Improving Euler Computations at Low Mach Numbers

B. Koren

CWI, Department of Numerical Mathematics
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

B. van Leer

The University of Michigan, Department of Aerospace Engineering
1320 Beal Avenue, Ann Arbor, Michigan 48109-2118, USA

Abstract

The paper consists of two parts, both dealing with conditioning techniques for low-Mach-number Euler-flow computations, in which a multigrid technique is applied.

In the first part, for subsonic flows and upwind-discretized, linearized 1-D Euler equations, the smoothing behavior of multigrid-accelerated point Gauss-Seidel relaxation is investigated. Error decay by convection over domain boundaries is also discussed. A fix to poor convergence rates at low Mach numbers is sought in replacing the point relaxation applied to unconditioned Euler equations, by locally implicit "time" stepping applied to preconditioned Euler equations. The locally implicit iteration step is optimized for good damping of high-frequency errors. Numerical inaccuracy at low Mach numbers is also addressed. In the present case it is not necessary to solve this accuracy problem.

In the second part, insight is given in the conditions of derivative matrices to be inverted in point-relaxation methods for 1-D and 2-D, upwind-discretized Euler equations. Speed regimes are found where ill-conditioning of these matrices occurs, 1-D flow equations appear to be less well conditioned than 2-D flow equations. Fixes to the ill conditions follow more or less directly, when thinking of adding regularizing matrices to the ill-conditioned derivative matrices. A smoothing analysis is made of point Gauss-Seidel relaxation applied to Euler equations conditioned by such an additive matrix. The method is successfully applied to a very low-subsonic, steady, 2-D stagnation flow.

Note: Major parts of this paper were already published in [1] and [2].

---

[1] KOREN, B.: Condition improvement for point relaxation in multigrid, subsonic Euler flow computations, *Applied Numerical Mathematics*, **16**, 457-469 (1995).

[2] KOREN, B. AND LEER, B. VAN: Analysis of preconditioning and multigrid for Euler flows with low-subsonic regions, *Advances in Computational Mathematics*, **4**, 127-144 (1995).

1 Preconditioning of Jacobians

1.1 Introduction

Mathematical theory of subsonic gas flows is relatively undeveloped in comparison with that of transonic, supersonic and hypersonic flows. An indication of this is the small amount of literature that is available on subsonic gas dynamics. Whereas various text books exist, exclusively dealing with the mathematics of either transonic, supersonic or hypersonic gas flows, for the subsonic case we only know a few book chapters (e.g. Chapters 2 and 3 from Bers [1], and Chapter 2 from Majda [6]). At present, research in the subsonic flow regime is at a rapid pace, particularly as far as it concerns numerical computations in the zero-Mach-number limit. Both sections from this paper contribute to this development. In the present section, the flows of interest are not flows with uniformly low Mach numbers (i.e. with $M \ll 1$ throughout almost the entire computational domain), but flows with *locally* low Mach numbers (flows with small stagnation regions and - particularly - for Navier-Stokes extensions: flows with thin boundary layers and wakes).

Since about a decade, various multigrid methods exist that give good convergence rates for steady Euler-flow computations at high-subsonic inflow Mach numbers (see Chapter 9 from Wesseling [14] for an overview). For decreasing inflow Mach numbers, or enlarging low-subsonic flow regions, convergence rates are known to deteriorate. This decrease is not specific for multigrid methods, but seems to hold for any solution method. The cause has to be sought in the continuous Euler equations, in their increasing stiffness (i.e. in their increasing disparity of wave speeds) at decreasing subsonic Mach numbers. With the application of single-grid, explicit time stepping schemes in mind, various fixes have been proposed already for this stiffness problem. See Turkel [11] for a review of this. An early research paper is Van Leer et al. [5]. In it, for the Jacobian of the 1-D Euler equations, the preconditioning matrix is given which equalizes the three wave speeds $u - c$, u and $u + c$. Further, the paper gives preconditioning matrices for the 2-D and 3-D Euler equations. Besides convergence problems, for decreasing Mach numbers also accuracy problems arise, independent of whether the discretization is central [13] or upwind [12].

In the present section, we will mainly focus on the stiffness problem. It is expected that solution methods other than explicit time stepping schemes may also profit from preconditioning matrices such as those proposed in Van Leer et al. [5]. Led by this expectation, we will optimize a multigrid-accelerated, locally implicit iteration method, applied to subsonic, preconditioned Euler equations. To start with, in Section 1.2, the continuous, unconditioned equations and their discretization are introduced. In Section 1.3, first a smoothing analysis is given of point Gauss-Seidel relaxation for the discrete equations, and next a discussion is made of error convection across domain boundaries. It is shown that for low Mach numbers, the convergence properties of this solution method are poor. In Section 1.4 it is made clear that for flows with uniformly low Mach numbers, numerical accuracy may be poor as well. Since the latter flows are not our interest, in Section 1.5 a 1-D preconditioning matrix is derived which is only meant for removing stiffness and not for also improving low-Mach-number inaccuracy. In Section 1.6, a simple way of implementing the preconditioning matrix is discussed. At the end of Section 1.6 we arrive at the discrete, preconditioned system to be solved. The system contains a free parameter: a locally implicit iteration step, which is optimized for smoothing. The optimization is done in Section 1.7, through local-mode analysis applied to the upwind-

discretized, linearized, preconditioned 1-D Euler equations. In Section 1.8, the error smoothing and error convection of the locally implicit iteration are verified.

1.2 Equations

1.2.1 Continuous equations

Consider the 1-D Euler equations

$$\frac{\partial Q}{\partial t} + \frac{\partial f(Q)}{\partial x} = 0, \tag{1a}$$

with Q the conservative state vector

$$Q = \begin{pmatrix} \rho \\ \rho u \\ \rho e \end{pmatrix}, \tag{1b}$$

$f(Q)$ the corresponding flux vector

$$f(Q) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho u(e + \frac{p}{\rho}) \end{pmatrix}, \tag{1c}$$

and e the internal energy, which for a perfect gas reads

$$e = \frac{1}{\gamma - 1}\frac{p}{\rho} + \frac{1}{2}u^2. \tag{1d}$$

Linearization of (1a) with respect to the conservative variables yields

$$\frac{\partial Q}{\partial t} + \frac{df}{dQ}\frac{\partial Q}{\partial x} = 0, \tag{2a}$$

$$\frac{df}{dQ} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{\gamma-3}{2}u^2 & (3-\gamma)u & \gamma-1 \\ \frac{\gamma-2}{2}u^3 - \frac{1}{\gamma-1}uc^2 & \frac{(3-2\gamma)}{2}u^2 + \frac{1}{\gamma-1}c^2 & \gamma u \end{pmatrix}, \tag{2b}$$

where $c^2 = \sqrt{\gamma p/\rho}$. To simplify the analysis, following Turkel [10], the transformation from conservative variables Q to non-conservative (entropy) variables q is made:

$$dq \equiv \begin{pmatrix} \frac{1}{\rho c}dp \\ du \\ dp - c^2 d\rho \end{pmatrix}. \tag{3}$$

The corresponding transformation matrix

$$\frac{dQ}{dq} = \begin{pmatrix} \frac{\rho}{c} & 0 & -\frac{1}{c^2} \\ \frac{\rho u}{c} & \rho & -\frac{u}{c^2} \\ \frac{1}{2}\frac{\rho u^2}{c} + \frac{1}{\gamma-1}\rho c & \rho u & -\frac{1}{2}\frac{u^2}{c^2} \end{pmatrix} \tag{4}$$

brings (1a) into the analytically much more tractable form

$$\frac{\partial q}{\partial t} + A\frac{\partial q}{\partial x} = 0, \tag{5a}$$

$$A = \frac{dq}{dQ}\frac{df}{dQ}\frac{dQ}{dq} = \begin{pmatrix} u & c & 0 \\ c & u & 0 \\ 0 & 0 & u \end{pmatrix}. \tag{5b}$$

1.2.2 Discrete equations

For simplicity we assume A to be constant and next make a first-order upwind, cell-centered finite-volume discretization of the space operator in (5a). Then, the semi-discrete equation in cell Ω_i (with mesh size h) reads

$$h\frac{\partial q_i}{\partial t} + A^+(q_i - q_{i-1}) + A^-(q_{i+1} - q_i) = 0, \tag{6}$$

with i running in positive x-direction, and with A^+ and A^- the matrices corresponding with the positive and negative eigenvalues of matrix A:

$$A^+ = R_A \Lambda_A^+ R_A^{-1}, \tag{7a}$$

$$A^- = R_A \Lambda_A^- R_A^{-1}. \tag{7b}$$

With $\Lambda_A = \text{diag}(u - c, u, u + c)$, it holds

$$R_A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{8}$$

and, hence, for subsonic flow in positive x-direction, $0 < u < c$:

$$A^+ = \frac{1}{2}\begin{pmatrix} u+c & u+c & 0 \\ u+c & u+c & 0 \\ 0 & 0 & 2u \end{pmatrix}, \tag{9a}$$

$$A^- = \frac{1}{2}\begin{pmatrix} u-c & c-u & 0 \\ c-u & u-c & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9b}$$

