3 Euler Equations

Recommended S-TWS Algorithm

As (4.42) for (2.56), Euler equations (2.38) can be re-written as

$$L(\rho): \quad \frac{\partial \rho}{\partial t} + v_j \frac{\partial \rho}{\partial x_j} + \rho \frac{\partial v_j}{\partial x_j} = 0 \tag{4.75}$$

$$L(m_i): \quad \frac{\partial m_i}{\partial t} + v_j \frac{\partial m_i}{\partial x_j} + m_i \frac{\partial v_j}{\partial x_j} + \frac{\partial p}{\partial x_i} = 0 \tag{4.76}$$

$$L(E): \quad \frac{\partial E}{\partial t} + v_j \frac{\partial}{\partial x_j}(E+p) + (E+p)\frac{\partial v_j}{\partial x_j} = 0 \tag{4.77}$$

which, from the S-TWS point of view, differs from (4.42) mainly in that it has in presence the pressure terms, and/hence, necessitates additional "selections".

Physically, pressure is the reason for the *wave* phenomenon in the flow-field, which is, on the one hand, *convected* with flow to the downstream, and on the other, *propagated* homogeneously to all fronts. Mathematically, the pressure terms differ, depending on where they are, in their reflecting the *convective* and *propagative* nature of flow. In the momentum equations, pressure is a force, wherein the *convectiveness* and *propagativeness* of flow are inseparable and they interact. In the energy equation, pressure is part of the total enthalpy $(E + p)$, which is purely convective (though in a compressible velocity field, as the q in (2.56) or (4.42)). The interaction between *convection* and propagation also varies, with the *convection* being more and more dominant as the flow speed rises, including to the extent that the "domain of dependence" is bounded (in the case of supersonics, shown in Figure 4.1). In addition, pressure is the sole variable coupling (4.75-4.76-4.77) after the presumed "frozen" velocity field (see § 4.2.1).

Based on the above understanding, and considering that indeed high speed flow is the interest of this study, I "select" to assess the the convective nature of the pressure terms in *full* yet a very limited fashion (in addition to the "selection" described in § 4.2.1):

A. Acknowledge *all* the pressure terms in the S-TWS procedure.

B. Do it only in so far as without involving cross-reference among the members of Euler system (4.75-4.76-4.77).

While **A** is self-evident, **B** is to ensure the diagonality of the ultimate dissiption scheme (the same thought as in the v-and-$v_j v_k$-form dissipation [17, 20, 21]).

The following is an outline of the so derived algorithm.

1. Initial "companion conservation laws".

$$\frac{\partial U}{\partial t} + \frac{\partial F_j}{\partial x_j} = \frac{\Delta t_*}{2}\frac{\partial^2 U}{\partial t^2} \tag{4.78}$$

or

$$\frac{\partial \rho}{\partial t} + v_j\frac{\partial \rho}{\partial x_j} + \rho\frac{\partial v_j}{\partial x_j} = \frac{\Delta t_*}{2}\frac{\partial^2 \rho}{\partial t^2} \tag{4.79}$$

$$\frac{\partial m_i}{\partial t} + v_j\frac{\partial m_i}{\partial x_j} + m_i\frac{\partial v_j}{\partial x_j} + \frac{\partial p}{\partial x_i} = \frac{\Delta t_*}{2}\frac{\partial^2 m_i}{\partial t^2} \tag{4.80}$$

$$\frac{\partial E}{\partial t} + v_j\frac{\partial}{\partial x_j}(E+p) + (E+p)\frac{\partial v_j}{\partial x_j} = \frac{\Delta t_*}{2}\frac{\partial^2 E}{\partial t^2} \tag{4.81}$$

2. Reenforcing the continuity equation.

$$\frac{\partial^2 \rho}{\partial t^2} = -\frac{\partial}{\partial x_j}\left(v_j\frac{\partial \rho}{\partial t} + \rho\frac{\partial v_j}{\partial t}\right)$$

$$
\begin{aligned}
&\simeq -\frac{\partial}{\partial x_j} v_j \frac{\partial \rho}{\partial t} \\
&= \frac{\partial}{\partial x_j}\left[v_j\left(v_k\frac{\partial \rho}{\partial x_k} + \rho\frac{\partial v_k}{\partial x_k}\right)\right] \\
&\simeq \frac{\partial}{\partial x_j} v_j v_k \frac{\partial \rho}{\partial x_k}
\end{aligned}
\tag{4.82}
$$

which is identical to the (4.44) for (4.42-4.43).

3. Reenforcing the momentum equation.

$$
\begin{aligned}
\frac{\partial^2 m_i}{\partial t^2} &= -\frac{\partial}{\partial x_j}\left(v_j\frac{\partial m_i}{\partial t} + m_i\frac{\partial v_j}{\partial t} + \delta_{ij}\frac{\partial p}{\partial t}\right) \\
&\simeq -\frac{\partial}{\partial x_j}\left(v_j\frac{\partial m_i}{\partial t} + \delta_{ij}\frac{\partial p}{\partial t}\right) \\
&\simeq -\frac{\partial}{\partial x_j}\left(v_j\frac{\partial m_i}{\partial t} + \delta_{ij}\frac{\partial p}{\partial m_{\underline{i}}}\frac{\partial m_{\underline{i}}}{\partial t}\right) \\
&= -\frac{\partial}{\partial x_j}\left[\left(v_j - (\gamma-1)\, v_{\underline{i}}\delta_{\underline{i}j}\right)\frac{\partial m_i}{\partial t}\right] \\
&= \frac{\partial}{\partial x_j}\left[\left(v_j - (\gamma-1)\, v_{\underline{i}}\delta_{\underline{i}j}\right)\left(v_k\frac{\partial m_i}{\partial x_k} + m_i\frac{\partial v_k}{\partial x_k} + \frac{\partial p}{\partial x_i}\right)\right] \\
&\simeq \frac{\partial}{\partial x_j}\left[\left(v_j - (\gamma-1)\, v_{\underline{i}}\delta_{\underline{i}j}\right)\left(v_k\frac{\partial m_i}{\partial x_k} + \delta_{ik}\frac{\partial p}{\partial x_k}\right)\right] \\
&\simeq \frac{\partial}{\partial x_j}\left[\left(v_j - (\gamma-1)\, v_{\underline{i}}\delta_{\underline{i}j}\right)\left(v_k\frac{\partial m_i}{\partial x_k} + \delta_{ik}\frac{\partial p}{\partial m_{\underline{i}}}\frac{\partial m_{\underline{i}}}{\partial x_k}\right)\right] \\
&= \frac{\partial}{\partial x_j}\left[\left(v_j - (\gamma-1)\, v_{\underline{i}}\delta_{\underline{i}j}\right)\left(v_k - (\gamma-1)\, v_{\underline{i}}\delta_{\underline{i}k}\right)\frac{\partial m_i}{\partial x_k}\right]
\end{aligned}
\tag{4.83}
$$

where "selection" **B** has been exercised *twice*, taking (for the *i*the momentum equation)

$$
\begin{aligned}
\frac{\partial p}{\partial t} &= \frac{\partial p}{\partial \rho}\frac{\partial \rho}{\partial t} + \frac{\partial p}{\partial m_j}\frac{\partial m_j}{\partial t} + \frac{\partial p}{\partial E}\frac{\partial E}{\partial t} \\
&\simeq \frac{\partial p}{\partial m_{\underline{i}}}\frac{\partial m_{\underline{i}}}{\partial t} \\
\frac{\partial p}{\partial x_k} &= \frac{\partial p}{\partial \rho}\frac{\partial \rho}{\partial x_k} + \frac{\partial p}{\partial m_j}\frac{\partial m_j}{\partial x_k} + \frac{\partial p}{\partial E}\frac{\partial E}{\partial x_k}
\end{aligned}
\tag{4.84}
$$

$$\simeq \frac{\partial p}{\partial m_i}\frac{\partial m_i}{\partial x_k} \tag{4.85}$$

with

$$\frac{\partial p}{\partial m_i} = (1-\gamma)\,v_i \tag{4.86}$$

which is derived in Appendix C.

4. Reenforcing the energy equation.

$$\begin{aligned}
\frac{\partial^2 E}{\partial t^2} &= -\frac{\partial}{\partial x_j}\left[v_j\frac{\partial}{\partial t}(E+p) + (E+p)\frac{\partial v_j}{\partial t}\right] \\
&\simeq -\frac{\partial}{\partial x_j}\left(v_j\frac{\partial E}{\partial t} + v_j\frac{\partial p}{\partial t}\right) \\
&\simeq -\frac{\partial}{\partial x_j}\left(v_j\frac{\partial E}{\partial t} + v_j\frac{\partial p}{\partial E}\frac{\partial E}{\partial t}\right) \\
&= \frac{\partial}{\partial x_j}\left[(v_j + (\gamma-1)\,v_j)\left(v_k\frac{\partial}{\partial x_k}(E+p) + (E+p)\frac{\partial v_k}{\partial x_k}\right)\right] \\
&\simeq \frac{\partial}{\partial x_j}\left[(v_j + (\gamma-1)\,v_j)\left(v_k\frac{\partial E}{\partial x_k} + v_k\frac{\partial p}{\partial x_k}\right)\right] \\
&\simeq \frac{\partial}{\partial x_j}\left[(v_j + (\gamma-1)\,v_j)\left(v_k\frac{\partial E}{\partial x_k} + v_k\frac{\partial p}{\partial E}\frac{\partial E}{\partial x_k}\right)\right] \\
&= \frac{\partial}{\partial x_j}\left[(v_j + (\gamma-1)\,v_j)(v_k + (\gamma-1)\,v_k)\frac{\partial E}{\partial x_k}\right] \\
&= \frac{\partial}{\partial x_j}\left[(\gamma v_j)(\gamma v_k)\frac{\partial E}{\partial x_k}\right]
\end{aligned} \tag{4.87}$$

Twice again, **B** has been applied, though here it is for E, i.e.,

$$\begin{aligned}
\frac{\partial p}{\partial t} &= \frac{\partial p}{\partial \rho}\frac{\partial \rho}{\partial t} + \frac{\partial p}{\partial m_j}\frac{\partial m_j}{\partial t} + \frac{\partial p}{\partial E}\frac{\partial E}{\partial t} \\
&\simeq \frac{\partial p}{\partial E}\frac{\partial E}{\partial t}
\end{aligned} \tag{4.88}$$

$$\begin{aligned}
\frac{\partial p}{\partial x_k} &= \frac{\partial p}{\partial \rho}\frac{\partial \rho}{\partial x_k} + \frac{\partial p}{\partial m_j}\frac{\partial m_j}{\partial x_k} + \frac{\partial p}{\partial E}\frac{\partial E}{\partial x_k} \\
&\simeq \frac{\partial p}{\partial E}\frac{\partial E}{\partial x_k}
\end{aligned} \tag{4.89}$$

where

$$\frac{\partial p}{\partial E} = \gamma - 1 \tag{4.90}$$

which is also derived in Appendix C.

5. End result.

Substituting (4.82), (4.83) and (4.87) into (4.78) or (4.79-4.80-4.81), we have the end "companion conservation laws" — a $w_j w_k$-form:

$$\frac{\partial U}{\partial t} + \frac{\partial F_j}{\partial x_j} - \frac{\partial}{\partial x_j}\frac{\Delta t_*}{2} w_j w_k \frac{\partial U}{\partial x_k} = 0 \tag{4.91}$$

where

$$\begin{aligned} w_j &= diag\left\{v_j,\ v_j - (\gamma-1)\, v_i \delta_{\underline{i}j},\ \gamma v_j\right\} \qquad (4.92) \\ &= diag\left\{v_j,\ v_j - (\gamma-1)\, v_j \delta_{i\underline{j}},\ \gamma v_j\right\} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1-(\gamma-1)\,\delta_{1\underline{j}} & 0 & 0 & 0 \\ 0 & 0 & 1-(\gamma-1)\,\delta_{2\underline{j}} & 0 & 0 \\ 0 & 0 & 0 & 1-(\gamma-1)\,\delta_{3\underline{j}} & 0 \\ 0 & 0 & 0 & 0 & \gamma \end{bmatrix} v_j \end{aligned}$$

or

$$w_1 = \begin{bmatrix} v_1 & 0 & 0 & 0 & 0 \\ 0 & (2-\gamma)\, v_1 & 0 & 0 & 0 \\ 0 & 0 & v_1 & 0 & 0 \\ 0 & 0 & 0 & v_1 & 0 \\ 0 & 0 & 0 & 0 & \gamma v_1 \end{bmatrix} \tag{4.93}$$

$$w_2 = \begin{bmatrix} v_2 & 0 & 0 & 0 & 0 \\ 0 & v_2 & 0 & 0 & 0 \\ 0 & 0 & (2-\gamma)\, v_2 & 0 & 0 \\ 0 & 0 & 0 & v_2 & 0 \\ 0 & 0 & 0 & 0 & \gamma v_2 \end{bmatrix} \tag{4.94}$$

$$w_3 = \begin{bmatrix} v_3 & 0 & 0 & 0 & 0 \\ 0 & v_3 & 0 & 0 & 0 \\ 0 & 0 & v_3 & 0 & 0 \\ 0 & 0 & 0 & (2-\gamma)\, v_3 & 0 \\ 0 & 0 & 0 & 0 & \gamma v_3 \end{bmatrix} \tag{4.95}$$

□

Note that, as shown in (4.83), the "pressure contribution" has actually *discounted* the momentum dissipation, a sharp opposition to the $v_j v_k$-form A-TWS algorithm (see (4.26) and (4.28-4.29-4.30)). Obviously, this should be referred to the afore-said *propagative* nature of flow that is *inseparably* seeded in the pressure terms in momentum equations.

Remarks

1. There are apparently other options in the above procedure. For example, one idea for *improvement* is to permit "cross-reference" among momentum equations, — if one is willing to pay the price of dealing with a *block* coefficient of dissipation (coupling the momentum equations).

2. The "linear analysis" can be pursued for (4.91), except that (a) the system has to be linearized, namely,

$$\frac{\partial U}{\partial t} + A_j \frac{\partial U}{\partial x_j} - \frac{\partial}{\partial x_j} \frac{\Delta t_*}{2} w_j w_k \frac{\partial U}{\partial x_k} = 0 \tag{4.96}$$

where A_j is the Jacobian matrix $\frac{\partial F_j}{\partial U}$, and (b) its "coefficient", i.e, A_j and w_j, must be frozen (at least locally).

3. About the extremely slow flow. Aymptotically, the thermodynamic dependence of pressure on $U = \{\rho, m_i, E\}^T$ becomes *trivialized* as flow gets extremely slow (a feature of incompressiblity), and therefore, (4.91-4.92) $\longrightarrow$

$$\frac{\partial U}{\partial t} + \frac{\partial F_j}{\partial x_j} - \frac{\partial}{\partial x_j} \frac{\Delta t_*}{2} v_j v_k \frac{\partial U}{\partial x_k} = 0 \tag{4.97}$$

It is interesting that such an "asymtote" is identical to a S-TWS with the pressure totally "selected" out, the way "extremely slow flow" ought to be treated (since pressure in extremely slow flow is totally *propagative*).

4.3 The Invariant Dissipation Length Algorithm

Dissipation Rescaling in General

By "dissipation rescaling" (DR), I mean to convert the dissipation *time* to a dissipation length ([17, 20, 21, 24, 25]). The reason DR is necessary is that the time step, as in most implicit computations, is often large, and that, therefore, there is a need of measuring/bounding the artificial dissipation to the spatial mesh size.

Symbolically, the general solution for dissipation rescaling is the dimension condition

$$[\Delta t_*] = \left[\frac{\Delta L}{V}\right] \tag{4.98}$$

where V is a velocity/speed. The question, however, is how such a symbolic solution should be carried out. We have to perform these tasks:

i. Choose V, — given that there are different velocities/speeds for choice, e.g., the velocity of flow/convection and the speed of wave/propagation.

ii. Interpret ΔL, — given that, on the one hand, being ruled by (4.98), it is generally incompatible with the *mesh size* (whatever it is), on the other, it must be "measured" on mesh.

Particularly, both face "dimension difficulty": In multi-D flow, not only are there multiple velocities/speeds for choice (let alone the different ways viewing them), but even the definition of *mesh size* is not unambiguous ([17, 20, 21, 24, 25]).

□

There have been several implementations of (4.98), such as ([21, 24, 25, 34])

$$\Delta t_* \simeq \frac{D_e^{*\frac{1}{n}}}{v} \tag{4.99}$$

$$\Delta t_* \simeq \frac{\Delta x_j^e}{v_j} \tag{4.100}$$

$$\Delta t_* \simeq \frac{\Delta x_j^e}{\rho\left(A_k A_l\right)} \tag{4.101}$$

whereby, v_j is the velocity of flow, and v its norm; $\rho\left(A_j A_k\right)$ is the maximal spectrum radius of $A_j A_k$, representing the *wave* speed; $D_e^{*\frac{1}{n}}$ in (4.99), with D_e^* being the

averaged determinent of the Jacobian matrix of the FE coordinate transform (see Chapter 3 for definition) and n the dimension of the problem, is a FE volume/area based mesh size; Δx_j^e in (4.100-4.101) is the size of FE mesh in jth direction, which was taken as ([21])

$$\begin{aligned} \Delta x_j^e &= \overline{\left|\frac{\partial x_j}{\partial \eta_i}\right|}^e \Delta \eta_i \\ &= \left(\overline{\left|\frac{\partial x_j}{\partial \eta_1}\right|}^e + \overline{\left|\frac{\partial x_j}{\partial \eta_2}\right|}^e + \overline{\left|\frac{\partial x_j}{\partial \eta_3}\right|}^e\right) \Delta \eta \end{aligned} \tag{4.102}$$

These implementations, while all being seemingly undisturbing to stability as per the reported computing ([21, 24, 25, 34]), share in common that the dissipation length has been *equated* to mesh size (though differently). As a result, the dissipation has been not only rescaled but *re-directed.* This is obvious for not only (4.100-4.101), which both are equivalent to "vectorizing" Δt_* as shown, but (4.99), which can be clearly spotted by, say, considering the 1-D inviscid case of (4.45) placed in a 2-D rectangular mesh: Presumably and as is the case prior to dissipation rescaling, solution q is independent of x_2, i.e., $q(t, x_1, x_2) \equiv q(t, x_1)$; but now, because of (4.99),

$$\begin{aligned} \Delta t_* &= \frac{D_e^{*\frac{1}{n}}}{v} \\ &= \frac{\sqrt{(\Delta x_1 \, \Delta x_2)^e}}{2\,v} \end{aligned} \tag{4.103}$$

and/hence, in each mesh cell,

$$\frac{\partial q}{\partial t} + \frac{\partial q v_j}{\partial x_j} - \frac{\partial}{\partial x_j} \frac{v_j v_k}{4v} \sqrt{(\Delta x_1 \, \Delta x_2)^e} \frac{\partial q}{\partial x_k} = 0 \tag{4.104}$$

which *is* x_2-independent only if Δx_2^e is *constant.*

Invariant Dissipation Length Algorithm in General

In the following, I will pursue for such a choice of (4.98):

$$dx_j = v_j\, dt \tag{4.105}$$

or

$$\frac{dx_j}{v_j} = dt \tag{4.106}$$

It is to be established that the dissipation length so dictated — one that is "streamlined", can be measured in the mesh *invariantly* in spite of its "incompatibility" (which is illustrated in Figure 4.2).

In fact,

$$dx_j = \frac{\partial x_j}{\partial \eta_i}\, d\eta_i \tag{4.107}$$

$$\begin{aligned} v_j &= \frac{\partial x_j}{\partial t} \\ &= \frac{\partial x_j}{\partial \eta_i}\frac{\partial \eta_i}{\partial t} \\ &= \frac{\partial x_j}{\partial \eta_i}\tilde{v}_i \end{aligned} \tag{4.108}$$

where η_i are the "natural" coordinates, and $\tilde{v}_i$ the "natural" velocities.

Equation (4.107-4.108) can be alternatvely writen as

$$d\mathbf{x} = [D]^T\, d\eta \tag{4.109}$$

$$\mathbf{v} = [D]^T\, \tilde{\mathbf{v}} \tag{4.110}$$

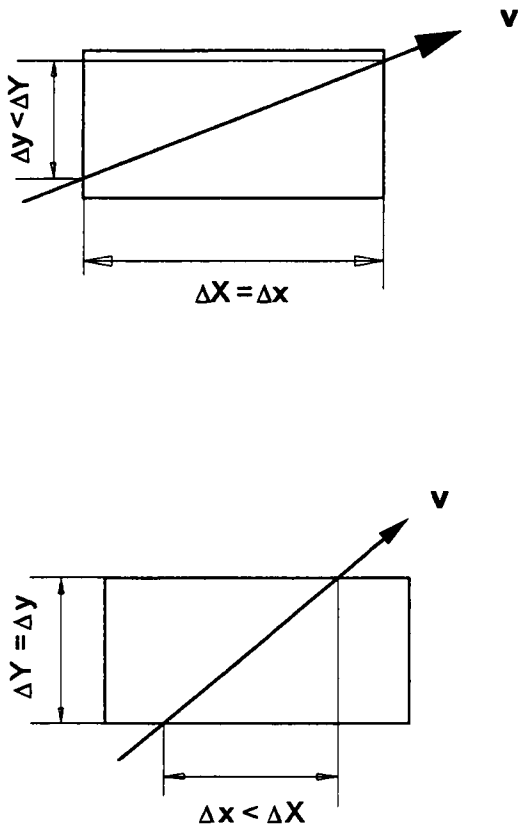


Figure 4.2: Illustration of "incompatibility" between disspation length and mesh size. Mesh: rectangular and sized $\Delta X \times \Delta Y$. Disspation length: $\Delta x_j = \int_{t_1}^{t_2} v_j dt = \bar{v}_j (t_2 - t_1)$, with $t_2 - t_1$ being such that vector is Δx_j is *just* enough to cross the mesh cell.

Applying them into (4.105) gives

$$d\eta_i = \tilde{v}_i \, dt \tag{4.111}$$

where

$$\tilde{v}_i = \frac{E_{im}}{D} v_m \tag{4.112}$$

Henceforth,

$$\frac{dx_i}{v_{\underline{i}}} = \frac{d\eta_i}{\tilde{v}_{\underline{i}}} = dt \tag{4.113}$$

in other words, $\frac{dx_i}{v_{\underline{i}}}$ is an *invariant across* not only diections but the FE coordinate transform.

Diffrentially, equation (4.113) defines how time should be measured *invariantly* in master-element on *streamline*. Integrating it, one has

$$\frac{\int_1^2 dx_i}{\bar{v}_{\underline{i}}} = \frac{\Delta x_i}{\bar{v}_{\underline{i}}} = t_2 - t_1 \tag{4.114}$$

$$\frac{\int_1^2 d\eta_i}{\bar{\tilde{v}}_{\underline{i}}} = \frac{\Delta \eta_i}{\bar{v}_{\underline{i}}} = t_2 - t_1 \tag{4.115}$$

where $\bar{v}_i$ and $\bar{\tilde{v}}_i$ are the mean value of v_i and $\tilde{v}_i$ over $[t_1, t_2]$ respectively. Particularly, if such t_1 and t_2 are considered that vector $\Delta\eta_i$ is just enough to *cross* the master-element (or equivalently Δx_i the *physical* mesh cell), then

$$\begin{aligned} \Delta t_* &= \frac{\Delta t_*}{t_2 - t_1} (t_2 - t_1) \\ &= \beta^* \frac{|\Delta \eta_i|}{\left|\bar{\tilde{v}}_{\underline{i}}\right|} \end{aligned} \tag{4.116}$$

where β^* is a parameter, which is defined as

$$\beta^* = \frac{\Delta t_*}{(t_2 - t_1)} \tag{4.117}$$

Note that $t_1 < t_2$ has been assumed in (4.116), which is why

$$t_2 - t_1 = \frac{\Delta\eta_i}{\bar{\tilde{v}}_{\underline{i}}} = \frac{|\Delta\eta_i|}{\left|\bar{\tilde{v}}_{\underline{i}}\right|} \tag{4.118}$$

could have been taken. Such an assumption, as shown in (4.117), is equivlent to $\beta^* > 0$.

□

Equation (4.116) is what I call the "invariant dissipation length" (IDL) algorithm. The invariant dissipation lengths are $\beta^* |\Delta\eta_i|$, with $|\Delta\eta_i|$ being the size of master-element (SME) along the "natural streamline" (in the ith direction). The *physical* dissipation lengths may be calculated based on their IDLs via the FE coordinate transform:

$$\begin{aligned}
\left|\Delta x_j^e\right| &= \left| \int_1^2 \left(\frac{\partial x_j}{\partial \eta_i}\right)^e d\eta_i \right| \\
&= \left| \overline{\frac{\partial x_j}{\partial \eta_i}} \int_1^2 d\eta_i \right|^e \\
&= \left| \bar{\tilde{v}}_i \overline{\frac{\partial x_j}{\partial \eta_i}} \frac{\Delta\eta_{\underline{i}}}{\bar{\tilde{v}}_{\underline{i}}} \right|^e \\
&= \left| \bar{\tilde{v}}_i \overline{\frac{\partial x_j}{\partial \eta_i}} \right|^e \left(\frac{|\Delta\eta_l|}{\left|\bar{\tilde{v}}_{\underline{l}}\right|}\right)^{\underline{e}}
\end{aligned} \tag{4.119}$$

Equation (4.119) may be further deduced:

$$\begin{aligned}
\left|\Delta x_j^e\right| &= \left| \bar{v}_m \frac{E^*_{im}}{D^*} \overline{\frac{\partial x_j}{\partial \eta_i}} \right|^e \left(\frac{|\Delta\eta_l|}{\left|\bar{\tilde{v}}_{\underline{l}}\right|}\right)^{\underline{e}} \\
&= \left| \bar{v}_m \overline{\frac{\partial x_j}{\partial x_m}} \right|^e \frac{\Delta t_*}{\beta^*} \\
&= |\bar{v}_m \delta_{jm}|^e \frac{\Delta t_*}{\beta^*} \\
&= \frac{1}{\beta^*} |\bar{v}_j|^e \Delta t_*
\end{aligned} \tag{4.120}$$

which is identical to what is obtained via (4.114) and (4.117) — the direct integration of (4.105) or (4.106) over dissipation time (well, a β^*th of it), and thus, a verification of the consistency of the IDL algorithm.

Invariant Dissipation Length Algorithms in Particular

Equation (4.116), albeit being the generally valid expression of IDL, has not yet been in the form that is ready to be numerically implemented. The problem with it is the same as on the physical domain, the "incompatibility" between the dssipation lengths and the common sense "mesh size" (see Figure 4.3). The difference, however, is that this is only a tiny obstacle to overcome given the *regular* geometry of the master-element, at least for the finite elements of most pratical values ([2]):

(**a**) quadrilateral and hexahedron elements, and

(**b**) triangle and tetrahedron elements.

□

For **a**, the master element is a square or square-box (see Figure 4.3), i.e.,

$$\Omega_* = [-\frac{\Delta\eta}{2}, \frac{\Delta\eta}{2}] \times [-\frac{\Delta\eta}{2}, \frac{\Delta\eta}{2}] \times \cdots$$

Accordingly, one knows that the $|\Delta\eta_i|$ which "is just enough to *cross* the master-element" must possess such a feature:

$$\max_i |\Delta\eta_i| = \Delta\eta \tag{4.121}$$

On the other hand, $\frac{\Delta\eta_i}{|\vec{v}_i|}$ is an invariant, i.e., $|\Delta\eta_l|$ attains its maximum in the direction

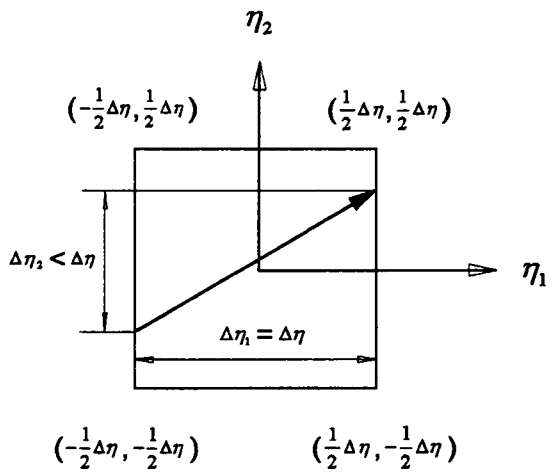


Figure 4.3: Illustration for the invariant dissipation length (IDL) in quadrilateral and hexahedron master-elements (citing the rectangle example); mesh size: $\Delta\eta$; IDL: $\Delta\eta_i = |\tilde{\bar{v}}_i| \frac{\Delta t_*}{\beta^*}$, which *just* crosses the mesh cell. IDL algorithm: $\Delta t_* = \beta^* \frac{\Delta\eta}{\max_i |\tilde{\bar{v}}_i|} = \beta^* \frac{\Delta\eta}{|\tilde{\bar{v}}_1|}$.

that $|\bar{\bar{v}}_i| = \max_l |\bar{\bar{v}}_l|$. Consequently,

$$\begin{aligned} \Delta t_* &= \beta^* \frac{|\Delta\eta_i|}{|\bar{\bar{v}}_i|} \\ &= \beta^* \frac{\Delta\eta}{\max_i |\bar{\bar{v}}_i|} \end{aligned} \tag{4.122}$$

□

For **b**, the master-element is an isosceles triangle or tetrahedron which is surrounded by (see Figure 4.4)

$$\eta_1 = 0 \tag{4.123}$$

$$\cdots$$

$$\eta_n = 0 \tag{4.124}$$

$$\sum_i \eta_i = \Delta\eta \tag{4.125}$$

where (4.123-4.124) are the axises or axis-planes of coordinates, and (4.125) is a plane or straight line in the first quadrant that intersects each axis at $\eta_i = \Delta\eta$.

Feature for the $|\Delta\eta_i|$ which "is just enough to *cross* the master-element" in this case, as shown in Figure 4.4, is that it defines a point lying in plane/straight-line (4.125), which which leads to the following in conjunction with (4.116),

$$\begin{aligned} \Delta\eta &= \sum_i |\Delta\eta_i| \\ &= \frac{\Delta t_*}{\beta^*} \sum_i |\bar{\bar{v}}_i| \end{aligned} \tag{4.126}$$

Consequently,

$$\Delta t_* = \beta^* \frac{\Delta\eta}{\sum_i |\bar{\bar{v}}_i|} \tag{4.127}$$

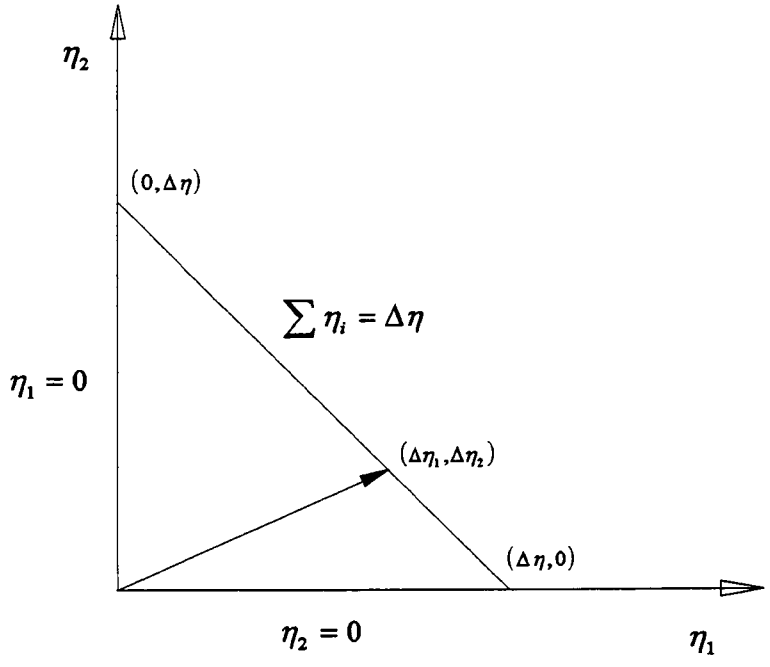


Figure 4.4: Illustration for the invariant dissipation length (IDL) in triangle and tetrahedron master-elements (citing the triangle example); mesh size: $\Delta\eta$; IDL: $\Delta\eta_i = |\bar{\tilde{v}}_i| \frac{\Delta t_*}{\beta^*}$, which *just* crosses the mesh cell and defines a point on plane/straight-line $\sum_i \eta_i = \Delta\eta$ if translated to being originated from the origin.

Equation (4.122) and (4.127) are the eventual results of the IDL algorithm. Later in this dissertation, they will be uniformly referred to as

$$\Delta t_* = \beta^* \frac{\Delta \eta}{\|\bar{\bar{v}}\|} \tag{4.128}$$

4.4 The Invariantly Rescaled Weak Statements

Having established IDL, I can now finalize the statements of the "to-be-solved" problems (with all the *weak* solutions being in $H^1(\Omega)$).

The Nonlinear Convection Problem

- Companion conservation law:

$$\frac{\partial q}{\partial t} + \frac{\partial C}{\partial t} + v_{0_j}\frac{\partial q}{\partial x_j} + \frac{\partial}{\partial x_j}\left(\alpha\beta^*\Delta\eta\frac{v_{0_j}}{\|\bar{\bar{v}}\|}\frac{\partial q}{\partial t} - \beta\beta^*\Delta\eta\frac{v_{0_j}v_{0_k}}{\|\bar{\bar{v}}\|}\frac{\partial q}{\partial x_k}\right) = 0 \qquad (4.129)$$

- Galerkin weak statement:

$$TWS(q): \quad \int_\Omega w\left(\frac{\partial q}{\partial t} + \frac{\partial C}{\partial t} + v_{0_j}\frac{\partial q}{\partial x_j}\right)d\Omega - \qquad (4.130)$$
$$\int_\Omega \frac{\partial w}{\partial x_j}\left(\alpha\,\beta^*\,\Delta\eta\,\frac{v_{0_j}}{\|\bar{\bar{v}}\|}\frac{\partial q}{\partial t} - \beta\,\beta^*\,\Delta\eta\,\frac{v_{0_j}v_{0_k}}{\|\bar{\bar{v}}\|}\frac{\partial q}{\partial x_k}\right)d\Omega +$$
$$\oint_{\partial\Omega} w\left(\alpha\,\beta^*\,\Delta\eta\,\frac{v_{0_j}}{\|\bar{\bar{v}}\|}\frac{\partial q}{\partial t} - \beta\,\beta^*\,\Delta\eta\,\frac{v_{0_j}v_{0_k}}{\|\bar{\bar{v}}\|}\frac{\partial q}{\partial x_k}\right)n_j d\sigma = 0$$

Compressible Potential Flow

- Companion conservation law:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho\frac{\partial \phi}{\partial x_j}\right) - \frac{\partial}{\partial x_j}\left[\max\left(0, 1 - \frac{1}{M^2}\right)\beta^*\Delta\eta\frac{v_j v_k}{\|\bar{\bar{v}}\|}\frac{\partial \rho}{\partial x_k}\right] = 0 \qquad (4.131)$$

$$\frac{\partial \phi}{\partial t} + \frac{1}{2}\frac{\partial \phi}{\partial x_j}\frac{\partial \phi}{\partial x_j} + \frac{\gamma}{\gamma - 1}\left(\frac{T_\infty}{\rho_\infty^{\gamma-1}}\rho^{\gamma-1} - T_{0\infty}\right) = 0 \qquad (4.132)$$

- Galerkin weak statement:

$$
\begin{aligned}
STWS(\rho): \quad & \int_\Omega \left(w \frac{\partial \rho}{\partial t} - \frac{\partial w}{\partial x_j} \rho \frac{\partial \phi}{\partial x_j} \right) d\Omega + \\
& \int_\Omega \frac{\partial w}{\partial x_j} \tau \beta^* \Delta\eta \frac{v_j v_k}{\|\bar{\tilde{v}}\|} \frac{\partial \rho}{\partial x_k} d\Omega + \\
& \oint_{\partial\Omega} w \rho \frac{\partial \phi}{\partial x_j} n_j d\sigma - \\
& \oint_{\partial\Omega} w \tau \beta^* \Delta\eta \frac{v_j v_k}{\|\bar{\tilde{v}}\|} \frac{\partial \rho}{\partial x_k} n_j d\sigma = 0 \qquad (4.133)
\end{aligned}
$$

$$
\begin{aligned}
STWS(\phi): \quad & \int_\Omega w \left(\frac{\partial \phi}{\partial t} + \frac{1}{2} \frac{\partial \phi}{\partial x_j} \frac{\partial \phi}{\partial x_j} \right) d\Omega + \\
& \frac{\gamma}{\gamma - 1} \int_\Omega w \left(\frac{T_\infty}{\rho_\infty^{\gamma-1}} \rho^{\gamma-1} - T_{0\infty} \right) d\Omega = 0 \qquad (4.134)
\end{aligned}
$$

where

$$
\tau = \max\left(0, 1 - \frac{1}{M^2}\right) \qquad (4.135)
$$

Euler Equations

- Companion conservation law:

$$
\frac{\partial U}{\partial t} + \frac{\partial F_j}{\partial x_j} - \frac{\partial}{\partial x_j} \beta^* \Delta\eta \frac{w_j w_k}{\|\bar{\tilde{v}}\|} \frac{\partial U}{\partial x_k} = 0 \qquad (4.136)
$$

- Galerkin weak statement:

$$
\begin{aligned}
STWS(U): \quad & \int_\Omega \left(w \frac{\partial U}{\partial t} - \frac{\partial w}{\partial x_j} F_j^c + w \frac{\partial F_j^p}{\partial x_j} \right) d\Omega + \\
& \int_\Omega \frac{\partial w}{\partial x_j} \beta^* \Delta\eta \frac{w_j w_k}{\|\bar{\tilde{v}}\|} \frac{\partial U}{\partial x_k} d\Omega + \\
& \oint_{\partial\Omega} w F_j^c n_j d\sigma - \\
& \oint_{\partial\Omega} w \beta^* \Delta\eta \frac{w_j w_k}{\|\bar{\tilde{v}}\|} \frac{\partial U}{\partial x_k} n_j d\sigma = 0 \qquad (4.137)
\end{aligned}
$$

in which

$$F_j = F_j^c + F_j^p \tag{4.138}$$

$$F_j^c = \{m_j,\, m_i v_j,\, (E+p)\, v_j\}^T$$

$$= \left\{m_j,\, \frac{m_i m_j}{\rho},\, \left[\gamma E - \frac{(\gamma-1)\, m_i^2}{2\rho}\right] \frac{m_j}{\rho}\right\}^T \tag{4.139}$$

$$F_j^p = \{0,\, p\delta_{ij},\, 0\}^T$$

$$= \left\{0,\, (\gamma-1)\left(E - \frac{m_i^2}{2\rho}\right)\delta_{ij},\, 0\right\}^T \tag{4.140}$$

Note that we have partitioned F_j^c whereas not F_j^p in (4.137). That is because F_j^c is a *flux*, which is often conditioned across at least part of the boundary (the Neuman boundary condition $(F_j^c n_j)|_{\partial\Omega}$), and F_j^p not, which is either totally free on boundary or Dirichlet given (such as on a subsonic outlet).

Remarks

Factor $\frac{1}{2}$ has been omitted in the artificial terms of both (4.129-4.130) and (4.136-4.137) (compare them with (4.6) and (4.91)). Obviously, this is equivalent to embodying it in β^*.

Chapter 5

Finite Element Implementations

Presented in this chapter is a *checklist* of the FE algorithms for the "to-be-solved" problems. The weak statements and FE procedures should be referred to § 4.4 and § 3.1.1 respectively. All FE matrices are metric data mean (MDM) based, unless unnecessary. Several linear algebra solvers are implemented, including conjugate gradient square (CGS, [65, 66]), tensor product factorization (TP, [21]), Gauss-Seidel (GS) and *line* Gauss-Seidel (LGS).

5.1 The Nonlinear Convection Problem

The considered "nonlinear convection problem" is the 1-D (and $\alpha = 0$) case of (4.130), i.e.,

$$\int_0^L w \left(\frac{\partial q}{\partial t} + \frac{\partial C}{\partial t} + v_0 \frac{\partial q}{\partial x} \right) dx +$$

$$\int_0^L \beta^* \Delta x \frac{dw}{dx} v_0 \frac{\partial q}{\partial x} dx - \left(\beta^* \Delta x \, w \, v_0 \frac{\partial q}{\partial x} \right)_0^L = 0 \tag{5.1}$$

where $[0, L]$ is the domain of definition of the solution. Note that here the dissipation length has been "recovered" in physical domain, which is because the problem is 1-D, and the physical mesh size is free of any ambiguity.

Finite Element Approximation

$$\begin{aligned} q(t,x) &\simeq q^h(t,x) \\ &= \bigcup_e q_\beta(t) \, N_\beta(\eta) \qquad (5.2) \\ C(t,x) &\simeq C^h(t,x) \\ &= \bigcup_e C_\beta(t) \, N_\beta(\eta) \qquad (5.3) \\ v_0(x) &\simeq v_0^h(x) \\ &= \bigcup_e v_{0_\beta} N_\beta(\eta) \qquad (5.4) \\ x^e(\eta) &= x_\beta^e \tilde{N}_\beta(\eta) \qquad (5.5) \end{aligned}$$

whereby, as in § 3.1.1, $\{N_\beta(\eta)\}$ are the basis of the FE subspace S^h, and $\tilde{N}_\beta(\eta)$ a set of independent shape functions that locally defines the FE coordinate transform.

Since the problem is 1-D, the FE coordinate transform (5.5) can and should always be linear, i.e.,

$$\{\tilde{N}_i(\varsigma)\} = \frac{1}{2} \begin{Bmatrix} (1-\varsigma) \\ (1+\varsigma) \end{Bmatrix} \tag{5.6}$$

As a consequence, metric data are elementally constant, and metric data mean (MDM) is hence unnecessary in the calculation of the FE matrices.