1.3 Convergence

1.3.1 Convergence through error smoothing

Applying point Gauss-Seidel relaxation to find the steady solution of (6), for successively a downstream and upstream relaxation sweep the iteration formulae are

$$|A|(q_i^{n+1} - q_i^n) = -A^+(q_i^n - q_{i-1}^{n+1}) - A^-(q_{i+1}^n - q_i^n), \tag{10a}$$

$$|A|(q_i^{n+2} - q_i^{n+1}) = -A^+(q_i^{n+1} - q_{i-1}^{n+1}) - A^-(q_{i+1}^{n+2} - q_i^{n+1}), \tag{10b}$$

with $|A| \equiv A^+ - A^-$ and n the relaxation sweep counter. To investigate the smoothing properties we introduce the local solution error

$$\delta_i^n = q_i^* - q_i^n, \tag{11a}$$

and the Fourier form

$$\delta_i^n = D^n e^{i\theta i}, \quad |\theta| \in \left[\frac{\pi}{2}, \pi\right], \tag{11b}$$

with q_i^* the exact local solution, D^n the amplitude vector (D_1^n, D_2^n, D_3^n) and $e^{i\theta i}$ the (scalar) mode. With (11a) and (11b), it follows for the amplification matrices $\mathcal{M}_{\text{downstream}}$ and $\mathcal{M}_{\text{upstream}}$:

$$\mathcal{M}_{\text{downstream}} = -\left(-e^{-i\theta}A^+ + |A|\right)^{-1} e^{i\theta}A^-, \tag{12a}$$

$$\mathcal{M}_{\text{upstream}} = \left(e^{i\theta}A^- + |A|\right)^{-1} e^{-i\theta}A^+. \tag{12b}$$

Substituting A^+ and A^- we find the solution-independent matrices

$$\mathcal{M}_{\text{downstream}} = \frac{1}{2}\begin{pmatrix} e^{i\theta} & -e^{i\theta} & 0 \\ -e^{i\theta} & e^{i\theta} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{13a}$$

$$\mathcal{M}_{\text{upstream}} = \frac{1}{2}\begin{pmatrix} e^{-i\theta} & e^{-i\theta} & 0 \\ e^{-i\theta} & e^{-i\theta} & 0 \\ 0 & 0 & 2e^{-i\theta} \end{pmatrix}, \tag{13b}$$

with spectral radii

$$\rho(\mathcal{M}_{\text{downstream}}) = |e^{i\theta}| = 1, \quad \forall |\theta| \in \left[\frac{\pi}{2}, \pi\right], \tag{14a}$$

$$\rho(\mathcal{M}_{\text{upstream}}) = |e^{-i\theta}| = 1, \quad \forall |\theta| \in \left[\frac{\pi}{2}, \pi\right]. \tag{14b}$$

(Since the matrices (13a) and (13b) are symmetric, the spectral norms, which determine the smoothing properties for $n = 1$, are identical to the spectral radii.) Note that in case of a symmetric sweep, according to this Fourier analysis, one has perfect smoothing: $\mathcal{M}_{\text{upstream}}\mathcal{M}_{\text{downstream}} = 0$. However, in case of subsonic flow with non-periodic boundary conditions, one generally has error reflections at the outflow boundary when still iterating. Therefore this theoretical, perfect smoothing result is not realistic and therefore we prefer to consider the downstream and upstream amplification matrices separately. However, for the two separate sweeps, the smoothing factors (14a) and (14b) are surprising as well. They are in contradiction with numerical findings; for e.g. standard, high-subsonic airfoil-flow computations, one generally observes good multigrid convergence. A first explanation of this contradictory result is that care has to be taken in interpreting (14a) and (14b); the frozen coefficients assumption generally loses its validity for high-subsonic Mach numbers. As opposed to this, for low-subsonic Mach numbers it seems a reasonable assumption (e.g., for $\lim_{M\downarrow 0}$, ρ becomes constant).

1.3.2 Convergence through error convection

A second explanation of the contradictory convergence estimate for high-subsonic flows in the general case of non-periodic boundary conditions is that for the downstream and upstream sweep separately, local-mode analysis solely is just too pessimistic. For non-periodic high-subsonic flow computations, additional error decay through advection over the domain boundaries may be of significant importance and may therefore not be neglected. Note herewith that point Gauss-Seidel relaxation can be interpreted as locally implicit time stepping at an infinitely large time step, which with non-zero wave propagation speeds $u - c$, u and $u + c$, implies a significant beneficial influence on convergence. This phenomenon of solution errors being expelled out of the computational domain by convection may next further explain the poor multigrid performance for low-subsonic flows. In spite of the infinitely large time step associated with point Gauss-Seidel relaxation, for $\lim_{M\downarrow 0}$, the propagation of entropy errors and therefore their expulsion, may well start to stagnate.

1.4 Accuracy

For $\lim_{M \downarrow 0}$ accuracy problems arise. The inaccuracy can be analyzed through the modified equation corresponding with: (5a)-(5b) discretized through e.g. a first-order accurate flux-difference splitting scheme (such as Osher's [7] or Roe's [9]). The corresponding modified equation reads

$$\frac{\partial q}{\partial t} + A\frac{\partial q}{\partial x} = \frac{h}{2}\frac{\partial}{\partial x}\left(|A|\frac{\partial q}{\partial x}\right) - \frac{h}{2}\frac{\partial}{\partial x}\left(\frac{dq}{dQ}\right)\frac{dQ}{dq}|A|\frac{\partial q}{\partial x}. \tag{15}$$

With

$$|A| = \begin{pmatrix} c & u & 0 \\ u & c & 0 \\ 0 & 0 & u \end{pmatrix}, \tag{16}$$

the numerical diffusion terms in the right-hand side of (15) can be written out as

$$\frac{h}{2}\left[\left\{\frac{\partial u}{\partial x}\begin{pmatrix} \frac{1}{M} & 1 & 0 \\ 1 & \frac{1}{M} & 0 \\ 0 & 0 & 1 \end{pmatrix} + u\frac{\partial}{\partial x}\begin{pmatrix} \frac{1}{M} & 1 & 0 \\ 1 & \frac{1}{M} & 0 \\ 0 & 0 & 1 \end{pmatrix}\right\}\frac{\partial q}{\partial x} + |A|\frac{\partial^2 q}{\partial x^2}\right] +$$

$$\frac{h}{2}\begin{pmatrix} \left(\frac{1}{\rho}\frac{\partial \rho}{\partial x} + \frac{1}{c}\frac{\partial c}{\partial x}\right)c + (\gamma-1)M\frac{\partial u}{\partial x} & \left(\frac{1}{\rho}\frac{\partial \rho}{\partial x} + \frac{1}{c}\frac{\partial c}{\partial x}\right)u + (\gamma-1)\frac{\partial u}{\partial x} & 0 \\ \frac{\partial u}{\partial x} + \frac{1}{\rho}\frac{\partial \rho}{\partial x}u & M\frac{\partial u}{\partial x} + \frac{1}{\rho}\frac{\partial \rho}{\partial x}c & -\frac{1}{\rho c}M\frac{\partial u}{\partial x} \\ \rho\left(2c\frac{\partial c}{\partial x} + \frac{1}{\gamma-1}u\frac{\partial u}{\partial x}\right) & \rho\left(2u\frac{\partial c}{\partial x} + \frac{1}{\gamma-1}c\frac{\partial u}{\partial x}\right) & -2M\frac{\partial c}{\partial x} \end{pmatrix}\frac{\partial q}{\partial x}. \tag{17}$$

It appears that for $\lim_{M \downarrow 0}$ and with h fixed and $\frac{\partial u}{\partial x}$ and $\frac{\partial p}{\partial x}$ non-zero, the first of the two numerical diffusion terms in system (15) becomes infinitely large for the first two equations in the system.

1.5 Preconditioning

1.5.1 Removing stiffness

For a detailed account of this topic we refer to Van Leer et al. [5]. For the condition K of A over the entire subsonic flow regime, it holds

$$K(A) = \max\left(\frac{1+M}{M}, \frac{1+M}{1-M}\right), \quad M \equiv \frac{|u|}{c} \in (0,1), \tag{18}$$

see also Figure 1. At $M = 0$ and $M = 1$, A is singular. Preconditioning A (by premultiplying it) with the 3×3-matrix P transforms (5a) into

$$\frac{\partial q}{\partial t} + PA\frac{\partial q}{\partial x} = 0. \tag{19}$$

For general P, the possibility of doing time-accurate calculations is lost. When solving steady problems, this is of no concern. P should at least be invertible and should remove the static and sonic singularity. In the ideal case, P leads to the situation: (i) that $K(PA) = 1$ over the entire subsonic Mach-number range, and (ii) that PA yields two downstream waves and one upstream wave. Satisfaction of the second property, conservation of the propagation directions of the three waves, avoids a change of numbers
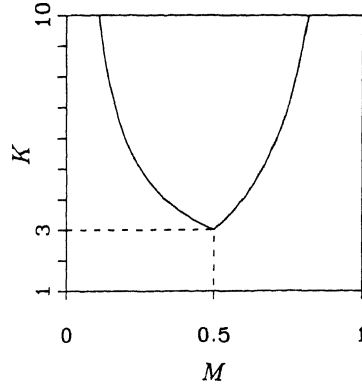
424

Figure 1: Condition of derivative matrix A as a function of the Mach number.

of boundary conditions to be imposed at in- and outlet. This property is satisfied by taking P positive definite, which implies that P must be symmetric.

A common choice for P is

$$P = \mathsf{v}|A|^{-1}, \tag{20}$$

with v some propagation speed that can still be chosen. With (20) one has $\Lambda_{PA} = \mathrm{diag}(-\mathsf{v}, \mathsf{v}, \mathsf{v})$. In multi-D, perfect subsonic preconditioning is not possible. For 2-D subsonic Euler flows and for $dq = (\frac{1}{\rho c}dp, du, dv, dp - c^2 d\rho)^T$, the following preconditioning matrix is proposed in Van Leer et al. [5]:

$$P = \begin{pmatrix} \frac{M^2}{\sqrt{1-M^2}} & \frac{-M}{\sqrt{1-M^2}} & 0 & 0 \\ \frac{-M}{\sqrt{1-M^2}} & \frac{1}{\sqrt{1-M^2}}+1 & 0 & 0 \\ 0 & 0 & \sqrt{1-M^2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{21}$$

The 3-D subsonic preconditioning matrix proposed in Van Leer et al. [5] is very much the same as (21). Our practical interest lies in doing 2-D and 3-D computations. However, since already in 2-D, local-mode analysis for the preconditioned, full Euler equations is hard and does not lead to transparent results, we do the analysis for the 1-D Euler equations, with as preconditioning matrix a 1-D version of 2-D P (21). We proceed by deriving such a 1-D P.

Striving for the almost diagonal form

$$PA = \begin{pmatrix} -\mathsf{v} & 0 & 0 \\ \hat{\mathsf{v}} & \mathsf{v} & 0 \\ 0 & 0 & \mathsf{v} \end{pmatrix}, \tag{22}$$

a symmetric 1-D version of (21) can be found. For $\mathsf{v} = u$, it follows

$$P = \begin{pmatrix} \frac{M^2}{1-M^2} & \frac{-M}{1-M^2} & 0 \\ \frac{-M}{1-M^2} & \frac{1}{1-M^2}+1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{23}$$

(and, unimportant, $\hat{\mathsf{v}} = 2c$). Since entropy propagates with the flow speed, just as in (21), the entropy equation is left unchanged. Note that P according to (23) is

positive-definite; for $M \in (0,1)$ its three eigenvalues are all positive: $\lambda_1 = 1$ and $\lambda_{2,3} = \frac{1 \pm \sqrt{1-M^2+M^4}}{1-M^2} > 0, \forall M \in (0,1)$. Also still note the freedom in the derivation of this preconditioning matrix. E.g., another v could have been chosen; $v = u + c$ would have yielded

$$P = \begin{pmatrix} \frac{M}{1-M} & \frac{-1}{1-M} & 0 \\ \frac{-1}{1-M} & \frac{2-M^2}{M(1-M)} & 0 \\ 0 & 0 & \frac{1+M}{M} \end{pmatrix}. \tag{24}$$

Moreover, instead of preconditioning, postconditioning could have been applied. The difference between pre- and postconditioning can be clarified by considering the auxiliary equation $A\frac{dq}{dx} = r$. Preconditioning this equation $(PA\frac{dq}{dx} = r)$ is identical to right-hand side transformation $(A\frac{dq}{dx} = P^{-1}r)$, whereas postconditioning $(AP\frac{dq}{dx} = r)$ can be interpreted as solution transformation. Postconditioning (5a) by a symmetric P such that

$$AP = \begin{pmatrix} -u & 0 & 0 \\ \hat{v} & u & 0 \\ 0 & 0 & u \end{pmatrix}, \tag{25}$$

leads to

$$P = \begin{pmatrix} -\frac{M^2-2}{M^2-1} & -\frac{M}{M^2-1} & 0 \\ -\frac{M}{M^2-1} & \frac{M^2}{M^2-1} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{26}$$

Interpreting this postconditioning matrix as a solution transformation matrix $\frac{d\tilde{q}}{dq}$, we get

$$d\tilde{q} = \begin{pmatrix} -\frac{M^2-2}{M^2-1}\frac{1}{\rho c}dp - \frac{M}{M^2-1}du \\ -\frac{M}{M^2-1}\frac{1}{\rho c}dp + \frac{M^2}{M^2-1}du \\ dp - c^2 d\rho \end{pmatrix}. \tag{27}$$

Physical interpretation of the first two components of $d\tilde{q}$ is not trivial. In the remainder we consider preconditioning according to (23).

1.5.2 Concerning inaccuracy

A partial fix to the discrete accuracy problem discussed in Section 1.4, is to make the discretization second-order accurate. (In practical computations, the discretization will be at least second-order accurate anyway.) Of course, as long as the two limits $M \downarrow 0$ and $h \downarrow 0$ are independent (and as long as the discretization method is not exact), formally the accuracy problem remains to exist. A subsequent partial fix would then be to take the mesh size appropriately dependent on the Mach number.

A real fix is to exploit the freedom still existing in the choice of the preconditioning matrices for removing the stiffness problem. By first preconditioning,

$$\frac{\partial q}{\partial t} + PA\frac{\partial q}{\partial x} = 0, \tag{28}$$

and next discretizing, one gets the influence of the preconditioning in the discretization error. For flow computations at uniformly low Mach numbers, the challenge is then to get rid of both the stiffness and the accuracy problem by a single preconditioning matrix P. Such double-edged preconditioning matrices are expected to become available soon [12]. Discretization of (28) requires the incorporation of a space discretization scheme

426

which is modified for the preconditioning (both at the interior and the boundary cell faces). Further, in multigrid contexts the residual transfer has to be reconsidered, in order to maintain the Galerkin property and hence good multigrid convergence [4, 3]. Since uniformly low-Mach-number flows are not our present interest, we will not apply the preconditioning in the form (28). (For computations in which the Mach number is not uniformly low, the accuracy problems occurring for $\lim_{M \downarrow 0}$ are local, and hence no reduction of global solution accuracy is expected to be found.)

1.6 Implementing preconditioning

By implementing the preconditioning as

$$P^{-1}\frac{\partial q}{\partial t} + A\frac{\partial q}{\partial x} = 0, \tag{29a}$$

with P^{-1} the inverse of (23):

$$P^{-1} = \begin{pmatrix} \frac{2-M^2}{M^2} & \frac{1}{M} & 0 \\ \frac{1}{M} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{29b}$$

the original space discretization scheme can still be applied (simply because the space operator is still original). Steady-state solutions will therefore be identical to those belonging to the unconditioned equations (5a) and (15). The conservative form corresponding with (29a) reads

$$\frac{dQ}{dq}P^{-1}\frac{\partial q}{\partial t} + \frac{\partial f(Q)}{\partial x} = 0. \tag{30}$$

Discretizing (30) by a first-order upwind finite-volume method, and denoting the numerical flux function which approximates cell-face flux $f(q_{i+\frac{1}{2}})$ by $F(q_i, q_{i+1})$, for cell Ω_i the semi-discrete equation reads

$$\int_{\Omega_i} \frac{dQ}{dq}P^{-1}\frac{\partial q}{\partial t}dx + (F(q_i, q_{i+1}) - F(q_{i-1}, q_i)) = 0. \tag{31}$$

Given the good smoothing properties of point Gauss-Seidel relaxation in the multigrid computation of high-subsonic, transonic and supersonic flows, in choosing the time discretization for (31) we deviate as little as possible from this trusty smoother, by applying: locally implicit time stepping in a Gauss-Seidel fashion. Hence, as fully discrete equation in cell Ω_i, for a downstream and upstream sweep respectively, it follows:

$$\left[\frac{h}{\Delta t}\frac{dQ}{dq}(q_i^n)P^{-1}(M_i^n) + \frac{\partial F(q_i^n, q_{i+1}^n)}{\partial q_i^n} - \frac{\partial F(q_{i-1}^{n+1}, q_i^n)}{\partial q_i^n}\right]\left(q_i^{n+1} - q_i^n\right) =$$
$$F(q_{i-1}^{n+1}, q_i^n) - F(q_i^n, q_{i+1}^n), \tag{32a}$$

$$\left[\frac{h}{\Delta t}\frac{dQ}{dq}(q_i^{n+1})P^{-1}(M_i^{n+1}) + \frac{\partial F(q_i^{n+1}, q_{i+1}^{n+2})}{\partial q_i^{n+1}} - \frac{\partial F(q_{i-1}^{n+1}, q_i^{n+1})}{\partial q_i^{n+1}}\right]\left(q_i^{n+2} - q_i^{n+1}\right) =$$
$$F(q_{i-1}^{n+1}, q_i^{n+1}) - F(q_i^{n+1}, q_{i+1}^{n+2}). \tag{32b}$$

The time step Δt (which due to the preconditioning is not identical to a physical time step) is still amenable to optimization. In the next section it will be optimized for smoothing.