Semi-Discretization

In S^h, (5.1) is semi-discrete:

$$\int_0^L N_\alpha \left(\frac{\partial q^h}{\partial t} + \frac{\partial C^h}{\partial t} + v_0^h \frac{\partial q^h}{\partial x} \right) dx + \int_0^L \beta^* \Delta x \frac{dN_\alpha}{dx} v_0^h \frac{\partial q^h}{\partial x} dx - \left(\beta^* \Delta x N_\alpha v_0 \frac{\partial q}{\partial x} \right)_0^L = 0 \tag{5.7}$$

which can be rewritten in the following mass-residual form.

$$M200_{\alpha\beta} \left(\dot{q} + \dot{C} \right)_\beta + R_\alpha = 0 \tag{5.8}$$

where

$$R_\alpha = \sum_e R_\alpha^e - \left(\beta^* \Delta x N_\alpha v_0 \frac{\partial q}{\partial x} \right)_0^L \tag{5.9}$$

$$R_\alpha^e = v_{0_\theta} q_\beta M3100_{\beta\theta\alpha}^e + \beta^* \Delta x_e v_{0_\theta} q_\beta M3101_{\beta\theta\alpha}^e \tag{5.10}$$

$$M200_{\alpha\beta}^e = \int_{-1}^1 N_\alpha N_\beta \, dx = \frac{\Delta x_e}{2} A200_{\alpha\beta} \tag{5.11}$$

$$M3100_{\beta\theta\alpha}^e = \int_{-1}^1 \frac{dN_\beta}{dx} N_\theta N_\alpha \, dx = A3100_{\beta\theta\alpha} \tag{5.12}$$

$$M3101_{\alpha\theta\beta}^e = \int_{\Omega^e} \frac{dN_\alpha}{dx} N_\theta \frac{dN_\beta}{dx} \, dx = \frac{2}{\Delta x_e} A3101_{\alpha\theta\beta} \tag{5.13}$$

Full Discretization

The full FE discretization of (5.1) results from further integrating the temporal ODE system (5.8), of which, as discussed in § 3.1.1, the general expression for the single step algorithms is the Θ-implicit scheme:

$$M200_{\alpha\beta}\,(\Delta q^n + \Delta C^n)_\beta + (1-\Theta)\,\Delta t\,R_\alpha^n + \Theta\,\Delta t\,R_\alpha^{n+1} = 0 \tag{5.14}$$

where Δt is the time step, and

$$\Delta q^n = q^{n+1} - q^n \tag{5.15}$$

$$\Delta C^n = C^{n+1} - C^n \tag{5.16}$$

The $\Theta = 0$ case of (5.14), i.e., the Euler explicit scheme

$$M200_{\alpha\beta}\,(\Delta q^n + \Delta C^n)_\beta + \Delta t\,R_\alpha^n = 0 \tag{5.17}$$

is what is actually used in this research.

Though being "explicit", (5.17) is a *nonlinear* algebra system, because of the dependence of C on q. Upon the completion of solving (5.17) for $\{\Delta q_\alpha^n + \Delta C_\alpha^n\}$ at each time step, the following nonlinear equation needs to be solved in order to determine q^{n+1} (see § 2.4):

$$q_\alpha^{n+1} + a\int_{q_\alpha^n}^{q_\alpha^{n+1}} f(q)\,dq = q_\alpha^n + (\Delta q^n + \Delta C^n)_\alpha \tag{5.18}$$

5.2 Compressible Potential Flow

Bilinear Isoparametric Finite Element Approximation

$$\begin{aligned} \rho(t,x_j) &\simeq \rho^h(t,x_j) \\ &= \bigcup_e \rho_\beta(t)\, N_\beta(\eta_i) \qquad (5.19) \\ \phi(t,x_j) &\simeq \phi^h(t,x_j) \\ &= \bigcup_e \phi_\beta(t)\, N_\beta(\eta_i) \qquad (5.20) \\ x^e_{j\beta}(\eta_i) &= x^e_{j\beta}\, \tilde{N}_\beta(\eta_i) \qquad (5.21) \end{aligned}$$

where

$$\begin{aligned} \{N_\alpha(\eta_j)\} &= \{\tilde{N}_\alpha(\eta_j)\} \\ &= \frac{1}{4}\begin{Bmatrix} (1-\eta_1)(1-\eta_2) \\ (1+\eta_1)(1-\eta_2) \\ (1+\eta_1)(1+\eta_2) \\ (1-\eta_1)(1+\eta_2) \end{Bmatrix} \qquad (5.22) \end{aligned}$$

The metric data for (5.21)$\bigcap$(5.22) have been formulated in Chapter 3, i.e.,

$$E_{ij} = \overline{E}^*_{ij} + (-1)^{i+j}\, \overline{E}^0_j\, \eta_i \qquad (5.23)$$

$$D = \overline{D}^*\,(1 + A\eta_1 + B\eta_2) \qquad (5.24)$$

where $\overline{E}^*_{ij}$ and $\overline{D}^*$ are respectively the elemental average of E_{ij} and D, defined in (3.68) and (3.71); $\overline{E}^0_j$, A and B are elemental constants, given in (3.69) and (3.72), which are all *trivial* in case the bilinear FE is a paralellgram.

Finite Element Semi-Discretization

Substituting (5.19-5.20-5.21) into (4.133-4.134), we have

$$\int_{\Omega^h} \left(N_\alpha \frac{\partial \rho^h}{\partial t} - \frac{\partial N_\alpha}{\partial x_j} \rho^h \frac{\partial \phi^h}{\partial x_j} \right) d\Omega +$$
$$\int_{\Omega^h} \left(\tau \frac{v_j^h v_k^h}{\|\tilde{\bar{v}}\|} \beta^* \Delta\eta \right) \frac{\partial N_\alpha}{\partial x_j} \frac{\partial \rho^h}{\partial x_k} d\Omega +$$
$$\oint_{\partial\Omega^h} N_\alpha \rho \frac{\partial \phi}{\partial x_j} n_j \, d\sigma -$$
$$\oint_{\partial\Omega^h} \left(\tau \frac{v_j v_k}{\|\tilde{\bar{v}}\|} \beta^* \Delta\eta \right) N_\alpha \frac{\partial \rho}{\partial x_k} n_j \, d\sigma = 0 \qquad (5.25)$$
$$\int_\Omega N_\alpha \left(\frac{\partial \phi^h}{\partial t} + \frac{1}{2} \frac{\partial \phi^h}{\partial x_j} \frac{\partial \phi^h}{\partial x_j} \right) d\Omega +$$
$$\frac{\gamma}{\gamma - 1} \int_\Omega N_\alpha \left(\frac{T_\infty}{\rho_\infty^{\gamma-1}} \left(\rho^h \right)^{\gamma-1} - T_{0\infty} \right) d\Omega = 0 \qquad (5.26)$$

which is the FE semi-discretion of the S-TWS for CPF, and in which (4.122) is the actual formula for IDL (since the FE is bilinear).

In mass-residual form, (5.26-5.26) read

$$M200_{\alpha\beta} \left\{ \dot{Q} \right\}_\beta + \{R\}_\alpha = 0 \qquad (5.27)$$

in which

$$\{Q\}_\alpha = \{\rho_\alpha, \phi_\alpha\}^T \qquad (5.28)$$

$$\{R\}_\alpha = \{R1_\alpha, R2_\alpha\}^T \qquad (5.29)$$

$$R1_\alpha^e = - \rho_\beta \phi_\theta M3J0J_{\theta\beta\alpha}^e$$
$$+ (\beta^* \Delta\eta) \overline{\left(\tau \frac{\tilde{v}_M \tilde{v}_N}{\|\tilde{\bar{v}}\|} \right)}^* \rho_\beta \overline{D}_{\underline{\alpha\beta}}^e B2NM_{\beta\alpha}$$
$$+ \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \rho \frac{\partial \phi}{\partial x_j} n_j \, d\sigma$$

$$- (\beta^* \Delta\eta) \oint_{\partial\Omega^e \cap \partial\Omega^h} \left(\tau \frac{v_j v_k}{\|\bar{\bar{v}}\|} \right) N_\alpha \frac{\partial \rho}{\partial x_k} n_j \, d\sigma \tag{5.30}$$

$$\begin{aligned} R2^e_\alpha &= \frac{1}{2} \phi_\theta \, \phi_\beta \, M3JJ0^e_{\theta\beta\alpha} + \\ & \frac{\gamma}{\gamma - 1} \frac{T_\infty}{\rho_\infty^{\gamma-1}} \rho_\beta^{\gamma-1} M200^e_{\beta\alpha} - \\ & T_{0\infty} M10^e_\alpha \end{aligned} \tag{5.31}$$

$$\begin{aligned} M200^e_{\alpha\beta} &= \int_{\Omega^*} N_\alpha \, N_\beta \, D^e \, d\Omega^* \\ &= \overline{D^e}_{\underline{\alpha\beta}} \, B200_{\alpha\beta} \\ &= \overline{D^*} \left(1 + A^e \, \overline{\varsigma 200}_{\underline{im}} + B^e \, \overline{\varsigma 200}_{\underline{jn}} \right) A200_{im} \, A200_{jn} \end{aligned} \tag{5.32}$$

$$\begin{aligned} M10^e_\alpha &= \int_{\Omega^*} N_\alpha \, D^e \, d\Omega^* \\ &= \overline{D^e}_{\underline{\alpha}} \, B10_\alpha \\ &= \overline{D^*} \left(1 + A^e \, \overline{\varsigma 10}_{\underline{i}} + B^e \, \overline{\varsigma 10}_{\underline{j}} \right) A10_i \, A10_j \end{aligned} \tag{5.33}$$

$$\begin{aligned} M3I0J^e_{\alpha\theta\beta} &= \int_{\Omega^*} \left(\frac{E_{MI} E_{NJ}}{D} \right)^e \frac{\partial N_\alpha}{\partial \eta_M} N_\theta \frac{\partial N_\beta}{\partial \eta_N} \, d\Omega^* \\ &= \left(\overline{\frac{E_{MI} E_{NJ}}{D}} \right)^e_1 B3M0N^1_{\alpha\theta\beta} + \\ & \left(\overline{\frac{E_{MI} E_{NJ}}{D}} \right)^e_2 B3M0N^2_{\alpha\theta\beta} \end{aligned} \tag{5.34}$$

Equation (5.32-5.33) are the "polynomial based MDM" algorithms, in which $\alpha \Leftrightarrow (i,j)$ and $\beta \Leftrightarrow (m,n)$. Equation (5.34) is the "submeshing based MDM", which, since $B3M0N\alpha\theta\beta \neq 0$ for *sure* in bilinear element, can also be expressed in the "mean value" form:

$$M3I0J^e_{\alpha\theta\beta} = \left(\overline{\frac{E_{MI} E_{NJ}}{D}} \right)^e_{\underline{\alpha\theta\beta}} B3M0N_{\alpha\theta\beta} \tag{5.35}$$

$$\left(\overline{\frac{E_{MI} E_{NJ}}{D}} \right)^e_{\alpha\theta\beta} = \left(\frac{E_{MI} E_{NJ}}{D} \right)^e_1 \frac{B3\underline{M0}\underline{N^1_{\alpha\theta\beta}}}{B3M0N_{\alpha\theta\beta}} +$$

$$\left(\frac{E_{MI}E_{NJ}}{D}\right)_2^e \frac{B3\underline{M}0\underline{N}^2_{\alpha\theta\underline{\beta}}}{B3M0N_{\alpha\theta\beta}} \tag{5.36}$$

$\overline{\left(\tau \frac{\tilde{v}_m\tilde{v}_n}{\|\tilde{\bar{v}}\|}\right)}^*$ in (5.30) is the elemental *average* of $\left(\tau \frac{\tilde{v}_m\tilde{v}_n}{\|\tilde{\bar{v}}\|}\right)$, i.e., $\left(\tau \frac{\tilde{v}_m\tilde{v}_n}{\|\tilde{\bar{v}}\|}\right)_{\eta_i=0}$, which we take since the S-TWS dissipation term is an error anyway. The derivation of the entire volumetric S-TWS dissipation term in (5.30) is as follows.

$$\begin{aligned}
&\int_{\Omega^e} \left(\tau \frac{v_j^h v_k^h}{\|\tilde{\bar{v}}\|} \beta^* \Delta\eta\right) \frac{\partial N_\alpha}{\partial x_j} \frac{\partial \rho^h}{\partial x_k} d\Omega \\
&= (\beta^* \Delta\eta)\, \rho_\beta \int_{\Omega^*} \left(\tau \frac{v_j^h v_k^h}{\|\tilde{\bar{v}}\|}\right) \left(\frac{E_{mj}}{D} \frac{E_{nk}}{D}\right)^e \frac{\partial N_\alpha}{\partial \eta_m} \frac{\partial N_\beta}{\partial \eta_n} D^e d\Omega^* \\
&= (\beta^* \Delta\eta)\, \rho_\beta \int_{\Omega^*} \left(\tau \frac{\tilde{v}_m^h \tilde{v}_n^h}{\|\tilde{\bar{v}}\|}\right) \frac{\partial N_\alpha}{\partial \eta_m} \frac{\partial N_\beta}{\partial \eta_n} D^e\, d\Omega^* \\
&\simeq (\beta^* \Delta\eta) \overline{\left(\tau \frac{\tilde{v}_m\tilde{v}_n}{\|\tilde{\bar{v}}\|}\right)}^* \rho_\beta \int_{\Omega^*} \frac{\partial N_\alpha}{\partial \eta_m} \frac{\partial N_\beta}{\partial \eta_n} D^e\, d\Omega^* \\
&= (\beta^* \Delta\eta) \overline{\left(\tau \frac{\tilde{v}_m\tilde{v}_n}{\|\tilde{\bar{v}}\|}\right)}^* \rho_\beta\, \overline{D}^e_{\alpha\underline{\beta}}\, B2NM_{\beta\alpha}
\end{aligned} \tag{5.37}$$

where $\overline{D}^e$ the MDM of D^e relative to matrix $[B2NM]$ (check (3.97) for the reason of TP indexing):

$$\begin{aligned}
\overline{D}^e_{\alpha\underline{\beta}}\, B2MN_{\alpha\beta} &= \begin{cases}
\overline{D}^e_{\alpha\underline{\beta}}\, B211_{\alpha\beta} \\
\overline{D}^e_{\alpha\underline{\beta}}\, B221_{\alpha\beta} \\
\overline{D}^e_{\alpha\underline{\beta}}\, B212_{\alpha\beta} \\
\overline{D}^e_{\alpha\underline{\beta}}\, B222_{\alpha\beta}
\end{cases} \\
&= \begin{cases}
\overline{D}^* \left(1 + A^e \overline{\varsigma 211}_{\underline{im}} + B^e \overline{\varsigma 200}_{\underline{jn}}\right) A211_{im} A200_{jn} \\
\overline{D}^* \left(1 + A^e \overline{\varsigma 201}_{\underline{im}} + B^e \overline{\varsigma 210}_{\underline{jn}}\right) A201_{im} A210_{jn} \\
\overline{D}^* \left(1 + A^e \overline{\varsigma 210}_{\underline{im}} + B^e \overline{\varsigma 201}_{\underline{jn}}\right) A210_{im} A201_{jn} \\
\overline{D}^* \left(1 + A^e \overline{\varsigma 200}_{\underline{im}} + B^e \overline{\varsigma 211}_{\underline{jn}}\right) A200_{im} A211_{jn}
\end{cases}
\end{aligned} \tag{5.38}$$

□

Note that, although

$$\left(\tau \frac{\tilde{v}_m \tilde{v}_n}{\|\bar{\tilde{v}}\|}\right) \simeq \overline{\left(\tau \frac{\tilde{v}_m \tilde{v}_n}{\|\bar{\tilde{v}}\|}\right)}^* = \left(\tau \frac{\tilde{v}_m \tilde{v}_n}{\|\bar{\tilde{v}}\|}\right)_{\eta_i=0} \tag{5.39}$$

is an approximation, it is not at all worse than, say, the "accurate" FE interpolation

—

$$\left(\tau \frac{\tilde{v}_m \tilde{v}_n}{\|\bar{\tilde{v}}\|}\right) = \left(\tau \frac{\tilde{v}_m \tilde{v}_n}{\|\bar{\tilde{v}}\|}\right)_\beta N_\beta(\eta_i) \tag{5.40}$$

And the reason is that the problem is CPF, where velocities have to be computed by differentiatiion $\frac{\partial \phi^h}{\partial \eta_i}$ which is second order "super convergent" at $\eta_i = 0$, whereas, only first order accurate at the nodes/corners of FE domain.

It is also noteworthy that we have "inconsistently" equated $\left(\rho^h\right)^{\gamma-1}$ to $(\rho^{\gamma-1})^h$ in (5.31). Supposedly — as according to (5.19), it should be

$$\left(\rho^h\right)^{\gamma-1} = (\rho_\beta N_\beta)^{\gamma-1} \tag{5.41}$$

And if this is indeed enforced, (5.31) should then become involving a partially *powered* mass matrix (which is what we tried to avoid):

$$\begin{aligned} R2^e_\alpha = & \frac{1}{2} \phi_\theta \phi_\beta M3JJ0^e_{\theta\beta\alpha} + \\ & \frac{\gamma}{\gamma-1} \frac{T_\infty}{\rho_\infty^{\gamma-1}} \rho_\beta^{\gamma-1} \int_{\Omega^e} N_\alpha N_\beta^{\gamma-1} d\Omega - \\ & T_{0\infty} M10^e_\alpha \end{aligned} \tag{5.42}$$

Θ-implicit Scheme

Temporally integrating (5.27) with a Θ-implicitness, one has

$$M200_{\alpha\beta}\,\{\Delta Q\}_\beta^n + (1-\Theta)\,\Delta t\,\{R\}_\alpha^n + \Theta\,\Delta t\,\{R\}_\alpha^{n+1} = 0 \qquad (5.43)$$

$$\{\Delta Q\}_\beta^n = \{Q\}_\beta^{n+1} - \{Q\}_\beta^n \qquad (5.44)$$

which is the eventual full discrete S-TWS FE algorithm for CPF. Hereby, while $\{R\}^n = \{R(Q^n)\}$ in complete, $\{R\}^{n+1} = \{R(Q^{n+1})\}$ with a locally frozen $\overline{\left(\tau\,\frac{\tilde{v}_m\tilde{v}_n}{\|\tilde{v}\|}\right)}^*$ — the S-TWS dissipation coefficient — to the previous time level (which was found stablizing the solution better).

Newton Linearized Iteration

Equation (5.43) is nonlinear, and its solution has to be done via successive linearization/iteration, such as the Newton algorithm:

$$\begin{aligned}
&\left[M200_{\alpha\beta} + \Theta\,\Delta t\,J_{\alpha\beta}^{p,n+1}\right]\{\delta Q\}_\beta^{p,n+1} \\
=\ & -M200_{\alpha\beta}\left(\{Q\}_\beta^{p,n+1} - \{Q\}_\beta^n\right) \\
& -\Theta\,\Delta t\,\{R\}_\alpha^{p,n+1} - (1-\Theta)\,\Delta t\,\{R\}_\alpha^n
\end{aligned} \qquad (5.45)$$

where $(\cdot)^{p,n+1}$ denotes the pth Newtonian approximatino of $(\cdot)^{n+1}$, and

$$\{Q\}_\beta^{p+1,n+1} = \{Q\}_\beta^{p,n+1} + \{\delta Q\}_\beta^{p,n+1} \qquad (5.46)$$

$$[J_{\alpha\beta}]^{n+1} = \left[\frac{\partial\{R\}_\alpha}{\partial\{Q\}_\beta}\right]^{n+1} = \begin{bmatrix} \frac{\partial R1_\alpha}{\partial \rho_\beta} & \frac{\partial R1_\alpha}{\partial \phi_\beta} \\ \frac{\partial R2_\alpha}{\partial \rho_\beta} & \frac{\partial R2_\alpha}{\partial \phi_\beta} \end{bmatrix}^{n+1}$$

$$= \begin{bmatrix} J11_{\alpha\beta} & J12_{\alpha\beta} \\ J21_{\alpha\beta} & J22_{\alpha\beta} \end{bmatrix}^{n+1} \tag{5.47}$$

$$\begin{aligned} J11^e_{\alpha\beta} &= -\phi_\theta\, M3J0J^e_{\alpha\beta\theta} \\ &\quad + (\beta^*\,\Delta\eta)\,\overline{\left(\tau\frac{\tilde{v}_M\tilde{v}_N}{\|\bar{\tilde{v}}\|}\right)}^{*}\,\overline{D}^e_{\underline{\alpha\beta}}\,B2MN_{\alpha\beta} \\ &\quad + \frac{\partial}{\partial\rho_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\,\rho\,\frac{\partial\phi}{\partial x_j}\,n_j\,d\sigma \\ &\quad - (\beta^*\,\Delta\eta)\oint_{\partial\Omega^e\cap\partial\Omega^h}\left(\tau\,\frac{v_j v_k}{\|\bar{\tilde{v}}\|}\right) N_\alpha\,\frac{\partial N_\beta}{\partial x_k}\,n_j\,d\sigma \end{aligned} \tag{5.48}$$

$$\begin{aligned} J12^e_{\alpha\beta} &= \rho_\theta\, M3J0J^e_{\alpha\theta\beta} + \\ &\quad \frac{\partial}{\partial\phi_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\,\rho\,\frac{\partial\phi}{\partial x_j}\,n_j\,d\sigma \end{aligned} \tag{5.49}$$

$$J21^e_{\alpha\beta} = \gamma\,\frac{T_\infty}{\rho_\infty^{\gamma-1}}\,\rho_{\underline{\beta}}^{\gamma-2}\,M200^e_{\beta\alpha} \tag{5.50}$$

$$J22^e_{\alpha\beta} = \phi_\theta\, M3J0J^e_{\theta\alpha\beta} \tag{5.51}$$

Note that $\frac{\partial}{\partial\rho_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\,\rho\,\frac{\partial\phi}{\partial x_j}\,n_j d\sigma$ and $\frac{\partial}{\partial\phi_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\,\rho\,\frac{\partial\phi}{\partial x_j}\,n_j d\sigma$ vary depending on the boundary conditions on $\rho\frac{\partial\phi}{\partial x_j}n_j$, and that is why they were not specified in (5.48-5.49). If $\partial\Omega^e\cap\partial\Omega^h$ happens to be where $\rho\frac{\partial\phi}{\partial x_j}n_j$ is *given*, they should be

$$\frac{\partial}{\partial\rho_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\,\rho\,\frac{\partial\phi}{\partial x_j}\,n_j d\sigma = 0 \tag{5.52}$$

$$\frac{\partial}{\partial\phi_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\,\rho\,\frac{\partial\phi}{\partial x_j}\,n_j d\sigma = 0 \tag{5.53}$$

Whereas, if $\left(\rho\frac{\partial\phi}{\partial x_j}n_j\right)_{\partial\Omega^e\cap\partial\Omega^h}$ is *free*,

$$\frac{\partial}{\partial\rho_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\,\rho\,\frac{\partial\phi}{\partial x_j}\,n_j d\sigma = \phi_\theta\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\,N_\beta\,\frac{\partial N_\theta}{\partial x_j}\,n_j d\sigma \tag{5.54}$$

$$\frac{\partial}{\partial\phi_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\,\rho\,\frac{\partial\phi}{\partial x_j}\,n_j d\sigma = \rho_\theta\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\,N_\theta\,\frac{\partial N_\beta}{\partial x_j}\,n_j d\sigma \tag{5.55}$$

5.3 Euler Equations

FE Approximation

$$
\begin{aligned}
U^h(t,x_j) &= \{\rho^h(t,x_j),\, m_i^h(t,x_j),\, E^h(t,x_j)\}^T \\
&= \bigcup_e \left\{ \begin{array}{c} \rho_\beta(t)\, N_\beta(\eta_k) \\ m_{i_\beta}(t)\, N_\beta(\eta_k) \\ E_\beta(t)\, N_\beta(\eta_k) \end{array} \right\} \qquad (5.56) \\
F_j^{c^h}(t,x_j) &= \{m_j^h(t,x_j),\, m_i^h(t,x_j)\, v_j^h(t,x_j),\, [E(t,x_j) + p(t,x_j)]^h\, v_j^h(t,x_j)\}^T \\
&= \bigcup_e \left\{ \begin{array}{c} m_{j_\beta}(t)\, N_\beta(\eta_k) \\ m_{i_\beta}(t)\, v_{j_\theta}(t)\, N_\beta(\eta_k)\, N_\theta(\eta_k) \\ [E(t) + p(t)]_\beta\, v_{j_\theta}(t)\, N_\beta(\eta_k)\, N_\theta(\eta_k) \end{array} \right\} \qquad (5.57) \\
F_j^{p^h}(t,x_j) &= \{0,\, p^h(t,x_j)\, \delta_{ij},\, 0\}^T \\
&= \bigcup_e \left\{ \begin{array}{c} 0 \\ p_\beta(t)\, N_\beta(\eta_k)\, \delta_{ij} \\ 0 \end{array} \right\} \qquad (5.58)
\end{aligned}
$$

where N_α and the FE coordinate transform that links x_j and η_j are the same as in CPF, i.e., (5.21), (5.22) and (5.23-5.24).

Semi-Discretization

$$
\begin{aligned}
&\int_\Omega \left(N_\alpha \frac{\partial U^h}{\partial t} - \frac{\partial N_\alpha}{\partial x_j} F_j^{c^h} + N_\alpha \frac{\partial F_j^{p^h}}{\partial x_j} \right) d\Omega + \\
&\int_\Omega \frac{\partial N_\alpha}{\partial x_j} \left(\beta^* \Delta\eta \frac{w_j^h w_k^h}{\|\tilde{\bar{v}}\|} \right) \frac{\partial U^h}{\partial x_k} d\Omega + \oint_{\partial\Omega} N_\alpha F_j^c n_j\, d\sigma - \\
&\oint_{\partial\Omega} N_\alpha \left(\beta^* \Delta\eta \frac{w_j w_k}{\|\tilde{\bar{v}}\|} \right) \frac{\partial U}{\partial x_k} n_j\, d\sigma = 0 \qquad (5.59)
\end{aligned}
$$

And again in mass-residual form,

$$M200_{\alpha\beta}\left\{\dot{U}\right\}_{\beta} + \{R\}_{\alpha} = 0 \tag{5.60}$$

$$\{U\}_{\alpha} = \{\rho_{\alpha}, m_{I_{\alpha}}, E_{\alpha}\}^{T} \tag{5.61}$$

$$\{R\}_{\alpha} = \{RR_{\alpha}, RI_{\alpha}, RE_{\alpha}\}^{T} \tag{5.62}$$

$$\begin{aligned} RR^{e}_{\alpha} &= -m_{J_{\beta}}\, M20J^{e}_{\beta\alpha} \\ &\quad + (\beta^{*}\,\Delta\eta)\,\overline{\left(\frac{\tilde{v}_{M}\tilde{v}_{N}}{\|\tilde{\bar{v}}\|}\right)}^{*}\rho_{\beta}\,\overline{D}^{e}_{\underline{\alpha\beta}}\,B2NM_{\beta\alpha} \\ &\quad + \oint_{\partial\Omega^{e}\cap\partial\Omega^{h}} N_{\alpha}\, m_{J}\, n_{J}\, d\sigma \\ &\quad - (\beta^{*}\,\Delta\eta)\oint_{\partial\Omega^{e}\cap\partial\Omega^{h}}\left(\frac{v_{J}v_{K}}{\|\tilde{\bar{v}}\|}\right) N_{\alpha}\,\frac{\partial\rho}{\partial x_{K}}\, n_{J}\, d\sigma \end{aligned} \tag{5.63}$$

$$\begin{aligned} RI^{e}_{\alpha} &= -m_{I_{\beta}}\, v_{J_{\theta}}\, M300J^{e}_{\beta\theta\alpha} + p_{\beta}\, M2I0^{e}_{\beta\alpha} \\ &\quad + (\beta^{*}\,\Delta\eta)\left(\Gamma_{\underline{IM}}\,\Gamma_{\underline{IN}}\right)\overline{\left(\frac{\tilde{v}_{M}\tilde{v}_{N}}{\|\tilde{\bar{v}}\|}\right)}^{*} m_{I_{\beta}}\,\overline{D}^{e}_{\underline{\alpha\beta}}\,B2NM_{\beta\alpha} \\ &\quad + \oint_{\partial\Omega^{e}\cap\partial\Omega^{h}} N_{\alpha}\, m_{I}\, v_{J}\, n_{J}\, d\sigma \\ &\quad - (\beta^{*}\,\Delta\eta)\left(\Gamma_{\underline{IJ}}\,\Gamma_{\underline{IK}}\right)\oint_{\partial\Omega^{e}\cap\partial\Omega^{h}}\left(\frac{v_{J}v_{K}}{\|\tilde{\bar{v}}\|}\right) N_{\alpha}\,\frac{\partial m_{I}}{\partial x_{K}}\, n_{J}\, d\sigma \end{aligned} \tag{5.64}$$

$$\begin{aligned} RE^{e}_{\alpha} &= -(E+p)_{\beta}\, v_{J_{\theta}}\, M300J^{e}_{\beta\theta\alpha} \\ &\quad + (\beta^{*}\,\Delta\eta)\,\gamma^{2}\,\overline{\left(\frac{\tilde{v}_{M}\tilde{v}_{N}}{\|\tilde{\bar{v}}\|}\right)}^{*} E_{\beta}\,\overline{D}^{e}_{\underline{\alpha\beta}}\,B2NM_{\beta\alpha} \\ &\quad + \oint_{\partial\Omega^{e}\cap\partial\Omega^{h}} N_{\alpha}\,(E+p)\, v_{J}\, n_{J}\, d\sigma \\ &\quad - (\beta^{*}\,\Delta\eta)\,\gamma^{2}\oint_{\partial\Omega^{e}\cap\partial\Omega^{h}}\left(\frac{v_{J}v_{K}}{\|\tilde{\bar{v}}\|}\right) N_{\alpha}\,\frac{\partial E}{\partial x_{K}}\, n_{J}\, d\sigma \end{aligned} \tag{5.65}$$

Herein, all the FE matrices are calculated via the "polynomial based MDM" algorithm, including (5.32) for $M200^{e}_{\alpha\beta}$, (5.38) for $\left(\overline{D}^{e}_{\underline{\alpha\beta}}B2MN_{\alpha\beta}\right)$, (3.76-3.79) for $M3J00^{e}_{\alpha\theta\beta}$, and $M2J0^{e}_{\alpha\beta}$ as this —

$$M2K0^{e}_{\alpha\beta} = (\overline{E_{MK}})^{e}_{\underline{\alpha\beta}}\,B2M0_{\alpha\beta}$$

$$
\begin{aligned}
&= (\overline{E_{1K}})^e_{\underline{\alpha\beta}}\, B210_{\alpha\beta} + (\overline{E_{2K}})^e_{\underline{\alpha\beta}}\, B220_{\alpha\beta} \\
&= (\overline{E_{1K}})^e_{\underline{im}}\, A210_{im}\, A200_{jn} + (\overline{E_{2K}})^e_{\underline{jn}}\, A200_{im}\, A210_{jln}
\end{aligned} \tag{5.66}
$$

where

$$
\begin{cases}
(\overline{E_{1K}})^e_{im} &= \overline{E}^*_{1K} + (-1)^{1+K}\, \overline{E}^0_K\, \overline{\varsigma 210}_{im} \\
(\overline{E_{2K}})^e_{jn} &= \overline{E}^*_{2K} + (-1)^{2+K}\, \overline{E}^0_K\, \overline{\varsigma 210}_{jn}
\end{cases} \tag{5.67}
$$

which are identical with (3.79) except that $\overline{\varsigma 210}$ are the MDM data referred to.

In addition, (4.92) has been applied, with Γ_{ij} denoting the coefficient of w_j corresponding to the ith momentum equation, i.e.,

$$
\Gamma_{ij} = 1 - (\gamma - 1)\, \delta_{ij} \tag{5.68}
$$

Collectively written, the volumetric part of the S-TWS dissipation terms in (5.63-5.65) comes along as follows.

$$
\begin{aligned}
&\int_{\Omega^e} \left(\beta^* \Delta\eta\, \frac{w^h_j w^h_k}{\|\tilde{\bar{v}}\|}\right) \frac{\partial N_\alpha}{\partial x_j}\, \frac{\partial U^h}{\partial x_k}\, d\Omega \\
&= (\beta^* \Delta\eta)\, \{U\}_\beta \int_{\Omega^*} \left(\frac{w^h_j w^h_k}{\|\tilde{\bar{v}}\|}\right) \left(\frac{E_{mj}}{D}\, \frac{E_{nk}}{D}\right)^e \frac{\partial N_\alpha}{\partial \eta_m}\, \frac{\partial N_\beta}{\partial \eta_n}\, D^{\underline{e}}\, d\Omega^* \\
&= (\beta^* \Delta\eta)\, \{U\}_\beta \int_{\Omega^*} \left(\frac{\tilde{w}^h_m \tilde{w}^h_n}{\|\tilde{\bar{v}}\|}\right) \frac{\partial N_\alpha}{\partial \eta_m}\, \frac{\partial N_\beta}{\partial \eta_n}\, D^e\, d\Omega^* \\
&\simeq (\beta^* \Delta\eta)\, \overline{\left(\frac{\tilde{w}_m \tilde{w}_n}{\|\tilde{\bar{v}}\|}\right)}^* \{U\}_\beta \int_{\Omega^*} \frac{\partial N_\alpha}{\partial \eta_m}\, \frac{\partial N_\beta}{\partial \eta_n}\, D^e\, d\Omega^* \\
&= (\beta^* \Delta\eta)\, \overline{\left(\frac{\tilde{w}_M \tilde{w}_N}{\|\tilde{\bar{v}}\|}\right)}^*\, \{U\}_\beta\, \overline{D}^e_{\underline{\alpha\beta}}\, B2NM_{\beta\alpha}
\end{aligned} \tag{5.69}
$$

Hereby, approximation $\left(\frac{\tilde{w}_m \tilde{w}_n}{\|\tilde{\bar{v}}\|}\right) = \overline{\left(\frac{\tilde{w}_m \tilde{w}_n}{\|\tilde{\bar{v}}\|}\right)}^* = \left(\frac{\tilde{w}_m \tilde{w}_n}{\|\tilde{\bar{v}}\|}\right)_{\eta_i=0}$, though being still second order "super convergent" and quite good enough, is no longer superior to a FE interpolation, i.e.,

$$
\left(\frac{\tilde{w}_m \tilde{w}_n}{\|\tilde{\bar{v}}\|}\right) = \left(\frac{\tilde{w}_m \tilde{w}_n}{\|\tilde{\bar{v}}\|}\right)_\beta N_\beta \tag{5.70}
$$

— since $\tilde{v}_i$ (hence $\tilde{w}_i$) now can be obtained directly/exactly (at every step of the numerical solution) on FE nodes. However, if (5.70) is indeed enforced, (5.69) should be modified into

$$\begin{aligned}
&\int_{\Omega^e} \left(\beta^* \Delta\eta \frac{w_j^h w_k^h}{\|\bar{\tilde{v}}\|}\right) \frac{\partial N_\alpha}{\partial x_j} \frac{\partial U^h}{\partial x_k} d\Omega \\
&= (\beta^* \Delta\eta) \{U\}_\beta \int_{\Omega^*} \left(\frac{\tilde{w}_m^h \tilde{w}_n^h}{\|\bar{\tilde{v}}\|}\right) \frac{\partial N_\alpha}{\partial \eta_m} \frac{\partial N_\beta}{\partial \eta_n} D^e \, d\Omega^* \\
&= (\beta^* \Delta\eta) \left(\frac{\tilde{w}_m \tilde{w}_n}{\|\bar{\tilde{v}}\|}\right)_\theta \{U\}_\beta \int_{\Omega^*} \frac{\partial N_\alpha}{\partial \eta_m} N_\theta \frac{\partial N_\beta}{\partial \eta_n} D^e \, d\Omega^* \\
&= (\beta^* \Delta\eta) \left(\frac{\tilde{w}_M \tilde{w}_N}{\|\bar{\tilde{v}}\|}\right)_\theta \{U\}_\beta \, \overline{D}^e_{\alpha\theta\underline{\beta}} \, B3N0M_{\beta\theta\alpha}
\end{aligned} \tag{5.71}$$

where

$$\overline{D}^e_{\alpha\theta\underline{\beta}} \, B3M0N_{\alpha\beta} = \begin{cases} \overline{D}^e_{\alpha\theta\underline{\beta}} \, B3101_{\alpha\theta\beta} \\ \overline{D}^e_{\alpha\theta\underline{\beta}} \, B3201_{\alpha\theta\beta} \\ \overline{D}^e_{\alpha\theta\underline{\beta}} \, B3102_{\alpha\theta\beta} \\ \overline{D}^e_{\alpha\theta\underline{\beta}} \, B3202_{\alpha\theta\beta} \end{cases} \tag{5.72}$$

$$= \begin{cases} \overline{D}^* \left(1 + A^e \overline{\varsigma 3101}_{\underline{ikm}} + B^e \overline{\varsigma 3000}_{\underline{jln}}\right) A3101_{ikm} A3000_{jln} \\ \overline{D}^* \left(1 + A^e \overline{\varsigma 3001}_{\underline{ikm}} + B^e \overline{\varsigma 3100}_{\underline{jln}}\right) A3001_{ikm} A3100_{jln} \\ \overline{D}^* \left(1 + A^e \overline{\varsigma 3100}_{\underline{ikm}} + B^e \overline{\varsigma 3001}_{\underline{jln}}\right) A3100_{ikm} A3001_{jln} \\ \overline{D}^* \left(1 + A^e \overline{\varsigma 3000}_{\underline{ikm}} + B^e \overline{\varsigma 3101}_{\underline{jln}}\right) A3000_{ikm} A3101_{jln} \end{cases}$$

Θ-implicit Scheme

The Θ-implicit scheme for (5.60) is

$$M200_{\alpha\beta} \{\Delta U\}^n_\beta + (1-\Theta) \Delta t \{R\}^n_\alpha + \Theta \Delta t \{R\}^{n+1}_\alpha = 0 \tag{5.73}$$

$$\{\Delta U\}^n_\beta = \{U\}^{n+1}_\beta - \{U\}^n_\beta \tag{5.74}$$

where, similar to CPF, the S-TWS coefficient $\overline{\left(\frac{\tilde{w}_m \tilde{w}_n}{\|\tilde{\bar{v}}\|}\right)}^*$ in $\{R\}_\alpha^{n+1}$ is "set back" to the previous time level (and for the same reason — a better stablized numerical process).