427

1.7 Optimizing locally implicit iteration step

For simplicity, smoothing optimization of Δt from (32a) and (32b) is done for the non-conservative, frozen-coefficient variants of both equations, i.e. for:

$$\left(\frac{h}{\Delta t}P^{-1} + |A|\right)(q_i^{n+1} - q_i^n) = -A^+(q_i^n - q_{i-1}^{n+1}) - A^-(q_{i+1}^n - q_i^n), \qquad (33a)$$

$$\left(\frac{h}{\Delta t}P^{-1} + |A|\right)(q_i^{n+2} - q_i^{n+1}) = -A^+(q_i^{n+1} - q_{i-1}^{n+1}) - A^-(q_{i+1}^{n+2} - q_i^{n+1}). \qquad (33b)$$

1.7.1 Qualitative optimization

From (33a) and (33b), with (11a) and (11b), in the same way as in Section 1.3.1 we derive:

$$\mathcal{M}_{\text{downstream}} = \left(\frac{h}{\Delta t}P^{-1} - e^{-i\theta}A^+ + |A|\right)^{-1}\left(\frac{h}{\Delta t}P^{-1} - e^{i\theta}A^-\right), \qquad (34a)$$

$$\mathcal{M}_{\text{upstream}} = \left(\frac{h}{\Delta t}P^{-1} + e^{i\theta}A^- + |A|\right)^{-1}\left(\frac{h}{\Delta t}P^{-1} + e^{-i\theta}A^+\right). \qquad (34b)$$

We proceed by considering the two highest error frequencies: $|\theta| = \pi$. For both frequencies, with

$$\sigma \equiv \frac{h}{c\Delta t}, \qquad (35)$$

(34a) and (34b) can be written out as:

$$\mathcal{M}_{\text{downstream}} = \begin{pmatrix} \sigma\frac{2-M^2}{M^2} + \frac{3}{2} + \frac{1}{2}M & \sigma\frac{1}{M} + \frac{1}{2} + \frac{3}{2}M & 0 \\ \sigma\frac{1}{M} + \frac{1}{2} + \frac{3}{2}M & \sigma + \frac{3}{2} + \frac{1}{2}M & 0 \\ 0 & 0 & \sigma + 2M \end{pmatrix}^{-1}$$
$$\begin{pmatrix} \sigma\frac{2-M^2}{M^2} - \frac{1}{2} + \frac{1}{2}M & \sigma\frac{1}{M} + \frac{1}{2} - \frac{1}{2}M & 0 \\ \sigma\frac{1}{M} + \frac{1}{2} - \frac{1}{2}M & \sigma - \frac{1}{2} + \frac{1}{2}M & 0 \\ 0 & 0 & \sigma \end{pmatrix}, \qquad (36a)$$

$$\mathcal{M}_{\text{upstream}} = \begin{pmatrix} \sigma\frac{2-M^2}{M^2} + \frac{3}{2} - \frac{1}{2}M & \sigma\frac{1}{M} - \frac{1}{2} + \frac{3}{2}M & 0 \\ \sigma\frac{1}{M} - \frac{1}{2} + \frac{3}{2}M & \sigma + \frac{3}{2} - \frac{1}{2}M & 0 \\ 0 & 0 & \sigma + M \end{pmatrix}^{-1}$$
$$\begin{pmatrix} \sigma\frac{2-M^2}{M^2} - \frac{1}{2} - \frac{1}{2}M & \sigma\frac{1}{M} - \frac{1}{2} - \frac{1}{2}M & 0 \\ \sigma\frac{1}{M} - \frac{1}{2} - \frac{1}{2}M & \sigma - \frac{1}{2} - \frac{1}{2}M & 0 \\ 0 & 0 & \sigma - M \end{pmatrix}. \qquad (36b)$$

The corresponding eigenvalues are:

$$(\lambda_1)_{\mathcal{M}_{\text{downstream}}} = \frac{\sigma}{\sigma + 2M}, \quad (\lambda_{2,3})_{\mathcal{M}_{\text{downstream}}} = \frac{\sigma + \sigma^2 - M^2 \pm \sqrt{4\sigma^2(1 - M^2) + M^4}}{3\sigma + \sigma^2 + 2M^2},$$
$$(37a)$$

$$(\lambda_1)_{\mathcal{M}_{\text{upstream}}} = \frac{\sigma - M}{\sigma + M}, \quad (\lambda_{2,3})_{\mathcal{M}_{\text{upstream}}} = \frac{\sigma + \sigma^2 - M^2 \pm \sqrt{4\sigma^2(1 - M^2) + M^4}}{3\sigma + \sigma^2 + 2M^2}. \qquad (37b)$$

428

Note that $(\lambda_{2,3})_{\mathcal{M}_{\text{downstream}}} = (\lambda_{2,3})_{\mathcal{M}_{\text{upstream}}}$. We proceed by considering the eigenvalues for $\lim_{M\downarrow 0}$. With σ a finite (positive) constant this yields

$$\lim_{M\downarrow 0} (\lambda_1, \lambda_2, \lambda_3)_{\mathcal{M}_{\text{downstream}}, \sigma=\text{constant}} = \left(1, 1, \frac{-\sigma + \sigma^2}{3\sigma + \sigma^2}\right), \tag{38a}$$

$$\lim_{M\downarrow 0} (\lambda_1, \lambda_2, \lambda_3)_{\mathcal{M}_{\text{upstream}}, \sigma=\text{constant}} = \left(1, 1, \frac{-\sigma + \sigma^2}{3\sigma + \sigma^2}\right). \tag{38b}$$

For $\sigma = \alpha M$ with α a finite (positive) constant it yields

$$\lim_{M\downarrow 0} (\lambda_1, \lambda_2, \lambda_3)_{\mathcal{M}_{\text{downstream}}, \sigma=\alpha M} = \left(\frac{\alpha}{\alpha + 2}, 1, -\frac{1}{3}\right), \tag{39a}$$

$$\lim_{M\downarrow 0} (\lambda_1, \lambda_2, \lambda_3)_{\mathcal{M}_{\text{upstream}}, \sigma=\alpha M} = \left(\frac{\alpha - 1}{\alpha + 1}, 1, -\frac{1}{3}\right). \tag{39b}$$

So the choice $\sigma = $ constant yields two maximum eigenvalues equal to one, for both the downstream and upstream sweep. For $\sigma = \alpha M$ with α constant, this number is only one, which probably implies smaller Frobenius matrix norms (see e.g. Chapter 2 from Golub and Van Loan [2]) and hence better smoothing when applying two, three, four, ... Gauss-Seidel sweeps. Note that no function $\sigma = \sigma(M)$ exists which makes the moduli of all three eigenvalues smaller than one for $\lim_{M\downarrow 0}$. We proceed with $\sigma = \alpha M$. In the next section the optimal value of α is derived.

1.7.2 Quantitative optimization

To optimize α from $\sigma = \alpha M$, we continue to apply Fourier analysis for the highest error frequencies $|\theta| = \pi$, where, as in Section 1.3.1, we look for spectral radii of the amplification matrices. To avoid Mach-number dependence of α, we consider the moduli of the amplification matrices' eigenvalues integrated over the entire subsonic Mach-number range. (Avoiding Mach-number dependence by taking $\lim_{M\downarrow 0}$ does not allow α-optimization; from (39a) and (39b) it appears that the corresponding spectral radii of both $\mathcal{M}_{\text{downstream}}$ and $\mathcal{M}_{\text{upstream}}$ equal one, for any α.) In Figure 2 the distributions of the aforementioned eigenvalue integrals are depicted over the α-range $[0, 10]$. (Note that since $(\lambda_{2,3})_{\mathcal{M}_{\text{downstream}}} = (\lambda_{2,3})_{\mathcal{M}_{\text{upstream}}}$, the corresponding integrals are the same.) From Figure 2 it can be seen that the optimal value of α follows from $\int_0^1 |(\lambda_1)_{\mathcal{M}_{\text{upstream}}}| dM = \int_0^1 |(\lambda_2)_{\mathcal{M}_{\text{upstream}}}| dM$ (dashed line in Figure 2b), i.e. (after some computer algebra) from:

$$\alpha \frac{-4 + 2\alpha - 2\alpha^2 + \alpha^3 + (5 - 2\alpha^2)\ln\left(\frac{(\alpha+2)^2}{\alpha(2\alpha+1)}\right) - 3\sqrt{1 - 4\alpha^2}\ln\left(\frac{\sqrt{1-4\alpha^2}+1}{2\alpha}\right)}{(\alpha^2 + 2)^2} + \frac{\alpha - 1}{\alpha + 1} = 0. \tag{40}$$

From (40), it follows by good approximation that $\alpha = \frac{2}{5}$, and thus as (approximately) optimal σ:

$$\sigma = \frac{2}{5}M. \tag{41}$$

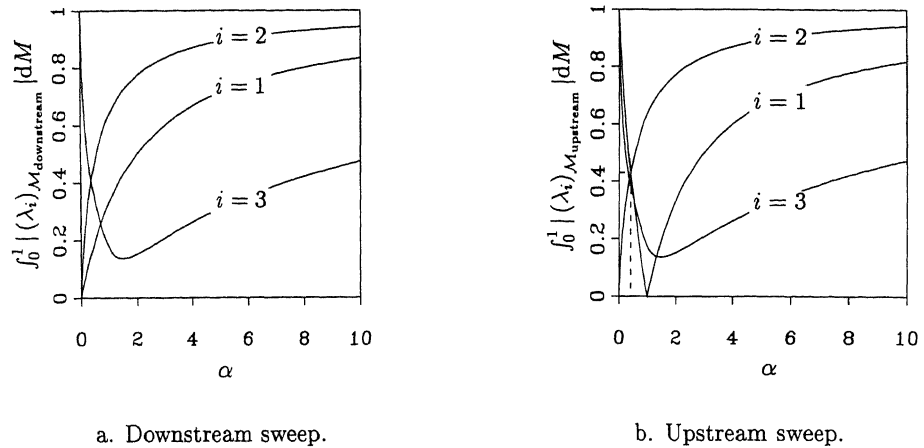a. Downstream sweep.   b. Upstream sweep.

Figure 2: Integrated moduli of eigenvalues of amplification matrices, for highest error frequencies.

1.8 Convergence for preconditioned equations

1.8.1 Error smoothing

Relation (41) implies as (approximately) optimal iteration step Δt:

$$\Delta t = \frac{5}{2} \frac{h}{|u|}, \tag{42}$$

i.e. CFL $= \frac{5}{2}$. We verify the smoothing behavior for this iteration step. This is done over the entire subsonic Mach-number range $(0,1)$, for the three error frequencies $\theta = \frac{\pi}{2}, \frac{3\pi}{4}$ and π. In Figure 3 the distributions of the corresponding spectral radii are depicted. Recalling from Section 1.3.1 that the spectral radii of downstream and upstream point Gauss-Seidel relaxation equal one over the entire subsonic Mach-number range, from Figure 3 it appears that the preconditioning does a good job.

1.8.2 Error convection

The locally implicit iteration applied to preconditioned Euler equations may be interpreted as physical time stepping. To do so, for simplicity we consider the common P according to (20) with $V = |u|$. Then, with CFL $= \frac{|u|\Delta t}{h}$, the iteration formulae (33a) and (33b) become

$$(1 + \text{CFL})|A|(q_i^{n+1} - q_i^n) = -A^+(q_i^n - q_{i-1}^{n+1}) - A^-(q_{i+1}^n - q_i^n), \tag{43a}$$

$$(1 + \text{CFL})|A|(q_i^{n+2} - q_i^{n+1}) = -A^+(q_i^{n+1} - q_{i-1}^{n+1}) - A^-(q_{i+1}^{n+2} - q_i^{n+1}). \tag{43b}$$

From (43a) and (43b) it appears that for this common P, the locally implicit time stepping can be directly interpreted as point Gauss-Seidel relaxation with underrelaxation factor $\omega = 1 + \text{CFL}$. I.e., even with CFL $= \mathcal{O}(1)$, (43a) and (43b) can still be interpreted as locally implicit physical time stepping at an infinitely large time step.

430
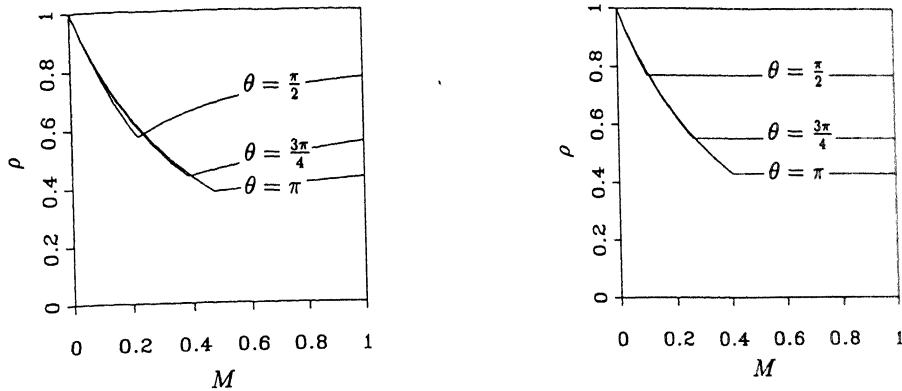
a. Downstream sweep.　　　　　b. Upstream sweep.

Figure 3: Spectral radii of amplification matrices, for downstream and upstream (approximately optimal) point Gauss-Seidel time-stepping, and three error frequencies.

2　Conditioning of absolute eigenvalue matrices

2.1　Introduction

In the zero-Mach-number limit, point-relaxation methods for solving discretized, steady Euler equations may suffer from ill-conditioning of the corresponding derivative matrices to be inverted. (For clarity: here derivative matrices are the local, absolute-eigenvalue matrices, so not the Jacobians.) To see this, we start by considering the perfect-gas, steady, 1-D Euler equations

$$\frac{\mathrm{d}f(Q)}{\mathrm{d}x} = 0, \tag{44}$$

with Q the conservative state vector (1b), $f(Q)$ the corresponding flux vector (1c) and e the internal energy (1d). Linearization of (44) with respect to the conservative variables and transformation from conservative variables Q to non-conservative (entropy) variables q according to (3), yields the steady, analytically tractable form

$$A\frac{\mathrm{d}q}{\mathrm{d}x} = 0, \tag{45}$$

with A according to (5b). Again we assume A to be constant and consider a first-order upwind, cell-centered finite-volume discretization of (45). Then, the discrete equation in cell Ω_i reads

$$A^+(q_i - q_{i-1}) + A^-(q_{i+1} - q_i) = 0, \tag{46}$$

with i running in positive x-direction, and with A^+ and A^- defined by (7a) and (7b), respectively. Applying point Gauss-Seidel relaxation to find the solution q_i of (46), for successively a downstream and upstream relaxation sweep we have the iteration formulae (10a) and (10b), respectively, with $|A|$ the matrix to be inverted;

$$|A| \equiv A^+ - A^- = \begin{pmatrix} c & u & 0 \\ u & c & 0 \\ 0 & 0 & u \end{pmatrix}. \tag{47}$$

431

For $0 < u < c$, it holds

$$\Lambda_{|A|} = \text{diag}(c - u, u, u + c), \tag{48}$$

and so $|A|$ has as condition over the entire subsonic flow regime (just as A):

$$K_{|A|}(M) = \max \left(\frac{1+M}{M}, \frac{1+M}{1-M} \right), \quad M \equiv \frac{|u|}{c} \in (0,1), \tag{49}$$

see also Figure 1. The best condition occurs at $M = \frac{1}{2}$; $K(M = \frac{1}{2}) = 3$, singularities occur at $M = 0$ and $M = 1$. Hence, in 1-D, in the neighborhood of the static and sonic flow conditions, when applying the iteration formulae (10a)-(10b), one may expect very large (too large) solution changes in case of very small right-hand sides only.

In 2-D numerical practice, ill-conditioning of derivative matrices to be inverted is not experienced in the neighborhood of $M = 1$, but only near $M = 0$. To get some evidence of this we also analyze the 2-D case. With $0 < u < c$, $0 < v < c$, a square finite volume, and j as additional running index in positive y-direction, one derives as iteration formulae for successively a downstream and upstream relaxation sweep:

$$(|A| + |B|)(q_{i,j}^{n+1} - q_{i,j}^n) = -A^+(q_{i,j}^n - q_{i-1,j}^{n+1}) - A^-(q_{i+1,j}^n - q_{i,j}^n) + \\ -B^+(q_{i,j}^n - q_{i,j-1}^{n+1}) - B^-(q_{i,j+1}^n - q_{i,j}^n), \tag{50a}$$

$$(|A| + |B|)(q_{i,j}^{n+2} - q_{i,j}^{n+1}) = -A^+(q_{i,j}^{n+1} - q_{i-1,j}^{n+1}) - A^-(q_{i+1,j}^{n+2} - q_{i,j}^{n+1}) + \\ -B^+(q_{i,j}^{n+1} - q_{i,j-1}^{n+1}) - B^-(q_{i,j+1}^{n+2} - q_{i,j}^{n+1}), \tag{50b}$$

with

$$A^+ = \frac{1}{2} \begin{pmatrix} u+c & u+c & 0 & 0 \\ u+c & u+c & 0 & 0 \\ 0 & 0 & 2u & 0 \\ 0 & 0 & 0 & 2u \end{pmatrix}, \quad A^- = \frac{1}{2} \begin{pmatrix} u-c & c-u & 0 & 0 \\ c-u & u-c & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{51a}$$

$$B^+ = \frac{1}{2} \begin{pmatrix} v+c & 0 & v+c & 0 \\ 0 & 2v & 0 & 0 \\ v+c & 0 & v+c & 0 \\ 0 & 0 & 0 & 2v \end{pmatrix}, \quad B^- = \frac{1}{2} \begin{pmatrix} v-c & 0 & c-v & 0 \\ 0 & 0 & 0 & 0 \\ c-v & 0 & v-c & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{51b}$$

So, in 2-D the matrix to be inverted is

$$|A| + |B| = A^+ - A^- + B^+ - B^- = \begin{pmatrix} 2c & u & v & 0 \\ u & v+c & 0 & 0 \\ v & 0 & u+c & 0 \\ 0 & 0 & 0 & u+v \end{pmatrix}. \tag{52}$$

For $\Lambda_{|A|+|B|}$ it holds when, without loss of generality, rotating in flow direction:

$$\Lambda_{|A|+|B|} = \text{diag} \left(\bar{u}, \frac{3}{2}c - \frac{1}{2}\sqrt{c^2 + 4\bar{u}^2}, \bar{u} + c, \frac{3}{2}c + \frac{1}{2}\sqrt{c^2 + 4\bar{u}^2} \right),$$

$$\bar{u} \equiv u \cos\phi + v \sin\phi, \quad \phi \equiv \arctan \frac{v}{u}, \tag{53}$$

and thus

$$K_{|A|+|B|}(M) = \max \left(\frac{3 + \sqrt{1 + 4M^2}}{2M}, \frac{3 + \sqrt{1 + 4M^2}}{3 - \sqrt{1 + 4M^2}} \right), \quad M \equiv \frac{|\bar{u}|}{c} \in (0,1), \tag{54}$$
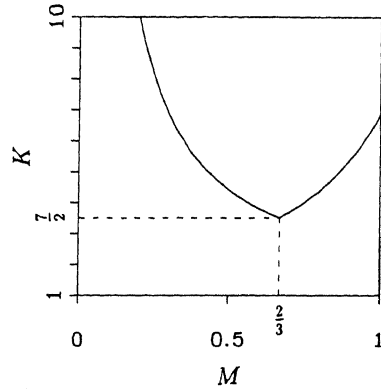
432

Figure 4: Condition of exact, 2-D absolute-eigenvalue matrix as a function of the Mach number.

see also Figure 4. We see that in 2-D the singularity at $M = 1$ no longer exists, which explains the aforementioned numerical experience. (The best condition occurs at $M = \frac{2}{3}$, $K(M = \frac{2}{3}) = \frac{7}{2}$.)