Newton Linearization/Iteration

Similar to CPF, the nonlinear algebra system (5.73) is solved via Newton iteration:

$$\begin{aligned} & \left[M200_{\alpha\beta} + \Theta\,\Delta t\, J_{\alpha\beta}^{p,n+1}\right] \{\delta U\}_\beta^{p,n+1} \\ = & -M200_{\alpha\beta}\left(\{U\}_\beta^{p,n+1} - \{U\}_\beta^n\right) \\ & -\Theta\,\Delta t\,\{R\}_\alpha^{p,n+1} - (1-\Theta)\,\Delta t\,\{R\}_\alpha^n \end{aligned} \tag{5.75}$$

where p dentoes a Newtonian step/approximation, and

$$\{U\}_\beta^{p+1,n+1} = \{U\}_\beta^{p,n+1} + \{\delta U\}_\beta^{p,n+1} \tag{5.76}$$

$$\begin{aligned} [J_{\alpha\beta}]^{n+1} &= \left[\frac{\partial \{R\}_\alpha}{\partial \{U\}_\beta}\right]^{n+1} \\ &= \begin{bmatrix} \frac{\partial RR_\alpha^e}{\partial \rho_\beta} & \frac{\partial RR_\alpha^e}{\partial m_{L_\beta}} & \frac{\partial RR_\alpha^e}{\partial E_\beta} \\ \frac{\partial RI_\alpha^e}{\partial \rho_\beta} & \frac{\partial RI_\alpha^e}{\partial m_{L_\beta}} & \frac{\partial RI_\alpha^e}{\partial E_\beta} \\ \frac{\partial RE_\alpha^e}{\partial \rho_\beta} & \frac{\partial RE_\alpha^e}{\partial m_{L_\beta}} & \frac{\partial RE_\alpha^e}{\partial E_\beta} \end{bmatrix}^{n+1} \\ &= \begin{bmatrix} JRR_{\alpha\beta} & JRL_{\alpha\beta} & JRE_{\alpha\beta} \\ JIR_{\alpha\beta} & JIL_{\alpha\beta} & JIE_{\alpha\beta} \\ JER_{\alpha\beta} & JEL_{\alpha\beta} & JEE_{\alpha\beta} \end{bmatrix}^{n+1} \end{aligned} \tag{5.77}$$

$$\begin{aligned} JRR_{\alpha\beta}^e &= (\beta^*\,\Delta\eta)\,\overline{\left(\frac{\tilde{v}_M \tilde{v}_N}{\|\tilde{\bar{v}}\|}\right)}^*\,\overline{D}_{\underline{\alpha\beta}}^e\, B2MN_{\alpha\beta} \\ & -(\beta^*\,\Delta\eta)\oint_{\partial\Omega^e\cap\partial\Omega^h}\left(\frac{v_J v_K}{\|\tilde{\bar{v}}\|}\right) N_\alpha \frac{\partial N_\beta}{\partial x_K}\, n_J\, d\sigma \end{aligned} \tag{5.78}$$

$$JRL_{\alpha\beta}^e = -\delta_{JL}\, M2J0_{\alpha\beta}^e + \frac{\partial}{\partial m_{L_\beta}}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\, m_J\, n_J\, d\sigma$$

$$= -M2L0^e_{\alpha\beta} + \frac{\partial}{\partial m_{L_\beta}} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, m_J \, n_J \, d\sigma \tag{5.79}$$

$$JRE^e_{\alpha\beta} = 0 \tag{5.80}$$

$$\begin{aligned} JIR^e_{\alpha\beta} &= -m_{I_\theta} \left(\frac{\partial v_J}{\partial \rho}\right)_{\underline{\beta}} M30J0^e_{\theta\alpha\beta} + \left(\frac{\partial p}{\partial \rho}\right)_{\underline{\beta}} M20I^e_{\alpha\beta} \\ &\quad + \frac{\partial}{\partial \rho_\beta} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, m_I \, v_J \, n_J \, d\sigma \\ &= m_{I_\theta} \left(\frac{v_J}{\rho}\right)_{\underline{\beta}} M30J0^e_{\theta\alpha\beta} + \frac{\gamma - 1}{2} \left(v_K^2\right)_{\underline{\beta}} M20I^e_{\alpha\beta} \\ &\quad + \frac{\partial}{\partial \rho_\beta} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, m_I \, v_J \, n_J \, d\sigma \end{aligned} \tag{5.81}$$

$$\begin{aligned} JIL^e_{\alpha\beta} &= \left(\frac{\partial p}{\partial m_L}\right)_{\underline{\beta}} M20I^e_{\alpha\beta} \\ &\quad - \left[\left(\frac{\partial m_I}{\partial m_L}\right)_{\underline{\beta}} v_{J_\theta} + m_{I_\theta} \left(\frac{\partial v_J}{\partial m_L}\right)_{\underline{\beta}}\right] M30J0^e_{\theta\alpha\beta} \\ &\quad + (\beta^* \, \Delta\eta) \left(\Gamma_{\underline{IM}} \, \Gamma_{\underline{IN}}\right) \overline{\left(\frac{\tilde{v}_M \tilde{v}_N}{\|\tilde{\bar{v}}\|}\right)}^* \left(\frac{\partial m_I}{\partial m_L}\right)_{\underline{\beta}} \overline{D}^e_{\underline{\alpha\beta}} \, B2MN_{\alpha\beta} \\ &\quad - (\beta^* \, \Delta\eta) \left(\Gamma_{\underline{IJ}} \, \Gamma_{\underline{IK}}\right) \left(\frac{\partial m_I}{\partial m_L}\right)_{\underline{\beta}} \oint_{\partial\Omega^e \cap \partial\Omega^h} \left(\frac{v_J v_K}{\|\tilde{\bar{v}}\|}\right) N_\alpha \frac{\partial N_\beta}{\partial x_K} \, n_J \, d\sigma \\ &\quad + \frac{\partial}{\partial m_{L_\beta}} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, m_I \, v_J \, n_J \, d\sigma \\ &= -(\gamma - 1) \, v_{L_{\underline{\beta}}} M20I^e_{\alpha\beta} \\ &\quad - \delta_{IL} \, v_{J_\theta} M30J0^e_{\theta\alpha\beta} - \frac{m_{I_\theta}}{\rho_{\underline{\beta}}} M30L0^e_{\theta\alpha\beta} \\ &\quad + (\beta^* \, \Delta\eta) \left(\Gamma_{\underline{IM}} \, \Gamma_{\underline{IN}}\right) \overline{\left(\frac{\tilde{v}_M \tilde{v}_N}{\|\tilde{\bar{v}}\|}\right)}^* \delta_{IL} \, \overline{D}^e_{\underline{\alpha\beta}} \, B2MN_{\alpha\beta} \\ &\quad - (\beta^* \, \Delta\eta) \left(\Gamma_{\underline{IJ}} \, \Gamma_{\underline{IK}}\right) \delta_{IL} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \left(\frac{v_J v_K}{\|\tilde{\bar{v}}\|}\right) \frac{\partial N_\beta}{\partial x_K} \, n_J \, d\sigma \\ &\quad + \frac{\partial}{\partial m_{L_\beta}} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, m_I \, v_J \, n_J \, d\sigma \end{aligned} \tag{5.82}$$

$$\begin{aligned} JIE^e_{\alpha\beta} &= \left(\frac{\partial p}{\partial E}\right)_{\underline{\beta}} M20I^e_{\alpha\beta} \\ &= (\gamma - 1) \, M20I^e_{\alpha\beta} \end{aligned} \tag{5.83}$$

$$
\begin{aligned}
JER^e_{\alpha\beta} &= -\left[\left(\frac{\partial p}{\partial \rho}\right)_{\underline{\beta}} v_{J_\theta} + (E+p)_\theta \left(\frac{\partial v_J}{\partial \rho}\right)_{\underline{\beta}}\right] M30J0^e_{\theta\alpha\beta} \\
&\quad + \frac{\partial}{\partial \rho_\beta} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, (E+p) \, v_J \, n_J \, d\sigma \\
&= -\left[\frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{\beta}} v_{J_\theta} - (E+p)_\theta \left(\frac{v_J}{\rho}\right)_{\underline{\beta}}\right] M30J0^e_{\theta\alpha\beta} \\
&\quad + \frac{\partial}{\partial \rho_\beta} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, (E+p) \, v_J \, n_J \, d\sigma \qquad (5.84) \\
JEL^e_{\alpha\beta} &= -\left[\left(\frac{\partial p}{\partial m_L}\right)_{\underline{\beta}} v_{J_\theta} + (E+p)_\theta \left(\frac{\partial v_J}{\partial m_L}\right)_{\underline{\beta}}\right] M30J0^e_{\theta\alpha\beta} \\
&\quad + \frac{\partial}{\partial m_{L_\beta}} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, (E+p) \, v_J \, n_J \, d\sigma \\
&= -(1-\gamma) \, v_{L_{\underline{\beta}}} v_{J_\theta} M30J0^e_{\theta\alpha\beta} + \frac{1}{\rho_{\underline{\beta}}} (E+p)_\theta \, M30L0^e_{\theta\alpha\beta} \\
&\quad + \frac{\partial}{\partial m_{L_\beta}} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, (E+p) \, v_J \, n_J \, d\sigma \qquad (5.85) \\
JEE^e_{\alpha\beta} &= -\left(1 + \frac{\partial p}{\partial E}\right)_{\underline{\beta}} v_{J_\theta} M30J0^e_{\theta\alpha\beta} \\
&\quad + (\beta^* \Delta\eta) \, \gamma^2 \overline{\left(\frac{\tilde{v}_M \tilde{v}_N}{\|\tilde{\bar{v}}\|}\right)^*} \, \overline{D}^e_{\underline{\alpha\beta}} B2MN_{\alpha\beta} \\
&\quad + \frac{\partial}{\partial E_\beta} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, (E+p) \, v_J \, n_J \, d\sigma \\
&\quad - (\beta^* \Delta\eta) \, \gamma^2 \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \left(\frac{v_J v_K}{\|\tilde{\bar{v}}\|}\right) \frac{\partial N_\beta}{\partial x_K} n_J \, d\sigma \\
&= -\gamma \, v_{J_\theta} M30J0^e_{\theta\alpha\beta} \\
&\quad + (\beta^* \Delta\eta) \, \gamma^2 \overline{\left(\frac{\tilde{v}_M \tilde{v}_N}{\|\tilde{\bar{v}}\|}\right)^*} \, \overline{D}^e_{\underline{\alpha\beta}} B2MN_{\alpha\beta} \\
&\quad + \frac{\partial}{\partial E_\beta} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \, (E+p) \, v_J \, n_J \, d\sigma \\
&\quad - (\beta^* \Delta\eta) \, \gamma^2 \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \left(\frac{v_J v_K}{\|\tilde{\bar{v}}\|}\right) \frac{\partial N_\beta}{\partial x_K} n_J \, d\sigma \qquad (5.86)
\end{aligned}
$$

It is also similar to CPF that boundary contribution $\frac{\partial}{\partial\{U\}_\beta} \oint_{\partial\Omega^e \cap \partial\Omega^h} N_\alpha \{F^c_j\} \, n_J d\sigma$ is "indefinite" in the sense of depending on boundary conditions. If $\partial\Omega^e \cap \partial\Omega^h$ is

where $(\{F_J^c\}\, n_J)$ is given,

$$\left[\frac{\partial}{\partial\{U\}_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\ \{F_J^c\}\ n_J\, d\sigma\right]\ =\ [0] \tag{5.87}$$

Whereas, if $(\{F_J^c\}\, n_J)_{\partial\Omega^e\cap\partial\Omega^h}$ is *free*, it is then

$$\begin{aligned}
&\left[\frac{\partial}{\partial\{U\}_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\ \{F_J^c\}\ n_J\, d\sigma\right]^{1r}\\
&= \oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha \left\{\begin{array}{c}\left(\frac{\partial m_J}{\partial\rho}\right)_{\underline{\beta}}\\ \left(\frac{\partial m_J}{\partial m_L}\right)_{\underline{\beta}}\\ \left(\frac{\partial m_J}{\partial E}\right)_{\underline{\beta}}\end{array}\right\}^T N_\beta\, n_J\, d\sigma\\
&= \oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha \left\{\begin{array}{c}0\\ \delta_{JL}\\ 0\end{array}\right\}^T N_\beta\, n_J\, d\sigma \qquad (5.88)\\
&\left[\frac{\partial}{\partial\{U\}_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\ \{F_J^c\}\ n_J\, d\sigma\right]^{1+ir}\\
&= \oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha \left\{\begin{array}{c}\left(\frac{\partial m_I}{\partial\rho}\right)_{\underline{\beta}} v_{J_\theta} + m_{I_\theta}\left(\frac{\partial v_J}{\partial\rho}\right)_{\underline{\beta}}\\ \left(\frac{\partial m_I}{\partial m_L}\right)_{\underline{\beta}} v_{J_\theta} + m_{I_\theta}\left(\frac{\partial v_J}{\partial m_L}\right)_{\underline{\beta}}\\ \left(\frac{\partial m_I}{\partial E}\right)_{\underline{\beta}} v_{J_\theta} + m_{I_\theta}\left(\frac{\partial v_J}{\partial E}\right)_{\underline{\beta}}\end{array}\right\}^T N_\beta\, N_\theta\, n_J\, d\sigma\\
&= \oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha \left\{\begin{array}{c}-m_{I_\theta}\left(\frac{v_J}{\rho}\right)_{\underline{\beta}}\\ \delta_{IL}\, v_{J_\theta} + \delta_{JL}\frac{m_{I_\theta}}{\rho_{\underline{\beta}}}\\ 0\end{array}\right\}^T N_\beta\, N_\theta\, n_J\, d\sigma \qquad (5.89)\\
&\left[\frac{\partial}{\partial\{U\}_\beta}\oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha\ \{F_J^c\}\ n_J\, d\sigma\right]^{2+nr}
\end{aligned}$$

$$
\begin{aligned}
&= \oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha \left\{ \begin{array}{c} \left(\frac{\partial p}{\partial \rho}\right)_{\underline{\beta}} v_{J_\theta} + (E+p)_\theta \left(\frac{\partial v_J}{\partial \rho}\right)_{\underline{\beta}} \\ \left(\frac{\partial p}{\partial m_L}\right)_{\underline{\beta}} v_{J_\theta} + (E+p)_\theta \left(\frac{\partial v_J}{\partial m_L}\right)_{\underline{\beta}} \\ \left[1 + \left(\frac{\partial p}{\partial E}\right)_{\underline{\beta}}\right] v_{J_\theta} + (E+p)_\theta \left(\frac{\partial v_J}{\partial E}\right)_{\underline{\beta}} \end{array} \right\}^T N_\beta N_\theta n_J \, d\sigma \\
&= \oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha \left\{ \begin{array}{c} \frac{\gamma-1}{2} (v_K^2)_{\underline{\beta}} v_{J_\theta} - (E+p)_\theta \left(\frac{v_J}{\rho}\right)_{\underline{\beta}} \\ (1-\gamma) v_{L_{\underline{\beta}}} v_{J_\theta} + \delta_{JL} \frac{(E+p)_\theta}{\rho_{\underline{\beta}}} \\ \gamma v_{J_\theta} \end{array} \right\}^T N_\beta N_\theta n_J \, d\sigma \qquad (5.90)
\end{aligned}
$$

where $\left[\frac{\partial}{\partial\{U\}_\beta} \oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha \{F_j^c\} n_J \, d\sigma\right]^{1r}$, $\left[\frac{\partial}{\partial\{U\}_\beta} \oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha \{F_j^c\} n_J \, d\sigma\right]^{1+ir}$ and $\left[\frac{\partial}{\partial\{U\}_\beta} \oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha \{F_j^c\} n_J \, d\sigma\right]^{2+nr}$ are respectively the first, $(1+i)$th and $(2+n)$th row of $\left[\frac{\partial}{\partial\{U\}_\beta} \oint_{\partial\Omega^e\cap\partial\Omega^h} N_\alpha \{F_j^c\} n_J \, d\sigma\right]$ (cooresponding to the continuity, ith momentum and energy equation).

Remarks

The above FE algorithm endures the same "inconsistency" as that of the $\rho^{\gamma-1}$ term in the energy equation of CPF. Theoretically, such an "inconsistency" may be simply overcome by enforcing $F_j^{c^h}$ and $F_j^{p^h}$ *strictly* in terms of $U = \{\rho, m_i, E\}^T$, i.e., letting (see (4.139-4.140))

$$
F_j^{c^h}(t, x_j) = \left\{ m_j^h, \frac{m_i^h m_j^h}{\rho^h}, \left[\gamma E^h - (\gamma-1)\frac{m_i^h m_i^h}{2\rho^h}\right] \frac{m_j^h}{\rho^h} \right\}^T \qquad (5.91)
$$

$$
F_j^{p^h}(t, x_j) = \left\{ 0, (\gamma-1)\left(E^h - \frac{m_i^h m_i^h}{2\rho^h}\right)\delta_{ij}, 0 \right\}^T \qquad (5.92)
$$

The problem, however, is whether this is numerically feasible, considering that many FE matrices would be thus *rationalized.*

5.4 Tensor Product Factorization

As described, the task of solving both (5.45) and (5.75) is to inverse Jacobian matrix $[M200_{\alpha\beta} + \Theta\Delta t J_{\alpha\beta}] = \sum_e \left[\overline{D}^e_{\underline{\alpha\beta}} B200_{\alpha\beta} + \Theta\Delta t J^e_{\alpha\beta}\right]$, which is *block* 9-diagonal in both cases. In this section, the tensor product (TP) factorization of such a matrix is to be established, which is possible because of the TP nature of the bilinear FE, and which makes possible that the solution may be pursued in an ADI fashion ([2, 20, 21]).

5.4.1 General Description

The general form of the TP factorization of $[M200_{\alpha\beta} + \Theta\Delta t J_{\alpha\beta}]$ is as follows (elementally written).

$$
\begin{aligned}
& \left[\overline{D}^e_{\underline{\alpha\beta}}\, B200_{\alpha\beta} + \Theta\,\Delta t\, J^e_{\alpha\beta}\right] \\
= \; & \left[\overline{D}^e_{\underline{ijmn}}\, A200_{im}\, A200_{jn} + \Theta\,\Delta t\, J^e_{\alpha\beta}\right] \\
\simeq \; & \left[\overline{D}^e_{\underline{ijmn}}\, A200_{im}\, A200_{jn} + \Theta\,\Delta t\, A200_{im}\, JY^{\underline{im}\,e}_{jn} + \Theta\,\Delta t\, A200_{jn}\, JX^{\underline{jn}\,e}_{im}\right] \\
\simeq \; & \overline{D}^e_{\underline{ijmn}} \left[A200_{im} + \frac{\Theta\,\Delta t}{\overline{D}^e_{\underline{ijmn}}} JX^{\underline{jn}\,e}_{im}\right] \left[A200_{jn} + \frac{\Theta\,\Delta t}{\overline{D}^e_{\underline{ijmn}}} JY^{\underline{im}\,e}_{jn}\right] \qquad (5.93)
\end{aligned}
$$

where $\alpha \Leftrightarrow (i,j)$ and $\beta \Leftrightarrow (m,n)$, and JY^e_{jn} and JX^e_{im} are the *approximate* factors of $J^e_{\alpha\beta}$, each associating with an 1-D mass matrix along the other direction, and both pending to be derived concretely per the problem concerned.

As shown, there are two "$\simeq$" steps in (5.93), in equating $J_{\alpha\beta}$ with $A200_{im}JY^e_{jn} + A200_{jn}JX^e_{im}$, and in the eventual "factorization". The error in the latter is obvious, i.e., a second order $\Theta^2\Delta t^2 A200_{im}JX^e_{im}A200_{jn}JX^e_{jn}$; the error in the former is first

order, resulting from disgarding the *cross* FE matrix and boundary integral terms ([2, 20, 21], see § 5.4.2 and § 5.4.3). However, because of these errors, the validity/effectiveness of TP factorization relies on a small Δt.

Considering further the following factorization of $\overline{D}^e_{ijmn}$ —

$$\begin{aligned} \overline{D}^e_{ijmn} &= \overline{D^*}\left(1 + A^e\,\overline{\varsigma 200}_{im} + B^e\,\overline{\varsigma 200}_{jn}\right) \\ &\simeq \overline{D^*}\left(1 + A^e\,\overline{\varsigma 200}_{im}\right)\left(1 + B^e\,\overline{\varsigma 200}_{jn}\right) \end{aligned} \tag{5.94}$$

plus the fact that $\overline{\varsigma 200}_{im} \neq 0$ only if $i = m$, we then have

$$\begin{aligned} & \left[\overline{D}^e_{\underline{\alpha\beta}}\,B200_{\alpha\beta} + \Theta\,\Delta t\,J_{\alpha\beta}\right] \\ \simeq\ & \overline{DX}^e_{\underline{i}}\left[A200_{im} + \frac{\Theta\,\Delta t}{\overline{D}^e_{\underline{ijmn}}}\,JX^{\underline{jn}\,e}_{\underline{im}}\right]\overline{DY}^e_{\underline{j}}\left[A200_{jn} + \frac{\Theta\,\Delta t}{\overline{D}^e_{\underline{ijmn}}}\,JY^{\underline{im}\,e}_{\underline{jn}}\right] \\ \simeq\ & \left[\overline{DX}^e_{\underline{i}}\,A200_{im} + \frac{\Theta\,\Delta t}{\overline{DY}^e_{\underline{j}}}\,JX^{\underline{jn}\,e}_{\underline{im}}\right]\left[\overline{DY}^e_{\underline{j}}\,A200_{jn} + \frac{\Theta\,\Delta t}{\overline{DX}^e_{\underline{i}}}\,JY^{\underline{im}\,e}_{\underline{jn}}\right] \end{aligned} \tag{5.95}$$

where

$$\overline{DX}^e_i = \overline{D^*}\left(1 + A^e\,\overline{\varsigma 200}_{i\underline{i}}\right) \tag{5.96}$$

$$\overline{DY}^e_j = \overline{D^*}\left(1 + B^e\,\overline{\varsigma 200}_{j\underline{j}}\right) \tag{5.97}$$

Equation (5.95) is the general form of the TP approximate factorization — one that is MDM based. Subsequent to it, (5.45) and (5.75), i.e.,

$$\left[M200_{\alpha\beta} + \Theta\,\Delta t\,J_{\alpha\beta}\right]\{\delta Q\}_\beta = -\{F\}_\alpha \tag{5.98}$$

where $\{F\}_\alpha$ represents the right-hand-sides of (5.45) and (5.75), becomes

$$\left[X200^{\underline{j}}_{\underline{im}} + \Theta\,\Delta t\,JDX^{\underline{j}}_{\underline{im}}\right]\left[Y200^{\underline{i}}_{\underline{jn}} + \Theta\,\Delta t\,JDY^{\underline{i}}_{\underline{jn}}\right]\{\delta Q\}_{mn} = -\{F\}_{ij} \tag{5.99}$$

where $\left[X200^{j}_{im} + \Theta\Delta t JDX^{j}_{im}\right]$ is the assembled $\left[\overline{DX^{e}_{\underline{i}}}A200_{im} + \frac{\Theta\Delta t}{\overline{DY^{e}_{\underline{j}}}}JX^{\underline{jn}\,e}_{\overline{im}}\right]$ along the η_1-direction to the jth mesh-line, and $\left[Y200^{i}_{jn} + \Theta\Delta t JDY^{i}_{jn}\right]$ the assembled $\left[\overline{DY^{e}_{\underline{j}}}A200_{jn} + \frac{\Theta\Delta t}{\overline{DX^{e}_{\underline{i}}}}JY^{\underline{im}\,e}_{\overline{jn}}\right]$ along the η_2-direction to ith mesh-line. Consequently, the solution of (5.98) can be *approximately* approached in a two-step sequence:

$$\left[X200^{j}_{im} + \Theta\,\Delta t\,JDX^{j}_{\overline{im}}\right]\{P\}^{j}_{m} = -\{F\}_{ij} \tag{5.100}$$

$$\left[Y200^{i}_{jn} + \Theta\,\Delta t\,JDY^{i}_{\overline{jn}}\right]\{\delta Q\}_{mn} = \{P\}^{j}_{m} \tag{5.101}$$

of which each involves inversing a *block* tri-diagonal matrix.

Remarks

Additional error $A^e B^{\underline{e}}\overline{\varsigma 200}_{im}\overline{\varsigma 200}_{jn}$ is invoked in (5.94). Correspondingly, the validity/effectiveness of procedure (5.100-5.101) relies on not only a small Δt but small A^e and B^e, which is the case when the *un-paralellgram-ness* of the bilinear FE mesh-cells is not severe (see chapter 3).

5.4.2 Compressible Potential Flow

Jacobian Entries

$$\begin{aligned}
J11^{e}_{\alpha\beta} &\simeq -\phi_\theta\, M3J0J^{e}_{\alpha\beta\theta} + (\beta^*\Delta\eta)\,\overline{\left(\tau\frac{\tilde{v}_M\tilde{v}_N}{\|\tilde{\tilde{v}}\|}\right)}^{*}\overline{D}^{e}\,B2MN_{\alpha\beta} \\
&= -\phi_\theta\left(\frac{E_{1J}E_{1J}}{D}\right)^{e} B3101_{\alpha\beta\theta} - \phi_\theta\left(\frac{E_{2J}E_{2J}}{D}\right)^{e} B3202_{\alpha\beta\theta} \\
&\quad + (\beta^*\Delta\eta)\,\overline{\left(\tau\frac{\tilde{v}_M\tilde{v}_N}{\|\tilde{\tilde{v}}\|}\right)}^{*}\overline{D}^{e}\,B2MN_{\alpha\beta} \\
&\simeq -\overline{\left(\frac{\partial\phi}{\partial\eta_1}\right)}^{\underline{e}}\left(\frac{E_{1J}E_{1J}}{D}\right)^{e} B210_{\alpha\beta}
\end{aligned}$$

$$
\begin{aligned}
& - \overline{\left(\frac{\partial \phi}{\partial \eta_2}\right)}^{\underline{e}} \left(\frac{E_{2J} E_{2J}}{D}\right)^{e} B220_{\alpha\beta} \\
& + (\beta^* \Delta\eta) \overline{\left(\tau \frac{\tilde{v}_1 \tilde{v}_1}{\|\tilde{\bar{v}}\|}\right)}^{*} \overline{D}^{e} B211_{\alpha\beta} \\
& + (\beta^* \Delta\eta) \overline{\left(\tau \frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)}^{*} \overline{D}^{e} B222_{\alpha\beta} \\
= & - \overline{\left(\frac{\partial \phi}{\partial \eta_1}\right)}^{\underline{e}} \left(\frac{E_{1J} E_{1J}}{D}\right)^{e} A210_{im} A200_{jn} \\
& - \overline{\left(\frac{\partial \phi}{\partial \eta_2}\right)}^{\underline{e}} \left(\frac{E_{2J} E_{2J}}{D}\right)^{e} A200_{im} A210_{jn} \\
& + (\beta^* \Delta\eta) \overline{\left(\tau \frac{\tilde{v}_1 \tilde{v}_1}{\|\tilde{\bar{v}}\|}\right)}^{*} \overline{D}^{e} A211_{im} A200_{jn} \\
& + (\beta^* \Delta\eta) \overline{\left(\tau \frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)}^{*} \overline{D}^{e} A200_{im} A211_{jn} \qquad (5.102)
\end{aligned}
$$

$$
\begin{aligned}
J12^{e}_{\alpha\beta} &\simeq \rho_\theta \, M3J0J^{e}_{\alpha\theta\beta} \\
&\simeq \rho_\theta \left(\frac{E_{1J} E_{1J}}{D}\right)^{e} B3101_{\alpha\theta\beta} + \rho_\theta \left(\frac{E_{2J} E_{2J}}{D}\right)^{e} B3202_{\alpha\theta\beta} \\
&= \overline{\rho^{\underline{e}}} \left(\frac{E_{1J} E_{1J}}{D}\right)^{e} B211_{\alpha\beta} + \overline{\rho^{\underline{e}}} \left(\frac{E_{2J} E_{2J}}{D}\right)^{e} B222_{\alpha\beta} \\
&= \overline{\rho^{\underline{e}}} \left(\frac{E_{1J} E_{1J}}{D}\right)^{e} A211_{im} A200_{jn} + \\
&\quad \overline{\rho^{\underline{e}}} \left(\frac{E_{2J} E_{2J}}{D}\right)^{e} A200_{im} A211_{jn} \qquad (5.103)
\end{aligned}
$$

$$
\begin{aligned}
J21^{e}_{\alpha\beta} &= \gamma \frac{T_\infty}{\rho_\infty^{\gamma-1}} \rho_{\underline{\beta}}^{\gamma-2} M200^{e}_{\alpha\beta} \\
&= \gamma \frac{T_\infty}{\rho_\infty^{\gamma-1}} \rho_{\underline{mn}}^{\gamma-2} \overline{D}^{e} A200_{im} A200_{jn} \\
&= \frac{\gamma}{2} \frac{T_\infty}{\rho_\infty^{\gamma-1}} \rho_{\underline{mn}}^{\gamma-2} \overline{D}^{e} A200_{im} A200_{jn} + \\
&\quad \frac{\gamma}{2} \frac{T_\infty}{\rho_\infty^{\gamma-1}} \rho_{\underline{mn}}^{\gamma-2} \overline{D}^{e} A200_{im} A200_{jn} \qquad (5.104)
\end{aligned}
$$

$$
\begin{aligned}
J22^{e}_{\alpha\beta} &= \phi_\theta \, M3J0J^{e}_{\theta\alpha\beta} \\
&\simeq \phi_\theta \left(\frac{E_{1J} E_{1J}}{D}\right)^{e} B3101_{\theta\alpha\beta} + \phi_\theta \left(\frac{E_{2J} E_{2J}}{D}\right)^{e} B3202_{\theta\alpha\beta}
\end{aligned}
$$

$$= \overline{\left(\frac{\partial\phi}{\partial\eta_1}\right)}^{\underline{e}} \left(\frac{E_{1J}E_{1J}}{D}\right)^{e} B201_{\alpha\beta} + \overline{\left(\frac{\partial\phi}{\partial\eta_2}\right)}^{\underline{e}} \left(\frac{E_{2J}E_{2J}}{D}\right)^{e} B202_{\alpha\beta}$$

$$= \overline{\left(\frac{\partial\phi}{\partial\eta_1}\right)}^{e} \left(\frac{E_{1J}E_{1J}}{D}\right)^{e} A201_{im}\, A200_{jn} +$$

$$\overline{\left(\frac{\partial\phi}{\partial\eta_2}\right)}^{e} \left(\frac{E_{2J}E_{2J}}{D}\right)^{e} A200_{im}\, A201_{jn} \qquad (5.105)$$

Subsequently,

$$JX11^{e}_{im} = -\overline{\left(\frac{\partial\phi}{\partial\eta_1}\right)}^{\underline{e}} \left(\frac{E_{1J}E_{1J}}{D}\right)^{e} A210_{im}$$

$$+ (\beta^* \Delta\eta) \overline{\left(\tau \frac{\tilde{v}_1\tilde{v}_1}{\|\tilde{\bar{v}}\|}\right)}^{*} \overline{D}^{e}\, A211_{im} \qquad (5.106)$$

$$JY11^{e}_{jn} = -\overline{\left(\frac{\partial\phi}{\partial\eta_2}\right)}^{\underline{e}} \left(\frac{E_{2J}E_{2J}}{D}\right)^{e} A210_{jn}$$

$$+ (\beta^* \Delta\eta) \overline{\left(\tau \frac{\tilde{v}_2\tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)}^{*} \overline{D}^{e}\, A211_{jn} \qquad (5.107)$$

$$JX12^{e}_{im} = \overline{\rho}^{\underline{e}} \left(\frac{E_{1J}E_{1J}}{D}\right)^{e} A211_{im} \qquad (5.108)$$

$$JY12^{e}_{jn} = \overline{\rho}^{\underline{e}} \left(\frac{E_{2J}E_{2J}}{D}\right)^{e} A211_{jn} \qquad (5.109)$$

$$JX21^{e}_{im} = \frac{\gamma}{2} \frac{T_\infty}{\rho_\infty^{\gamma-1}} \rho_{\underline{mn}}^{\gamma-2} \overline{D}^{e}\, A200_{im} \qquad (5.110)$$

$$JY21^{e}_{\alpha\beta} = \frac{\gamma}{2} \frac{T_\infty}{\rho_\infty^{\gamma-1}} \rho_{\underline{mn}}^{\gamma-2} \overline{D}^{e}\, A200_{jn} \qquad (5.111)$$

$$JX22^{e}_{\alpha\beta} = \overline{\left(\frac{\partial\phi}{\partial\eta_1}\right)}^{\underline{e}} \left(\frac{E_{1J}E_{1J}}{D}\right)^{e} A201_{im} \qquad (5.112)$$

$$JY22^{e}_{\alpha\beta} = \overline{\left(\frac{\partial\phi}{\partial\eta_2}\right)}^{\underline{e}} \left(\frac{E_{2J}E_{2J}}{D}\right)^{e} A201_{jn} \qquad (5.113)$$

Remarks

1. In addition to "extracting out" FE metric data, the MDM algorithm has been

used *contracting* dummy indexes in the above, such as the following process in (5.102):

$$\phi_\theta \, B3101_{\alpha\beta\theta} \;=\; \overline{\left(\frac{\partial \phi}{\partial \eta_1}\right)}^{e} \, B210_{\alpha\beta}$$

2. The entire TP factorization algorithm for CPF is as follows.

$$\left(X200^{\underline{j}}_{im} \;+\; \Theta\,\Delta t \left[\begin{array}{cc} JDX11 & JDX12 \\ JDX21 & JDX22 \end{array} \right]^{\underline{j}}_{im} \right)$$

$$\left(Y200^{\underline{i}}_{jn} \;+\; \Theta\,\Delta t \left[\begin{array}{cc} JDY11 & JDY12 \\ JDY21 & JDY22 \end{array} \right]^{\underline{i}}_{jn} \right) \left\{ \begin{array}{c} \rho \\ \phi \end{array} \right\}_{mn} \;=\; -\{F\}_{ij} \qquad (5.114)$$

5.4.3 Euler System

Jacobian Entries

$$\begin{aligned}
JRR^{e}_{\alpha\beta} \;&\simeq\; (\beta^* \Delta\eta) \, \overline{\left(\frac{\tilde{v}_M \tilde{v}_N}{\|\tilde{\bar{v}}\|}\right)^*} \, \overline{D}^{e} \, B2MN_{\alpha\beta} \\
&\simeq\; (\beta^* \Delta\eta) \, \overline{\left(\frac{\tilde{v}_1 \tilde{v}_1}{\|\tilde{\bar{v}}\|}\right)^*} \, \overline{D}^{e} \, B211_{\alpha\beta} \;+\; (\beta^* \Delta\eta) \, \overline{\left(\frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)^*} \, \overline{D}^{e} \, B222_{\alpha\beta} \\
&=\; (\beta^* \Delta\eta) \, \overline{\left(\frac{\tilde{v}_1 \tilde{v}_1}{\|\tilde{\bar{v}}\|}\right)^*} \, \overline{D}^{e} \, A211_{im} \, A200_{jn} \;+ \\
&\quad\; (\beta^* \Delta\eta) \, \overline{\left(\frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)^*} \, \overline{D}^{e} \, A200_{im} \, A211_{jn} \qquad (5.115)
\end{aligned}$$

$$\begin{aligned}
JRL^{e}_{\alpha\beta} \;&\simeq\; -\,M2L0^{e}_{\alpha\beta} \\
&=\; -\,(\overline{E_{1L}})^{e} \, B210_{\alpha\beta} \;-\; (\overline{E_{2L}})^{e} \, B220_{\alpha\beta} \\
&=\; -\,(\overline{E_{1L}})^{e} \, A210_{im} \, A200_{jn} \;-\; (\overline{E_{2L}})^{e} \, A200_{im} \, A210_{jn} \qquad (5.116)
\end{aligned}$$

$$JRE^{e}_{\alpha\beta} \;=\; 0 \qquad (5.117)$$

$$JIR^{e}_{\alpha\beta} \;\simeq\; m_{I_\theta} \left(\frac{v_J}{\rho}\right)_{\underline{\beta}} M30J0^{e}_{\theta\alpha\beta} \;+\; \frac{\gamma - 1}{2} \left(v_K^2\right)_{\underline{\beta}} M20I^{e}_{\alpha\beta}$$

$$
\begin{aligned}
&= m_{I_\theta}\left(\frac{v_J}{\rho}\right)_{\underline{\beta}}(\overline{E_{1J}})^e\,B3010_{\theta\alpha\beta} + m_{I_\theta}\left(\frac{v_J}{\rho}\right)_{\underline{\beta}}(\overline{E_{2J}})^e\,B3020_{\theta\alpha\beta} + \\
&\quad \frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{\beta}}(\overline{E_{1I}})^e\,B201_{\alpha\beta} + \frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{\beta}}(\overline{E_{2I}})^e\,B202_{\alpha\beta} \\
&= \overline{(m_I)}^e\left(\frac{v_J}{\rho}\right)_{\underline{\beta}}(\overline{E_{1J}})^e\,B210_{\alpha\beta} + \overline{(m_I)}^e\left(\frac{v_J}{\rho}\right)_{\underline{\beta}}(\overline{E_{2J}})^e\,B220_{\alpha\beta} + \\
&\quad \frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{\beta}}(\overline{E_{1I}})^e\,B201_{\alpha\beta} + \frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{\beta}}(\overline{E_{2I}})^e\,B202_{\alpha\beta} \\
&= \overline{(m_I)}^e\left(\frac{v_J}{\rho}\right)_{\underline{mn}}(\overline{E_{1J}})^e\,A210_{im}\,A200_{jn} + \\
&\quad \overline{(m_I)}^e\left(\frac{v_J}{\rho}\right)_{\underline{mn}}(\overline{E_{2J}})^e\,A200_{im}\,A210_{jn} + \\
&\quad \frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{mn}}(\overline{E_{1I}})^e\,A201_{im}\,A200_{jn} + \\
&\quad \frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{mn}}(\overline{E_{2I}})^e\,A200_{im}\,A201_{jn} \qquad (5.118)
\end{aligned}
$$

$$
\begin{aligned}
JIL^e_{\alpha\beta} &\simeq -(\gamma-1)\,v_{L_{\underline{\beta}}}\,M20I^e_{\alpha\beta} \\
&\quad -\delta_{IL}\,v_{J_\theta}\,M30J0^e_{\theta\alpha\beta} - \frac{m_{I_\theta}}{\rho_{\underline{\beta}}}\,M30L0^e_{\theta\alpha\beta} \\
&\quad +(\beta^*\,\Delta\eta)\left(\Gamma_{\underline{IM}}\,\Gamma_{\underline{IN}}\right)\overline{\left(\frac{\tilde{v}_M\tilde{v}_N}{\|\tilde{\bar{v}}\|}\right)}^*\delta_{IL}\,\overline{D}^e_{\underline{\alpha\beta}}\,B2MN_{\alpha\beta} \\
&\simeq -(\gamma-1)\,v_{L_{\underline{\beta}}}(\overline{E_{1I}})^e\,B201_{\alpha\beta} - (\gamma-1)\,v_{L_{\underline{\beta}}}(\overline{E_{2I}})^e\,B202_{\alpha\beta} \\
&\quad -\delta_{IL}\,v_{J_\theta}(\overline{E_{1J}})^e\,B3010_{\theta\alpha\beta} - \delta_{IL}\,v_{J_\theta}(\overline{E_{2J}})^e\,B3020_{\theta\alpha\beta} \\
&\quad -\frac{m_{I_\theta}}{\rho_{\underline{\beta}}}(\overline{E_{1L}})^e\,B3010_{\theta\alpha\beta} - \frac{m_{I_\theta}}{\rho_{\underline{\beta}}}(\overline{E_{2L}})^e\,B3020_{\theta\alpha\beta} \\
&\quad +(\beta^*\,\Delta\eta)\,\Gamma^2_{\underline{I1}}\,\delta_{IL}\overline{\left(\frac{\tilde{v}_1\tilde{v}_1}{\|\tilde{\bar{v}}\|}\right)}^*\overline{D}^e\,B211_{\alpha\beta} \\
&\quad +(\beta^*\,\Delta\eta)\,\Gamma^2_{\underline{I2}}\,\delta_{IL}\overline{\left(\frac{\tilde{v}_2\tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)}^*\overline{D}^e\,B222_{\alpha\beta} \\
&= -(\gamma-1)\,v_{L_{\underline{\beta}}}(\overline{E_{1I}})^e\,B201_{\alpha\beta} - (\gamma-1)\,v_{L_{\underline{\beta}}}(\overline{E_{2I}})^e\,B202_{\alpha\beta} \\
&\quad -\delta_{IL}\,\overline{v}^e_J(\overline{E_{1J}})^e\,B210_{\alpha\beta} - \delta_{IL}\,\overline{v}^e_J(\overline{E_{2J}})^e\,B220_{\alpha\beta}
\end{aligned}
$$

$$
\begin{aligned}
& - \frac{\overline{m_I^e}}{\rho_{\underline{\beta}}} (\overline{E_{1L}})^e \, B210_{\alpha\beta} - \frac{\overline{m_I^e}}{\rho_{\underline{\beta}}} (\overline{E_{2L}})^e \, B220_{\alpha\beta} \\
& + (\beta^* \, \Delta\eta) \, \Gamma^2_{\underline{I}1} \, \delta_{IL} \overline{\left(\frac{\tilde{v}_1 \tilde{v}_1}{\|\tilde{\bar{v}}\|} \right)^*} \, \overline{D}^e \, B211_{\alpha\beta} \\
& + (\beta^* \, \Delta\eta) \, \Gamma^2_{\underline{I}2} \, \delta_{IL} \overline{\left(\frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|} \right)^*} \, \overline{D}^e \, B222_{\alpha\beta} \\
= & - (\gamma - 1) \, v_{L_{\underline{mn}}} (\overline{E_{1I}})^e \, A201_{im} \, A200_{jn} \\
& - (\gamma - 1) \, v_{L_{\underline{mn}}} (\overline{E_{2I}})^e \, A200_{im} \, A201_{jn} \\
& - \delta_{IL} \, \overline{v_J^e} \, (\overline{E_{1J}})^e \, A210_{im} \, A200_{jn} \\
& - \delta_{IL} \, \overline{v_J^e} \, (\overline{E_{2J}})^e \, A200_{im} \, A210_{jn} \\
& - \frac{\overline{m_I^e}}{\rho_{\underline{mn}}} (\overline{E_{1L}})^e \, A210_{im} \, A200_{jn} \\
& - \frac{\overline{m_I^e}}{\rho_{\underline{mn}}} (\overline{E_{2L}})^e \, A200_{im} \, A210_{jn} \\
& + (\beta^* \, \Delta\eta) \, \Gamma^2_{\underline{I}1} \, \delta_{IL} \overline{\left(\frac{\tilde{v}_1 \tilde{v}_1}{\|\tilde{\bar{v}}\|} \right)^*} \, \overline{D}^e \, A211_{im} \, A200_{jn} \\
& + (\beta^* \, \Delta\eta) \, \Gamma^2_{\underline{I}2} \, \delta_{IL} \overline{\left(\frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|} \right)^*} \, \overline{D}^e \, A200_{im} \, A211_{jn} \qquad (5.119)
\end{aligned}
$$

$$
\begin{aligned}
JIE^e_{\alpha\beta} &= (\gamma - 1) \, M20I^e_{\alpha\beta} \\
&= (\gamma - 1) \, (\overline{E_{1I}})^e \, B201_{\alpha\beta} + (\gamma - 1) \, (\overline{E_{2I}})^e \, B202_{\alpha\beta} \\
&= (\gamma - 1) \, (\overline{E_{1I}})^e \, A201_{im} \, A200_{jn} + \\
&\quad (\gamma - 1) \, (\overline{E_{2I}})^e \, A200_{im} \, A201_{jn} \qquad (5.120)
\end{aligned}
$$

$$
JER^e_{\alpha\beta} \simeq - \left[\frac{\gamma - 1}{2} \left(v_K^2 \right)_{\underline{\beta}} v_{J_\theta} - (E + p)_\theta \left(\frac{v_J}{\rho} \right)_{\underline{\beta}} \right] M30J0^e_{\theta\alpha\beta} \qquad (5.121)
$$

$$
\begin{aligned}
= & - \left[\frac{\gamma - 1}{2} \left(v_K^2 \right)_{\underline{\beta}} v_{J_\theta} - (E + p)_\theta \left(\frac{v_J}{\rho} \right)_{\underline{\beta}} \right] (\overline{E_{1J}})^e \, B3010_{\theta\alpha\beta} \\
& - \left[\frac{\gamma - 1}{2} \left(v_K^2 \right)_{\underline{\beta}} v_{J_\theta} - (E + p)_\theta \left(\frac{v_J}{\rho} \right)_{\underline{\beta}} \right] (\overline{E_{2J}})^e \, B3020_{\theta\alpha\beta}
\end{aligned}
$$

$$
\begin{aligned}
&= -\left[\frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{\beta}}\overline{v_J}^e - \overline{(E+p)}^e\left(\frac{v_J}{\rho}\right)_{\underline{\beta}}\right](\overline{E_{1J}})^e\, B210_{\alpha\beta} \\
&\quad -\left[\frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{\beta}}\overline{v_J}^e - \overline{(E+p)}^e\left(\frac{v_J}{\rho}\right)_{\underline{\beta}}\right](\overline{E_{2J}})^e\, B220_{\alpha\beta} \\
&= -\left[\frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{mn}}\overline{v_J}^e - \overline{(E+p)}^e\left(\frac{v_J}{\rho}\right)_{\underline{mn}}\right](\overline{E_{1J}})^e\, A210_{im}\, A200_{jn} \\
&\quad -\left[\frac{\gamma-1}{2}\left(v_K^2\right)_{\underline{mn}}\overline{v_J}^e - \overline{(E+p)}^e\left(\frac{v_J}{\rho}\right)_{\underline{mn}}\right](\overline{E_{2J}})^e\, A200_{im}\, A210_{jn}
\end{aligned}
$$

$$
\begin{aligned}
JEL^e_{\alpha\beta} &\simeq -(1-\gamma)\, v_{L_{\underline{\beta}}}\, v_{J_\theta}\, M30J0^e_{\theta\alpha\beta} + \frac{1}{\rho_{\underline{\beta}}}(E+p)_\theta\, M30L0^e_{\theta\alpha\beta} \\
&= -(1-\gamma)\, v_{L_{\underline{\beta}}}\, v_{J_\theta}\left[(\overline{E_{1J}})^e\, B3010_{\theta\alpha\beta} + (\overline{E_{2J}})^e\, B3020_{\theta\alpha\beta}\right] \\
&\quad + \frac{1}{\rho_{\underline{\beta}}}(E+p)_\theta\left[(\overline{E_{1L}})^e\, B3010_{\theta\alpha\beta} + (\overline{E_{2L}})^e\, B3020_{\theta\alpha\beta}\right] \\
&= -(1-\gamma)\, v_{L_{\underline{\beta}}}\,\overline{v_J}^e\left[(\overline{E_{1J}})^e\, B210_{\alpha\beta} + (\overline{E_{2J}})^e\, B220_{\alpha\beta}\right] \\
&\quad + \frac{1}{\rho_{\underline{\beta}}}\overline{(E+p)}^e\left[(\overline{E_{1L}})^e\, B210_{\alpha\beta} + (\overline{E_{2L}})^e\, B220_{\alpha\beta}\right] \\
&= -(1-\gamma)\, v_{L_{\underline{mn}}}\,\overline{v_J}^e\,(\overline{E_{1J}})^e\, A210_{im}\, A200_{jn} \\
&\quad -(1-\gamma)\, v_{L_{\underline{mn}}}\,\overline{v_J}^e\,(\overline{E_{2J}})^e\, A200_{im}\, A210_{jn} \\
&\quad + \frac{1}{\rho_{\underline{mn}}}\overline{(E+p)}^e(\overline{E_{1L}})^e\, A210_{im}\, A200_{jn} \\
&\quad + \frac{1}{\rho_{\underline{mn}}}\overline{(E+p)}^e(\overline{E_{2L}})^e\, A200_{im}\, A210_{jn} \qquad (5.122)
\end{aligned}
$$

$$
\begin{aligned}
JEE^e_{\alpha\beta} &\simeq -\gamma\, v_{J_\theta}\, M30J0^e_{\theta\alpha\beta} + (\beta^*\,\Delta\eta)\,\gamma^2\,\overline{\left(\frac{\tilde{v}_M\tilde{v}_N}{\|\bar{\tilde{v}}\|}\right)}^*\,\overline{D}^e_{\underline{\alpha\beta}}\, B2MN_{\alpha\beta} \\
&= -\gamma\, v_{J_\theta}\left[(\overline{E_{1J}})^e\, B3010_{\theta\alpha\beta} + (\overline{E_{2J}})^e\, B3020_{\theta\alpha\beta}\right] \\
&\quad + (\beta^*\,\Delta\eta)\,\gamma^2\,\overline{\left(\frac{\tilde{v}_M\tilde{v}_N}{\|\bar{\tilde{v}}\|}\right)}^*\,\overline{D}^e_{\underline{\alpha\beta}}\, B2MN_{\alpha\beta} \\
&\simeq -\gamma\,\overline{v_J}^e\left[(\overline{E_{1J}})^e\, B210_{\alpha\beta} + (\overline{E_{2J}})^e\, B220_{\alpha\beta}\right] \\
&\quad + (\beta^*\,\Delta\eta)\,\gamma^2\,\overline{\left(\frac{\tilde{v}_1\tilde{v}_1}{\|\bar{\tilde{v}}\|}\right)}^*\,\overline{D}^e\, B211_{\alpha\beta}
\end{aligned}
$$

$$
\begin{aligned}
&+ (\beta^* \Delta\eta)\, \gamma^2 \overline{\left(\frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)}^* \overline{D}^e\, B222_{\alpha\beta} \\
= \; &- \gamma\, \overline{v_J}^e \left[(\overline{E_{1J}})^e\, A210_{im}\, A200_{jn} + (\overline{E_{2J}})^e\, A200_{im}\, A210_{jn}\right] \\
&+ (\beta^* \Delta\eta)\, \gamma^2 \overline{\left(\frac{\tilde{v}_1 \tilde{v}_1}{\|\tilde{\bar{v}}\|}\right)}^* \overline{D}^e\, A211_{im}\, A200_{jn} \\
&+ (\beta^* \Delta\eta)\, \gamma^2 \overline{\left(\frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)}^* \overline{D}^e\, A200_{im}\, A211_{jn} \qquad (5.123)
\end{aligned}
$$

Subsequently,

$$
JXRR^e_{im} = (\beta^* \Delta\eta) \overline{\left(\frac{\tilde{v}_1 \tilde{v}_1}{\|\tilde{\bar{v}}\|}\right)}^* \overline{D}^e\, A211_{im} \qquad (5.124)
$$

$$
JYRR^e_{jn} = (\beta^* \Delta\eta) \overline{\left(\frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)}^* \overline{D}^e\, A211_{jn} \qquad (5.125)
$$

$$
JXRL^e_{im} = -(\overline{E_{1L}})^e\, A210_{im} \qquad (5.126)
$$

$$
JYRL^e_{jn} = (\overline{E_{2L}})^e\, A210_{jn} \qquad (5.127)
$$

$$
JXRE^e_{im} = 0 \qquad (5.128)
$$

$$
JYRE^e_{jn} = 0 \qquad (5.129)
$$

$$
\begin{aligned}
JXIR^e_{im} = \; &\overline{(m_I)}^e \left(\frac{v_J}{\rho}\right)_{\underline{mn}} (\overline{E_{1J}})^e\, A210_{im} + \\
&\frac{\gamma - 1}{2} \left(v_K^2\right)_{\underline{mn}} (\overline{E_{1I}})^e\, A201_{im} \qquad (5.130)
\end{aligned}
$$

$$
\begin{aligned}
JYIR^e_{jn} = \; &\overline{(m_I)}^e \left(\frac{v_J}{\rho}\right)_{\underline{mn}} (\overline{E_{2J}})^e\, A210_{jn} + \\
&\frac{\gamma - 1}{2} \left(v_K^2\right)_{\underline{mn}} (\overline{E_{2I}})^e\, A201_{jn} \qquad (5.131)
\end{aligned}
$$

$$
\begin{aligned}
JXIL^e_{im} = \; &-(\gamma - 1)\, v_{L_{\underline{mn}}} (\overline{E_{1I}})^e\, A201_{im} \\
&- \delta_{IL}\, \overline{v_J}^e\, (\overline{E_{1J}})^e\, A210_{im} \\
&- \frac{\overline{m_I}^e}{\rho_{\underline{mn}}} (\overline{E_{1L}})^e\, A210_{im}
\end{aligned}
$$

$$+ (\beta^* \Delta\eta)\, \Gamma^2_{\underline{I}1}\, \delta_{IL} \overline{\left(\frac{\tilde{v}_1 \tilde{v}_1}{\|\tilde{\bar{v}}\|}\right)^*}\, \overline{D}^e\, A211_{im} \tag{5.132}$$

$$\begin{aligned} JYIL^e_{jn} &= -(\gamma-1)\, v_{L_{\underline{mn}}}\, (\overline{E_{2I}})^e\, A201_{jn} \\ &\quad - \delta_{IL}\, \overline{v^e_J}\, (\overline{E_{2J}})^e\, A210_{jn} \\ &\quad - \frac{\overline{m^e_I}}{\rho_{\underline{mn}}}\, (\overline{E_{2L}})^e\, A210_{jn} \\ &\quad + (\beta^* \Delta\eta)\, \Gamma^2_{\underline{I}2}\, \delta_{IL} \overline{\left(\frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)^*}\, \overline{D}^e\, A211_{jn} \end{aligned} \tag{5.133}$$

$$JXIE^e_{im} = (\gamma-1)\, (\overline{E_{1I}})^e\, A201_{im} \tag{5.134}$$

$$JYIE^e_{jn} = (\gamma-1)\, (\overline{E_{2I}})^e\, A201_{jn} \tag{5.135}$$

$$\begin{aligned} JXER^e_{im} &= -\frac{\gamma-1}{2} \left(v^2_K\right)_{\underline{mn}}\, \overline{v^e_J}\, A210_{im} \\ &\quad + \overline{(E+p)^e} \left(\frac{v_J}{\rho}\right)_{\underline{mn}} (\overline{E_{1J}})^e\, A210_{im} \end{aligned} \tag{5.136}$$

$$\begin{aligned} JYER^e_{jn} &= -\frac{\gamma-1}{2} \left(v^2_K\right)_{\underline{mn}}\, \overline{v^e_J}\, A210_{jn} \\ &\quad + \overline{(E+p)^e} \left(\frac{v_J}{\rho}\right)_{\underline{mn}} (\overline{E_{2J}})^e\, A210_{jn} \end{aligned} \tag{5.137}$$

$$\begin{aligned} JXEL^e_{im} &= -(1-\gamma)\, v_{L_{\underline{mn}}}\, \overline{v^e_J}\, (\overline{E_{1J}})^e\, A210_{im} \\ &\quad + \frac{1}{\rho_{\underline{mn}}}\, \overline{(E+p)^e}\, (\overline{E_{1L}})^e\, A210_{im} \end{aligned} \tag{5.138}$$

$$\begin{aligned} JYEL^e_{jn} &= -(1-\gamma)\, v_{L_{\underline{mn}}}\, \overline{v^e_J}\, (\overline{E_{2J}})^e\, A210_{jn} \\ &\quad + \frac{1}{\rho_{\underline{mn}}}\, \overline{(E+p)^e}\, (\overline{E_{2L}})^e\, A210_{jn} \end{aligned} \tag{5.139}$$

$$\begin{aligned} JXEE^e_{im} &= -\gamma\, \overline{v^e_J}\, (\overline{E_{1J}})^e\, A210_{im} \\ &\quad + (\beta^* \Delta\eta)\, \gamma^2 \overline{\left(\frac{\tilde{v}_1 \tilde{v}_1}{\|\tilde{\bar{v}}\|}\right)^*}\, \overline{D}^e\, A211_{im} \end{aligned} \tag{5.140}$$

$$\begin{aligned} JYEE^e_{jn} &= -\gamma\, \overline{v^e_J}\, (\overline{E_{2J}})^e\, A210_{jn} \\ &\quad + (\beta^* \Delta\eta)\, \gamma^2 \overline{\left(\frac{\tilde{v}_2 \tilde{v}_2}{\|\tilde{\bar{v}}\|}\right)^*}\, \overline{D}^e\, A211_{jn} \end{aligned} \tag{5.141}$$

Remarks

1. Like in the case of compressible potential flow, the MDM algorithm has been used here to not only "extract" FE metric data but "contract" dummy indexes (described in § 5.4.2).

2. The following is the collective expression of TP factorization for the Euler equatinos.

$$\left(X200^{\underline{j}}_{im} + \Theta\,\Delta t \begin{bmatrix} JXRR & JXR1 & JXR2 & JXRE \\ JX1R & JX11 & JX12 & JX1E \\ JX2R & JX21 & JX22 & JX2E \\ JXER & JXE1 & JXE2 & JXEE \end{bmatrix}^{\underline{j}}_{im} \right)$$

$$\left(Y200^{\underline{i}}_{jn} + \Theta\,\Delta t \begin{bmatrix} JYRR & JYR1 & JYR2 & JYRE \\ JY1R & JY11 & JY12 & JY1E \\ JY2R & JY21 & JY22 & JY2E \\ JYER & JYE1 & JYE2 & JYEE \end{bmatrix}^{\underline{i}}_{jn} \right)$$

$$\begin{Bmatrix} \rho \\ m_1 \\ m_2 \\ E \end{Bmatrix}_{mn} = -\{F\}_{ij} \tag{5.142}$$

5.5 Numerical Linear Algebra Solvers

In addition to the Gauss elimination method used in solving the TP factorized "sub-matrices", the following algorithms are implemented and tesed as solvers in this study.

- Line Gauss-Seidel relaxation (LGS).
- Gauss-Seidel (GS) relaxation.
- Conjugate gradient square (CGS).
- Preconditioned CGS.

In this section, these solvers will be defined. For the convenience of discussion, the linear algebra system will be generally referred to as

$$[A]\,\{x\} \;=\; \{b\} \tag{5.143}$$

and its member either

$$A_{\alpha\beta}\,x_{\beta} \;=\; b_{\alpha} \tag{5.144}$$

or

$$\begin{aligned} &A_{ij}^{i-1j+1}\,x_{i-1j+1} \;+\; A_{ij}^{ij+1}\,x_{ij+1} \;+\; A_{ij}^{i+1j+1}\,x_{i+1j+1} \;+ \\ &A_{ij}^{i-1j}\,x_{i-1j} \;+\; A_{ij}^{ij}\,x_{ij} \;+\; A_{ij}^{i+1j}\,x_{i+1j} \;+ \\ &A_{ij}^{i-1j-1}\,x_{i-1j-1} \;+\; A_{ij}^{ij-1}\,x_{ij-1} \;+\; A_{ij}^{i+1j-1}\,x_{i+1j-1} \;=\; b_{ij} \end{aligned} \tag{5.145}$$

which respectively correspond to node α and node (i,j) — the FD style.

5.5.1 Line Gauss-Seidel Relaxation

The *line* Gaussian-Seidel (LGS) iterative algorithm may be run either vertically orhorizontally, and in each case, one has the option of advancing the "relaxation" from either side of the *line*. These options are referred to as the "modes" of LGS in this dissertation, which are the following (a complete list).

- Relaxing from *left* to right, lined *vertically* (LV).
- Relaxing from *right* to left, lined *vertically* (RV).
- Relaxing from *bottom* up, lined *horizontally* (BH).
- Relaxing from *top* down, lined *horizontally* (TH).

These "modes" represent different "domain of dependence", as illustrated in Figure 5.1, which can be formulated mathematically as follows.

The LV Mode of LGS:

$$\begin{aligned} A_{ij}^{i-1j+1}\, x_{i-1j+1}^{n+1} &+ A_{ij}^{ij+1}\, x_{ij+1}^{n+1} + A_{ij}^{i+1j+1}\, x_{i+1j+1}^{n} + \\ A_{ij}^{i-1j}\, x_{i-1j}^{n+1} &+ A_{ij}^{ij}\, x_{ij}^{n+1} + A_{ij}^{i+1j}\, x_{i+1j}^{n} + \\ A_{ij}^{i-1j-1}\, x_{i-1j-1}^{n+1} &+ A_{ij}^{ij-1}\, x_{ij-1}^{n+1} + A_{ij}^{i+1j-1}\, x_{i+1j-1}^{n} = b_{ij} \end{aligned} \tag{5.146}$$

where n denote the iteration step.

The RV Mode of LGS:

$$\begin{aligned} A_{ij}^{i-1j+1}\, x_{i-1j+1}^{n} &+ A_{ij}^{ij+1}\, x_{ij+1}^{n+1} + A_{ij}^{i+1j+1}\, x_{i+1j+1}^{n+1} + \\ A_{ij}^{i-1j}\, x_{i-1j}^{n} &+ A_{ij}^{ij}\, x_{ij}^{n+1} + A_{ij}^{i+1j}\, x_{i+1j}^{n+1} + \\ A_{ij}^{i-1j-1}\, x_{i-1j-1}^{n} &+ A_{ij}^{ij-1}\, x_{ij-1}^{n+1} + A_{ij}^{i+1j-1}\, x_{i+1j-1}^{n+1} = b_{ij} \end{aligned} \tag{5.147}$$

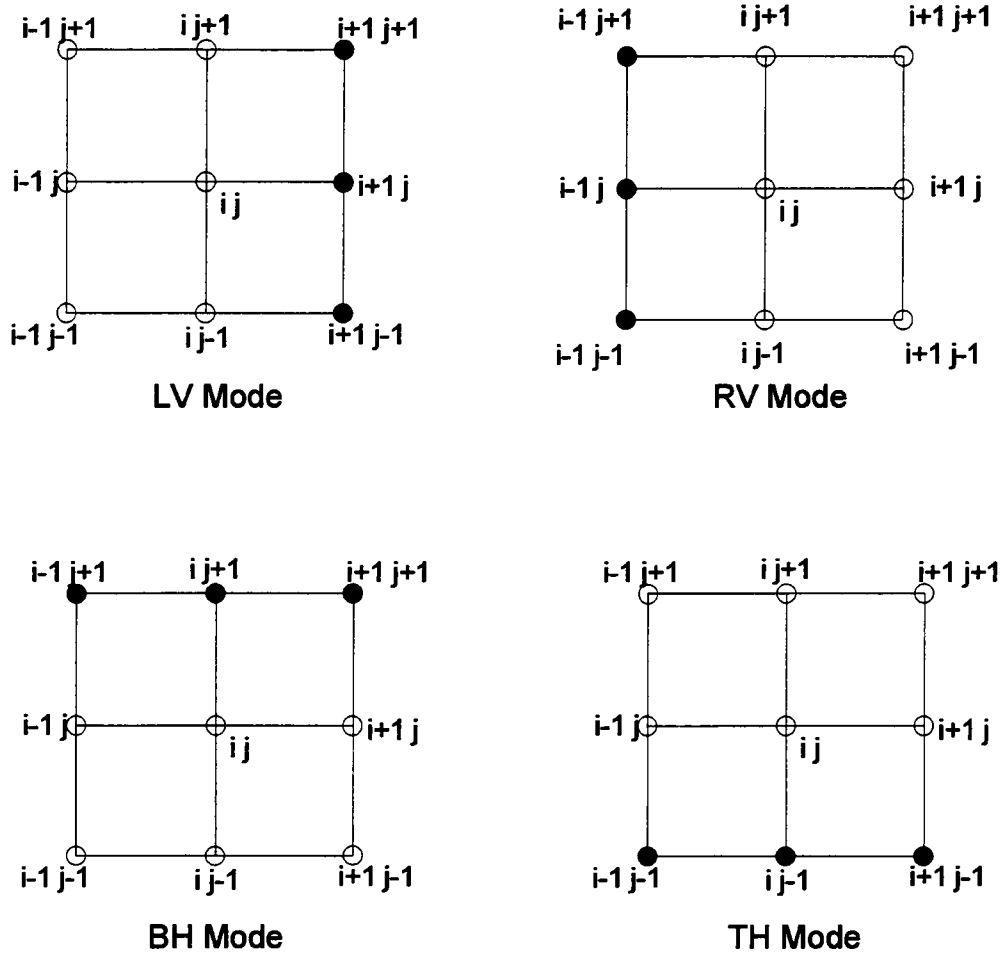


Figure 5.1: Illustration for the "domain of dependence" of the line Gauss-Seidel algorithm, with • and o being respectively the locations of previous and current. LV mode: relaxation from *left* to right, lined *vertically*; RV mode: relaxation from *right* to left, lined *vertically*; BH mode: relaxation from *bottom* up, lined *horizontally*; TH mode: relaxation from *top* down, lined *horizontally*.

The BH Mode of LGS:

$$\begin{aligned} & A_{ij}^{i-1j+1}\, x_{i-1j+1}^{n} + A_{ij}^{ij+1}\, x_{ij+1}^{n} + A_{ij}^{i+1j+1}\, x_{i+1j+1}^{n} + \\ & A_{ij}^{i-1j}\, x_{i-1j}^{n+1} + A_{ij}^{ij}\, x_{ij}^{n+1} + A_{ij}^{i+1j}\, x_{i+1j}^{n+1} + \\ & A_{ij}^{i-1j-1}\, x_{i-1j-1}^{n+1} + A_{ij}^{ij-1}\, x_{ij-1}^{n+1} + A_{ij}^{i+1j-1}\, x_{i+1j-1}^{n+1} = b_{ij} \end{aligned} \tag{5.148}$$

The TH Mode of LGS:

$$\begin{aligned} & A_{ij}^{i-1j+1}\, x_{i-1j+1}^{n+1} + A_{ij}^{ij+1}\, x_{ij+1}^{n+1} + A_{ij}^{i+1j+1}\, x_{i+1j+1}^{n+1} + \\ & A_{ij}^{i-1j}\, x_{i-1j}^{n+1} + A_{ij}^{ij}\, x_{ij}^{n+1} + A_{ij}^{i+1j}\, x_{i+1j}^{n+1} + \\ & A_{ij}^{i-1j-1}\, x_{i-1j-1}^{n} + A_{ij}^{ij-1}\, x_{ij-1}^{n} + A_{ij}^{i+1j-1}\, x_{i+1j-1}^{n} = b_{ij} \end{aligned} \tag{5.149}$$

5.5.2 Gauss-Seidel Iteration

There are different modes of the Gauss-Seidel (GS) solver, which differ from each other in the *route* along which the solution is relaxed (pointwisely). The implemented GS modes in this study are these:

- started from the *left/bottom* corner, marching *horizontally* line by line up.
- started from the *left/bottom* corner, marching *vertically* line by line to the right.
- started from the *right/top* corner, marching *horizontally* line by line down.
- started from the *right/top* corner, marching *vertically* line by line to the left.

which I call respectively the LBH, LBV, RTH and RTV of GS, and of which the "domain of dependence" are shown in Figure 5.2. The following are their formulas.

The LBH Mode of GS:

$$\begin{aligned} &A_{ij}^{i-1j+1}\, x_{i-1j+1}^{n} + A_{ij}^{ij+1}\, x_{ij+1}^{n} + A_{ij}^{i+1j+1}\, x_{i+1j+1}^{n} + \\ &A_{ij}^{i-1j}\, x_{i-1j}^{n+1} + A_{ij}^{ij}\, x_{ij}^{n+1} + A_{ij}^{i+1j}\, x_{i+1j}^{n} + \\ &A_{ij}^{i-1j-1}\, x_{i-1j-1}^{n+1} + A_{ij}^{ij-1}\, x_{ij-1}^{n+1} + A_{ij}^{i+1j-1}\, x_{i+1j-1}^{n+1} = b_{ij} \end{aligned} \tag{5.150}$$

The LBV Mode of GS:

$$\begin{aligned} &A_{ij}^{i-1j+1}\, x_{i-1j+1}^{n+1} + A_{ij}^{ij+1}\, x_{ij+1}^{n} + A_{ij}^{i+1j+1}\, x_{i+1j+1}^{n} + \\ &A_{ij}^{i-1j}\, x_{i-1j}^{n+1} + A_{ij}^{ij}\, x_{ij}^{n+1} + A_{ij}^{i+1j}\, x_{i+1j}^{n} + \\ &A_{ij}^{i-1j-1}\, x_{i-1j-1}^{n+1} + A_{ij}^{ij-1}\, x_{ij-1}^{n+1} + A_{ij}^{i+1j-1}\, x_{i+1j-1}^{n} = b_{ij} \end{aligned} \tag{5.151}$$

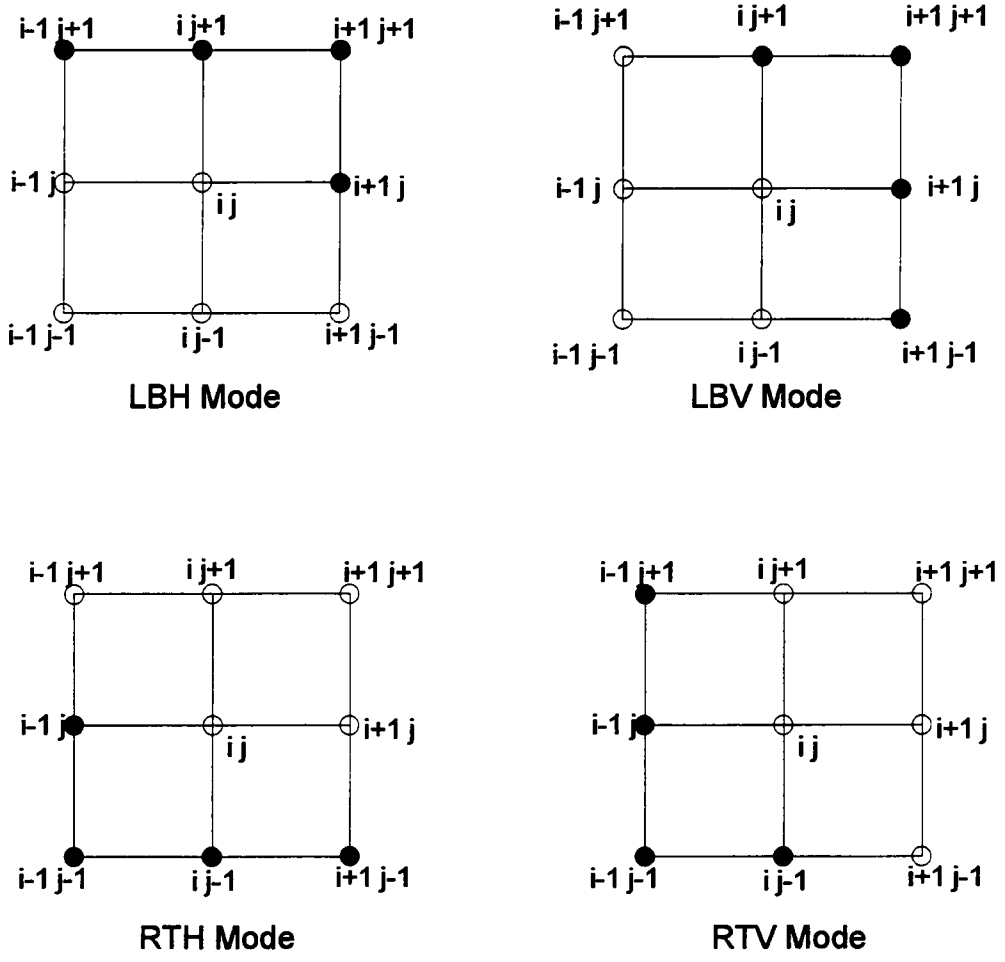


Figure 5.2: Illustration for the "domain of dependence" of the Gauss-Seidel algorithm, with • and ○ being respectively the locations of previous and current. LBH: solution marching *horizontally* line by line up from the *left/bottom* corner; LBV: solution marching *vertically* line by line to the right from the *left/bottom* corner; RTH: solution marching *horizontally* line by line down from the *right/top* corner; RTV: solution marching *vertically* line by line to the left from the *right/top* corner.

The RTH Mode of GS:

$$\begin{aligned} & A_{ij}^{i-1j+1}\, x_{i-1j+1}^{n+1} + A_{ij}^{ij+1}\, x_{ij+1}^{n+1} + A_{ij}^{i+1j+1}\, x_{i+1j+1}^{n+1} + \\ & A_{ij}^{i-1j}\, x_{i-1j}^{n} + A_{ij}^{ij}\, x_{ij}^{n+1} + A_{ij}^{i+1j}\, x_{i+1j}^{n+1} + \\ & A_{ij}^{i-1j-1}\, x_{i-1j-1}^{n} + A_{ij}^{ij-1}\, x_{ij-1}^{n} + A_{ij}^{i+1j-1}\, x_{i+1j-1}^{n} = b_{ij} \end{aligned} \tag{5.152}$$

The RTV Mode of GS:

$$\begin{aligned} & A_{ij}^{i-1j+1}\, x_{i-1j+1}^{n} + A_{ij}^{ij+1}\, x_{ij+1}^{n+1} + A_{ij}^{i+1j+1}\, x_{i+1j+1}^{n+1} + \\ & A_{ij}^{i-1j}\, x_{i-1j}^{n} + A_{ij}^{ij}\, x_{ij}^{n+1} + A_{ij}^{i+1j}\, x_{i+1j}^{n+1} + \\ & A_{ij}^{i-1j-1}\, x_{i-1j-1}^{n} + A_{ij}^{ij-1}\, x_{ij-1}^{n} + A_{ij}^{i+1j-1}\, x_{i+1j-1}^{n} = b_{ij} \end{aligned} \tag{5.153}$$

Remarks

The above algorithms differ from the normally described GS scheme, i.e.,

$$\begin{cases} A_{11}\, x_1^{n+1} + A_{12}\, x_2^{n} + \cdots + A_{1N}\, x_N^{n} = b_1 \\ A_{21}\, x_1^{n+1} + A_{22}\, x_2^{n+1} + \cdots + A_{2N}\, x_N^{n} = b_2 \\ \cdots \\ A_{N1}\, x_1^{n+1} + A_{N2}\, x_2^{n+1} + \cdots + A_{NN}\, x_N^{n+1} = b_N \end{cases} \tag{5.154}$$

only in that, in (5.154), the solution relaxation is *routed* along the way nodes are numbered, whereas in (5.150-5.150-5.152-5.153), it has been *ruled* independently.

5.5.3 Conjugate Gradient Square Algorithm

The conjugate gradient square (CGS) algorithm, as it is named, means a *squared* CG, which is a generalization of the latter to non-SPD system ([65]). Translated numerically, CGS is composed the following procedures.

- Initialization:

  Choose $\{x^0\}$;

  $\{r^0\} := \{b\} - [A]\{x^0\}$;

  $\{q^0\} := \{p\} := \{0\}$;

  $\rho^{-1} := 1$;

  Suitably choose $\{\bar{r}^0\}$;

  $n := 0$.

- CGS Loop Body:

  do while residual > tolerance

  begin

  $\rho^n := \{\bar{r}^0\}^T \{r^n\}$;

  $\beta^n := \frac{\rho^n}{\rho^{n-1}}$;

  $\{u^n\} := \{r^n\} + \beta^n \{q^{n-1}\}$;

  $\{p^n\} := \{u^n\} + \beta^n (\{q^n\} + \beta^n \{p^{n-1}\})$;

  $\{v^n\} := [A]\{p^n\}$;

  $\sigma^n := \{\bar{r}^0\}^T \{v^n\}$;

$\alpha^n := \frac{\rho^n}{\sigma^n}$;

$\{q^{n+1}\} := \{u^n\} - \alpha^n \{v^n\}$;

$\{r^{n+1}\} := \{r^n\} - \alpha^n [A](\{u^n\} + \{q^{n+1}\})$;

$\{x^{n+1}\} := \{x^n\} - \alpha^n (\{u^n\} + \{q^{n+1}\})$;

$n := n + 1$.

end

Herein, the intial "conjugate" residual $\{\bar{r}^0\}$, as noted, should be "suitably" chosen. Theoretically, in order for CGS to be finite (assuming absent roundoff [65]), $\{\bar{r}^0\}$ must not be orthogonal to any invariant of the Krylov subspace associated with $[A]$ and $\{r^0\}$. However, this is not severe constraint numerically, and a randomly given $\{\bar{r}^0\}$ is usually "suitable" enough ([65, 66]).

5.5.4 Preconditioned CGS

Suppose that

$$\left[\tilde{A}\right]\{x\} = \left\{\tilde{b}\right\} \tag{5.155}$$

is the preconditioned system of (5.143), where $[B]^{-1}$ is the preconditioner, $[\tilde{A}] = [B]^{-1}[A]$, and $\{\tilde{b}\} = [B]^{-1}\{b\}$.

A preconditioned CGS is nothing but the CGS for a preconditoined system, in other words, the following procedure:

- Initialization:

$\{\tilde{r}^0\} := \{b\} - \left[\tilde{A}\right]\{x^0\};$

$\{q^0\} := \{p\} := \{0\};$

$\rho^{-1} := 1;$

Suitably choose $\left\{\tilde{\bar{r}}^0\right\};\ \ n := 0.$

- Preconditioned CGS Loop Body:

  do while residual > tolerance

  begin

  $\rho^n := \left\{\tilde{\bar{r}}^0\right\}^T \{\tilde{r}^n\};$

  $\beta^n := \frac{\rho^n}{\rho^{n-1}};$

  $\{u^n\} := \{\tilde{r}^n\} + \beta^n \{q^{n-1}\};$

  $\{p^n\} := \{u^n\} + \beta^n (\{q^n\} + \beta^n \{p^{n-1}\});$

  $\{v^n\} := \left[\tilde{A}\right]\{p^n\};$

  $\sigma^n := \left\{\tilde{\bar{r}}^0\right\}^T \{v^n\};$

  $\alpha^n := \frac{\rho^n}{\sigma^n};$

  $\{q^{n+1}\} := \{u^n\} - \alpha^n \{v^n\};$

  $\{\tilde{r}^{n+1}\} := \{\tilde{r}^n\} - \alpha^n \left[\tilde{A}\right](\{u^n\} + \{q^{n+1}\});$

  $\{x^{n+1}\} := \{x^n\} - \alpha^n (\{u^n\} + \{q^{n+1}\});$

  $n := n + 1.$

  end

□

Note that it is usually an easier task to "suitably choose" $\{\tilde{\bar{r}}^0\}$ here than in the plain CGS. And the reason is that $[\tilde{A}]$ is closer to SPD than $[A]$ (since preconditioner $[B]^{-1}$ by definition is an approximation of $[A]^{-1}$, no matter how bad it is). In fact, the recommended $\{\tilde{\bar{r}}^0\} = \{\tilde{r}^0\}$ ([65, 66]) proves being good enough for all the preconditioners in the computation of this study (see Chapter 6).

In addition, it should also be noted that there are only two steps in the "loop body" of the preconditioned CGS involving the preconditioner: the calculation of $\{v^n\}$ and $\{\tilde{r}^{n+1}\}$.

Chapter 6

Numerical Results

6.1 Introduction

The numerical analysis of this dissertation involves several aspects, including:

I. The FE solution of convection system (2.58-2.59), which was performed to confirm the nonlinearity/shock mechanism of the system that has been theoretically established in Chapter 2.

II. The FE solution of compressible potential flow (CPF) and Euler equations, which was conducted serving to examine the metric data mean (MDM), *selective* Taylor weak statement (S-TWS) and invariant dissipation length (IDL) algorithms.

III. The behavior of linear algebra solvers, which is part of the proposed commitment ([1]) that initiated this dissertation.

The details of these algorithms is "checklisted" in Chapter 5. Documented in this chapter are the numerical results benchmarking them.

6.2 The Nonlinear Convection Problem

Problem Statement

• Nonlinear Configuration. The following is the concrete setting of (2.58-2.59) used in numerical analysis.

$$\frac{\partial q}{\partial t} + v \frac{\partial q}{\partial x_j} = 0 \tag{6.1}$$

$$v = \frac{v_0}{1 + a f(q)} \tag{6.2}$$

$$f(q) = \frac{1}{(1 + bq)^2} \tag{6.3}$$

Here, (6.3), in which $b > 0$, actually defines the nonlinearity of the system. Physically, such a $f(q)$ analogizes the Langmuir isotherm of nonlinear chromatography ([107]), although, from the view point of this study, what is important is not what it physically analogizes but that it meets the conditions described in Chapter 2 (i.e., (2.60) and being "monotonely decreasing").

• Initial Condition. A ground-flat profile of q is taken as the initial condition:

$$q(0, x) = 0 \tag{6.4}$$

• Boundary Conditions. The solution of (6.1-6.2-6.3) is considered to be defined on $[0, L]$, with $x = 0$ and $x = L$ being the inlet and outlet. While the outlet is set

free (as being characteristically enforced), Dirichlet boundary condition (BC) applies at the inlet. Concretely, two cases are computed:

$$q(t,0) = \begin{cases} 0, & as\ t = 0^- \\ q_0, & as\ 0 \leq t \leq t_p \\ 0, & as\ t > t_p \end{cases} \tag{6.5}$$

$$q(t,0) = \begin{cases} 0, & as\ t = 0^- \\ q_0, & as\ 0 \leq t \end{cases} \tag{6.6}$$

which are respectively a *square-cap* (of width t_p), and a *step*, both being of height q_0.

Numerical Results

- Parameters. The above system is solved using the parameters in table 6.1.

Parameters	a	b	v_0	q_0	t_p	L	$Cour$	β^*
Values	4.50	0.18	0.60	24.00	5.00	30.00	1.00	0.499

Table 6.1: The parameters of the nonlinear convection problem.

The pick of these numbers are mostly *casual*, and they can be summarized as follows if consideration has to be offered:

1. a and b: b is the parameter controlling the nonlinearity in $f(q)$, which was chosen to be *weak*; a is the parameter controlling the impact of the entire $f(q)$, which was chosen to be *strong*. The net nonlinearity thus attained is such that

$$v_{min} = \frac{v_0}{1 + a f(0)} \simeq 0.182\, v_0 \tag{6.7}$$

$$v_{max} = \frac{v_0}{1 + a\, f(q_0)} \simeq 0.863\, v_0 \tag{6.8}$$

2. q_0 and t_p: As said, they are respectively the amplitude and width of the inlet plateau. A *high* q_0 is given so that the shock is *strong*. A *narrow* t_p is input so that *soon* the characteristics emitted from the rear side of the inlet cap can catch up with and start to interact with the shock on the front.

3. *Cour* and β^*: *Cour* is the linear velocity based Courant number, i.e., $\frac{v_0 \Delta t}{\Delta x}$, which is set *large* (within its permissible range), so that the time step could be large; β^* is the "size" of dissipation, which is chosen to be *close* to $\frac{1}{2}Cour$, so that the effective dissipation can be *small* (valid only for the Euler explicit scheme).

Note that, although $Cour = 1$ has been taken, it does not mean that the system is in "critical" state ([2]). In fact, the real Courant number $\frac{v \Delta t}{\Delta x}$, as according to (6.7-6.8), is between 0.182 and 0.863.

• Case I. Solution in this case corresponds to BC (6.5), and it has been run for a series of meshes in order to assess the convergence (or divergence). The results are shown in Table 6.2, Figure 6.1-6.2, and Figure 6.3.

Table 6.2 is the convergence history of the solution in terms of net conservation:

$$\epsilon = \frac{\int_0^\infty q\,(t, L)\; dt}{\int_0^\infty q\,(t, 0)\; dt} \tag{6.9}$$

where $\int_0^\infty q\,(t, 0)\; dt$ is the total "mass" of q analytically given at the inlet, and $\int_0^\infty q\,(t, L)\; dt$ is what has been numerically received at the outlet. The reason for

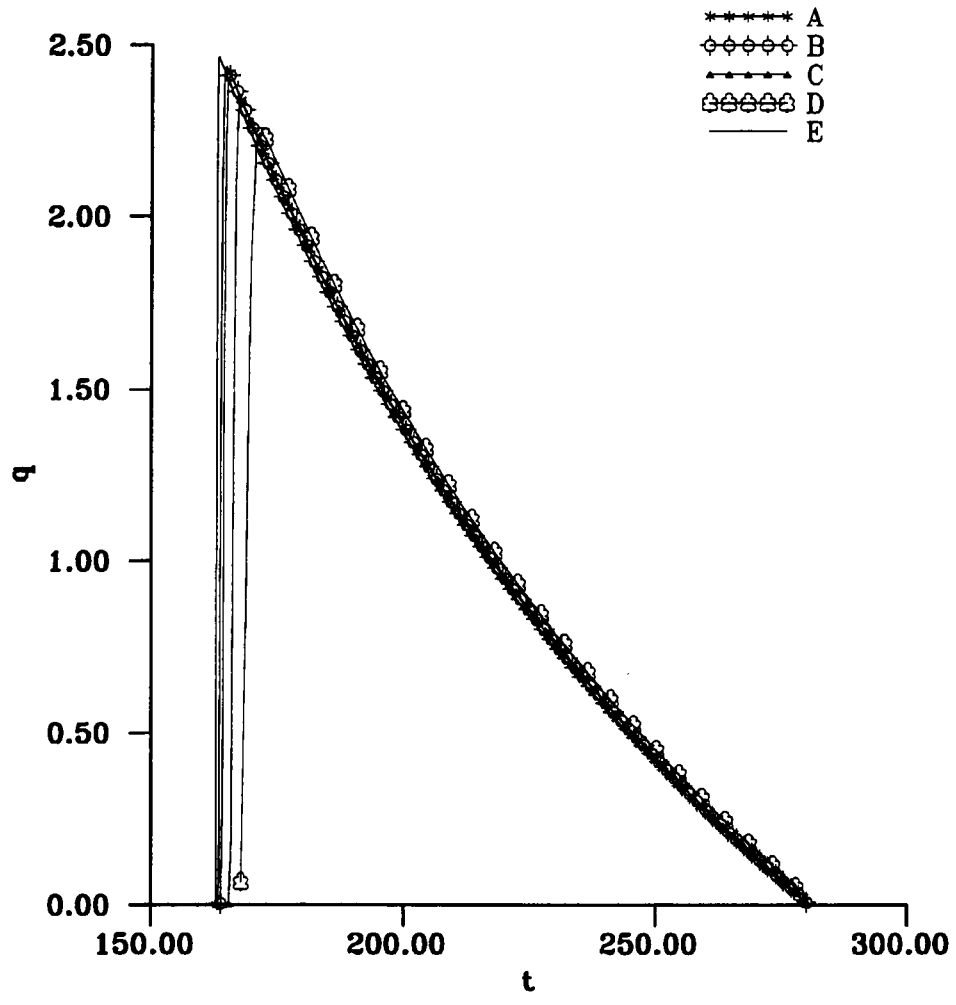


Figure 6.1: Outlet profiles **I** of the nonlinear convection problem with a square-cap inlet BC. Mesh numbers: A — 1000; B — 900; C — 700; D — 300; E — 5000.

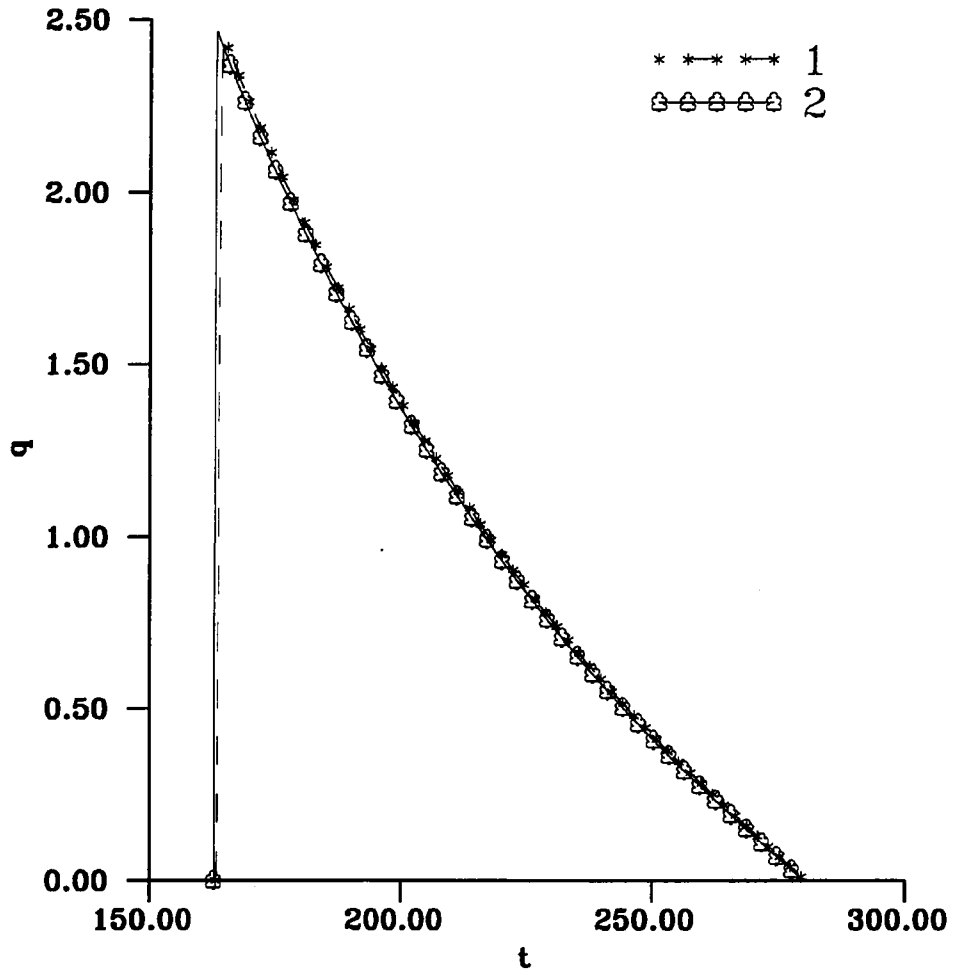


Figure 6.2: Outlet profiles **II** of the nonlinear convection problem with a square-cap inlet BC. Mesh number: A — 1000; B — 5000.