In the remainder of this paper, we discuss possible fixes to the 1-D and 2-D ill-conditionings of the absolute-eigenvalue matrices (Section 2.2), analyze the multigrid smoothing properties of a favorite fix (Section 2.3) and do some numerical experiments (Section 2.4). The first, already mentioned difference with the work presented in Section 1 is that whereas in Section 1 the condition of a non-absolute-eigenvalue matrix was improved, here the conditions of absolute-eigenvalue matrices are improved. Moreover, whereas in Section 1 the improvement was only made for 1-D, here it is done for 1-D and 2-D.

2.2 Fixes to ill conditions subsonic, absolute-eigenvalue matrices

2.2.1 Trimming 2-D singular matrix

For 2-D low-Mach-number flows, equations (50a)-(50b) can be simply regularized by (locally) dropping the entropy-equation part, and by replacing that, in case of e.g. (50a), by either the homentropic iteration formula

$$s_{i,j}^{n+1} - s_{i,j}^n \equiv \left(p_{i,j}^{n+1} - p_{i,j}^n\right) - \left(c_{i,j}^n\right)^2 \left(\rho_{i,j}^{n+1} - \rho_{i,j}^n\right) = 0, \qquad (55a)$$

or - alternatively - the incompressible formula

$$s_{i,j}^{n+1} - s_{i,j}^n = p_{i,j}^{n+1} - p_{i,j}^n. \qquad (55b)$$

Dropping the entropy equation from system (50a)-(50b), the corresponding derivative matrix to be inverted reduces to

$$|A| + |B| = \begin{pmatrix} 2c & u & v \\ u & v+c & 0 \\ v & 0 & u+c \end{pmatrix}, \qquad (56)$$

433

with, rotating again in flow direction

$$\Lambda_{|A|+|B|} = \mathrm{diag}\left(\frac{3}{2}c - \frac{1}{2}\sqrt{c^2 + 4\bar{u}^2}, \bar{u} + c, \frac{3}{2}c + \frac{1}{2}\sqrt{c^2 + 4\bar{u}^2}\right), \tag{57}$$

and thus

$$K_{|A|+|B|}(M) = \frac{3 + \sqrt{1 + 4M^2}}{3 - \sqrt{1 + 4M^2}}, \quad M \in (0, 1), \tag{58}$$

see also Figure 5. A difficulty of splitting off the singular part from the iteration formulae
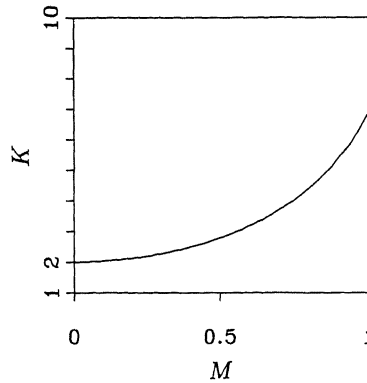

Figure 5: Condition of trimmed, 2-D, absolute-eigenvalue matrix as a function of the Mach number.

in case of general subsonic flows is that it requires the introduction of a monitor for switching on and off homentropy or incompressibility, i.e. (55a) or (55b). Rigorous formulae for setting thresholds for the monitors are hard to derive. Therefore we refrain from applying these reduced derivative matrices.

2.2.2 Adding 1-D and 2-D regularizing matrices

Considering the 1-D absolute-eigenvalue matrix (47), it can be regularized by adding a matrix R to it, leading to the approximate derivative matrix:

$$|A|_R \equiv |A| + R. \tag{59}$$

Taking

$$R = \beta \begin{pmatrix} 0 & -u & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c - u \end{pmatrix} \quad \text{or} \quad R = \beta \begin{pmatrix} 0 & 0 & 0 \\ -u & 0 & 0 \\ 0 & 0 & c - u \end{pmatrix}, \tag{60}$$

for any constant $\beta \in (0, 1]$ the singularities at $M = 0$ and $M = 1$ are removed. For both R's from (60) we find

$$\Lambda_{|A|_R} = \mathrm{diag}(c - \sqrt{1 - \beta}u, \beta c + (1 - \beta)u, c + \sqrt{1 - \beta}u). \tag{61}$$

However, since $|A|_R$ is not symmetric, its condition number does not equal the ratio of its largest and smallest eigenvalues. Here we use the general formula

434

$K_{|A|_R} = \parallel |A|_R \parallel \parallel |A|_R^{-1} \parallel$. Taking the L_∞-norm as the matrix norm, for both R's in (60) we derive

$$K_{|A|_R}(M) = \max\left(\frac{1+M}{\beta + (1-\beta)M}, \frac{(1+M)^2}{1-(1-\beta)M^2}\right), \quad M \in (0,1), \quad \beta \in (0,1], \quad (62)$$

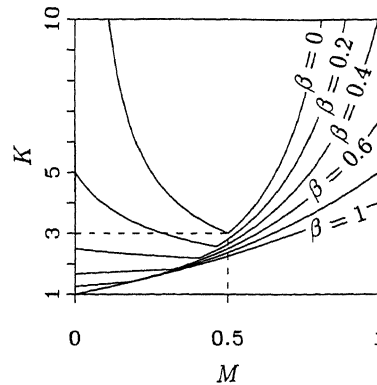see also Figure 6. For $\beta = 1$, $|A|_R$ is best conditioned over the entire subsonic Mach-

Figure 6: Condition of through addition regularized, 1-D absolute-eigenvalue matrix, as a function of the Mach number.

number range, whereas the corresponding approximate derivative matrix $|A|_R$ will generally be rather close to the exact derivative matrix $|A|$. A convergence requirement to be satisfied is that the eigenvalues of $|A|_R$ are positive. This requirement is met by both R's from (60), for any $\beta \in (0,1]$.

In 2-D, where no sonic singularity exists, to regularize (52) we may take

$$R = \beta \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c-u-v \end{pmatrix}, \quad \beta \in (0,1]. \quad (63)$$

So, in 2-D we have the advantage of remaining closer to the exact derivative matrix than in 1-D. For the corresponding eigenvalue-matrix $\Lambda_{(|A|+|B|)_R}$, where $(|A|+|B|)_R \equiv |A| + |B| + R$, it holds

$$\Lambda_{(|A|+|B|)_R} = \mathrm{diag}\left(\beta c + (1-\beta)\overline{u}, \frac{3}{2}c - \frac{1}{2}\sqrt{c^2 + 4\overline{u}^2}, \overline{u} + c, \frac{3}{2}c + \frac{1}{2}\sqrt{c^2 + 4\overline{u}^2}\right), \quad (64)$$

and hence

$$K_{(|A|+|B|)_R}(M) = \max\left(\frac{3 + \sqrt{1+4M^2}}{2(\beta + (1-\beta)M)}, \frac{3 + \sqrt{1+4M^2}}{3 - \sqrt{1+4M^2}}\right), \quad M \in (0,1), \quad \beta \in (0,1], \quad (65)$$

see also Figure 7. As opposed to the preceding 1-D conditioning, which is perfect for $\beta = 1$, perfect 2-D conditioning through (63) is not possible. In the next section, for 2-D flows, we will investigate the multigrid smoothing (high-frequency damping) properties of point Gauss-Seidel relaxation when applying additive conditioning. For reasons of transparency, smoothing properties are investigated for the 1-D equation.
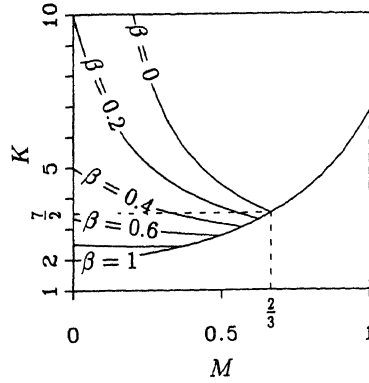
435

Figure 7: Condition of through addition regularized, 2-D absolute-eigenvalue matrix as a function of the Mach number.