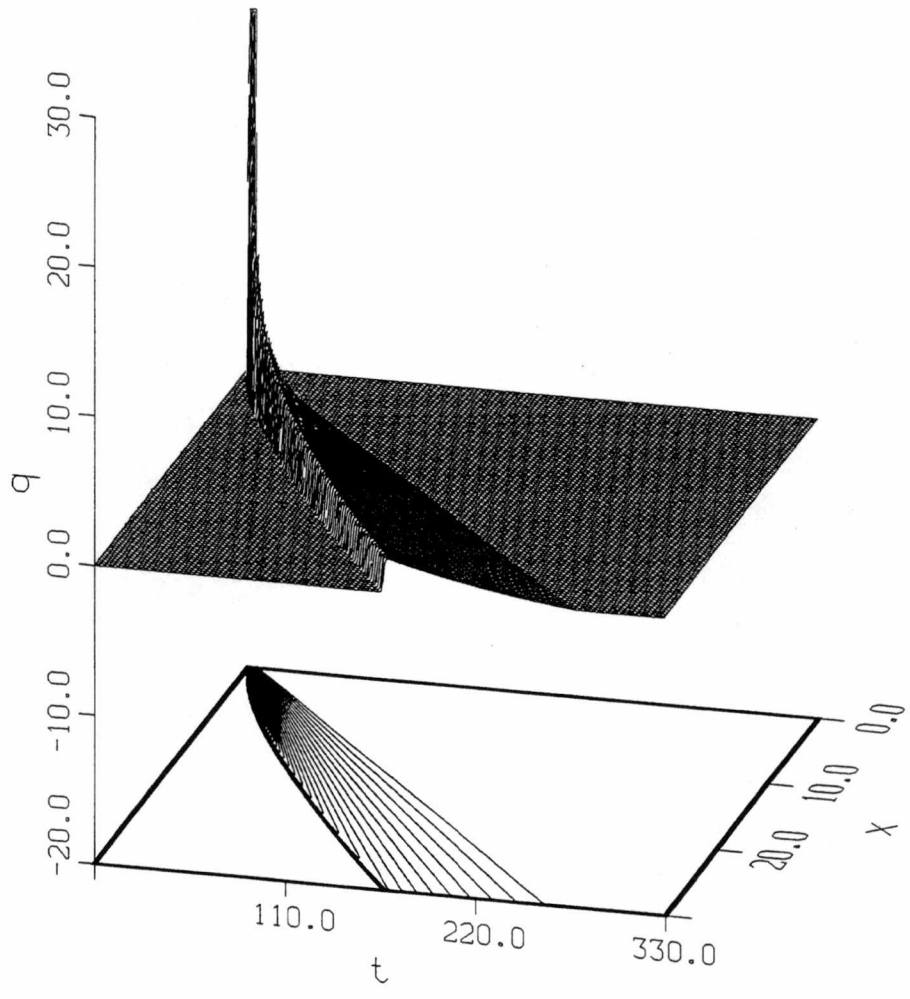


Figure 6.3: 3-D display for the FE solution of the nonlinear convection problem with a square-cap inlet BC ($N = 300$).

such being taken as the measure of convergence is that (6.1-6.2) is a system in "non-conservative" form, which is unable to strictly maintain conservation. So indeed, conservation has not been strictly observed in numerical solution, and it is part of the converging process.

Mesh Number N	300	700	900	950	1000	5000
Net Convergence ϵ	0.917	.943	.978	.984	0.990	0.990

Table 6.2: The convergence of convection in terms of conservation.

Figure 6.1 and Figure 6.2 are the profiles detected at the outlet of these solutions (except for $N = 950$). The $N = 5000$ one is the sufficiently converged solution (which occues actually after $N > 2600$). The pattern of these profiles is that they all arrive in the shape of single-sided "roof"— with a shock front. The shock fronts originate from the front "down-side" of the inlet square-cap (since they arrive first), and the slopes of the "roof" are *stretched* rear "up-side", both being as asserted in Chapter 2. A better view of this is Figure 6.3, where, in addition to that the front "down-side" of the inlet square-cap remains being a shock and the rear "up-side" keeps being *stretched*, it is also obvious that

A. the front shock, except at the very early stage of the convection (before the plateau is "submerged"), is constantly "overtaken on the top " by the characteristics from behind, and

B. the shock is slowed down everytime it is caught up with by a characteristics (which is emittedd from "up-side", of the inlet cap, hence, dilute and reduce the amplitude of the shock at every encountering).

A verifies the asserted "rule of thumb" for the interaction between shock and characteristics (see Chapter 2). **B** coincides with the fact that the shock speed, as according to (2.63) or (2.69), should indeed be decreasing as Δq decreases for ground-fixed shock (since ratio $\frac{\Delta C}{\Delta q}$ increases as Δq decreases, as a result of that $C(q)$ is increasing and convex, see Figure 2.4).

• Case II. To further confirm the shock speed formula (2.69), the solution was run under BC (6.6). The solution is shown in Figure 6.4. The numerical shock speed is found to be 0.3243, which is the slope of the shock on the contour plane in Figure 6.4, in comparison with the analytical

$$\begin{aligned} v_s &= \frac{v_0}{1 + \frac{\Delta C}{\Delta q}} \\ &= \frac{v_0}{1 + \frac{1}{q_0} a \int_0^{q_0} f(q)\, dq} \\ &= \frac{v_0}{1 + \frac{a}{1 + b\, q_0}} \\ &= \frac{0.6}{1 + \frac{4.5}{1 + 0.18 \times 24.0}} \\ &\simeq 0.3251 \end{aligned}$$

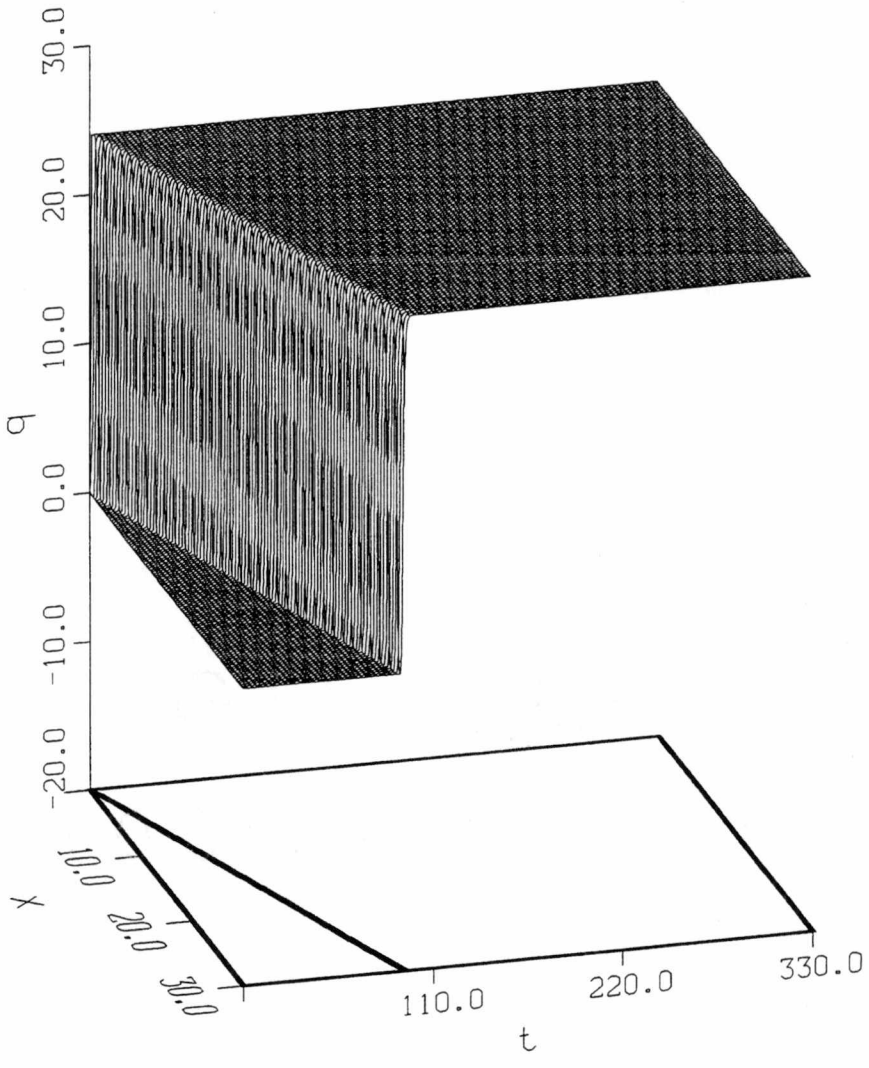


Figure 6.4: 3-D display for the FE solution of the nonlinear convection problem with step function inlet BC ($N = 300$).

6.3 Aerodynamic Flows

6.3.1 General Statements

Configuration

The benchmark problem for the aerodyamic flows in this dissertation is an arc bump, which is of a 10% height-secant ratio shown in Figure 6.5 (a). The bump is "attacked" in zero angle by air flow from left to right, with a homogeneous far upstream. Air is assumed to be at "ordinary" temperature, taking $\gamma = 1.4$.

The appealingness of this problem is that, in addition to the normal aerodynamic features, it has two *singular*-like points, the front and rear edges of the arc-bump, where the streamline has indefinite direction, and the velocity hence must be zero (since only zero vector has indefinite direction). It is not the intention of this research to really "mess around" with these "singularities". Their being in presence, however, provides extra challenge to the stability of the developed algorithms.

General Boundary Conditions (BC)

By "general BC", I mean the ones described in *primitive* variables, i.e., the ones used by the Euler equations. Illustration for the "general BC" is in Figure 6.5 (b). In summary, there are three types of boundaries for the above system, and they are conditioned as follows as far as the numerical examples of this dissertation are concerned.

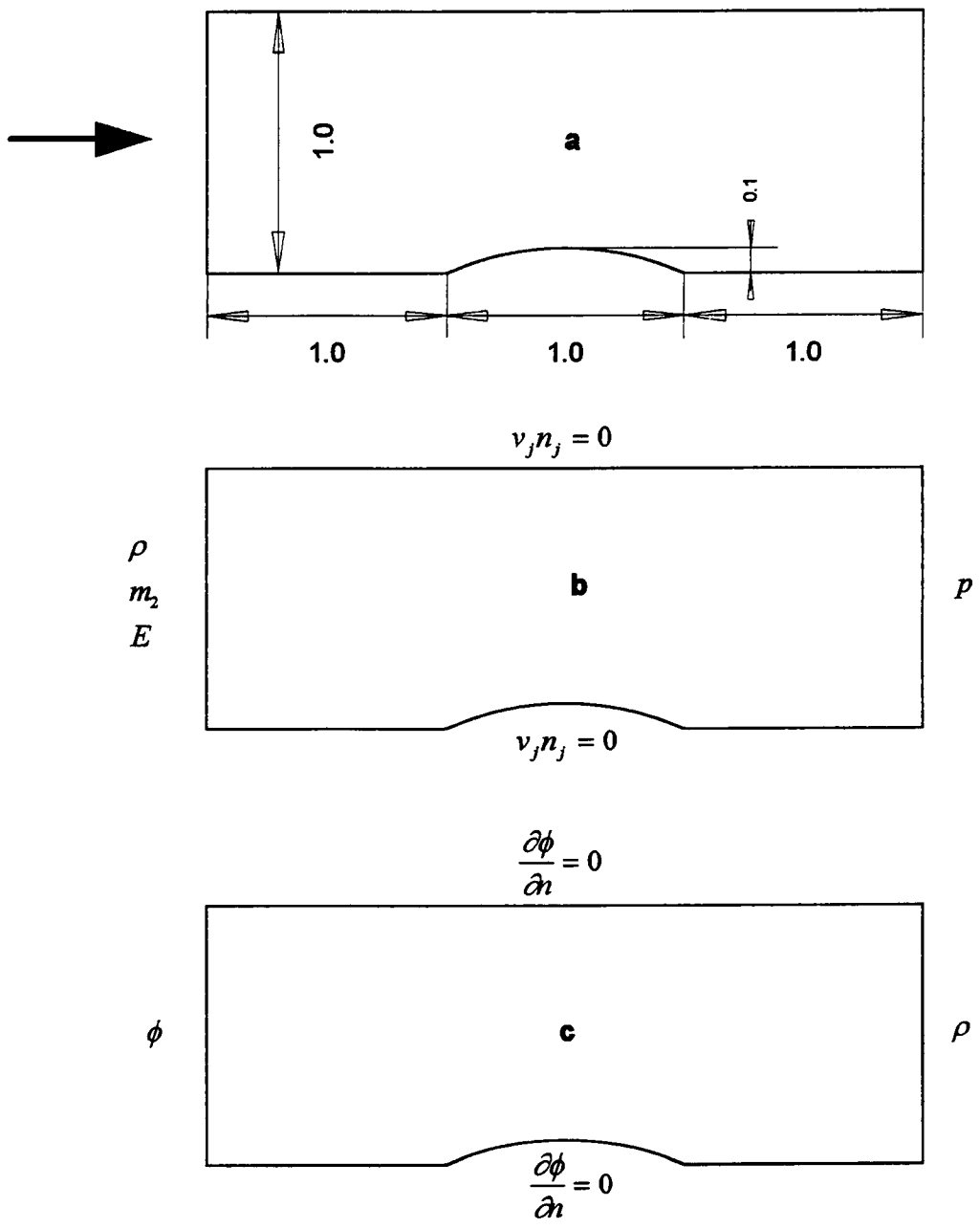


Figure 6.5: Configuration of the arc-bump flow problem: (**a**) geometry; (**b**) general boundary conditions; (**c**) potential boundary conditions.

- The *walls*, which include the *straight* top and the "bumped" bottom edges of the solution domain, and where a $v_j n_j = 0$ is enforced.

- The *inlet* and *outlet*, which both are *incompletely* Dirichlet conditioned, being given $\{\rho_{in}, v_{2_{in}}, E_{in}\}$ and p_{out} respectively.

The wall condition $v_j n_j = 0$ is self-evident and generally valid. It is nothing but an expression of the fact that the boundary where it is imposed on is (i) a *wall*, which by definition is not penetratable, and (ii) *inviscid*, which permits tangential flow. The inlet and outlet conditions, on the other hand, is valid assuming that flow enters and exits subsonically (in terms of the normal velocity component), which is when the flow endures $dim + 1$ characteristic "regularities" from the upstream (i.e., *three*, in the case of 2-D), whereas, *one* from the downstream ([20]). If flow enters/exits supersonically, the inlet should then be modified into being fully conditioned, and the outlet left totally free.

It should be noted that, although the number of BC on the inlet and outlet are characteristically regulated, the concrete BC are not. For example, one may well choose to give $\{\rho_{in}, v_{1_{in}}, E_{in}\}$ instead of $\{\rho_{in}, v_{2_{in}}, E_{in}\}$ on the inlet. In addition, the $v_{2_{in}}$ is always set *zero*, since what is considered is only the cases when the bump gets "attacked in zero angle".

Potential Boundary Conditions

There has been no serious effort in this study with regards to the boundary condition of compressible potential flow (CPF). What has been done is to *translate* the

"general BC" — in the way that is (**a**) coping with the irrotational and isentropic nature of CPF, and (**b**) not contradicting the characteristic nature of flow. The "translated" BC are shown in Figure 6.5 (c), and the "translations" are these:

- The wall, where the "general BC" is unambiguous, and it can only be "potentially" described as $\frac{\partial \phi}{\partial n} = 0$.

- The outlet, where the "general BC" is also unambiguous, since a known p_{out}, given the isentropicity of CPF, can only be "translated" into a known ρ_{out}.

- Finally the inlet, which is where CPF is confused and where one has to *pick* among the "general BC". The following are the understanding.

  i. The E-condition should be beyond choice, since the total energy equation has been disgarded in CPF (upon Cauchy-Lagrange integration, see chapter 2).

  ii. The remaining option hence is the Dirichlet value of either ρ_{in} or $v_{2_{in}}$ $(= 0)$ (note that it is *not* both, since the subsonic inlet should not be fully regulated).

  iii. What has been picked in this study is the latter of **ii**, as shown in Figure 6.5 (c), which, in CPF terms, means a Dirichlet and constant value of ϕ across the inlet, for these considerations: First, ϕ needs to be Dirichlet conditioned somewhere anyway (since otherwise a constant shift would be permitted in it by definition, see Chapter 2). Second, a *given* ρ_{in} in general

would be inconsistent with the ρ_{out} at the downstream, since the latter, together with T_∞ and ρ_∞ (which define the "reference state in the CPF energy equation), have already *constrained* the distribution of ρ_{in} (at least its average value).

Initialization

Both the Euler and CPF algorithms are initialized by solving the incompressible potential flow, i.e., the Laplacian model

$$\frac{\partial^2 \tilde{\phi}}{\partial x_j^2} = 0 \tag{6.10}$$

where $\tilde{\phi}$ is the incompressible potential. The boundary conditions used on $\tilde{\phi}$ are identical to those on ϕ in CPF, except that it has been complemented at the outlet that $\left(\frac{\partial \tilde{\phi}}{\partial n}\right)_{out} = v_{1_{out}}$, in which $v_{1_{out}}$ is isentropically derived from either p_{out} or ρ_{out} (in the case of Euler and CPF respectively, both assuming $v_{2_{out}} = 0$), in conjunction with a known or estimated state of stagnation. Under such BC, the Galerkin FE algorithm for (6.10) is as follows.

$$\tilde{\phi}_\beta \int_\Omega \frac{\partial N_\alpha}{\partial x_j} \frac{\partial N_\alpha}{\partial x_j} d\Omega = \int_{\partial\Omega_{out}} N_\alpha \, v_{1_{out}} \, \sigma \tag{6.11}$$

i.e.,

$$M2JJ_{\alpha\beta} \, \tilde{\phi}_\beta = \int_{\partial\Omega_{out}} N_\alpha \, v_{1_{out}} \, \sigma \tag{6.12}$$

where α covers all but the Dirichlet inlet nodes. Equation (6.12) is SPD, and it may be solved by any of the implemented linear algebra algorithms.

The solution of (6.10) is channelled to the S-TWS FE algorithms either *directly* being part of the initial condition — as for ϕ in CPF, or *indirectly* via

$$v_j^0 = \frac{\partial \tilde{\phi}}{\partial x_j} \tag{6.13}$$

— as for all other variables. In the latter case, the resulting v_j^0 is augmented isentropically to "all other variables" (the same way as the outlet BC for $\tilde{\phi}$ is derived).

Equation (6.13) is also integrated through the Galerkin FE:

$$\int_\Omega N_\alpha \, v_j^0 \, d\Omega = \int_\Omega N_\alpha \frac{\partial \tilde{\phi}}{\partial x_j} \, d\Omega \tag{6.14}$$

or

$$M200_{\alpha\beta} \, v_{j_\beta}^0 = M20J_{\alpha\beta} \, \tilde{\phi}_\beta \tag{6.15}$$

System (6.15) is also SPD, and it can also be solved by any of the implemented solvers. Particularly, the TP solver has always been used for it, and non-iteratively, since its coefficient $[M200]$ is almost analytically TP factorizable (as shown in Chapter 5).

□

Compared with the commonly used quasi-1D initialization, the major advantage of the above procedure is that it is feasible to general geometry. And it is not very costly, needing only to solve a Laplacian which is no big deal nowadays.

6.3.2 Numerical Results

The numerical examples presented in this section are the results running the $\Theta = 1$ cases of the S-TWS FE algorithms documented in Chapter 5. The solution accuracy ϵ_n are all set around/below 10^{-6}, measured in residual $\{R\}^n_\alpha$. The Newton relaxation (NR) is limited to $p \leq 11$, which is a cap that is enough for NR to be converged to the level of $10\epsilon_n$ on all but the first few time stations (where the solution "correction" is large). $\beta^* = \frac{1}{4}$ has been taken for the numerical dissipation, which is a casual input — just so that the dissipation at convergence corresponds to a fourth of a fully upwinding scheme. The linear algebra iterations are always driven at least two decimal points below ϵ_n, to the effect that all Newton intermediate solutions are virtually accurate whichever solver used.

Subsonic Solution for Euler Equations

The flow in this case is for $M_{in} = 0.5$, using a $(21 + 31 + 21) \times 21$ three block grid system as shown in Figure 6.6. The grid is transfinitely generated in such a way:

- In the two side-blocks, mesh is stretched progressively at ratio 0.901 in both directions — towards the ends of the "bump".
- In the middle-block, while the mesh is stretched at the same ratio as in its neighbor blocks along the vertical direction, it is spread horizontally to the sides so that the mesh-cell at the "bump" ends could be the *squarest* possible.

The solution is shown in Figure 6.7, which is the result at the 97th time step,

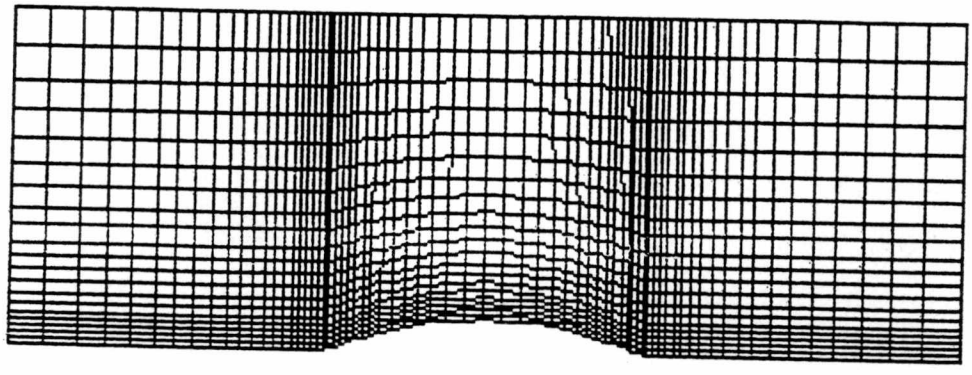


Figure 6.6: Mesh for Subsonic flow.

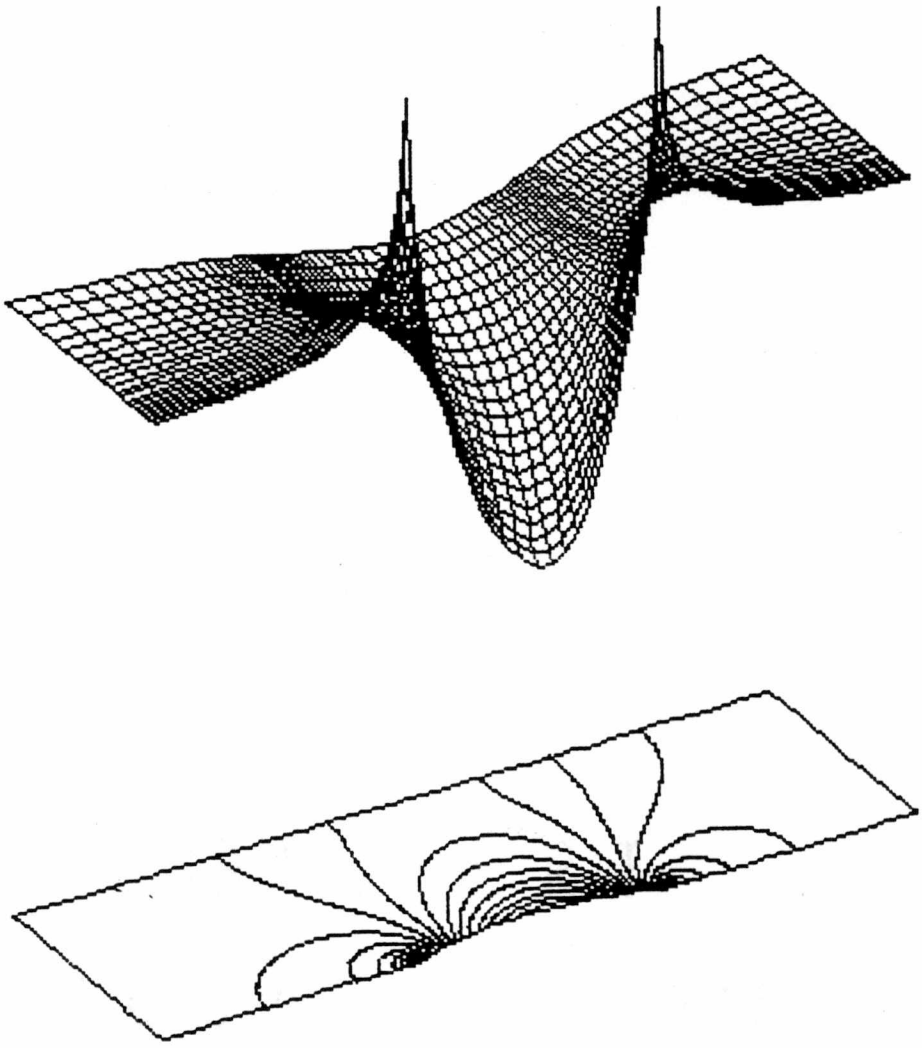


Figure 6.7: The subsonic solution of the S-TWS FE algorithm for Euler equations. The surface and contour are respectively for pressure and Mach number.

bearing a 0.9997×10^{-6} residual. Clearly, the "singularities" at the bump-ends have been caught, at least graphically, and without any oscillation. The maximum Courant number at convergence is 40.271 (measured based on the mesh size along the streamline, i.e., a β^*th of the invariant dissipation length).

The Boundary Closure of Dissipation

• About the v-form Assessment. Having got the above, the problem is re-executed to test about the "boundary closure" issue of TWS-FE algorithm ([1]) — whether it is necessary to include the the boundary integral terms of the numerical dissipation, i.e., $\Delta t_* \oint_{\partial\Omega} N_\alpha \mu_{jk} \frac{\partial U}{\partial x_k} n_j d\sigma$, where μ_{jk} stands for the coefficient of the numerical dissipation (the part other than Δt_*). The re-execution is under exactly the same conditions except that the v-form A-TWS dissipation has been used (meaning to take $\mu_{jk} = v^2 \delta_{jk}$ instead of the $w_j w_k$ as in S-TWS, see Chapter 4). The sole reason for such a v-form assessment is that the boundary on which $\frac{\partial U}{\partial x_k}$ is significant in the considered problem is a streamline (the bump), where the boundary flux of the S-TWS $w_j w_k$-form dissipation vanishes by default (see § 5.2 and § 5.3). — And it would not be necessary if the system possesses a piece of boundary which is not a streamline and where $\frac{\partial U}{\partial x_k}$ is significant.

• Verification. The numerical results are shown in Figure 6.8 and Figure 6.9, which are respectively the Mach numbers of a boundary *unclosed* and *closed* solution. It is obvious that the boundary *unclosed* solution suffers severe damping across the boundary, which confirms the asserted *adiabatic* effect of neglecting the boundary flux

of the numerical dissipation ([1]).

□

Remarkably, the Gauss-Seidel (GS) iterative scheme has been failed in the boundary *closed* solution (see Figure 6.10 and Figure 6.11), albeit no visible modification has been detected on the behavior of other solvers (LGS, TP and CGS). So, the boundary closure of dissipation is significant also in terms of its impact on he matrix-conditioning of the system, at least in the sense of the performance of the GS solver.

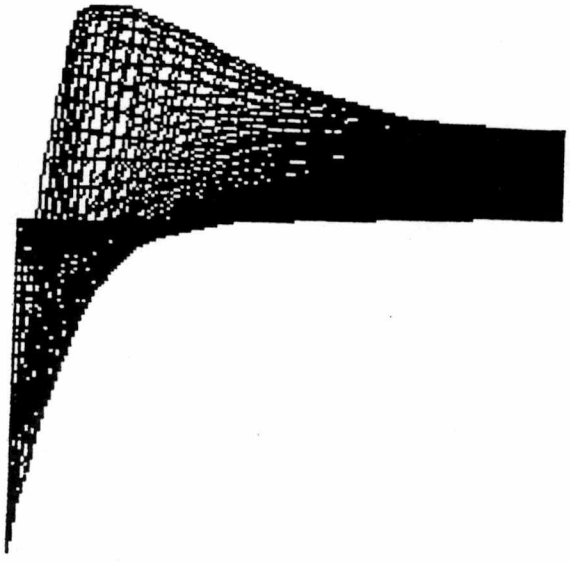


Figure 6.8: The boundary *unclosed* solution of Euler equations for subsonic flow — an upwinding side-view of the Mach number ("assessed" using the v-form approximate TWS dissipation).

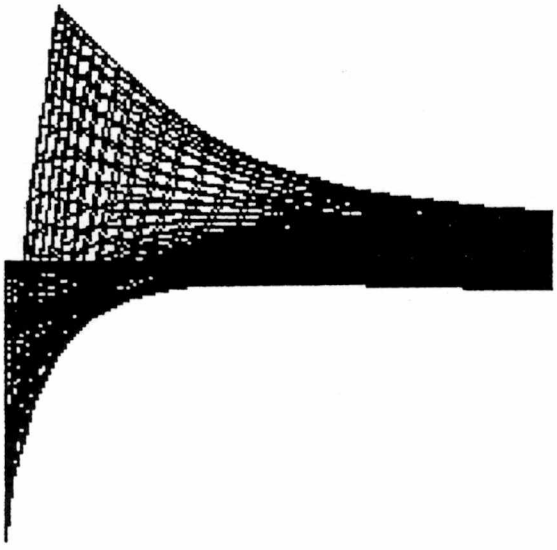


Figure 6.9: The boundary *closed* solution of Euler equations for subsonic flow — an upwinding side-view of the Mach number ("assessed" using the v-form approximate TWS dissipation).

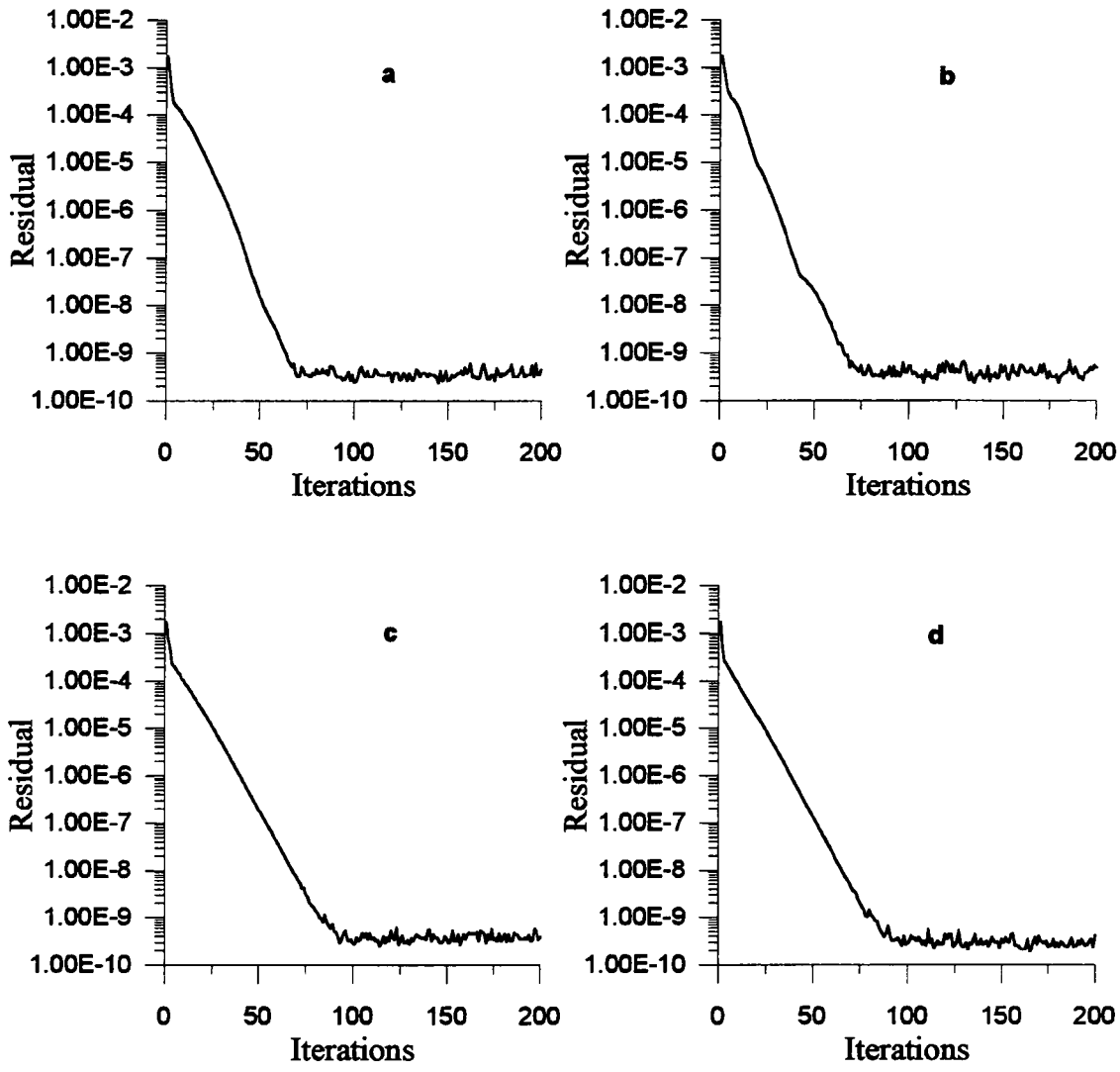


Figure 6.10: The convergence history of GS, in boundary unclosed FE solution of Euler equations. (**a**) LBH mode; (**b**) LBV mode; (**c**) RTH mode; (**d**) RTV mode.

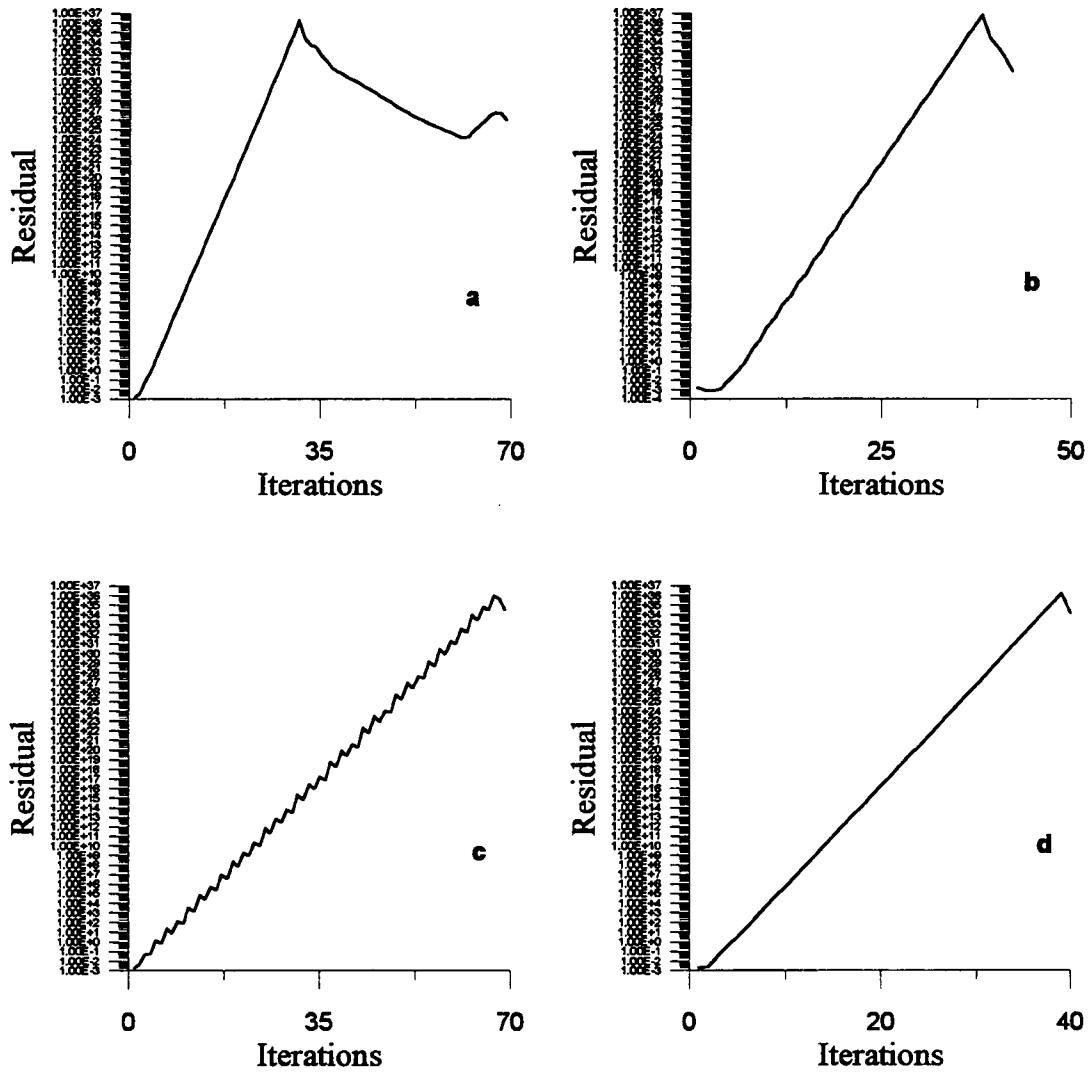


Figure 6.11: The divergence history of GS, in boundary closed FE solution of Euler equations. (**a**) LBH mode; (**b**) LBV mode; (**c**) RTH mode; (**d**) RTV mode.

Transonic Potential Flow

The verification case for transonic potential flow is for a 0.675 inlet Mach number. The solution is first driven for 30 steps from initialization using the subsonic mesh Figure 6.6, and then for another 60 steps using the refined mesh as shown in Figure 6.12. The refinement of mesh is basically a shift of the original one, adding only three vertical grid-lines but being a lot more densified in the downward part of the bump where shock is detected. The solution bears an residual of 1.237×10^{-6}. The maximum Courant number at convergence is 152.609.

The resulting Mach number is displayed in Figure 6.13 and Figure 6.14. As shown, unlike the previous subsonic case where flow get accelerated and deccelerated symmetrically along the bump, here a *shock* breaks the symmetricity and deccelerates the flow abruptly. The shock is located at about 0.79 of the bump-secant (measured from the bump-front), and with a 1.38 maximum shock Mach number, both matching the reported results ([12]) to the range of 0.5%.

While the quality of the numerically captured shock is obvious (within 3 meshes), overshooting does occur in the vincinity of its rear side, reflected in that the flow has to accelerate a bit before continuing to deccelerate. And it has a good reason: the flow is *subsonic* on the rear side of the shock, where thus the numerical dissipation has been switched off. Perhaps, a complete transonic potential algorithm should be one combining the "Mach switch" and the "shock operator" — with the latter being to extend the numerical dissipation to cover the entire shock (and only).

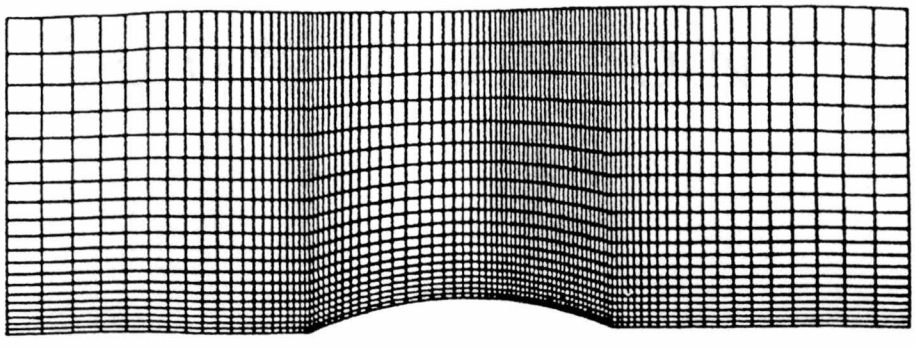


Figure 6.12: Refined mesh for Supersonic flow.

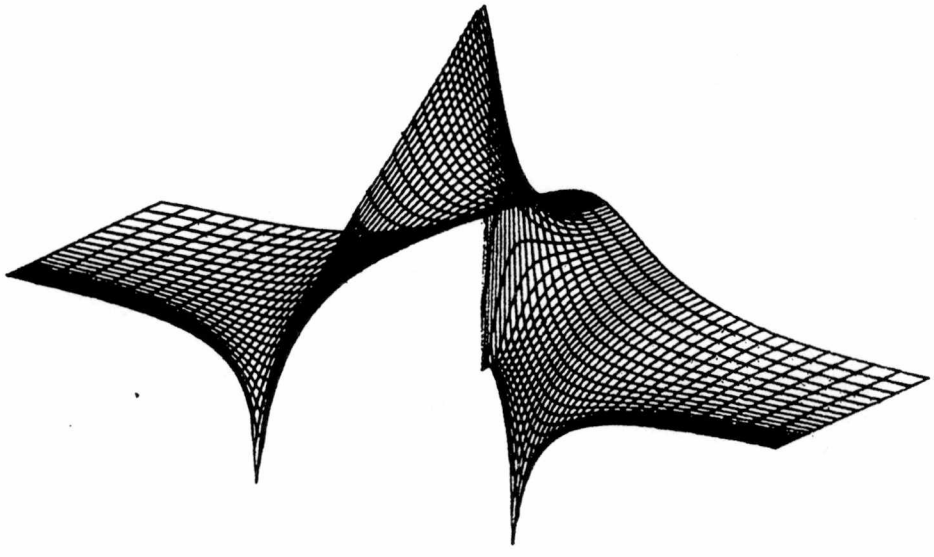


Figure 6.13: 3-D display for the Mach number of transonic potential flow, which is the solution using the S-TWS $v_j v_k$-form numerical dissipation (with $M_{in} = 0.675$).

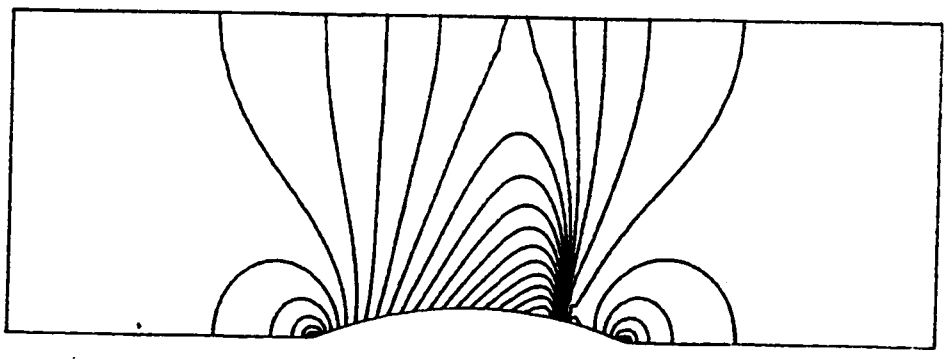


Figure 6.14: Contour display for the Mach number of the transonic potential flow, which is the solution using the S-TWS $v_j v_k$-form numerical dissipation (with $M_{in} = 0.675$).

Transonic Solution for Euler Equations

Finally, the previous transonic flow problem has been run for the Euler system S-TWS algorithm (the $w_j w_k$-form dissipation). Mesh refinement is the same as in the case of CPF, i.e., using the subsonic mesh (Figure 6.6) for the initial 30 steps, and the refined supersonic (Figure 6.12) in the next 60. The residual at convergence is 1.276×10^{-6}, with a maximum Courant number 159.138.

The converged Mach number is shown in Figure 6.15-6.16. The maximum shock Mach number is 1.43, which enjoys the same comparison with the reported solution ([12]) as that of CPF (i.e., also a range of 0.5%). The detected shock location is at about 0.83 of the bump-secant, which I regard as an 1.37% differnece from the reported 0.82 ([12]), considering that (a) the 0.83 shock location of this study is actually rounded from 0.8286..., and (b) the worst possible about the reported 0.82 is that it is rounded from 0.81499.... The worsened shock location comparison should probably be caused by that the reported data cited here is a full-upwinding solution (via Lax-Wendroff), which is three times more dissipative than the executed S-TWS of $\beta^* = 0.25$ (see Page 174).

The shock captured differs from its CPF counterpart mainly in being stronger, evidenced in its larger entering Mach number, more downstream location and being extended *deeper* into the flow channel as shown in Figure 6.15-6.16 (where even the back side of the profile has been "bumped" up). And it is quite reasonable, considering that the shock here is for *real*, whereas the shock in CPF is *isentropic*.

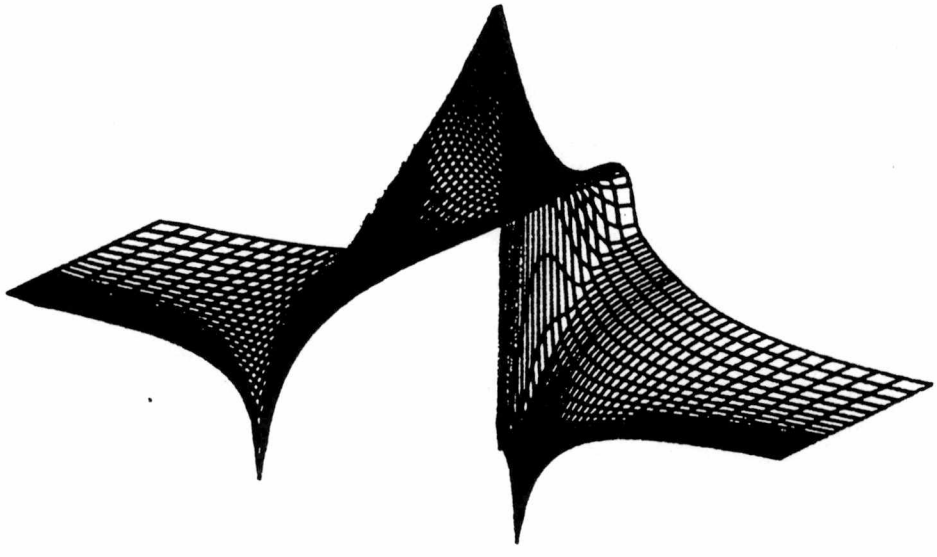


Figure 6.15: 3-D display for the transonic solution (Mach number) of Euler equations, which is the solution using the S-TWS $w_j w_k$-form numerical dissipation (with $M_{in} = 0.675$).

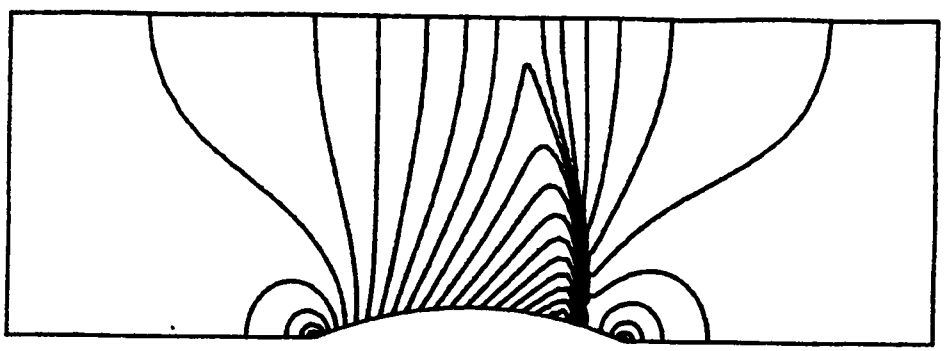


Figure 6.16: Contour display for the transonic solution (Mach number) of Euler euqaiotns, Euler equations, which is the solution using the S-TWS $w_j w_k$-form numerical dissipation (with $M_{in} = 0.675$).

6.4 Iterative Behavior

The implemented solvers have been tested at many states of the numerical solutions (for both the CPF and the Euler system). Reported in this section is the behavior of these solvers in the Euler solution as $n = 31$ and $p = 1$, i.e., the first Newton step of the first time station right after the mesh refinement, which is where (a) shock has been sufficiently developed, (b) the residual is still severe, and (c) the matrix conditioning has been worsened due to the mesh refinement. The iterations are all initiated setting

$$\{x\}_\alpha^0 = \{\Delta U\}_\alpha^{n-1} \tag{6.16}$$

where $\{\Delta U\}_\alpha^{n-1}$ is the net increment of $\{U\}_\alpha$ on the previous time station (which needs to be projected onto the refined mesh).

The Behavior of LGS and TP

The convergence history of the line Gauss-Seidel (LGS) relaxation algorithm and the tensor product (TP) approximate factorization are shown in Figure 6.17 and Figure 6.18 respectively. It is obvious that the LGS is far superior. Typically, all LGS solvers are fully converged in about 125 steps, reaching the "bottom line" (at about 10^{-9}), whereas the TP stays above 10^{-7} after 300 iterations. The TP scheme is in fact even worse than the GS solver in the case of "boundary unclosed" solution (see Section 6.3.2), although it is not "boundary closure" sensitive. Further, the superiority of LGS over TP is actually larger than it looks, since its per step work

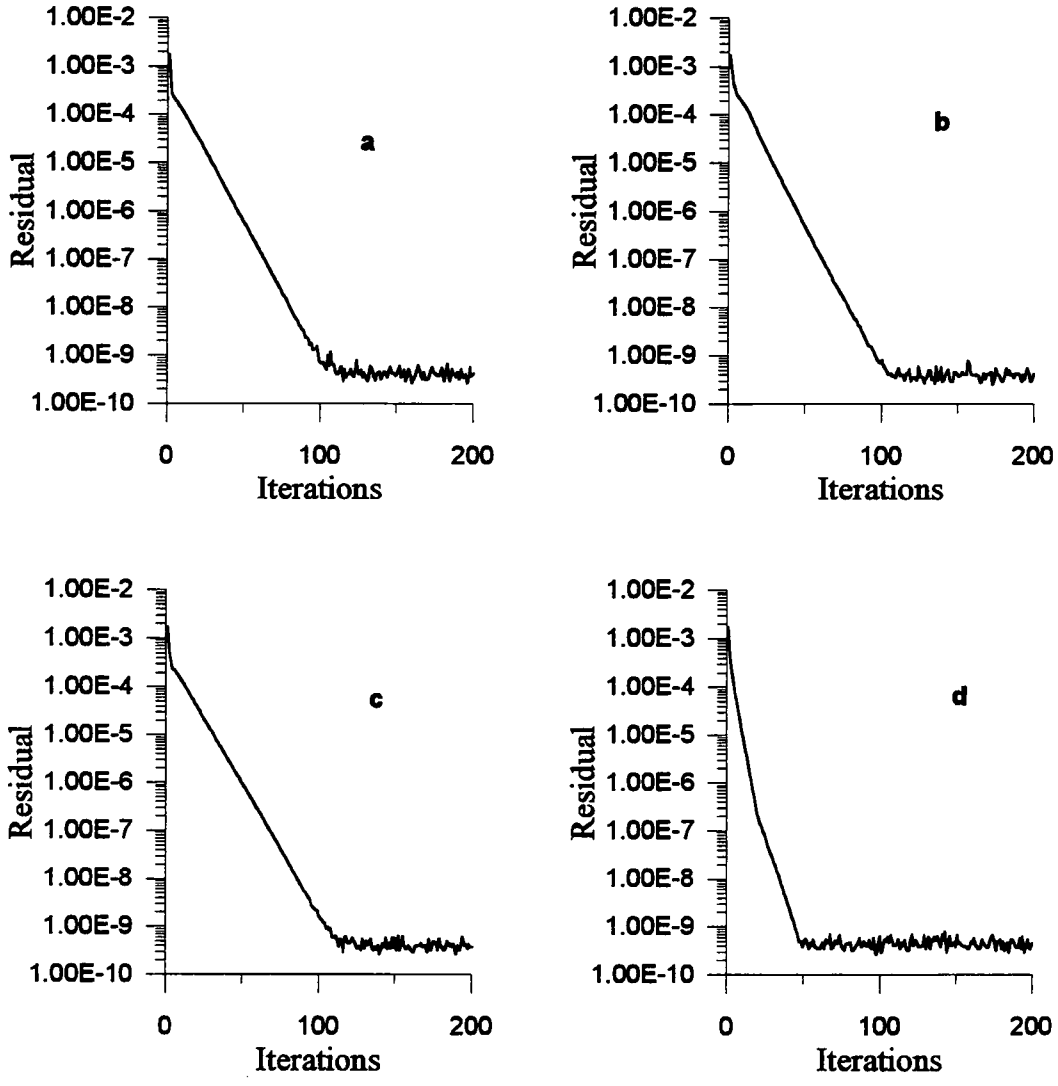


Figure 6.17: The convergence history of line Gauss-Seidel relaxation: (**a**) LV mode; (**b**) RV mode; (**c**) BH mode; (**d**) TH mode.

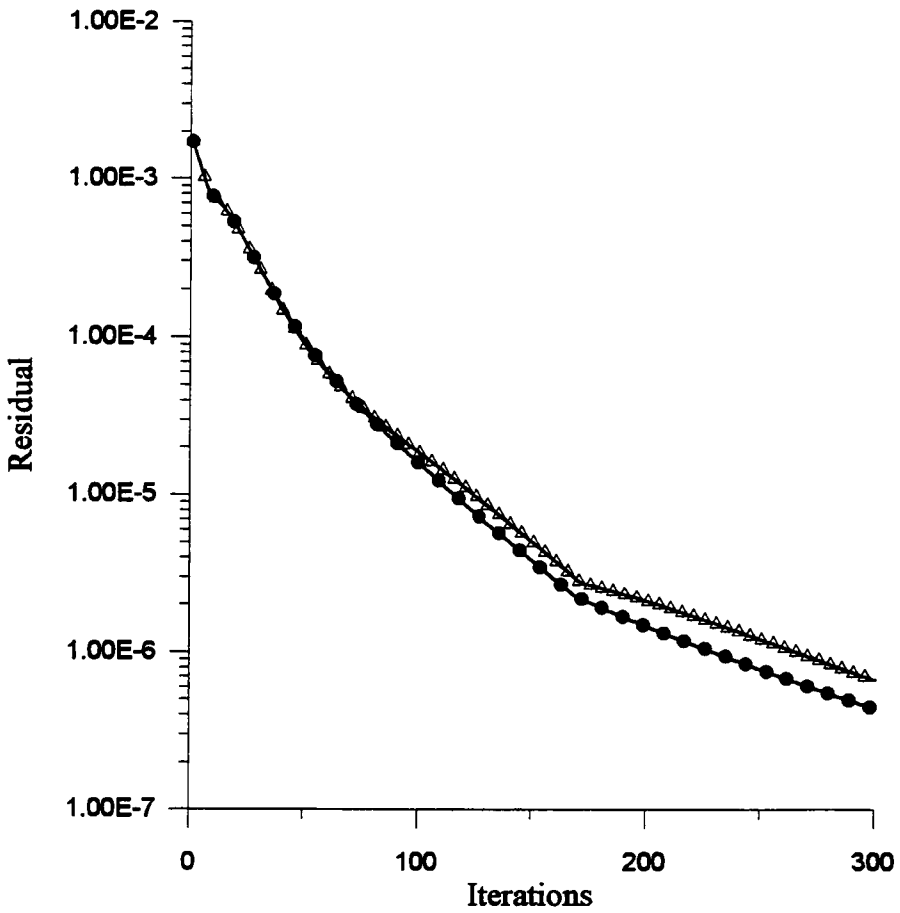


Figure 6.18: The convergence history of the tesor product factorization solver. • $J1 \otimes J2$; ∘ $J2 \otimes J1$.

is about half of the TP (consuming to inverse *line*-matrix in only *one* direction).

The reason for the ill performance of TP is probably the *large* Courant number (i.e., time step Δt) of the tested circumstance, which proportions the inaccuracy of the TP factorization (see Chapter ??).

The Composite LGS Solvers

Since the convergence rate of the LGS solver varies depending on the "mode", as shown in Figure 6.17, several "composite LGS solvers" are tested. The results are shown in Figure 6.19. The LV-RV pair is fully converged in about 125 steps, being about the same as each standing alone. The BH-TH pair takes about 50 steps, being comparable with the TH mode — the better one between them. However, the best performance is generated by a "cycled" LV-BH-RV-TH, which reaches the "bottom line" in approximately 35 steps.

The Conjugate Gradient Square Solvers

The tests of the CGS solvers are documented in Figure 6.20-6.21-6.22, which are respectively the convergence history of a plain, TP and LGS preconditioned CGS. The plain CGS takes about 200 steps to drive the residual to the level of 10^{-8}, much worse than any of the LGS modes (and their "composition") and just slightly better than the TP. The TP preconditioners, as in Figure 6.21, show to be either basically effectless (the $J1 \otimes J2$ case) or actually worsening (the $J2 \otimes J1$ case) the numerical efficiency of CGS, which confirms the earlier assertion, i.e., the TP factorization is a very inaccurate approximation for this large Courant number linear algebra system.

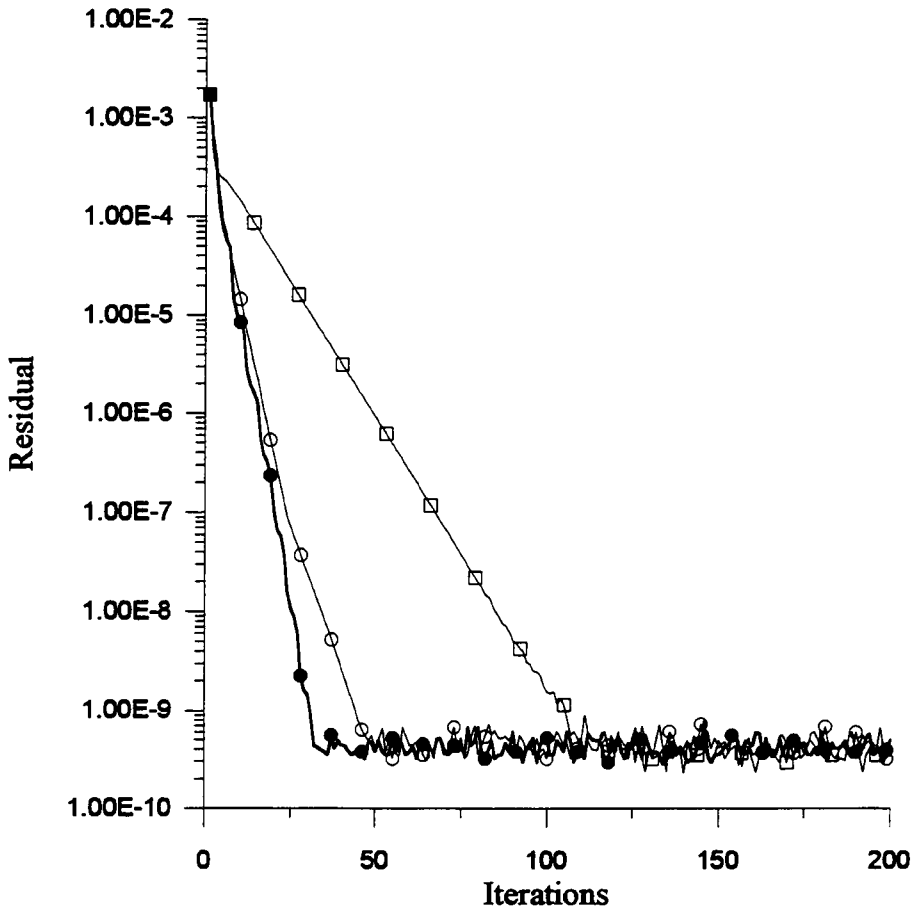


Figure 6.19: The convergence history of the "composite" line Gauss-Seidel solvers; □ — the LV-RV pair; o — the BH-TH pair; • — the cycled LV-BH-RV-TH.

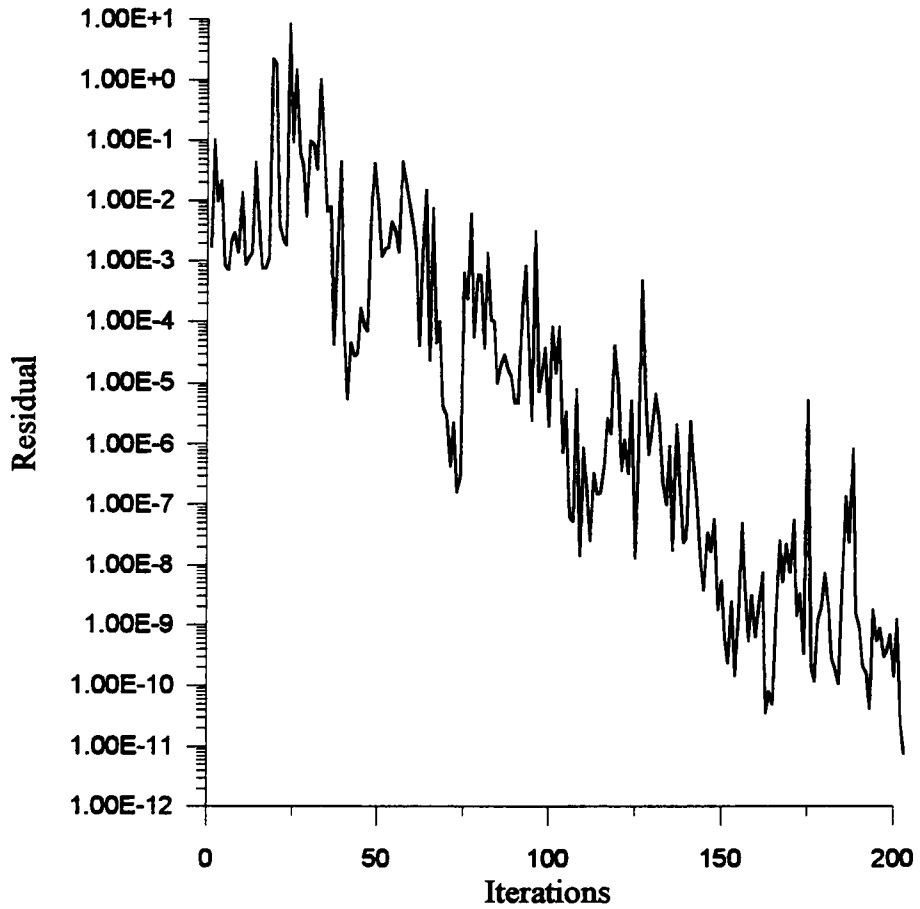


Figure 6.20: The convergence history of the plain conjugate gradient sqaure solver.

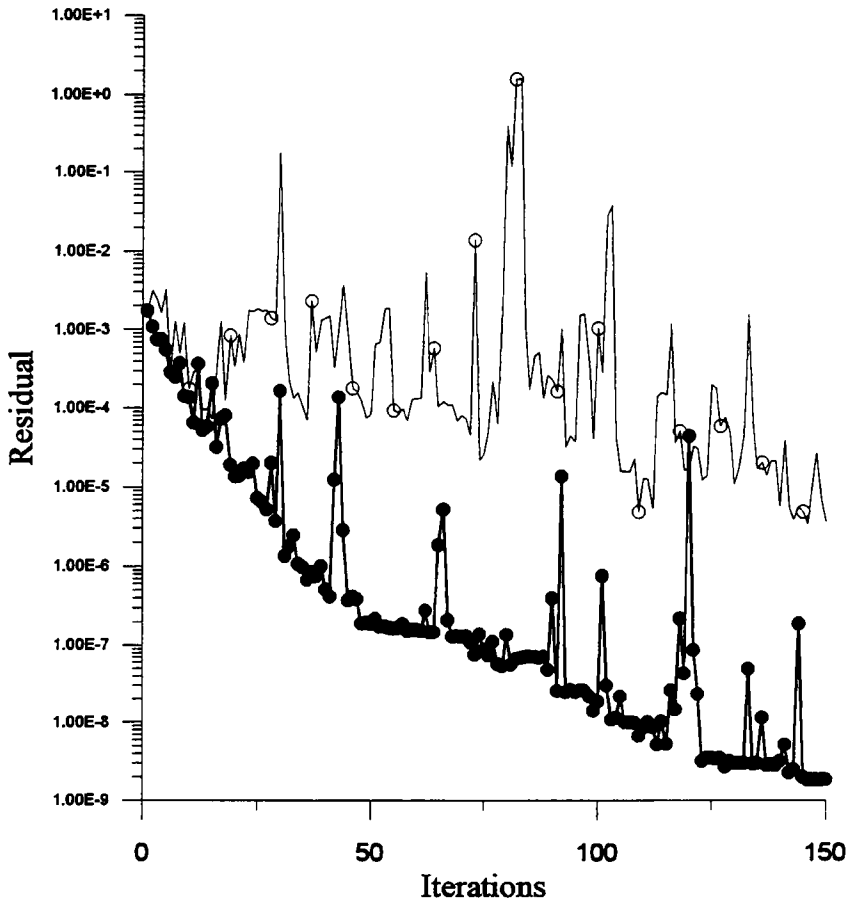


Figure 6.21: The convergence history of the TP preconditioned conjugate gradient square method. Preconditioner: • — $J1 \otimes J2$; ∘ $J2 \otimes J1$

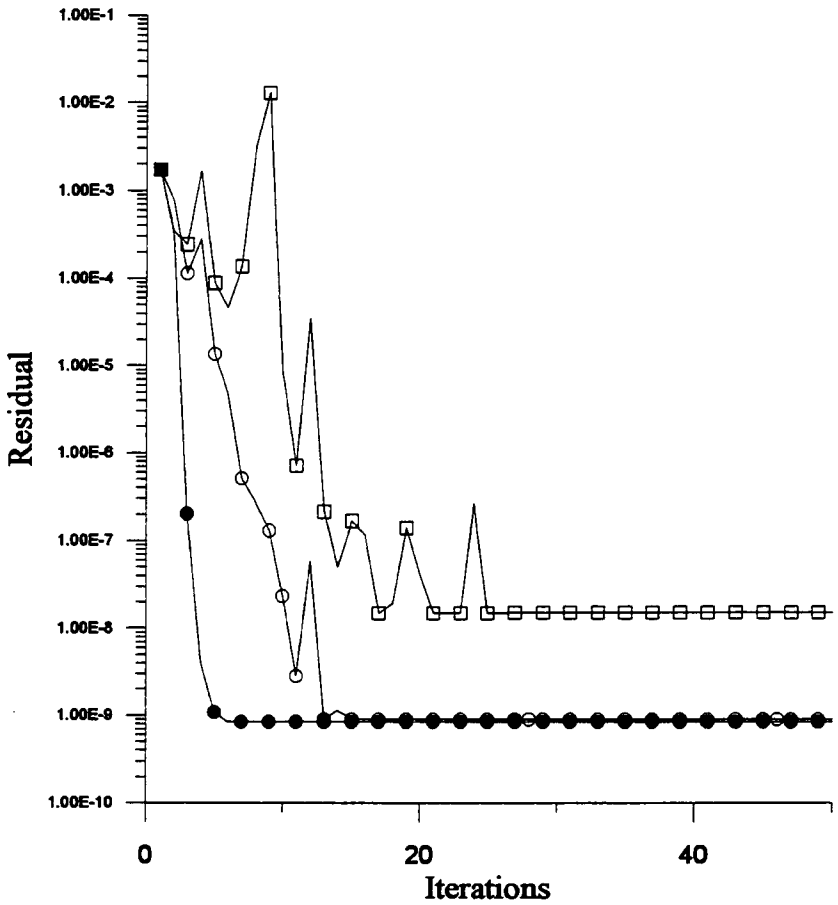


Figure 6.22: The convergence history of the LGS preconditioned conjugate gradient square method. Preconditioners: □ — the LV-RV pair; o — the BH-TH pair; • — the cycled LV-BH-RV-TH.

In comparison, the "composite" LGS preconditioned CGS solvers yield huge improvement, as shown in Figure 6.22. The *best* among them is the one with the LV-BH-RV-TH "cycled" LGS as the preconditioner, which is fully converged in just 7 steps (in terms of residual). And even the *worst*, the one alternatively using the LV-RV mode of LGS as the preconditioner, is able to reach the "bottom line" at $N = 23$.

□

Finally, the "residual" and "preconditioned" residual of the LV-BH-RV-TH "cycled" LGS preconditioned CGS algorithm are plotted in Figure 6.23. What is interesting about this picture is that, as "r" becomes stagnated, "pr" continues to converge and to a much lower level. Since

$$\begin{aligned} \{pr\} &= [B]^{-1}\{r\} \\ &= [B]^{-1}\left(\{b\} - [A]\{x\}\right) \\ &\simeq [A]^{-1}\{b\} - \{x\} \end{aligned}$$

where $[B]^{-1} \simeq [A]^{-1}$ is the preconditioner, the conclusion should be that it is the solution which continues to converge and to a much lower level.

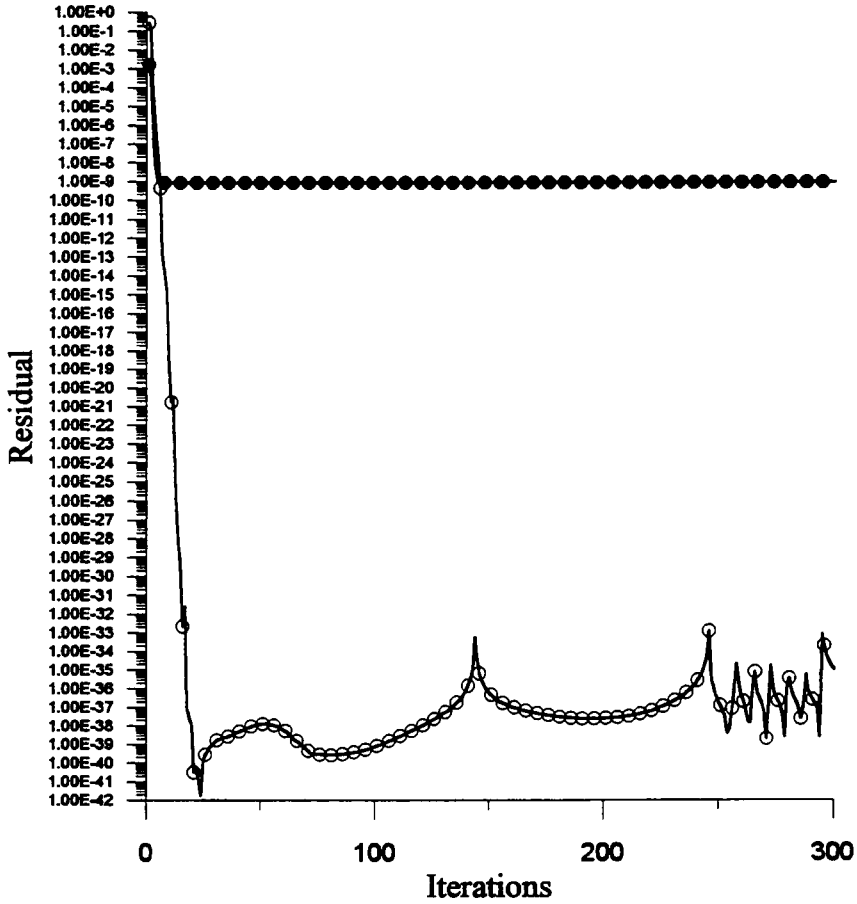


Figure 6.23: The convergence history of conjugate gradient square method with the LV-BH-RV-TH cycled LGS as the preconditioner; • — residual; o — preconditioned residual.

Chapter 7

Conclusions

The following summarizes the research of this dissertation, which fulfills not only all the proposed commitments ([1], marked by •) but a few extra developments that were found meaningful in the process:

- A metric data mean (MDM) algorithm has been developed for finite element CFD method, to extract out the metric data in FE integrals, and thus to admit in memory storage the FE matrices on a localized data base. Major application of MDM is to facilitate quadrature-free FE computing, so that the efficiency of the FE-CFD algorithm may become basically comparable to finite difference (FD) algorithm. Application has also been found in tensor product factorization (for contracting the dummy indexes).

- A *selective* Taylor Weak Statement (S-TWS) algorithm has been developed. S-TWS for compressible potential flow is the *just* solution for the problem which a

plain TWS endures — that the "companion conservation laws" are third order differential and thus necessitates the use of C^1 Hermite elements, yielding a positive-definite dissipation scheme. S-TWS for Euler/Navier-Stokes equations is a platform offering insight and choice for the simplification of numerical dissipation, which, under certain *selections*, leads to a diagonalized "$w_j w_k$-form" dissipation algorithm (w_j being a diagonal matrix in which v_j, the jth component of velocity is the entry at all but the jth row/column, where $(2-\gamma)v_j$ takes place).

- An invariant dissipation length (IDL) algorithm has been established for dissipation "rescaling". Theoretically, what the IDL algorithm reflects is nothing but the fact the dissipation time (of TWS, or whatever else *temporal* dissipation procedures) is a *scalar*, which neither discriminates nor privileges any direction in the flow field regardless to being "rescaled" or not. The ultimate results of the IDL algorithm are two master-element based IDL formulas, one for quadrilateral/hexahedron elements, and one for triangle/tetrahedron (both being appropriate regardless to the *order* of the elements).

- To verify the above algorithms, compressible potential flow and Euler equations have been solved, benchmarked for the airflow around an arc bump, observing stiff stability/shock. Observed in this process is also the necessity for the "boundary closure" of FE numerical dissipation, in terms of its impact on both the solution accuracy and the matrix-condition of the numerical linear algebra

system.

- Several linear algebra algorithms have been implementated solving the finite element systems of CPF and Euler equations, including the Gauss-Seidel iteration (GS), the tensor product approximate factorization (TP), the line Gauss-Seidel relaxation (LGS), the conjugate gradient square (CGS), and the GS, TP and LGS preconditioned CGS. The tests over these methods, in the case of rather large Courant numbers (between 100 and 150), demonstrate the LGS preconditioned CGS being the best solver, and the "cycled composite LGS" the best line relaxation solver (far superior to TP).

- A special single variable convection system of clear yet analyzable nonlinearity/shock mechanism has been established. One-dimensional TWS-FE confirmation has been performed in this case. The fact that the system is in "non-conservative" form proves being of no effect to the stability of its TWS algorithm, in spite of the presence of very strong shock, albeit it is indeed true that the "conservation" of the system has not strictly maintained in numerical solution, and it is part of the convergence process.

- The positive definiteness of the numerical dissipation for CPF has been theoretically justified, being a necessary condition derived from the Second Law of thermodynamics. Remarkably, such a common sense result is not even obeyed by some "mainstream" schemes, which put the S-TWS algorithm built in this study in an unique position.

- Finally, an incompressible potential flow based solution initialization procedure has been developed. The advantage of this algorithm is its feasibility in generic geometry, at the price of solving the Laplacian incompressible potential equation.

BIBLIOGRAPHY

Bibliography

[1] Jing Zhang "Metric Data Mean Technique for Finite Element Algorithm and TWS Computation for Compressible Flow Finite Element Computational Fluid Dynamics", A dissertation proposal, University of Tennessee, 1992.

[2] A. J. Baker, "Finite Element Computational Fluid Dynamics", Hemisphere, Washington DC, 1983.

[3] Eric B. Becker, Graham F. Carey and J. Tinsley Oden, "Finite Elements, *An Introduction*", Pentice-Hall, Inc., Englewood Cliffs, New Jersey, 1981.

[4] G. Strang and G. J. Fix, "An Analysis of the Finite Element Method", Pentice-Hall, Inc., 1973.

[5] O. C. Zienkiewicz, "The Finite Element Method in Engineering Science", Mcgraw-Hill, New York, 1971.

[6] R. Courant and D. Hilbert, "Methods of Mathematical Physics", John Wiley, New York, 1989.

[7] P. R. Garabedian, "Partial Differential Equations", Chelsea, New York, 1986.

[8] S. Whitaker, "Introduction to Fluid Mechanics", Prentice-Hall, Englewood Cliffs, NJ, 1968.

[9] Mech. Dept, Qinghua Univ., "Introduction to Fluid Dynamics", Mechanical Print, Beijing, December, 1980.

[10] M. E. O'Neil and F. Chorlton, "Viscous and Compressible Fluid Dynamics", Prentice-Hall, Englewood Cliffs, NJ, 1968.

[11] F. M. White, "Viscous Flow", McGraw-Hill, New York, 1978.

[12] Sheng Zhou and Mengyu Shen, "Computational Fluid Dynamics", Beijing Aerospace University, January, 1984.

[13] I. Prigogine, "Introduction to Non-equilibrium Thermodynamics", Science Print, Beijing, 1960.

[14] J. Donea, "A Taylor-Galerkin Method for Convective Transport Problems", Int. J. of Num. Methods Eng., No. 20, 1984

[15] A. J. Baker and J. W. Kim, "Taylor Weak Statement Algorithm for Hyperbolic Conservation Laws", Int. J. Num. Mtds. Fluids, No. 7, 1987.

[16] A. J. Baker, "Progress on A Taylor Weak Statement Finite Element CFD Algorithm for Configuration Aerodynamics Analysis", COMCO TR 88-20.0, 1988.

[17] G. S. Iannelli and A. J. Baker, "An Efficient Solution-Adaptive Implicit Finite Element CFD Navier-Stokes Algorithm", AIAA-90-0400, 1990.

[18] G. S. Iannelli and A. J. Baker, "An Implicit and Stiffly Stable Finite Element CFD Algorithm for Unsteady Aerodynamics", AIAA-90-0400, 1990.

[19] G. S. Iannelli and A. J. Baker, "An Intrinstically N-dimensional Generalized Flux Vector Splitting Implicit Finite Element Euler Algorithm", AIAA-91-0123, 1990.

[20] G. S. Iannelli, "A Globally Well-posed Accurate and Efficient Finite Element Finite CFD Algorithm for Comressible Aerodynamics", Ph.D thesis, The University of Tennessee, Knoxville, 1991.

[21] James D. Freels, "A Taylor Weak Statement Finite Element Algorithm for Real-Gas Compressible Navier-Stokes Simulation"Ph.D thesis, The University of Tennessee, Knoxville, 1992.

[22] J. W. Kim, "On Taylor Weak Statement Finite Element Computation in Gas Dynamics", Ph.D thesis, The University of Tennessee, Knoxville, 1988.

[23] A. J. Baker, G. S. Iannelli and J. D. Freels, "Recent Advances on a Finite element Algorithm for Computational Aerodynamics: Trasonics-Hypersonics", 4th International Symposium on Computational Fluid Dynamics, pp 48-53, The University of California, Davis, September, 1981.

[24] Wilbert. P. Noronha, "Accuracy, Convergence, and Stability of Finite Element Algorithm for Incompressible Flow", Ph.D thesis, The University of Tennessee, Knoxville, 1989.

[25] Paul T. Williams, "A Three-Dimensional Time-Accurate Incompressible Navier-Stokes Finite Element CFD Algorithm", Ph.D thesis, The University of Tennessee, Knoxville, 1993.

[26] A. J. Baker and M. O. Soliman, "A Finite Element Algorithm for Computational Fluid Dynamics", AIAA Journal, Vol. 21, pp 816-827, June, 1983.

[27] T. E. Tezduyar and T. J. R. Hughes, "Development of Time-accurate Finite Element Techniques for First-order Hyperbolic System with Particular Emphasis on Compressible Euler Equations", Report Prepared under NASA-Ames University Consortium Interchange, No. NCA2-OR745-104, 1982.

[28] T. J. R. Hughes, L. P. Franca and M. Mallet, "A New Finite Element Formulation for Computational Fluid Dynamics: I. Symetric Forms of the Compressible Euler and Navier-Stokes Equations and the Second Law of Thermodynamics", Computer Methods in Applied Mathematics and Engineering, 54, pp 223-234, 1986.

[29] T. J. R. Hughes, M. Mallet and Akira Mizukami, "A New Finite Element Formulation for Computational Fluid Dynamics: II. Beyond SUPG", Computer Methods in Applied Mathematics and Engineering, 54, pp 341-355, 1986.

[30] T. J. R. Hughes, M. Mallet and Akira Mizukami, "A New Finite Element Formulation for Computational Fluid Dynamics: III. The Generalized Streamline Operator for Multidimensional Advective-Diffusive System", Computer Methods in Applied Mathematics and Engineering, 58, pp 305-328, 1986.

[31] T. J. R. Hughes, M. Mallet and Akira Mizukami, "A New Finite Element Formulation for Computational Fluid Dynamics: IV. A Discontinuity-Capturing Operator for Multidimensional Advective-Diffusive System", Computer Methods in Applied Mathematics and Engineering, 58, pp 329-339, 1986.

[32] T. J. R. Hughes and T. E. Tezduyar, "Finite Element Methods for First-order Hyperbolic Systems with Particular Emphasis on the Compressible Euler Equations", Computer Methods in Applied Mathematics and Engineering, 45, pp 217-284, 1984.

[33] T. E. Tezduyar and T. J. R. Hughes, "Finite Element Formulations for Convection Dominated Flows with Particular Emphasis on the Compressible Euler Equations", AIAA Paper 83-0125, 1983.

[34] A. N. Brooks and T. J. R. Hughes, "Streamline Upwind/Petrov-Galerkin Formulations Convection Dominated FLows with Particular Emphasis on the Incompressible Navier-Stokes Equations", Computer Methods in Applied Mathematics and Engineering, 32, pp 199-259, 1982.

[35] G. J. LeBeau, S. E. Ray, S. K. Aliabadi and T. E. Tezduyar, "SUPG Finite Element Computation of Compressible Flows with The Entropy and Conservation Variables Formulation", UMSI 92/26, University of Minnesota Supercomputer Institute, 1992.

[36] K. W. Morton and A. K. Parrott, "Generalized Galerkin Method for First Order Hyperbolic Equations", Journal for Computational Physics, Vol 36, pp 249-270, 1980.

[37] K. W. Morton, "Characteristics Galerkin Methods for Hyperbolic Flow", Proceedings of the Fifth GAMM Conference on Numerical Methods in Fluid Dynamics, Rome, 1983.

[38] G. F. Garey, J. T. Oden and Paresh Pattani, "An h-p Adaptive Finite Element Method for the numerical simulation of Compressible Flow", Computer Methods in Applied Mechanics and Engineering, 70(2), pp203-235, 1988.

[39] Ph. Devloo, J. T. Oden and Paresh Pattani, "Adaptive Finite Element Methods for Compressible Flow Problems", Applied Mechanics Division Symposia Series, Vol. 78, pp 115-126, 1986.

[40] J. T. Oden, T. Strouboulis and Ph. Devloo "Adaptive Finite Element Methods for High-speed Compressible Flow Problems", International Journal for Numerical Methods in Fluids, 7:1211-1228, 1987.

[41] W. G. Habashi, M. Robichaud, V. N. Nguyen, W. S. Ghaly and M. Fortin, "Large-scale Computational Fluid Dynamics by Finite Element Method", AIAA Paper 91-0120, 1991.

[42] W. G. Habashi, G. Baruzzi and M. Hafez, "A Second Order Finite Element Formulation for Viscous Transonic Flows", 4th International Symposium on Computational Fluid Dynamics, Davis, CA, September 1991.

[43] M. O. Bristeau, R. Glowinski, B. Mantel, J. Periaux and P. Perrier, "Numerical Methods for the Time Dependent Compressible Navier-Stokes Equations", Proceedings of the Sixth International Symposium for Computational Methods in Applied Sciences and Engineering, pp 565-583, 1983.

[44] M. O. Bristeau, R. Glowinski, J. Periaux and H. Viviand, "Numerical Simulation of Compressible Navier-Stokes Flows", Notes on Numerical Fluid Mechanics, Vol. 18, pp 349, 1987.

[45] M. O. Bristeau and J. Periaux, "Finite Element Methods for the Calculation of Compressible Viscous Flows Using Self-adaptive Mesh Refinement", Lectures on Computational Fluid Dynamics, Vol. 2, pp 108; Von Karman Institute for Fluid Dynamics, 1987.

[46] M. O. Bristeau, R. Glowinski and J. Periaux, "Numerical Methods for Navier-Stokes Equations, Applications to the Simulation of Compressible and Incom-

pressible Viscous Flow", Finite Element in Physics, pp 73-187, American Mathematics Society, North-Holland, 1986.

[47] M. O. Bristeau, R. Glowinski, J. Periaux and G. Roge, "Compressible Viscous Flow Calculations Using Compatible Finite Element Approximations", Proceedings of the Seventh International Conference on Finite Element Methods in Flow Problems, pp 322-329, April, 1989.

[48] Bo-Nan Jiang and C. L. Chang "Least-Square Finite Elements for Stokes Problem", NASA TM-101308, ICOMP-88-16, December, 1988.

[49] G. F. Carey and Bo-Nan Jiang "Least-Square Finite Elements for First Order Hyperbolic System", International Journal for Numerical Methods in Engineering, Vol. 26, pp 81-93, 1988. , April, 1989.

[50] Bo-Nan Jiang and L. A. Povinelli, "Least-Square Finite Element Method for Fluid Dynamics", Proceedings of the Seventh International Conference on Finite Element Methods in Flow Problems, pp 105-110, April, 1989.

[51] R. Glowinski, "Viscous Flow Simulation by Finite Element Methods and Related Numerical Technique", Progress and Supercomputing in Computational Fluid Dynamics, pp 173-210, December, 1984.

[52] R. Glowinski and Periaux, "Finite Element, Least Squares and Domain Decomposition Methods for the Numerical Solution of Nonlinear Problems in Fluid

Dynamics", Numerical Method in Fluid Dynamics, pp 1-114, Spring-Verlag, New York, 1985.

[53] L. W. Spradley, J. F. Stalnaker and A. W. Ratliff, "Computation of Three-Dimensional Viscous Flows with Navier-Stokes Equations", AIAA Paper 80-1348, 1980.

[54] Jing Zhang, "Pointwise Finite Element Method and Its Application to Compressible Flows", Proceedings of the 5th International Conference for Computational Fluid Dynamics, Bejing, 1985.

[55] P. Lax and B. Wendroff, "System of Conservation Laws", Communication on Pure and Applied Mathematics, Vol. XIII, pp 217-237, 1960.