2.3 Smoothing for additive conditioning

Consider the downstream iteration formula

$$(R + |A|)(q_i^{n+1} - q_i^n) = -A^+(q_i^n - q_{i-1}^{n+1}) - A^-(q_{i+1}^n - q_i^n), \tag{66a}$$

and the upstream formula

$$(R + |A|)(q_i^{n+2} - q_i^{n+1}) = -A^+(q_i^{n+1} - q_{i-1}^{n+1}) - A^-(q_{i+1}^{n+2} - q_i^{n+1}), \tag{66b}$$

where R is the 1-D equivalent of the 2-D additive matrix (63):

$$R = \beta \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c - u \end{pmatrix}. \tag{67}$$

To investigate the smoothing properties, the local solution error (11a) and its Fourier form (11b) are introduced again. Keeping the coefficient matrices in (66a) and (66b) frozen, with (11a) and (11b), from (66a) and (66b) it follows for the corresponding amplification matrices $\mathcal{M}_{\text{downstream}}$ and $\mathcal{M}_{\text{upstream}}$:

$$\mathcal{M}_{\text{downstream}} = \left(R + (1 - e^{-i\theta})A^+ - A^-\right)^{-1}\left(R - e^{i\theta}A^-\right), \tag{68a}$$

$$\mathcal{M}_{\text{upstream}} = \left(R + A^+ - (1 - e^{i\theta})A^-\right)^{-1}\left(R + e^{-i\theta}A^+\right). \tag{68b}$$

In both $\mathcal{M}_{\text{downstream}}$ and $\mathcal{M}_{\text{upstream}}$ the influence of β is confined to a single eigenvalue per matrix only:

$$\lambda_{\mathcal{M}_{\text{downstream}}}(\beta) = \frac{\beta(1 - M)}{\beta(1 - M) + (1 - e^{-i\theta})M}, \tag{69a}$$

$$\lambda_{\mathcal{M}_{\text{upstream}}}(\beta) = \frac{\beta(1 - M) + e^{-i\theta}M}{\beta(1 - M) + M}. \tag{69b}$$

It can be seen that for $\beta = 1$, it still holds

$$|\lambda_{\mathcal{M}_{\text{downstream}}}(\beta = 1)| \leq 1 \quad \text{and} \quad |\lambda_{\mathcal{M}_{\text{upstream}}}(\beta = 1)| \leq 1, \quad \forall |\theta| \in \left[\frac{\pi}{2}, \pi\right], \quad \forall M \in (0, 1). \tag{70}$$

436

We assume that $\beta = 1$ is an acceptable choice in 2-D as well. In the next section the conservative implementation of the 2-D additive conditioning is discussed and some numerical results, obtained for a 2-D stagnation flow, are given.

2.4 Applying additive conditioning

2.4.1 2-D conservative implementation

Discretizing the steady, 2-D, conservative Euler equations by a first-order upwind, cell-centered finite-volume method, and denoting the numerical flux functions which approximate the cell-face fluxes in x- and y-direction ($f(q_{i+\frac{1}{2},j})$ and $g(q_{i,j+\frac{1}{2}})$) by $F(q_{i,j}, q_{i+1,j})$ and $G(q_{i,j}, q_{i,j+1})$, respectively, the conservative upstream and downstream relaxation sweeps read:

$$\left[\frac{\mathrm{d}Q}{\mathrm{d}q}(q_{i,j}^n)R(q_{i,j}^n)\bar{h}_{i,j} + \frac{\partial F(q_{i,j}^n, q_{i+1,j}^n)}{\partial q_{i,j}^n}h_{i+\frac{1}{2},j} - \frac{\partial F(q_{i-1,j}^{n+1}, q_{i,j}^n)}{\partial q_{i,j}^n}h_{i-\frac{1}{2},j}+\right.$$
$$\left.+\frac{\partial G(q_{i,j}^n, q_{i,j+1}^n)}{\partial q_{i,j}^n}h_{i,j+\frac{1}{2}} - \frac{\partial G(q_{i,j-1}^{n+1}, q_{i,j}^n)}{\partial q_{i,j}^n}h_{i,j-\frac{1}{2}}\right]\left(q_{i,j}^{n+1} - q_{i,j}^n\right) =$$
$$F(q_{i-1,j}^{n+1}, q_{i,j}^n)h_{i-\frac{1}{2},j} - F(q_{i,j}^n, q_{i+1,j}^n)h_{i+\frac{1}{2},j}$$
$$+G(q_{i,j-1}^{n+1}, q_{i,j}^n)h_{i,j-\frac{1}{2}} - G(q_{i,j}^n, q_{i,j+1}^n)h_{i,j+\frac{1}{2}}, \qquad (71a)$$

$$\left[\frac{\mathrm{d}Q}{\mathrm{d}q}(q_{i,j}^{n+1})R(q_{i,j}^{n+1})\bar{h}_{i,j} + \frac{\partial F(q_{i,j}^{n+1}, q_{i+1,j}^{n+2})}{\partial q_{i,j}^{n+1}}h_{i+\frac{1}{2},j} - \frac{\partial F(q_{i-1,j}^{n+1}, q_{i,j}^{n+1})}{\partial q_{i,j}^{n+1}}h_{i-\frac{1}{2},j}+\right.$$
$$\left.+\frac{\partial G(q_{i,j}^{n+1}, q_{i,j+1}^{n+2})}{\partial q_{i,j}^{n+1}}h_{i,j+\frac{1}{2}} - \frac{\partial G(q_{i,j-1}^{n+1}, q_{i,j}^{n+1})}{\partial q_{i,j}^{n+1}}h_{i,j-\frac{1}{2}}\right]\left(q_{i,j}^{n+2} - q_{i,j}^{n+1}\right) =$$
$$F(q_{i-1,j}^{n+1}, q_{i,j}^{n+1})h_{i-\frac{1}{2},j} - F(q_{i,j}^{n+1}, q_{i+1,j}^{n+2})h_{i+\frac{1}{2},j}$$
$$+G(q_{i,j-1}^{n+1}, q_{i,j}^{n+1})h_{i,j-\frac{1}{2}} - G(q_{i,j}^{n+1}, q_{i,j+1}^{n+2})h_{i,j+\frac{1}{2}}, \qquad (71b)$$

where $\bar{h}_{i,j}$ is a cell-averaged mesh width, e.g. $\bar{h}_{i,j} = \frac{1}{4}\left(h_{i-\frac{1}{2},j} + h_{i+\frac{1}{2},j} + h_{i,j-\frac{1}{2}} + h_{i,j+\frac{1}{2}}\right)$, and where

$$\frac{\mathrm{d}Q}{\mathrm{d}q} = \begin{pmatrix} \frac{\rho}{c} & 0 & 0 & -\frac{1}{c^2} \\ \frac{\rho u}{c} & \rho & 0 & -\frac{u}{c^2} \\ \frac{\rho v}{c} & 0 & \rho & -\frac{v}{c^2} \\ \frac{1}{2}\frac{\rho(u^2+v^2)}{c} + \frac{1}{\gamma-1}\rho c & \rho u & \rho v & -\frac{1}{2}\frac{u^2+v^2}{c^2} \end{pmatrix}, \qquad (72)$$

yielding

$$\frac{\mathrm{d}Q}{\mathrm{d}q}R = \begin{pmatrix} 0 & 0 & 0 & -\frac{c-u-v}{c^2} \\ 0 & 0 & 0 & -\frac{u(c-u-v)}{c^2} \\ 0 & 0 & 0 & -\frac{v(c-u-v)}{c^2} \\ 0 & 0 & 0 & -\frac{1}{2}\frac{(u^2+v^2)(c-u-v)}{c^2} \end{pmatrix}. \qquad (73)$$

In 3-D, we have

$$\frac{\mathrm{d}Q}{\mathrm{d}q}R = \begin{pmatrix} 0 & 0 & 0 & 0 & -\frac{c-u-v-w}{c^2} \\ 0 & 0 & 0 & 0 & -\frac{u(c-u-v-w)}{c^2} \\ 0 & 0 & 0 & 0 & -\frac{v(c-u-v-w)}{c^2} \\ 0 & 0 & 0 & 0 & -\frac{w(c-u-v-w)}{c^2} \\ 0 & 0 & 0 & 0 & -\frac{1}{2}\frac{(u^2+v^2+w^2)(c-u-v-w)}{c^2} \end{pmatrix}. \qquad (74)$$

2.4.2 Numerical results

A suitable test case is steady, 2-D stagnation flow normal to a flat plate (Figure 8). A favorable property of this test case is the direct availability of good approximate
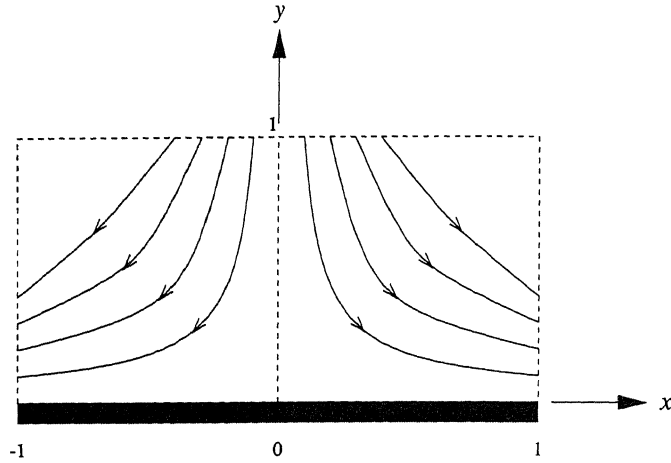

Figure 8: Stagnation flow against a flat plate.

boundary conditions (because of the availability of an exact, incompressible potential flow solution, see e.g. Chapter X from Prandtl and Tietjens [8]). For computational efficiency, we only compute the half problem ($x \geq 0$). (We remark that exact solutions of subsonic flows along a kinked wall have a singularity at the kink for all kink angles δ except for $\delta = 0$ and $\delta = \frac{\pi}{2}$, which latter case is identical to the present normal stagnation flow.) Introducing as known quantities in the point $(x, y) = (1, 1)$: a reference speed V_{ref}, a reference density ρ_{ref}, and a reference Mach number M_{ref}, the boundary conditions imposed are:

- at the inflow boundary, assuming homenthalpy:

$$u(x, y = 1) = V_{ref}x, \tag{75}$$

$$v(x, y = 1) = -V_{ref}, \tag{76}$$

$$c(x, y = 1) = \sqrt{\frac{V_{ref}^2}{M_{ref}^2} + \frac{\gamma - 1}{2}(V_{ref}^2 - u^2(x, y = 1) - v^2(x, y = 1))}, \tag{77}$$