[56] P. Lax and B. Wendroff, "Difference Schemes for Hyperbolic Equations with High Order of Accuracy", Communication on Pure and Applied Mathematics, Vol. XVII, pp 381-398, 1964.

[57] K. C. Reddy and J. N. Reddy, "A Locally Implicit Finite Element Analysis Algorithm for 3D Compressible Flows", Proceedings of the 7th International Conference on Finite Element Methods in Flow Problems, pp 339-344, Huntsville, April 1989.

[58] C. A. J. Fletcher, "An Alternative Direction Implicit Finite Element Method for Compressible viscous Flow", Seventh International Conference on Numerical Methods in Fluid Dynamics, pp 182-187, Berlin, 1981.

[59] W. R. Dykson, "Tensor Product Generalized ADI Methods for Separable Elliptic Problem", SIAM J. Num. Anal., Vol. 24, 1987.

[60] L. J. Hayes, "Galerkin Alternative-Direction Methods on Non-Rectangular Regions Using Patch Approximations", SIAM J. Num. Anal., Vol. 24, 1987.

[61] L. J. Hayes, "Implementation of Finite Element Alternative-Direction Methods on Non-Rectangular Regions", SIAM J. Num. Anal., Vol. 24, 1987.

[62] H. L. Stone, "Iterative Solution of Implicit Approximations of Multidimensional Partial Differential Equations", SIAM J. Num. Anal., Vol. 5, No. 3, 1987.

[63] G. E. Schneider and M. Zedan, "A Modified Strongly Implicit Procedure for the Numerical Solution of Field Problems", Numerical Heat Transfer, No. 4, pp 457-461, 1981.

[64] A. J. Baker, W. P. Noronha and E. L. Wachspress, "Consistent Sparse Factorization of Elliptic Difference Equations", Applied Mathematics Letters, 2:215-220, 1989.

[65] P. Sonneveld, "CGS, A Fast Lanczos-type Solver for Non-symmetric Linear System", SIAM J of Sci. and Stat. Comput., Vol. 10, No. 1, January 1986.

[66] M. Vitaletti, "A Preconditioned CGS Algorithm for the Solution of Unfactorized Implicit Schemes in 2D Inviscid Aerodynamics", IBM European Center for Sci. and Eng. Comput., Via Giorgione 159, 00147 Rome, Italy.

[67] R. Fletcher "Conjugate Gradient Methods for Indefinite System", Lecture Notes in Mathematics 506, Spring-Verlag, Berlin, Heidelberg, New York, 1976, pp. 73-89.

[68] Y. Saad and M. H. Schultz, "GMRES: A Generalized Minimal Residual Algorithm for Solving Non-Symmetric Linear System", SIAM J of Sci. and Stat. Comput., Vol. 7, No. 3, July 1986.

[69] O. Axelsson, "A Generalized Conjugate Direction Method and Its Application on a Singular Perturbation Problem", Lecture Notes in Mathematics 773, Spring-Verlag, Berlin, Heidelberg, New York, 1980, pp. 1-11.

[70] Kumar K. Tamma and Kong C. Saw, "Hierarchical p-version Finite Elements and Adaptive A-Posteriori Computational Formulations for Two-dimensional Thermal Analysis", Computers and Structures, 32, pp. 1183-1194, 1989.

[71] O. C. Zienkiewicz and A. W. Craig, "A-Posteriori Error Estimation and Adaptive Mesh Refinement in Finite Element Method, *The Mathematical Basis of Finite Element Methods*", Editted by David F. Griffiths, Clarendon Press, Oxford, 1984.

[72] O. C. Zienkiewicz, J. P. Gago, S. R. De and D. W. Kelly, "The Hierchical Concept in Finite Element Analysis", Computers and Structures, 16, pp. 53-65, 1983.

[73] G. S. Dulikravich, "Analysis of Artificial Dissipation Models for the Transonic Full Potential Equations", AIAA Journal, Vol. 26, No. 10, 1988.

[74] G. S. Dulikravich, "A Physically Consistent Model for Artificial Dissipation in Transonic Potential Computations", NASA TM 100846, ICOMP-88-6, May 1988.

[75] N. L. Sankar, J. B. Malone and Y. Tassa "An Implicit Conservative Algorithm for Steady and Unsteady Three-Dimensional Transonic Potential Flows", AIAA-81-1016, 1981.

[76] W. G. Habashi and M. Hafez, "Finite Element Solutions of Transonic Flow Problems", AIAA Journal, Vol. 20, No. 10, 1981.

[77] M. Hafez and E. Murman, "Artificial Compresiibility Mthods for Numerical Solutions of Transonic Full Potential Equation", AIAA Journal, Vol. 17, No. 8, 1979.

[78] A. Jameson and D. Caughey, "A Finite Volume Scheme for Transonic Potential Flow Calculation", AIAA, New York, pp. 35-54, 1977.

[79] A. Jameson, "Transonic Potential Flow Calculation Using Conservation Form", Proceedings of the AIAA Second Computational Fluid Dynamics Conference, Hartford, June, 1975, pp.148-161

[80] A. Jameson, "Iterative Solution of Transonic FLows Over Airfoil and Wings", Communications on Pure ans Applied Mathematics, Vol. 2, 1974, pp. 283-309.

[81] E. M. Murman and J. D. Cole, "Calculation of Plane steady Transonic Flow", AIAA Journal, Vol. 9, No. 1, 1971.

[82] J. Steinhoff and A. Jameson, "Multiple Solution of the Transonic Potential Flow Equation", AIAA Journal, Vol. 20, pp 1521-1525, November, 1982.

[83] Zhuquan Guan and Sheng Zhou, "On Plane Steady Inviscid Flow and Minimal Entropy Production Principle (I)", China Science, April, 1985.

[84] Zhuquan Guan and Sheng Zhou, "On Plane Steady Inviscid Flow and Minimal Entropy Production Principle", China Science, April, 1985.

[85] B. van Leer, "Towards the Ultimate Conservation Difference Scheme, II Monotonicity and Conservation Combined in a Second Order Scheme", Journal of Computational Physics, Vol. 14, pp 361-370, 1974.

[86] B. van Leer, "Flux-Vector Splitting for the Euler Equations"Lecture Notes in Physics, Vol. 170, pp 507-512, 1982.

[87] W. K. Anderson, J. L. Thomas and B. van Leer, "A Comparison of Finite Volume Flux-Vector Splittings for the Euler Equations"AIAA Paper 85-0122, 23rd Aerospace Science Meeting, Reno, NV, 1985.

[88] J. L. Thomas, B. van Leer and R. W. Walters, "Implicit Flux-Split Schemes for the Euler Equations"AIAA Paper 85-1680, 18th Fluid Dynamics, Plasmadynamics and Lasers Conference, Cincinnati, OH, 1985.

[89] M. S. Liou and B. van Leer, "Choice of Implicit and Explicit Operators for Upwinding Differencing Method" AIAA Paper 88-0624, 26th Aerospace Science Meeting, Reno, NV, 1988.

[90] R. M. Beam and R. F. Warming, "An Implicit Finite-Difference Algorithm Hyperbolic Systems in Conservation-Law Form", Journal of Computational Physics, Vol. 22, pp 87-110, 1976.

[91] J. L. Steger and R. F. Warming, "Flux Vector Splitting of Inviscid Gasdynamics Equations with Application to Finite Difference Methods" Journal of Computational Physics, Vol. 40, pp 263-293, 1981.

[92] T. H. Pulliam, "Efficient Solution Methods for the Navier-Stokes Equations" Lecture Notes, Von Karman Institute Lecture Series, Numerical Techniques for Viscous Flow Computation in Turbomachinery Bladings, Brussels, Belgium, 1986.

[93] G.M.Johnson, "Relaxation Solution of Full Euler Equations", NASA TM 82889, 1982.

[94] V. Yakhot and S. A. Orszag, "Renormalization Group Analysis of Turbulence. I. Basic Theory", Journal of Scientific Computing, Vol. 1, 1986.

[95] L. Martinelli and V. Yakhot, "RNG-Based Turbulence Transport Approximation with Application to Transonic Flows", AIAA-89-1950, 1989.

[96] V. C. Patel and et. al, "Turbulence Models for Near Wall and Low Reynolds Number Flows: A Review", AIAA Journal, Vol.23, No.9, 1984.

[97] K. Y. Chien, "Predictions of Channel ans Boundary-Layer Flows with a Low-Reynolds-Number Turbulence Model", AIAA Journal, Vol.20, pp. 33-38, 1984.

[98] D. Dutoya and P. Michard, "A Program for Calculating Boundary Layers along Compressor and Turbine Blades", Numerical Methods in Heat Transfer, edited by R.W. Lewis and et. al, New York, 1981.

[99] S. Hassid and M. Poreh, "A Turbulent Energy Dissipation Model for FLows with Drag Reduction", Journal of Fluids Engeering, Vol. 100, pp. 107-112, 1978.

[100] G. H. Hoffmann, "Improved Form of the Low-Reynolds Number $k-\epsilon$ Turbulence Model", Physics of Fluids, Vol. 18, pp. 309-312, 1975.

[101] C. K. G. Lam and K. Bremhorst, "A Modified Form of $k-\epsilon$ Model for Predicting Wall Turbulence", Transaction of ASME, Vol.103, Septempber 1981.

[102] B. E. Launder and B. I. Sharma, "A Modified Form of $k-\epsilon$ Model for Predicting Wall Turbulence", Letters in Heat and Mass Transfer, Vol. 1, pp. 131-138, 1974.

[103] W. C. Reynolds, "Computation of Turbulent Flows", Annual Review of Fluid Mechanics, pp. 183-208, 1976.

[104] D. C. Wilcox and W. M. Rubesin, "Progress in Turbulence Modeling for Complex Flow Field Including Effects of Compressibility", NASA Tech. Paper 1517, 1980

[105] S. W. Kim, "Calculation of Reattaching Shear Layers in Divergent Channel with a Multi-Time-Scale Turbulence Model", NASA TM 102293, ICOMP-89-18, 1989

[106] Joe F. Thompson, Z. U. A. Warsi and C. Wayne Mastin, "Numerical Grid Generation, — fundations and applications", North-Holland, New York, 1985.

[107] Torgny Fornstedt and Gorge Guiochon, "Comparison between Experimental and Ttheoretical Profiles of High Concentration Elution Bands and Large System Peaks in Nonlinear chromatography", Analytical Chemistry, Vol. 66, No. 17, September, 1994.

APPENDICES

Appendix A

The Derivation of Aerodynamic Navier-Stokes System

A.1 Navier-Stokes Equations

The laws governing the motion of fluid are the conservation of mass, momentum and energy. Under the condition that body force and radiation are neglectable, these conservation laws can be described as below ([8, 9]).

$$\frac{D}{Dt}\int_{\Omega(t)} \rho \, d\Omega = 0 \tag{A.1}$$

$$\frac{D}{Dt}\int_{\Omega(t)} m_i \, d\Omega = \oint_{\partial\Omega(t)} \sigma_{ji} n_j \, d\sigma \tag{A.2}$$

$$\frac{D}{Dt}\int_{\Omega(t)} E \, d\Omega = \oint_{\partial\Omega(t)} \left(\sigma_{ji} v_i + \lambda \frac{\partial T}{\partial x_j}\right) n_j \, d\sigma \tag{A.3}$$

where $\frac{D}{Dt}$ is the Lagrangian time differentiator following the motion of fluid *system*; $\Omega(t)$ is the space taken by the fluid *system*, with n_j being the outward unit vector

$\partial\Omega(t)$; ρ is the density, and T the temperature; v_i the ith component of the velocity, and $m_i = \rho v_i$ the momentum; σ_{ij} and λ are respectively the stress tensor and thermal conductivity; E is the total energy per unit volume,

$$E = \rho \left(e + \frac{1}{2} v_i^2 \right) \tag{A.4}$$

in which e is the internal energy per unit mass.

The stress tensor σ_{ij} can be partitioned into

$$\sigma_{ij} = -p\delta_{ij} + \tau_{ij} \tag{A.5}$$

in which τ_{ij} is the shear stress, and p the pressure, i.e., the first invariant of σ_{ij}:

$$p = -\frac{1}{3} \sigma_{ii} \tag{A.6}$$

For Newtonian fluid, which is the case of most aerodynamic flows, the shear stress τ_{ij} can further be linearly related to the deformation of the velocity field,

$$\tau_{ij} = \nu \left[\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) - \frac{2}{3} \frac{\partial v_j}{\partial x_j} \delta_{ij} \right] \tag{A.7}$$

where ν is the viscosity ([8]).

□

Using the Reynolds transportation formula (RTF, [9]) —

$$\frac{D}{Dt} \int_{\Omega(t)} q \, d\Omega = \int_{\Omega} \frac{\partial q}{\partial t} \, d\Omega + \oint_{\partial\Omega(t)} q v_j n_j \, d\sigma \tag{A.8}$$

where q is a transported variable in the flow-field, we can convert the Lagrangian described conservation laws (A.1-A.2-A.3) from the fluid *system* onto the *control volume* (by letting $q = \rho$, m_i, and E respectively):

$$\int_\Omega \frac{\partial \rho}{\partial t}\, d\Omega + \oint_{\partial\Omega} \rho\, v_j\, n_j\, d\sigma = 0 \qquad \text{(A.9)}$$

$$\int_\Omega \frac{\partial m_i}{\partial t}\, d\Omega + \oint_{\partial\Omega} (m_i v_j + p\delta_{ji} - \tau_{ji})\, n_j\, d\sigma = 0 \qquad \text{(A.10)}$$

$$\int_\Omega \frac{\partial E}{\partial t}\, d\Omega + \oint_{\partial\Omega} \left[(E+p)\, v_j - \tau_{ji} v_i - \lambda \frac{\partial T}{\partial x_j}\right] n_j\, d\sigma = 0 \qquad \text{(A.11)}$$

Here, the boundary fluxes, according to the Gauss formula ([9]), account for the volumetric collection of their divergences, in other words,

$$\int_\Omega \left(\frac{\partial \rho}{\partial t} + \frac{\partial \rho v_j}{\partial x_j}\right) d\Omega = 0 \qquad \text{(A.12)}$$

$$\int_\Omega \left(\frac{\partial m_i}{\partial t} + \frac{\partial m_i v_j}{\partial x_j} + \frac{\partial p}{\partial x_i} - \frac{\partial \tau_{ji}}{\partial x_j}\right) d\Omega = 0 \qquad \text{(A.13)}$$

$$\int_\Omega \left(\frac{\partial E}{\partial t} + \frac{\partial (E+p) v_j}{\partial x_j} - \frac{\partial \tau_{ji} v_i}{\partial x_j} - \frac{\partial}{\partial x_j}\lambda \frac{\partial T}{\partial x_j}\right) d\Omega = 0 \qquad \text{(A.14)}$$

Consequently, because of the arbitrariness of the *control volume*, there must be

$$L(\rho): \quad \frac{\partial \rho}{\partial t} + \frac{\partial \rho v_j}{\partial x_j} = 0 \qquad \text{(A.15)}$$

$$L(m_i): \quad \frac{\partial m_i}{\partial t} + \frac{\partial m_i v_j}{\partial x_j} + \frac{\partial p}{\partial x_i} - \frac{\partial \tau_{ji}}{\partial x_j} = 0 \qquad \text{(A.16)}$$

$$L(E): \quad \frac{\partial E}{\partial t} + \frac{\partial (E+p) v_j}{\partial x_j} - \frac{\partial \tau_{ji} v_i}{\partial x_j} - \frac{\partial}{\partial x_j}\lambda \frac{\partial T}{\partial x_j} = 0 \qquad \text{(A.17)}$$

and these are the Navier-Stokes equations for fluid dynamics.

Remarks:

- The proof that the "the arbitrariness of the control volume" leads to (A.12-A.13-A.14) in their form of (A.15-A.16-A.17) is similar to that of the Fundamental Theorem of Variational Principle ([6]): Suppose (A.15-A.16-A.17) is untrue,

say $L(E) > 0$ (or < 0) at a point $\mathbf{x}_0$ in the flow field. Then, since $L(E)$ is continuous, there is an $r, > 0$ such that $L(E)$ will be > 0 (or < 0) in the neighborhood sphere $S = \{\mathbf{x}, \mid \mathbf{x} - \mathbf{x}_0 \mid < r\}$, wherein, therefore, (A.12-A.13-A.14) are violated.

- The Gaussian flux formula applied in (A.9-A.10-A.11) $\Longrightarrow$ (A.12-A.13-A.14) is not unconditional, and the condition is that the fluxes are differentiable. Although such a condition is safely warranted in *real* flow, it is not in the *idealized/inviscid* flow, where, due to the absence of viscosity, the upmost *smoothness* that flow-field has to meet is what is required by the conservation laws themselves, i.e., the *integratability* as shown in (A.1-A.2-A.3) and (A.9-A.10-A.11), which permits even *discontinuity* (see Chapter 2).

A.2 Constitutive Equations

"Constutive equation" (CE) is a term referring to the *closure* conditions that are needed in order for the Navier-Stokes equations to be solution definite. Mathematically, the necessity of CE *closure* can be spotted by counting the Navier-Stokes system. In fact, if we do so, we immediately find that, for the Newtonian Navier-Stokes flow, *four* CEs should be *just* meeting such a necessity, since (A.15-A.16-A.17)$\cap$(A.7) contains $2 + dim$ equations ($L(\rho)$, $L(m_i)$ and $L(E)$) whereas $6 + dim$ unknowns/variables (ρ, p, T, e, ν, λ and m_i), — assuming that CEs induce no unknowns/variables. Fur-

thermore, we can also generally *figure* the format of CEs:

$$\nu = \nu(\cdot) \tag{A.18}$$

$$\lambda = \lambda(\cdot) \tag{A.19}$$

$$F_1(\rho, p, T, e) = 0 \tag{A.20}$$

$$F_2(\rho, p, T, e) = 0 \tag{A.21}$$

since, while it is apparent that ν and λ must be given in order for the Navier-Stokes equations to be solution definite, $\{\rho, p, T, e\}$ are the variables of state among which only two are independent.

□

What is considered filling-in the above format in this dissertation is the aerodynamic flow of (**i**) perfect gas, and (**ii**) constant heat capacities. As a result, (A.20-A.21) are substantiated as:

$$p = \rho R T \tag{A.22}$$

$$e = \frac{R T}{\gamma - 1} \tag{A.23}$$

where R is the gas constant, γ the heat-capacity ratio, (A.22) the equation of state for perect gas, and (A.23) the internal energy for perfect gas under the condition of constant heat capacities. Identical to (A.23),

$$h = \frac{\gamma R T}{\gamma - 1} \tag{A.24}$$

where h is the enthalpy ([9]).

The *perfect gas* assumption also provides models ν and λ ([11, 17, 20]). However, it is usually reasonable to permit ν and λ not to behave in full accordance with the molecular model of gas ([8, 9]), since usually the molecularly caused variation of ν and λ is far less important than their *typical* values (which decide Reynolds and Prandtl numbers, see Section 2.1) in determining the characteristics of flow.

Remarks:

1. Constitutive equation (CE) is a *relative* concept, in the sense that the contents of the CE system depends on the *ground* where we "constitute" the *closure* for our conservation laws. For example, if our *ground* is (A.15-A.16-A.17) instead of (A.15-A.16-A.17)$\cap$(A.7), then the Newton stress formula (A.7) should be regarded as part of the CEs. Another example is turbulence, where, if we apply (A.15-A.16-A.17)$\cap$(A.7) to the mean-flow, models for "eddy" viscosity and heat-conductivity ([94, 97, 98, 99, 100, 101, 102, 103, 104, 105]) will then have to be included accounting for the turbulent fluctuations in addtion to the molecular ones.

2. Perfect gas does not by default possess constant heat-capacities, although it is often applied so. In fact, the theoretical formula of the heat-capacity ratio for perfect gas is

$$\gamma = \frac{2+m}{m} \tag{A.25}$$

where m is the DOF of the gas molecule. Consequently, constant heat capacities occur or are good approximations only under these circumstances:

- Single-atom-molecule gas. Gas molecules in this case are spherical particles, of

which, thus, $m \equiv 3$ (for the 3-D transplation), and/hence, $\gamma \equiv 1.66667$.

- Multi-atom-molecule gas at *low* temperature — so *low* that the automic vibration within the gas molecules can be neglected. Gas molecules in this case move as if they are rigid bodies, of which, thus, $m \equiv 3 + 2$ (with 2 accounting for the two independent components of rotation), and/hence, $\gamma \equiv 1.4$.

- Multi-atom-molecule gas at *high* temperature — so *high* that the atomic vibration are totally activated. In this case, the DOF of a gas molecule consist of DOF of its as a rigid body, i.e., 5, plus one for each atomic vibration within. For example, two-atom-molecular gas at high temperature has $m \equiv 5 + 1$, since there is only one atomic vibration DOF, and/hence, $\gamma \equiv 1.33333$.

On the other hand, multi-atom-molecule gases at *intermediate* temperature, which is the case of flows in jet-engine, endure unneglectable yet not fully activated atomic vibration. As a result, their heat-capacities are temperature dependent.

Appendix B

Metric Data Mean Data Base

What is listed in this appendix are the tensor-product factorized data base of the MDM algorithm, which is as defined in Chapter 3, e.g.,

$$A(\varsigma^q)3101_{\alpha\gamma\beta} = \int_{\Omega^*} \varsigma^q \frac{dN_\alpha(\varsigma)}{d\varsigma} N_\gamma(\varsigma) \frac{dN_\beta(\varsigma)}{d\varsigma} d\varsigma \tag{B.1}$$

$$\overline{\varsigma 3101^q}_{ijk} = \frac{A(\varsigma^q)3101_{ijk}}{A3101_{\underline{ijk}}} \tag{B.2}$$

The data base is arrayed in the *column-first* manner, i.e., by the FORTRAN rule. For instance, the data in $[A3000_{ikm}]$ for an element with two nodes (which is the case of linear element) are meant to be

$$[A3000_{ikm}] = \begin{bmatrix} a_{111} & a_{121} & a_{112} & a_{122} \\ a_{211} & a_{221} & a_{212} & a_{222} \end{bmatrix} \tag{B.3}$$

B.1 Quasi-linear Element

Shape Functions

$$\{N_i(\varsigma)\} = \frac{1}{2}\begin{Bmatrix} (1-\varsigma) \\ (1+\varsigma) \end{Bmatrix} \tag{B.4}$$

Data for General MDM Procedure

$$\{A10_i\} = \{\ 1\ \ 1\ \}^T \tag{B.5}$$

$$\{A(\varsigma)10_i\} = \frac{1}{3}\{\ -1\ \ 1\ \}^T \tag{B.6}$$

$$\{A(\varsigma^2)10_i\} = \frac{1}{3}\{\ 1\ \ 1\ \}^T \tag{B.7}$$

$$[A200_{im}] = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{B.8}$$

$$[A(\varsigma)200_{im}] = \frac{1}{3}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \tag{B.9}$$

$$\left[A(\varsigma^2)200_{im}\right] = \frac{1}{15}\begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} \tag{B.10}$$

$$[A210_{im}] = \frac{1}{2}\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \tag{B.11}$$

$$[A(\varsigma)210_{im}] = \frac{1}{6}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{B.12}$$

$$\left[A(\varsigma^2)210_{im}\right] = \frac{1}{6}\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \tag{B.13}$$

$$[A201_{im}] = \frac{1}{2}\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \tag{B.14}$$

$$[A(\varsigma)201_{im}] = \frac{1}{6}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{B.15}$$

$$\left[A(\varsigma^2)201_{im}\right] = \frac{1}{6}\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \tag{B.16}$$

$$[A211_{im}] = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{B.17}$$

$$[A(\varsigma)211_{im}] = [0] \tag{B.18}$$

$$\left[A(\varsigma^2)211_{im}\right] = \frac{1}{6}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{B.19}$$

$$[A3000_{ikm}] = \frac{1}{6}\begin{bmatrix} 3 & 1 & 1 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix} \tag{B.20}$$

$$[A(\varsigma)3000_{ikm}] = \frac{1}{30}\begin{bmatrix} -9 & -1 & -1 & 1 \\ -1 & 1 & 1 & 9 \end{bmatrix} \tag{B.21}$$

$$\left[A(\varsigma^2)3000_{ikm}\right] = \frac{1}{30}\begin{bmatrix} 7 & 1 & 1 & 1 \\ 1 & 1 & 1 & 7 \end{bmatrix} \tag{B.22}$$

$$[A3100_{ikm}] = \frac{1}{6}\begin{bmatrix} -2 & -1 & -1 & -2 \\ 2 & 1 & 1 & 2 \end{bmatrix} \tag{B.23}$$

$$[A(\varsigma)3100_{ikm}] = \frac{1}{6}\begin{bmatrix} 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \tag{B.24}$$

$$\left[A(\varsigma^2)3100_{ikm}\right] = \frac{1}{30}\begin{bmatrix} -4 & -1 & -1 & -4 \\ 4 & 1 & 1 & 4 \end{bmatrix} \tag{B.25}$$

$$[A3010_{ikm}] = \frac{1}{6}\begin{bmatrix} -2 & 2 & -1 & 1 \\ -1 & 1 & -2 & 2 \end{bmatrix} \tag{B.26}$$

$$[A(\varsigma)3010_{ikm}] = \frac{1}{6}\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \tag{B.27}$$

$$\left[A(\varsigma^2)3010_{ikm}\right] = \frac{1}{30}\begin{bmatrix} -4 & 4 & -1 & 1 \\ -1 & 1 & -4 & 4 \end{bmatrix} \tag{B.28}$$

$$[A3001_{ikm}] = \frac{1}{6}\begin{bmatrix} -2 & -1 & 2 & 1 \\ -1 & -2 & 1 & 2 \end{bmatrix} \tag{B.29}$$

$$[A(\varsigma)3001_{ikm}] = \frac{1}{6}\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \tag{B.30}$$

$$\left[A(\varsigma^2)3001_{ikm}\right] = \frac{1}{30}\begin{bmatrix} -4 & -1 & 4 & 1 \\ -1 & -4 & 1 & 4 \end{bmatrix} \tag{B.31}$$

$$[A3110_{ikm}] = \frac{1}{4}\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix} \tag{B.32}$$

$$[A(\varsigma)3110_{ikm}] = \frac{1}{12}\begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{B.33}$$

$$\left[A(\varsigma^2)3110_{ikm}\right] = \frac{1}{12}\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix} \tag{B.34}$$

$$[A3101_{ikm}] = \frac{1}{4}\begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \tag{B.35}$$

$$[A(\varsigma)3101_{ikm}] = \frac{1}{12}\begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{B.36}$$

$$\left[A(\varsigma^2)3101_{ikm}\right] = \frac{1}{12}\begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \tag{B.37}$$

$$[A3011_{ikm}] = \frac{1}{4}\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{B.38}$$

$$[A(\varsigma)3011_{ikm}] = \frac{1}{12}\begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{B.39}$$

$$\left[A(\varsigma^2)3011_{ikm}\right] = \frac{1}{12}\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{B.40}$$

$$[A3111_{ikm}] = \frac{1}{4}\begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{B.41}$$

$$[A(\varsigma)3111_{ikm}] = [0] \tag{B.42}$$

$$\left[A(\varsigma^2)3111_{ikm}\right] = \frac{1}{12}\begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{B.43}$$

Data for Mean Form MDM Procedure

$$\{\overline{\varsigma 10_i}\} = \frac{1}{3}\{\ -1\ \ 1\ \}^T \tag{B.44}$$

$$\{\overline{\varsigma 10^2{}_i}\} = \frac{1}{3}\{\ 1\ \ 1\ \}^T \tag{B.45}$$

$$\left[\overline{\varsigma 200}_{im}\right] = \frac{1}{2}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \tag{B.46}$$

$$\left[\overline{\varsigma 200^2}_{im}\right] = \frac{1}{5}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{B.47}$$

$$\left[\overline{\varsigma 210}_{im}\right] = \frac{1}{3}\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \tag{B.48}$$

$$\left[\overline{\varsigma 210^2}_{im}\right] = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{B.49}$$

$$\left[\overline{\varsigma 201}_{im}\right] = \frac{1}{3}\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \tag{B.50}$$

$$\left[\overline{\varsigma 201^2}_{im}\right] = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{B.51}$$

$$\left[\overline{\varsigma 211}_{im}\right] = [0] \tag{B.52}$$

$$\left[\overline{\varsigma 211^2}_{im}\right] = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{B.53}$$

$$\left[\overline{\varsigma 3000}_{ikm}\right] = \frac{1}{5}\begin{bmatrix} -3 & -1 & -1 & 1 \\ -1 & 1 & 1 & 3 \end{bmatrix} \tag{B.54}$$

$$\left[\overline{\varsigma 3000^2}_{ikm}\right] = \frac{1}{15}\begin{bmatrix} 7 & 3 & 3 & 3 \\ 3 & 3 & 3 & 7 \end{bmatrix} \tag{B.55}$$

$$\left[\overline{\varsigma 3100}_{ikm}\right] = \frac{1}{2}\begin{bmatrix} -1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \tag{B.56}$$

$$\left[\overline{\varsigma 3100^2}_{ikm}\right] = \frac{1}{5}\begin{bmatrix} 2 & 1 & 1 & 2 \\ 2 & 1 & 1 & 2 \end{bmatrix} \tag{B.57}$$

$$\left[\overline{\varsigma 3010}_{ikm}\right] = \frac{1}{2}\begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \tag{B.58}$$

$$\left[\overline{\varsigma 3010^2}_{ikm}\right] = \frac{1}{5}\begin{bmatrix} 2 & 2 & 1 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix} \tag{B.59}$$

$$\left[\overline{\varsigma 3001}_{ikm}\right] = \frac{1}{2}\begin{bmatrix} -1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \tag{B.60}$$

$$\left[\overline{\varsigma 3001^2}_{ikm}\right] = \frac{1}{5}\begin{bmatrix} 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \end{bmatrix} \tag{B.61}$$

$$\left[\overline{\varsigma 3110}_{ikm}\right] = \frac{1}{3}\begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \tag{B.62}$$

$$\left[\overline{\varsigma 3110^2}_{ikm}\right] = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{B.63}$$

$$\left[\overline{\varsigma 3101}_{ikm}\right] = \frac{1}{3}\begin{bmatrix} -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix} \tag{B.64}$$

$$\left[\overline{\varsigma 3101^2}_{ikm}\right] = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{B.65}$$

$$\left[\overline{\varsigma 3011}_{ikm}\right] = \frac{1}{3}\begin{bmatrix} -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{B.66}$$

$$\left[\overline{\varsigma 3011^2}_{ikm}\right] = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{B.67}$$

$$\left[\overline{\varsigma 3111}_{ikm}\right] = [0] \tag{B.68}$$

$$\left[\overline{\varsigma 3111^2}_{ikm}\right] = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{B.69}$$

B.2 Quasi-quadratic Element

Shape Functions

$$\{N_i(\varsigma)\} = \begin{Bmatrix} \frac{1}{2}\varsigma(\varsigma-1) \\ 1-\varsigma^2 \\ \frac{1}{2}\varsigma(\varsigma+1) \end{Bmatrix} \tag{B.70}$$

Data for General MDM Procedure

$$\{A10_i\} = \frac{1}{3}\{\ 1\ \ 4\ \ 1\ \}^T \tag{B.71}$$

$$\{A(\varsigma)10_i\} = \frac{1}{3}\{\ -1\ \ 0\ \ 1\ \}^T \tag{B.72}$$

$$\{A(\varsigma^2)10_i\} = \frac{1}{15}\{\ 3\ \ 4\ \ 3\ \}^T \tag{B.73}$$

$$[A200_{im}] = \frac{1}{15}\begin{bmatrix} 4 & 2 & -1 \\ 2 & 16 & 2 \\ -1 & 2 & 4 \end{bmatrix} \tag{B.74}$$

$$[A(\varsigma)200_{im}] = \frac{1}{15}\begin{bmatrix} -3 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & 3 \end{bmatrix} \tag{B.75}$$

$$\left[A(\varsigma^2)200_{im}\right] = \frac{1}{105}\begin{bmatrix} 18 & 6 & -3 \\ 6 & 16 & 6 \\ -3 & 6 & 18 \end{bmatrix} \tag{B.76}$$

$$[A210_{im}] = \frac{1}{6}\begin{bmatrix} -3 & -4 & 1 \\ 4 & 0 & -4 \\ -1 & 4 & 3 \end{bmatrix} \tag{B.77}$$

$$[A(\varsigma)210_{im}] = \frac{1}{30}\begin{bmatrix} 11 & 8 & 1 \\ -12 & -16 & -12 \\ 1 & 8 & 11 \end{bmatrix} \tag{B.78}$$

$$\left[A(\varsigma^2)210_{im}\right] = \frac{1}{30}\begin{bmatrix} -9 & -4 & 3 \\ 12 & 0 & -12 \\ -3 & 4 & 9 \end{bmatrix} \tag{B.79}$$

$$[A201_{im}] = \frac{1}{6}\begin{bmatrix} -3 & 4 & -1 \\ -4 & 0 & 4 \\ 1 & -4 & 3 \end{bmatrix} \tag{B.80}$$

$$[A(\varsigma)201_{im}] = \frac{1}{30}\begin{bmatrix} 11 & -12 & 1 \\ 8 & -16 & 8 \\ 1 & -12 & 11 \end{bmatrix} \tag{B.81}$$

$$\left[A(\varsigma^2)201_{im}\right] = \frac{1}{30}\begin{bmatrix} -9 & 12 & -3 \\ -4 & 0 & 4 \\ 3 & -12 & 9 \end{bmatrix} \tag{B.82}$$

$$[A211_{im}] = \frac{1}{6}\begin{bmatrix} 7 & -8 & 1 \\ -8 & 16 & -8 \\ 1 & -8 & 7 \end{bmatrix} \tag{B.83}$$

$$[A(\varsigma)211_{im}] = \frac{1}{3}\begin{bmatrix} -2 & 2 & 0 \\ 2 & 0 & -2 \\ 0 & -2 & 2 \end{bmatrix} \tag{B.84}$$

$$\left[A(\varsigma^2)211_{im}\right] = \frac{1}{30}\begin{bmatrix} 17 & -24 & 7 \\ -24 & 48 & -24 \\ 7 & -24 & 17 \end{bmatrix} \tag{B.85}$$

$$[A3000_{ikm}]^T = \frac{1}{210}\begin{bmatrix} 39 & 20 & -3 \\ 20 & 16 & -8 \\ -3 & -8 & -3 \\ 20 & 16 & -8 \\ 16 & 192 & 16 \\ -8 & 16 & 20 \\ -3 & -8 & -3 \\ -8 & 16 & 20 \\ -3 & 20 & 39 \end{bmatrix} \tag{B.86}$$

$$[A(\varsigma)3000_{ikm}]^T = \frac{1}{210}\begin{bmatrix} -33 & -12 & 3 \\ -12 & -16 & 0 \\ 3 & 0 & -3 \\ -12 & -16 & 0 \\ -16 & 0 & 16 \\ 0 & 16 & 12 \\ 3 & 0 & -3 \\ 0 & 16 & 12 \\ -3 & 12 & 33 \end{bmatrix} \tag{B.87}$$

$$\left[A(\varsigma^2)3000_{ikm}\right]^T = \frac{1}{630}\begin{bmatrix} 85 & 28 & -5 \\ 28 & 16 & -8 \\ -5 & -8 & -5 \\ 28 & 16 & -8 \\ 16 & 64 & 16 \\ -8 & 16 & 28 \\ -5 & -8 & -5 \\ -8 & 16 & 28 \\ -5 & 28 & 85 \end{bmatrix} \tag{B.88}$$

$$[A3100_{ikm}]^T = \frac{1}{3\ 0}\begin{bmatrix} -10 & 12 & -2 \\ -6 & 8 & -2 \\ 1 & 0 & -1 \\ -6 & 8 & -2 \\ -16 & 0 & 16 \\ 2 & -8 & 6 \\ 1 & 0 & -1 \\ 2 & -8 & 6 \\ 2 & -12 & 10 \end{bmatrix} \tag{B.89}$$

$$[A(\varsigma)3100_{ikm}]^T = \frac{1}{210}\begin{bmatrix} 57 & -72 & 15 \\ 26 & -24 & -2 \\ -6 & 12 & -6 \\ 26 & -24 & -2 \\ 32 & -64 & 32 \\ -2 & -24 & 26 \\ -6 & 12 & -6 \\ -2 & -24 & 26 \\ 15 & -72 & 57 \end{bmatrix} \tag{B.90}$$

$$\left[A(\varsigma^2)3100_{ikm}\right]^T = \frac{1}{210}\begin{bmatrix} -48 & 60 & -12 \\ -18 & 24 & -6 \\ 3 & 0 & -3 \\ -18 & 24 & -6 \\ -16 & 0 & 16 \\ 6 & -24 & 18 \\ 3 & 0 & -3 \\ 6 & -24 & 18 \\ 12 & -60 & 48 \end{bmatrix} \tag{B.91}$$

$$[A3010_{ikm}]^T = \frac{1}{3\,0}\begin{bmatrix} -10 & -6 & 1 \\ 12 & 8 & 0 \\ -2 & -2 & -1 \\ -6 & -16 & 2 \\ 8 & 0 & -8 \\ -2 & 16 & 6 \\ 1 & 2 & 2 \\ 0 & -8 & -12 \\ -1 & 6 & 10 \end{bmatrix} \tag{B.92}$$

$$[A(\varsigma)3010_{ikm}]^T = \frac{1}{210}\begin{bmatrix} 57 & 26 & -6 \\ -72 & -24 & 12 \\ 15 & -2 & -6 \\ 26 & 32 & -2 \\ -24 & -64 & -24 \\ -2 & 32 & 26 \\ -6 & -2 & 15 \\ 12 & -24 & -72 \\ -6 & 26 & 57 \end{bmatrix} \tag{B.93}$$

$$\left[A(\varsigma^2)3010_{ikm}\right]^T = \frac{1}{210}\begin{bmatrix} -48 & -18 & 3 \\ 60 & 24 & 0 \\ -12 & -6 & -3 \\ -18 & -16 & 6 \\ 24 & 0 & -24 \\ -6 & 16 & 18 \\ 3 & 6 & 12 \\ 0 & -24 & -60 \\ -3 & 18 & 48 \end{bmatrix} \tag{B.94}$$

$$[A3001_{ikm}]^T = \frac{1}{3\,0}\begin{bmatrix} -10 & -6 & 1 \\ -6 & -16 & 2 \\ 1 & 2 & 2 \\ 12 & 8 & 0 \\ 8 & 0 & -8 \\ 0 & -8 & -12 \\ -2 & -2 & -1 \\ -2 & 16 & 6 \\ -1 & 6 & 10 \end{bmatrix} \tag{B.95}$$

$$[A(\varsigma)3001_{ikm}]^T = \frac{1}{210}\begin{bmatrix} 57 & 26 & -6 \\ 26 & 32 & -2 \\ -6 & -2 & 15 \\ -72 & -24 & 12 \\ -24 & -64 & -24 \\ 12 & -24 & -72 \\ 15 & -2 & -6 \\ -2 & 32 & 26 \\ -6 & 26 & 57 \end{bmatrix} \tag{B.96}$$

$$\left[A(\varsigma^2)3001_{ikm}\right]^T = \frac{1}{210}\begin{bmatrix} -48 & -18 & 3 \\ -18 & -16 & 6 \\ 3 & 6 & 12 \\ 60 & 24 & 0 \\ 24 & 0 & -24 \\ 0 & -24 & -60 \\ -12 & -6 & -3 \\ -6 & 16 & 18 \\ -3 & 18 & 48 \end{bmatrix} \tag{B.97}$$

$$[A3110_{ikm}]^T = \frac{1}{60}\begin{bmatrix} 37 & -44 & 7 \\ -44 & 48 & -4 \\ 7 & -4 & -3 \\ 36 & -32 & -4 \\ -32 & 64 & -32 \\ -4 & -32 & 36 \\ -3 & -4 & 7 \\ -4 & 48 & -44 \\ 7 & -44 & 37 \end{bmatrix} \tag{B.98}$$

$$[A(\varsigma)3110_{ikm}]^T = \frac{1}{60}\begin{bmatrix} -29 & 36 & -7 \\ 36 & -48 & 12 \\ -7 & 12 & -5 \\ -16 & 16 & 0 \\ 16 & 0 & -16 \\ 0 & -16 & 16 \\ 5 & -12 & 7 \\ -12 & 48 & -36 \\ 7 & -36 & 29 \end{bmatrix} \tag{B.99}$$

$$\left[A(\varsigma^2)3110_{ikm}\right]^T = \frac{1}{420}\begin{bmatrix} 165 & -204 & 39 \\ -204 & 240 & -36 \\ 39 & -36 & -3 \\ 76 & -96 & 20 \\ -96 & 192 & -96 \\ 20 & -96 & 76 \\ -3 & -36 & 39 \\ -36 & 240 & -204 \\ 39 & -204 & 165 \end{bmatrix} \tag{B.100}$$

$$[A3101_{ikm}]^T = \frac{1}{6\ 0}\begin{bmatrix} 37 & -44 & 7 \\ 36 & -32 & -4 \\ -3 & -4 & 7 \\ -44 & 48 & -4 \\ -32 & 64 & -32 \\ -4 & 48 & -44 \\ 7 & -4 & -3 \\ -4 & -32 & 36 \\ 7 & -44 & 37 \end{bmatrix} \tag{B.101}$$

$$[A(\varsigma)3101_{ikm}]^T = \frac{1}{6\,0}\begin{bmatrix} -29 & 36 & -7 \\ -16 & 16 & 0 \\ 5 & -12 & 7 \\ 36 & -48 & 12 \\ 16 & 0 & -16 \\ -12 & 48 & -36 \\ -7 & 12 & -5 \\ 0 & -16 & 16 \\ 7 & -36 & 29 \end{bmatrix} \tag{B.102}$$

$$\left[A(\varsigma^2)3101_{ikm}\right]^T = \frac{1}{420}\begin{bmatrix} 165 & -204 & 39 \\ 76 & -96 & 20 \\ -3 & -36 & 39 \\ -204 & 240 & -36 \\ -96 & 192 & -96 \\ -36 & 240 & 204 \\ 39 & -36 & -3 \\ 20 & -96 & 76 \\ 39 & -204 & 165 \end{bmatrix} \tag{B.103}$$

$$[A3011_{ikm}]^T = \frac{1}{60}\begin{bmatrix} 37 & 36 & -3 \\ -44 & -32 & -4 \\ 7 & -4 & 7 \\ -44 & -32 & -4 \\ 48 & 64 & 48 \\ -4 & -32 & -44 \\ 7 & -4 & 7 \\ -4 & -32 & -44 \\ -3 & 36 & 37 \end{bmatrix} \tag{B.104}$$

$$[A(\varsigma)3011_{ikm}]^T = \frac{1}{60}\begin{bmatrix} -29 & -16 & 5 \\ 36 & 16 & -12 \\ -7 & 0 & 7 \\ 36 & 16 & -12 \\ -48 & 0 & 48 \\ 12 & -16 & -36 \\ -7 & 0 & 7 \\ 12 & -16 & -36 \\ -5 & 16 & 29 \end{bmatrix} \tag{B.105}$$

$$\left[A(\varsigma^2)3011_{ikm}\right]^T = \frac{1}{420}\begin{bmatrix} 165 & 76 & -3 \\ -204 & -96 & -36 \\ 39 & 20 & 39 \\ -204 & -96 & -36 \\ 240 & 192 & 240 \\ -36 & -96 & -204 \\ 39 & 20 & 39 \\ -36 & -96 & -204 \\ -3 & 76 & 165 \end{bmatrix} \tag{B.106}$$

$$[A3111_{ikm}]^T = \frac{1}{1\,2}\begin{bmatrix} -15 & 16 & -1 \\ 16 & -16 & 0 \\ -1 & 0 & 1 \\ 16 & -16 & 0 \\ -16 & 0 & 16 \\ 0 & 16 & -16 \\ -1 & 0 & 1 \\ 0 & 16 & -16 \\ 1 & -16 & 15 \end{bmatrix} \tag{B.107}$$

$$[A(\varsigma)3111_{ikm}]^T = \frac{1}{3\,0}\begin{bmatrix} 27 & -34 & 7 \\ -34 & 48 & -14 \\ 7 & -14 & 7 \\ -34 & 48 & -14 \\ 48 & -96 & 48 \\ -14 & 48 & -34 \\ 7 & -14 & 7 \\ -14 & 48 & -34 \\ 7 & -34 & 27 \end{bmatrix} \tag{B.108}$$

$$\left[A(\varsigma^2)3111_{ikm}\right]^T = \frac{1}{6\,0}\begin{bmatrix} -41 & 48 & -7 \\ 48 & -48 & 0 \\ -7 & 0 & 7 \\ 48 & -48 & 0 \\ -48 & 0 & 48 \\ 0 & 48 & -48 \\ -7 & 0 & 7 \\ 0 & 48 & -48 \\ 7 & -48 & 41 \end{bmatrix} \tag{B.109}$$

B.3 Quasi-cubic Element

Shape Functions

$$\{N_i(\varsigma)\} = \left\{ \begin{array}{c} -\frac{1}{16}(3\varsigma+1)(3\varsigma-1)(\varsigma-1) \\ \frac{9}{16}(\varsigma+1)(3\varsigma-1)(\varsigma-1) \\ -\frac{9}{16}(\varsigma+1)(3\varsigma+1)(\varsigma-1) \\ \frac{1}{16}(\varsigma+1)(3\varsigma+1)(3\varsigma-1) \end{array} \right\} \tag{B.110}$$

Data for General MDM Procedure

$$\{A10_i\} = \frac{1}{4}\{\ 1 \quad 3 \quad 3 \quad 1\ \}^T \tag{B.111}$$

$$\{A(\varsigma)10_i\} = \frac{1}{60}\{\ -11 \quad -27 \quad 27 \quad 11\ \}^T \tag{B.112}$$

$$\{A(\varsigma^2)10_i\} = \frac{1}{60}\{\ 11 \quad 9 \quad 9 \quad 11\ \}^T \tag{B.113}$$

$$[A200_{im}] = \frac{1}{840}\begin{bmatrix} 128 & 99 & -36 & 19 \\ 99 & 648 & -81 & -36 \\ -36 & -81 & 648 & 99 \\ 19 & -36 & 99 & 128 \end{bmatrix} \tag{B.114}$$

$$[A(\varsigma)200_{im}] = \frac{1}{840}\begin{bmatrix} -109 & -90 & 45 & 0 \\ -90 & -243 & 0 & -45 \\ 45 & 0 & 243 & 90 \\ 0 & -45 & 90 & 109 \end{bmatrix} \tag{B.115}$$

$$\left[A(\varsigma^2)200_{im}\right] = \frac{1}{840}\begin{bmatrix} 98 & 63 & -18 & 11 \\ 63 & 162 & -81 & -18 \\ -18 & -81 & 162 & 63 \\ 11 & -18 & 63 & 98 \end{bmatrix} \tag{B.116}$$

$$[A210_{im}] = \frac{1}{80}\begin{bmatrix} -40 & -57 & 24 & -7 \\ 57 & 0 & -81 & 24 \\ -24 & 81 & 0 & -57 \\ 7 & -24 & 57 & 40 \end{bmatrix} \tag{B.117}$$

$$[A(\varsigma)210_{im}] = \frac{1}{1680}\begin{bmatrix} 712 & 783 & -216 & -19 \\ -981 & -648 & 81 & 288 \\ 288 & 81 & -648 & -981 \\ -19 & -216 & 783 & 712 \end{bmatrix} \tag{B.118}$$

$$\left[A(\varsigma^2)210_{im}\right] = \frac{1}{1680}\begin{bmatrix} -622 & -531 & 198 & -109 \\ 891 & 486 & -243 & 378 \\ -378 & 243 & -486 & -891 \\ 109 & -198 & 531 & 622 \end{bmatrix} \tag{B.119}$$

$$[A201_{im}] = \frac{1}{80}\begin{bmatrix} -40 & 57 & -24 & 7 \\ -57 & 0 & 81 & -24 \\ 24 & -81 & 0 & 57 \\ -7 & 24 & -57 & 40 \end{bmatrix} \tag{B.120}$$

$$[A(\varsigma)201_{im}] = \frac{1}{1680}\begin{bmatrix} 712 & -981 & 288 & -19 \\ 783 & -648 & 81 & -216 \\ -216 & 81 & -648 & 783 \\ -19 & 288 & -981 & 712 \end{bmatrix} \tag{B.121}$$

$$\left[A(\varsigma^2)201_{im}\right] = \frac{1}{1680}\begin{bmatrix} -622 & 891 & -378 & 109 \\ -531 & 486 & 243 & -198 \\ 198 & -243 & -486 & 531 \\ -109 & 378 & -891 & 622 \end{bmatrix} \tag{B.122}$$

$$[A211_{im}] = \frac{1}{80}\begin{bmatrix} 148 & -189 & 54 & -13 \\ -189 & 432 & -297 & 54 \\ 54 & -297 & 432 & -189 \\ -13 & 54 & -189 & 148 \end{bmatrix} \tag{B.123}$$

$$[A(\varsigma)211_{im}] = \frac{1}{40}\begin{bmatrix} -57 & 69 & -12 & 0 \\ 69 & -81 & 0 & 12 \\ -12 & 0 & 81 & -69 \\ 0 & 12 & -69 & 57 \end{bmatrix} \tag{B.124}$$

$$\left[A(\varsigma^2)211_{im}\right] = \frac{1}{1680}\begin{bmatrix} 2080 & -2943 & 1242 & -379 \\ -2943 & 4860 & -3159 & 1242 \\ 1242 & -3159 & 4860 & -2943 \\ -379 & 1242 & -2943 & 2080 \end{bmatrix} \tag{B.125}$$

$$[A3000_{ikm}]^T = \frac{1}{3360}\begin{bmatrix} 357 & 216 & -81 & 20 \\ 216 & 324 & -162 & 18 \\ -81 & -162 & 81 & 18 \\ 20 & 18 & 18 & 20 \\ 216 & 324 & -162 & 18 \\ 324 & 2187 & 0 & 81 \\ -162 & 0 & 0 & -162 \\ 18 & 81 & -162 & -81 \\ -81 & -162 & 81 & 18 \\ -162 & 0 & 0 & -162 \\ 81 & 0 & 2187 & 324 \\ 18 & -162 & 324 & 216 \\ 20 & 18 & 18 & 20 \\ 18 & 81 & -162 & -81 \\ 18 & -162 & 324 & 216 \\ 20 & -81 & 216 & 357 \end{bmatrix} \quad \text{(B.126)}$$

$$[A(\varsigma)3000_{ikm}]^T = \frac{1}{36960}\begin{bmatrix} -3555 & -1908 & 819 & -152 \\ -1908 & -2592 & 810 & -270 \\ 819 & 810 & 81 & 270 \\ -152 & -270 & 270 & 152 \\ -1908 & -2592 & 810 & -270 \\ -2592 & -10935 & 2916 & -81 \\ 810 & 2916 & -2916 & -810 \\ -270 & -81 & -810 & -819 \\ 819 & 810 & 81 & 270 \\ 810 & 2916 & -2916 & -810 \\ 81 & -2916 & 10935 & 2592 \\ 270 & -810 & 2592 & 1908 \\ -152 & -270 & 270 & 152 \\ -270 & -81 & -810 & -819 \\ 270 & -810 & 2592 & 1908 \\ 152 & -819 & 1908 & 3555 \end{bmatrix} \quad \text{(B.127)}$$

$$\left[A(\varsigma^2)3000_{ikm}\right]^T = \frac{1}{36960}\begin{bmatrix} 3251 & 1548 & -639 & 152 \\ 1548 & 1944 & -810 & 90 \\ -639 & -810 & 567 & 90 \\ 152 & 90 & 90 & 152 \\ 1548 & 1944 & -810 & 90 \\ 1944 & 5589 & -972 & 567 \\ -810 & -972 & -972 & -810 \\ 90 & 567 & -810 & -639 \\ -639 & -810 & 567 & 90 \\ -810 & -972 & -972 & -810 \\ 567 & -972 & 5589 & 1944 \\ 90 & -810 & 1944 & 1548 \\ 152 & 90 & 90 & 152 \\ 90 & 567 & -810 & -639 \\ 90 & -810 & 1944 & 1548 \\ 152 & -639 & 1548 & 3251 \end{bmatrix} \quad \text{(B.128)}$$

$$[A3100_{ikm}]^T = \frac{1}{6720}\begin{bmatrix} -2240 & 3222 & -1260 & 278 \\ -1611 & 2025 & -405 & -9 \\ 630 & -648 & -162 & 180 \\ -139 & 189 & -189 & 139 \\ -1611 & 2025 & -405 & -9 \\ -4050 & 0 & 4374 & -324 \\ 1053 & -2187 & 2187 & -1053 \\ -180 & 162 & 648 & -630 \\ 630 & -648 & -162 & 180 \\ 1053 & -2187 & 2187 & -1053 \\ 324 & -4374 & 0 & 4050 \\ 9 & 405 & -2025 & 1611 \\ -139 & 189 & -189 & 139 \\ -180 & 162 & 648 & -630 \\ 9 & 405 & -2025 & 1611 \\ -278 & 1260 & -3222 & 2240 \end{bmatrix} \quad \text{(B.129)}$$

$$[A(\varsigma)3100_{ikm}]^T = \frac{1}{6720}\begin{bmatrix} 2002 & -2934 & 1170 & -238 \\ 1251 & -1539 & 405 & -117 \\ -504 & 648 & -324 & 180 \\ 99 & -99 & -99 & 99 \\ 1251 & -1539 & 405 & -117 \\ 2430 & -1458 & -1458 & 486 \\ -729 & 729 & 729 & -729 \\ 180 & -324 & 648 & -504 \\ -504 & 648 & -324 & 180 \\ -729 & 729 & 729 & -729 \\ 486 & -1458 & -1458 & 2430 \\ -117 & 405 & -1539 & 1251 \\ 99 & -99 & -99 & 99 \\ 180 & -324 & 648 & -504 \\ -117 & 405 & -1539 & 1251 \\ -238 & 1170 & -2934 & 2002 \end{bmatrix} \quad \text{(B.130)}$$

$$\left[A(\varsigma^2)3100_{ikm}\right]^T = \frac{1}{73920}\begin{bmatrix} -19900 & 29538 & -12312 & 2674 \\ -10953 & 13851 & -3159 & 261 \\ 4518 & -5832 & 486 & 828 \\ -1033 & 1647 & -1647 & 1033 \\ -10953 & 13851 & -3159 & 261 \\ -17334 & 14580 & 1458 & 1296 \\ 5751 & -6561 & 6561 & -5751 \\ -828 & -486 & 5832 & -4518 \\ 4518 & -5832 & 486 & 828 \\ 5751 & -6561 & 6561 & -5751 \\ -1296 & -1458 & -14580 & 17334 \\ -261 & 3159 & -13851 & 10953 \\ -1033 & 1647 & -1647 & 1033 \\ -828 & -486 & 5832 & -4518 \\ -261 & 3159 & -13851 & 10953 \\ -2674 & 12312 & -29538 & 19900 \end{bmatrix} \quad \text{(B.131)}$$

$$[A3010_{ikm}]^T = \frac{1}{6720}\begin{bmatrix} -2240 & -1611 & 630 & -139 \\ 3222 & 2025 & -648 & 189 \\ -1260 & -405 & -162 & -189 \\ 278 & -9 & 180 & 139 \\ -1611 & -4050 & 1053 & -180 \\ 2025 & 0 & -2187 & 162 \\ -405 & 4374 & 2187 & 648 \\ -9 & -324 & -1053 & -630 \\ 630 & 1053 & 324 & 9 \\ -648 & -2187 & -4374 & 405 \\ -162 & 2187 & 0 & -2025 \\ 180 & -1053 & 4050 & 1611 \\ -139 & -180 & 9 & -278 \\ 189 & 162 & 405 & 1260 \\ -189 & 648 & -2025 & -3222 \\ 139 & -630 & 1611 & 2240 \end{bmatrix} \quad \text{(B.132)}$$

$$[A(\varsigma)3010_{ikm}]^T = \frac{1}{6720}\begin{bmatrix} 2002 & 1251 & -504 & 99 \\ -2934 & -1539 & 648 & -99 \\ 1170 & 405 & -324 & -99 \\ -238 & -117 & 180 & 99 \\ 1251 & 2430 & -729 & 180 \\ -1539 & -1458 & 729 & -324 \\ 405 & -1458 & 729 & 648 \\ -117 & 486 & -729 & -504 \\ -504 & -729 & 486 & -117 \\ 648 & 729 & -1458 & 405 \\ -324 & 729 & -1458 & -1539 \\ 180 & -729 & 2430 & 1251 \\ 99 & 180 & -117 & -238 \\ -99 & -324 & 405 & 1170 \\ -99 & 648 & -1539 & -2934 \\ 99 & -504 & 1251 & 2002 \end{bmatrix} \quad \text{(B.133)}$$

$$\left[A(\varsigma^2)3010_{ikm}\right]^T = \frac{1}{73920}\begin{bmatrix} -19900 & -10953 & 4518 & -1033 \\ 29538 & 13851 & -5832 & 1647 \\ -12312 & -3159 & 486 & -1647 \\ 2674 & 261 & 828 & 1033 \\ -10953 & -17334 & 5751 & -828 \\ 13851 & 14580 & -6561 & -486 \\ -3159 & 1458 & 6561 & 5832 \\ 261 & 1296 & -5751 & -4518 \\ 4518 & 5751 & -1296 & -261 \\ -5832 & -6561 & -1458 & 3159 \\ 486 & 6561 & -14580 & -13851 \\ 828 & -5751 & 17334 & 10953 \\ -1033 & -828 & -261 & -2674 \\ 1647 & -486 & 3159 & 12312 \\ -1647 & 5832 & -13851 & -29538 \\ 1033 & -4518 & 10953 & 19900 \end{bmatrix} \quad \text{(B.134)}$$

$$[A3001_{ikm}]^T = \frac{1}{6720}\begin{bmatrix} -2240 & -1611 & 630 & -139 \\ -1611 & -4050 & 1053 & -180 \\ 630 & 1053 & 324 & 9 \\ -139 & -180 & 9 & -278 \\ 3222 & 2025 & -648 & 189 \\ 2025 & 0 & -2187 & 162 \\ -648 & -2187 & -4374 & 405 \\ 189 & 162 & 405 & 1260 \\ -1260 & -405 & -162 & -189 \\ -405 & 4374 & 2187 & 648 \\ -162 & 2187 & 0 & -2025 \\ -189 & 648 & -2025 & -3222 \\ 278 & -9 & 180 & 139 \\ -9 & -324 & -1053 & -630 \\ 180 & -1053 & 4050 & 1611 \\ 139 & -630 & 1611 & 2240 \end{bmatrix} \quad \text{(B.135)}$$

$$[A(\varsigma)3001_{ikm}]^T = \frac{1}{6720}\begin{bmatrix} 2002 & 1251 & -504 & 99 \\ 1251 & 2430 & -729 & 180 \\ -504 & -729 & 486 & -117 \\ 99 & 180 & -117 & -238 \\ -2934 & -1539 & 648 & -99 \\ -1539 & -1458 & 729 & -324 \\ 648 & 729 & -1458 & 405 \\ -99 & -324 & 405 & 1170 \\ 1170 & 405 & -324 & -99 \\ 405 & -1458 & 729 & 648 \\ -324 & 729 & -1458 & -1539 \\ -99 & 648 & -1539 & -2934 \\ -238 & -117 & 180 & 99 \\ -117 & 486 & -729 & -504 \\ 180 & -729 & 2430 & 1251 \\ 99 & -504 & 1251 & 2002 \end{bmatrix} \quad \text{(B.136)}$$

$$\left[A(\varsigma^2)3001_{ikm}\right]^T = \frac{1}{73920}\begin{bmatrix} -19900 & -10953 & 4518 & -1033 \\ -10953 & -17334 & 5751 & -828 \\ 4518 & 5751 & -1296 & -261 \\ -1033 & -828 & -261 & -2674 \\ 29538 & 13851 & -5832 & 1647 \\ 13851 & 14580 & -6561 & -486 \\ -5832 & -6561 & -1458 & 3159 \\ 1647 & -486 & 3159 & 12312 \\ -12312 & -3159 & 486 & -1647 \\ -3159 & 1458 & 6561 & 5832 \\ 486 & 6561 & -14580 & -13851 \\ -1647 & 5832 & -13851 & -29538 \\ 2674 & 261 & 828 & 1033 \\ 261 & 1296 & -5751 & -4518 \\ 828 & -5751 & 17334 & 10953 \\ 1033 & -4518 & 10953 & 19900 \end{bmatrix} \quad \text{(B.137)}$$

$$[A3110_{ikm}]^T = \frac{1}{4480}\begin{bmatrix} 4795 & -6753 & 2481 & -523 \\ -6753 & 9585 & -3699 & 867 \\ 2481 & -3699 & 1971 & -753 \\ -523 & 867 & -753 & 409 \\ 4539 & -4131 & -567 & 159 \\ -4131 & 8505 & -4617 & 243 \\ -567 & -4617 & 4131 & 1053 \\ 159 & 243 & 1053 & -1455 \\ -1455 & 1053 & 243 & 159 \\ 1053 & 4131 & -4617 & -567 \\ 243 & -4617 & 8505 & -4131 \\ 159 & -567 & -4131 & 4539 \\ 409 & -753 & 867 & -523 \\ -753 & 1971 & -3699 & 2481 \\ 867 & -3699 & 9585 & -6753 \\ -523 & 2481 & -6753 & 4795 \end{bmatrix} \quad \text{(B.138)}$$

$$[A(\varsigma)3110_{ikm}]^T = \frac{1}{13440}\begin{bmatrix} -12613 & 18099 & -6903 & 1417 \\ 18099 & -26487 & 10449 & -2061 \\ -6903 & 10449 & -3645 & 99 \\ 1417 & -2061 & 99 & 545 \\ -9855 & 10611 & -1053 & 297 \\ 10611 & -15309 & 6561 & -1863 \\ -1053 & 6561 & -10935 & 5427 \\ 297 & -1863 & 5427 & -3861 \\ 3861 & -5427 & 1863 & -297 \\ -5427 & 10935 & -6561 & 1053 \\ 1863 & -6561 & 15309 & -10611 \\ -297 & 1053 & -10611 & 9855 \\ -545 & -99 & 2061 & -1417 \\ -99 & 3645 & -10449 & 6903 \\ 2061 & -10449 & 26487 & -18099 \\ -1417 & 6903 & -18099 & 12613 \end{bmatrix} \quad \text{(B.139)}$$

$$\left[A(\varsigma^2)3110_{ikm}\right]^T = \frac{1}{13440}\begin{bmatrix} 11245 & -16443 & 6615 & -1417 \\ -16443 & 24543 & -10449 & 2349 \\ 6615 & -10449 & 5589 & -1755 \\ -1417 & 2349 & -1755 & 823 \\ 7389 & -8505 & 1215 & -99 \\ -8505 & 10935 & -2187 & -243 \\ 1215 & -2187 & -2187 & 3159 \\ -99 & -243 & 3159 & -2817 \\ -2817 & 3159 & -243 & -99 \\ 3159 & -2187 & -2187 & 1215 \\ -243 & -2187 & 10935 & -8505 \\ -99 & 1215 & -8505 & 7389 \\ 823 & -1755 & 2349 & -1417 \\ -1755 & 5589 & -10449 & 6615 \\ 2349 & -10449 & 24543 & -16443 \\ -1417 & 6615 & -16443 & 11245 \end{bmatrix} \quad \text{(B.140)}$$

$$[A3101_{ikm}]^T = \frac{1}{4480}\begin{bmatrix} 4795 & -6753 & 2481 & -523 \\ 4539 & -4131 & -567 & 159 \\ -1455 & 1053 & 243 & 159 \\ 409 & -753 & 867 & -523 \\ -6753 & 9585 & -3699 & 867 \\ -4131 & 8505 & -4617 & 243 \\ 1053 & 4131 & -4617 & -567 \\ -753 & 1971 & -3699 & 2481 \\ 2481 & -3699 & 1971 & -753 \\ -567 & -4617 & 4131 & 1053 \\ 243 & -4617 & 8505 & -4131 \\ 867 & -3699 & 9585 & -6753 \\ -523 & 867 & -753 & 409 \\ 159 & 243 & 1053 & -1455 \\ 159 & -567 & -4131 & 4539 \\ -523 & 2481 & -6753 & 4795 \end{bmatrix} \quad \text{(B.141)}$$

$$[A(\varsigma)3101_{ikm}]^T = \frac{1}{13440}\begin{bmatrix} -12613 & 18099 & -6903 & 1417 \\ -9855 & 10611 & -1053 & 297 \\ 3861 & -5427 & 1863 & -297 \\ -545 & -99 & 2061 & -1417 \\ 18099 & -26487 & 10449 & -2061 \\ 10611 & -15309 & 6561 & -1863 \\ -5427 & 10935 & -6561 & 1053 \\ -99 & 3645 & -10449 & 6903 \\ -6903 & 10449 & -3645 & 99 \\ -1053 & 6561 & -10935 & 5427 \\ 1863 & -6561 & 15309 & -10611 \\ 2061 & -10449 & 26487 & -18099 \\ 1417 & -2061 & 99 & 545 \\ 297 & -1863 & 5427 & -3861 \\ -297 & 1053 & -10611 & 9855 \\ -1417 & 6903 & -18099 & 12613 \end{bmatrix} \quad \text{(B.142)}$$

$$\left[A(\varsigma^2)3101_{ikm}\right]^T = \frac{1}{13440}\begin{bmatrix} 11245 & -16443 & 6615 & -1417 \\ 7389 & -8505 & 1215 & -99 \\ -2817 & 3159 & -243 & -99 \\ 823 & -1755 & 2349 & -1417 \\ -16443 & 24543 & -10449 & 2349 \\ -8505 & 10935 & -2187 & -243 \\ 3159 & -2187 & -2187 & 1215 \\ -1755 & 5589 & -10449 & 6615 \\ 6615 & -10449 & 5589 & -1755 \\ 1215 & -2187 & -2187 & 3159 \\ -243 & -2187 & 10935 & -8505 \\ 2349 & -10449 & 24543 & -16443 \\ -1417 & 2349 & -1755 & 823 \\ -99 & -243 & 3159 & -2817 \\ -99 & 1215 & -8505 & 7389 \\ -1417 & 6615 & -16443 & 11245 \end{bmatrix} \quad \text{(B.143)}$$

$$[A3011_{ikm}]^T = \frac{1}{4480}\begin{bmatrix} 4795 & 4539 & -1455 & 409 \\ -6753 & -4131 & 1053 & -753 \\ 2481 & -567 & 243 & 867 \\ -523 & 159 & 159 & -523 \\ -6753 & -4131 & 1053 & -753 \\ 9585 & 8505 & 4131 & 1971 \\ -3699 & -4617 & -4617 & -3699 \\ 867 & 243 & -567 & 2481 \\ 2481 & -567 & 243 & 867 \\ -3699 & -4617 & -4617 & -3699 \\ 1971 & 4131 & 8505 & 9585 \\ -753 & 1053 & -4131 & -6753 \\ -523 & 159 & 159 & -523 \\ 867 & 243 & -567 & 2481 \\ -753 & 1053 & -4131 & -6753 \\ 409 & -1455 & 4539 & 4795 \end{bmatrix} \quad \text{(B.144)}$$

$$[A(\varsigma)3011_{ikm}]^T = \frac{1}{13440}\begin{bmatrix} -12613 & -9855 & 3861 & -545 \\ 18099 & 10611 & -5427 & -99 \\ -6903 & -1053 & 1863 & 2061 \\ 1417 & 297 & -297 & -1417 \\ 18099 & 10611 & -5427 & -99 \\ -26487 & -15309 & 10935 & 3645 \\ 10449 & 6561 & -6561 & -10449 \\ -2061 & -1863 & 1053 & 6903 \\ -6903 & -1053 & 1863 & 2061 \\ 10449 & 6561 & -6561 & -10449 \\ -3645 & -10935 & 15309 & 26487 \\ 99 & 5427 & -10611 & -18099 \\ 1417 & 297 & -297 & -1417 \\ -2061 & -1863 & 1053 & 6903 \\ 99 & 5427 & -10611 & -18099 \\ 545 & -3861 & 9855 & 12613 \end{bmatrix} \quad \text{(B.145)}$$

$$\left[A(\varsigma^2)3011_{ikm}\right]^T = \frac{1}{13440}\begin{bmatrix} 11245 & 7389 & -2817 & 823 \\ -16443 & -8505 & 3159 & -1755 \\ 6615 & 1215 & -243 & 2349 \\ -1417 & -99 & -99 & -1417 \\ -16443 & -8505 & 3159 & -1755 \\ 24543 & 10935 & -2187 & 5589 \\ -10449 & -2187 & -2187 & -10449 \\ 2349 & -243 & 1215 & 6615 \\ 6615 & 1215 & -243 & 2349 \\ -10449 & -2187 & -2187 & -10449 \\ 5589 & -2187 & 10935 & 24543 \\ -1755 & 3159 & -8505 & -16443 \\ -1417 & -99 & -99 & -1417 \\ 2349 & -243 & 1215 & 6615 \\ -1755 & 3159 & -8505 & -16443 \\ 823 & -2817 & 7389 & 11245 \end{bmatrix} \quad \text{(B.146)}$$

$$[A3111_{ikm}]^T = \frac{1}{1120}\begin{bmatrix} -4006 & 5319 & -1728 & 415 \\ 5319 & -7641 & 3294 & -972 \\ -1728 & 3294 & -2538 & 972 \\ 415 & -972 & 972 & -415 \\ 5319 & -7641 & 3294 & -972 \\ -7641 & 8748 & -3645 & 2538 \\ 3294 & -3645 & 3645 & -3294 \\ -972 & 2538 & -3294 & 1728 \\ -1728 & 3294 & -2538 & 972 \\ 3294 & -3645 & 3645 & -3294 \\ -2538 & 3645 & -8748 & 7641 \\ 972 & -3294 & 7641 & -5319 \\ 415 & -972 & 972 & -415 \\ -972 & 2538 & -3294 & 1728 \\ 972 & -3294 & 7641 & -5319 \\ -415 & 1728 & -5319 & 4006 \end{bmatrix} \quad \text{(B.147)}$$

$$[A(\varsigma)3111_{ikm}]^T = \frac{1}{4480}\begin{bmatrix} 13617 & -18795 & 6315 & -1137 \\ -18795 & 27135 & -9477 & 1137 \\ 6315 & -9477 & 2025 & 1137 \\ -1137 & 1137 & 1137 & -1137 \\ -18795 & 27135 & -9477 & 1137 \\ 27135 & -38637 & 9477 & 2025 \\ -9477 & 9477 & 9477 & -9477 \\ 1137 & 2025 & -9477 & 6315 \\ 6315 & -9477 & 2025 & 1137 \\ -9477 & 9477 & 9477 & -9477 \\ 2025 & 9477 & -38637 & 27135 \\ 1137 & -9477 & 27135 & -18795 \\ -1137 & 1137 & 1137 & -1137 \\ 1137 & 2025 & -9477 & 6315 \\ 1137 & -9477 & 27135 & -18795 \\ -1137 & 6315 & -18795 & 13617 \end{bmatrix} \quad \text{(B.148)}$$

$$\left[A(\varsigma^2)3111_{ikm}\right]^T = \frac{1}{6720}\begin{bmatrix} -17867 & 25569 & -10017 & 2315 \\ 25569 & -38151 & 17253 & -4671 \\ -10017 & 17253 & -11907 & 4671 \\ 2315 & -4671 & 4671 & -2315 \\ 25569 & -38151 & 17253 & -4671 \\ -38151 & 59049 & -32805 & 11907 \\ 17253 & -32805 & 32805 & -17253 \\ -4671 & 11907 & -17253 & 10017 \\ -10017 & 17253 & -11907 & 4671 \\ 17253 & -32805 & 32805 & -17253 \\ -11907 & 32805 & -59049 & 38151 \\ 4671 & -17253 & 38151 & -25569 \\ 2315 & -4671 & 4671 & -2315 \\ -4671 & 11907 & -17253 & 10017 \\ 4671 & -17253 & 38151 & -25569 \\ -2315 & 10017 & -25569 & 17867 \end{bmatrix} \quad \text{(B.149)}$$

Appendix C

The Jacobian Matrices of Aerodynamic Flux-vectors

C.1 General Formulation

Defintions

$$A_j^{mn} = \frac{\partial F_j^{(m)}}{U^{(n)}} \tag{C.1}$$

$$U = \{\rho, m_i, E\}^T \tag{C.2}$$

$$F_j = \{m_j, \frac{m_i m_j}{\rho} + p\delta_{ij}, (E+p)\frac{m_j}{\rho}\}^T \tag{C.3}$$

$$E = \frac{m_i^2}{2\rho} + \frac{p}{\gamma - 1} \tag{C.4}$$

where U, F_j and E are respectively the independent variables, the flux-vectors and total energy per unit volume of the aerodynamic Euler or Navier-Stokes equations, as being defined in Chapter 2; the ith components of U and F_j correspond to the ith

momentum equation, with i ranging from 1 to n which is the dimension; A_j is the Jacobian matrix of F_j, of which A_j^{mn} is the entry at the mth row and nth column.

Pressure Derivatives

$$\frac{\partial p}{\partial \rho} = \frac{\gamma-1}{2}\frac{m_i^2}{\rho^2} = \frac{\gamma-1}{2} v_i^2 \tag{C.5}$$

$$\frac{\partial p}{\partial m_i} = (1-\gamma)\frac{m_i}{\rho} = (1-\gamma)\, v_i \tag{C.6}$$

$$\frac{\partial p}{\partial E} = \gamma - 1 \tag{C.7}$$

which are the results of differentiating (C.4).

Derivation

Considering (C.5-C.6-C.7), we can carry out definition (C.1) as follows.

- The 1st Row:

$$A_j^{1\,1} = \frac{\partial m_j}{\partial \rho} = 0 \tag{C.8}$$

$$A_j^{1\,1+i} = \frac{\partial m_j}{\partial m_i} = \delta_{ji} \tag{C.9}$$

$$A_j^{1\,2+n} = \frac{\partial m_j}{\partial E} = 0 \tag{C.10}$$

- The $(1+i)$th Row:

$$\begin{aligned} A_j^{1+i\,1} &= \frac{\partial}{\partial \rho}\left(\frac{m_j\, m_i}{\rho} + p\,\delta_{ji}\right) \\ &= \frac{\partial p}{\partial \rho}\delta_{ji} - \frac{m_j\, m_i}{\rho^2} \\ &= \frac{\gamma-1}{2} v_l^2\,\delta_{ji} - v_j\, v_i \end{aligned} \tag{C.11}$$

(C.12)

$$
\begin{aligned}
A_j^{1+i\,1+k} &= \frac{\partial}{\partial m_k}\left(\frac{m_j\,m_i}{\rho} + p\,\delta_{ji}\right) \\
&= \frac{\partial p}{\partial m_k}\,\delta_{ji} + \frac{m_i}{\rho}\,\delta_{jk} + \frac{m_j}{\rho}\,\delta_{ik} \\
&= (1-\gamma)\,v_k\,\delta_{ji} + v_i\,\delta_{jk} + v_j\,\delta_{ik} \qquad \text{(C.13)} \\
A_j^{1+i\,2+n} &= \frac{\partial}{\partial E}\left(\frac{m_j\,m_i}{\rho} + p\,\delta_{ji}\right) \\
&= \frac{\partial p}{\partial E}\,\delta_{ji} \\
&= (\gamma-1)\,\delta_{ji} \qquad \text{(C.14)}
\end{aligned}
$$

- The $(2+n)$th Row:

$$
\begin{aligned}
A_j^{2+n\,1} &= \frac{\partial}{\partial \rho}\left[(E+p)\,\frac{m_j}{\rho}\right] \\
&= \frac{\partial p}{\partial \rho}\,\frac{m_j}{\rho} - (E+p)\,\frac{m_j}{\rho^2} \\
&= \frac{\gamma-1}{2}\,v_l^2\,v_j - \left(\frac{m_l^2}{2\,\rho} + \frac{\gamma}{\gamma-1}\,p\right)\frac{m_j}{\rho^2} \\
&= \frac{\gamma-1}{2}\,v_j\,v_l^2 - \frac{1}{2}\,v_j\,v_l^2 - \frac{\gamma}{\gamma-1}\,\frac{p}{\rho}\,v_j \\
&= \left[\left(\frac{\gamma}{2}-1\right)v_l^2 - \frac{a^2}{\gamma-1}\right]v_j \qquad \text{(C.15)} \\
A_j^{2+n\,1+i} &= \frac{\partial}{\partial m_i}\left[(E+p)\,\frac{m_j}{\rho}\right] \\
&= \frac{E+p}{\rho}\,\delta_{ji} + \frac{\partial p}{\partial m_i}\,v_j \\
&= \left(\frac{v_l^2}{2} + \frac{\gamma}{\gamma-1}\,\frac{p}{\rho}\right)\delta_{ji} + (1-\gamma)\,v_i\,v_j \\
&= \left(\frac{v_l^2}{2} + \frac{a^2}{\gamma-1}\right)\delta_{ji} + (1-\gamma)\,v_i\,v_j \qquad \text{(C.16)} \\
A_j^{2+n\,2+n} &= \frac{\partial}{\partial E}\left[(E+p)\,\frac{m_j}{\rho}\right] \\
&= \left(1 + \frac{\partial p}{\partial E}\right)v_j
\end{aligned}
$$

$$= \gamma v_j \tag{C.17}$$

where a in (C.15) and (C.16) is the sonic speed, i.e., $a = \sqrt{\gamma \frac{p}{\rho}}$.

In summary,

$$\{A_j^{1\beta}\} = \{0,\ \delta_{j1},\ \delta_{j2},\ \delta_{j3},\ 0\} \tag{C.18}$$

$$\{A_j^{2\beta}\} = \begin{Bmatrix} \frac{\gamma-1}{2} v_l^2 \delta_{j1} - v_j v_1 \\ (1-\gamma) v_1 \delta_{j1} + v_1 \delta_{j1} + v_j \\ (1-\gamma) v_2 \delta_{j1} + v_1 \delta_{j2} \\ (1-\gamma) v_3 \delta_{j1} + v_1 \delta_{j3} \\ (\gamma - 1)\ \delta_{j1} \end{Bmatrix}^T \tag{C.19}$$

$$\{A_j^{3\beta}\} = \begin{Bmatrix} \frac{\gamma-1}{2} v_l^2 \delta_{j2} - v_j v_2 \\ (1-\gamma) v_1 \delta_{j2} + v_2 \delta_{j1} \\ (1-\gamma) v_2 \delta_{j2} + v_2 \delta_{j2} + v_j \\ (1-\gamma) v_3 \delta_{j2} + v_2 \delta_{j3} \\ (\gamma - 1)\ \delta_{j2} \end{Bmatrix}^T \tag{C.20}$$

$$\{A_j^{4\beta}\} = \begin{Bmatrix} \frac{\gamma-1}{2} v_l^2 \delta_{j3} - v_j v_3 \\ (1-\gamma) v_1 \delta_{j3} + v_3 \delta_{j1} \\ (1-\gamma) v_2 \delta_{j3} + v_3 \delta_{j2} \\ (1-\gamma) v_3 \delta_{j3} + v_3 \delta_{j3} + v_j \\ (\gamma - 1)\ \delta_{j3} \end{Bmatrix}^T \tag{C.21}$$

$$\{A_j^{5\beta}\} = \left\{ \begin{array}{c} \left[\left(\frac{\gamma}{2}-1\right)v_l^2 - \frac{a^2}{\gamma-1}\right] v_j \\ \left(\frac{v_l^2}{2} + \frac{a^2}{\gamma-1}\right)\delta_{j1} + (1-\gamma)\, v_1\, v_j \\ \left(\frac{v_l^2}{2} + \frac{a^2}{\gamma-1}\right)\delta_{j2} + (1-\gamma)\, v_2\, v_j \\ \left(\frac{v_l^2}{2} + \frac{a^2}{\gamma-1}\right)\delta_{j3} + (1-\gamma)\, v_3\, v_j \\ \gamma\, v_j \end{array} \right\}^T \qquad \text{(C.22)}$$

C.2 Concrete Formulas

The following are the contcrete form of (C.18-C.22) for flows of different dimensions.

3-D Flow

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{\gamma-1}{2}v_l^2 - v_1^2 & (3-\gamma)v_1 & (1-\gamma)v_2 & (1-\gamma)v_3 & \gamma-1 \\ -v_1v_2 & v_2 & v_1 & 0 & 0 \\ -v_1v_3 & v_3 & 0 & v_1 & 0 \\ \left[\left(\frac{\gamma}{2}-1\right)v_l^2 - \frac{a^2}{\gamma-1}\right]v_1 & \left(\frac{v_l^2}{2}+\frac{a^2}{\gamma-1}\right) + (1-\gamma)\,v_1^2 & (1-\gamma)\,v_1v_2 & (1-\gamma)\,v_1v_3 & \gamma v_1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ -v_1v_2 & v_2 & v_1 & 0 & 0 \\ \frac{\gamma-1}{2}v_l^2 - v_2^2 & (1-\gamma)v_1 & (3-\gamma)v_2 & (1-\gamma)v_3 & \gamma-1 \\ -v_2v_3 & 0 & v_3 & v_2 & 0 \\ \left[\left(\frac{\gamma}{2}-1\right)v_l^2 - \frac{a^2}{\gamma-1}\right]v_2 & (1-\gamma)\,v_1v_2 & \left(\frac{v_l^2}{2}+\frac{a^2}{\gamma-1}\right) + (1-\gamma)\,v_2^2 & (1-\gamma)\,v_2v_3 & \gamma v_2 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ -v_3 v_1 & v_3 & 0 & v_1 & 0 \\ -v_3 v_2 & 0 & v_3 & v_2 & 0 \\ \frac{\gamma-1}{2} v_l^2 - v_3^2 & (1-\gamma) v_1 & (1-\gamma) v_2 & (3-\gamma) v_3 & \gamma - 1 \\ \left[\left(\frac{\gamma}{2} - 1 \right) v_l^2 - \frac{a^2}{\gamma-1} \right] v_3 & (1-\gamma)\, v_1 v_3 & (1-\gamma)\, v_2 v_3 & \left(\frac{v_l^2}{2} + \frac{a^2}{\gamma-1} \right) + (1-\gamma)\, v_3^2 & \gamma v_3 \end{bmatrix}$$

2-D Flow

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{\gamma-1}{2} v_l^2 - v_1^2 & (3-\gamma) v_1 & (1-\gamma) v_2 & \gamma - 1 \\ -v_1 v_2 & v_2 & v_1 & 0 \\ \left[\left(\frac{\gamma}{2} - 1 \right) v_l^2 - \frac{a^2}{\gamma-1} \right] v_1 & \left(\frac{v_l^2}{2} + \frac{a^2}{\gamma-1} \right) + (1-\gamma)\, v_1^2 & (1-\gamma)\, v_1 v_2 & \gamma v_1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ -v_1 v_2 & v_2 & v_1 & 0 \\ \frac{\gamma-1}{2} v_l^2 - v_2^2 & (1-\gamma) v_1 & (3-\gamma) v_2 & \gamma - 1 \\ \left[\left(\frac{\gamma}{2} - 1 \right) v_l^2 - \frac{a^2}{\gamma-1} \right] v_2 & (1-\gamma)\, v_1 v_2 & \left(\frac{v_l^2}{2} + \frac{a^2}{\gamma-1} \right) + (1-\gamma)\, v_2^2 & \gamma v_2 \end{bmatrix}$$

1-D Flow

$$A = \begin{bmatrix} 0 & 1 & 0 \\ \frac{\gamma-1}{2} v^2 - v^2 & (3-\gamma) v & \gamma - 1 \\ \left[\left(\frac{\gamma}{2} - 1 \right) v^2 - \frac{a^2}{\gamma-1} \right] v & \left(\frac{v^2}{2} + \frac{a^2}{\gamma-1} \right) + (1-\gamma)\, v^2 & \gamma v \end{bmatrix}$$

VITA

Jing Zhang was born on March 4, 1959, in Qingdao, China. He attended Qingdao Fen-Yang Road Primary School in 1966, and then Qingdao 33rd Middle School in 1972. He graduated from 33rd in 1975, and became employed with Qingdao Fertilizer Factory, being an electrician in the next three years. In 1978, he was admitted and went to Xi'an Jiaotong University, majoring in turbomachinery. In 1982, he was awarded the Bachelor of Science Degree. He moved on to work towards a M.S., enrolling a joint program between The Department of Mathematics and The Department of Energy and Power Engineering, Xi'an Jiaotong University (XJU). Upon completing the M.S. study in 1984, his thesis, "A Pointwise Finite Element Method And Its Application in Aerodynamics", was approved qualifying for the application of a Doctor of Philosophy Degree (which was never carried out). Jing Zhang continued his scientist career at XJT until 1987, when he accepted a Teaching Assistantship from The University of Tennessee, Knoxville. He flew over the Pacific, and has been *lost* ever since (perhaps, until now :-).

Jing Zhang has published eleven technical papers.

Jing Zhang was a member of the first term congress of the Independent Federation of Chinese Students and Scholars (IFCSS).