- at the outflow boundary, assuming homentropy and homenthalpy:

$$p(x = 1, y) = \left(1 + \frac{\gamma - 1}{2}\frac{w^2(x = 1, y)}{c^2(x = 1, y)}\right)^{-\frac{\gamma}{\gamma - 1}} p_t, \tag{78a}$$

where

$$w^2(x = 1, y) = V_{ref}^2(1 + y^2), \tag{78b}$$

$$c^2(x = 1, y) = \frac{V_{ref}^2}{M_{ref}^2} + \frac{\gamma - 1}{2}\left(V_{ref}^2 - w^2(x = 1, y)\right), \tag{78c}$$

438

$$p_t = \left(1 + \frac{\gamma - 1}{2} M_{\text{ref}}^2\right)^{\frac{\gamma}{\gamma - 1}} p_{\text{ref}}, \quad p_{\text{ref}} = \frac{1}{\gamma} \frac{V_{\text{ref}}^2}{M_{\text{ref}}^2} \rho_{\text{ref}}, \tag{78d}$$

- at the vertical-wall boundary:

$$u(x = 0, y) = 0, \tag{79}$$

- at the lower-wall boundary:

$$v(x, y = 0) = 0. \tag{80}$$

In Figure 9, for two low-subsonic (though not yet very low-subsonic) values of M_{ref}, we give the convergence behaviors of the point relaxation in some arbitrary cell, at some arbitrary instant in the iteration process. (The residual considered is that of the energy equation.) From the results it appears that the additive conditioning does a good job.
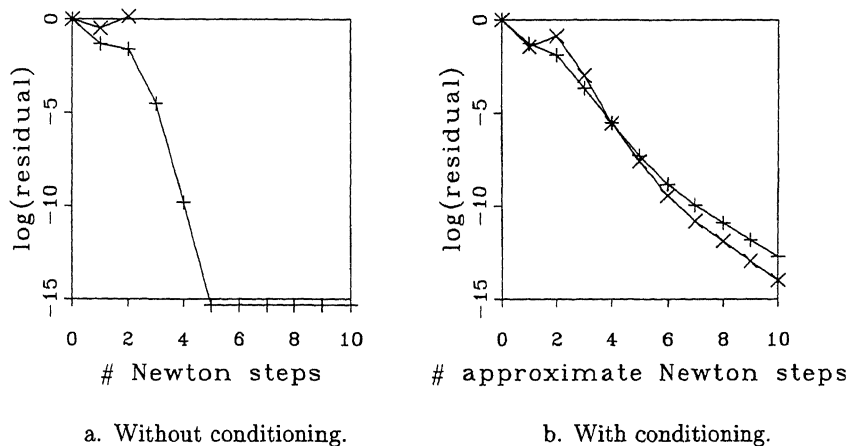

| a. Without conditioning. | b. With conditioning. |

Figure 9: Convergence behaviors point relaxation, +: $M_{\text{ref}} = 0.1$, ×: $M_{\text{ref}} = 0.075$.

Though quadratic convergence is lost, the divergence that occurs at $M_{\text{ref}} = 0.075$ (Figure 9a) has disappeared by applying the conditioning (Figure 9b).

In Figure 10, convergence results are presented, as obtained through the conditioned relaxation method accelerated by nonlinear multigrid. (The residual considered is the L_∞-norm of the energy equation's residual field.) The Mach-number sequence considered is: $M_{\text{ref}} = 0.5, 0.05, 0.005$. Note that the method does not break down, but still converges in the very low-subsonic case $M_{\text{ref}} = 0.005$.

3   Conclusions

- Poor convergence of multigrid-accelerated point Gauss-Seidel relaxation at low Mach numbers is explained by the relaxation's poor smoothing at low Mach numbers and by the likewise poor entropy-error expulsion across domain boundaries.

- Poor solution accuracy known to occur at low Mach numbers can be explained by means of the modified equation for the 1-D Euler equations, discretized by a
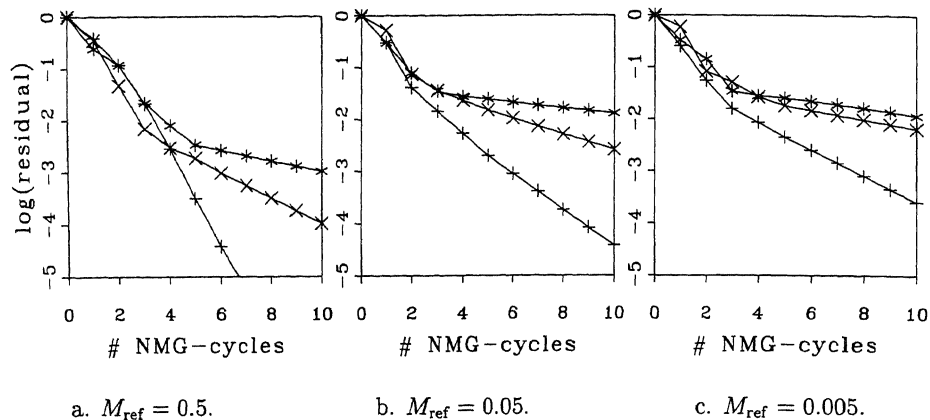
a. $M_{ref} = 0.5$.   b. $M_{ref} = 0.05$.   c. $M_{ref} = 0.005$.

Figure 10: Convergence behaviors nonlinear multigrid iteration, +: $h = \frac{1}{8}$, ×: $h = \frac{1}{16}$, *: $h = \frac{1}{32}$.

first-order accurate flux-difference splitting scheme. For flows with uniformly low Mach numbers, a fix to this inaccuracy is a necessity. For flows of which the global solution error is not affected by the occurrence of low-subsonic flow regions, it may not be necessary.

- For the latter flows, implementation of preconditioning in a locally implicit time stepping method with the inverse of the preconditioning matrix working on the time operator, may be practical. It allows the application of an off-the-shelf space discretization method.

- Local-mode analysis shows that optimal high-frequency damping for locally implicit "time" stepping in a Gauss-Seidel way, is obtained for CFL $\approx \frac{5}{2}$. (When preconditioning with the 1-D matrix $P = |u||A|^{-1}$, the locally implicit "time" stepping boils down to point Gauss-Seidel relaxation with underrelaxation factor $1 + \text{CFL}$.)

- Given the direct availability of the 2-D and 3-D extensions of the 1-D preconditioning matrix analyzed, the present improved solution method is directly extendible to multi-D.

- Two methods have been proposed for removing singularities in local, absolute-eigenvalue matrices of upwind-discretized Euler equations: (i) elimination of the entropy-equation part from the exact, 2-D derivative matrix, (ii) addition of a singular matrix (which is very close to the zero matrix) to the singular, exact derivative matrix. The first fix does not work in 1-D. Another drawback is that its successful application requires tuning. The second fix is free of tuning parameters and may remove the ill-conditioning without deteriorating too much the quadratic convergence rate of exact Newton iteration. The latter fix has been successfully applied to a steady, 2-D, low-subsonic stagnation flow.

440

References

[1] BERS, L.: *Mathematical Aspects of Subsonic and Transonic Gas Dynamics* (Wiley, London, 1958).

[2] GOLUB, G.H. AND VAN LOAN, C.F.: *Matrix Computations* (The John Hopkins University Press, Baltimore, 1989).

[3] HEMKER, P.W. AND KOREN, B.: Defect correction and nonlinear multigrid for the steady Euler equations, in *Solution Techniques for Large-Scale CFD Problems* (HABASHI, W.G., ED.), 273-291 (Wiley, Chichester, 1995).

[4] HEMKER, P.W. AND SPEKREIJSE, S.P.: Multiple grid and Osher's scheme for the efficient solution of the steady Euler equations, *Appl. Numer. Math.*, **2**, 475-493 (1986).

[5] LEER, B. VAN, LEE, W.-T. AND ROE, P.L.: Characteristic time-stepping or local preconditioning of the Euler equations, *AIAA Paper 91-1552* (1991).

[6] MAJDA, A.: *Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables* (Springer, Berlin, 1984).

[7] OSHER, S. AND SOLOMON, F.: Upwind difference schemes for hyperbolic systems of conservation laws, *Math. Comput.*, **38**, 339-374 (1982).

[8] PRANDTL, L. AND TIETJENS, O.G.: *Fundamentals of Hydro- and Aeromechanics* (Dover, New York, 1957).

[9] ROE, P.L.: Approximate Riemann solvers, parameter vectors, and difference schemes, *J. Comput. Phys.*, **43**, 357-372 (1981).

[10] TURKEL, E.: Preconditioned methods for solving the incompressible and low speed compressible equations, *J. Comput. Phys.*, **72**, 277-298 (1987).

[11] TURKEL, E.: Review of preconditioning methods for fluid dynamics, *Appl. Numer. Math.*, **12**, 257-284 (1993).

[12] TURKEL, E., FITERMAN, A. AND LEER, B. VAN: Preconditioning and the limit of the compressible to the incompressible flow equations, for finite difference schemes, in *Computing the Future: Advances and Prospects for Computational Aerodynamics* (HAFEZ, M.M. AND CAUGHEY, D.A., EDS.), 215-234 (Wiley, Chichester, 1994).

[13] VOLPE, G.: Performance of compressible codes at low Mach numbers, *AIAA J.*, **31**, 49-56 (1993).

[14] WESSELING, P.: *An Introduction to Multigrid Methods* (Wiley, Chichester, 1992).