A PRECONDITIONING MASS MATRIX TO ACCELERATE
THE CONVERGENCE TO THE STEADY EULER SOLUTIONS
USING EXPLICIT SCHEMES

M. A. STORTI†
C. E. BAUMANN†
S. R. IDELSOHN ‡

Mechanical Laboratory of INTEC
(Universidad Nacional del Litoral and CONICET)
Güemes 3450 - 3000 Santa Fe - Argentina

ABSTRACT

When explicit time marching algorithms are used to reach the steady state of problems governed by the Euler eqns, the rate of convergence is strongly impaired both in the zones with low Mach number and in the zones with transonic flow, let say $\text{Mach} \leq \alpha$ and $|\text{Mach} - 1| \leq \alpha$, with $\alpha \leq .2$. The rate of convergence becomes slower as $\alpha$ diminishes.

We show in this paper, with analytical and numerical results, how the use of a preconditioning mass matrix accelerates the convergence in the aforementioned ranges of Mach numbers.

The Preconditioning Mass Matrix (PMM) we advocate in this paper can be applied to any FEM/FVM that uses an explicit time-marching scheme to find the steady state. The method's rate of convergence to the steady state is studied, and results for the one- and two-dimensional cases are presented.

In section 1, using the one-dimensional Euler eqns, we first explain why there exists a slow rate of convergence when the plain lumping of mass is used. Then the convergence rate to steady solutions is analyzed from its two constituents, that is, convergence by absorption at the boundaries and by damping in the domain. Next we give the natural solution to this problem, and with several examples we show the effectiveness of the proposed mass matrix when compared with the plain scheme.

In section 2 we give the multidimensional version of the preconditioning mass matrix. We make a stability analysis and compare the group velocities and damping with and without the new mass matrix. To finish, we show the velocity of convergence for a common test problem.

† Graduate Research Assistant
‡ Professor and Scientific Researcher
1. ONE DIMENSIONAL CASE

1.1 Rate of convergence, statement of the problem:

The 1-D Euler eqns can be written in conservation form as

\[ U_{t} + F_{x} = U_{x} + A U_{x} = 0, \]

(1)

where \( U \) is the vector of conservation variables, \( F \) the flux vector, and \( A \) the Jacobian of the flux vector.

For the sake of simplicity, we use in this explanation the Steger-Warming splitting spatially differenced according to the first-order one-sided upwind scheme; nevertheless, the results that are to be drawn hold for any flux splitting, let say Van Leer's for example, and for any consistent spatial differencing, let say MUSCL-type or first- and second-order one-sided differences, or any other [1-2]. Using the forward Euler discretization in time and one-sided first-order differentiation in space, for a uniform grid spacing \( h \), we get

\[ U_{j}^{n+1} = U_{j}^{n} - \frac{\Delta t}{h} (F_{j}^{+} - F_{j-1}^{-} + F_{j+1}^{-} - F_{j}^{-}), \]

(2)

in which

\[ F_{j}^{\pm} = (SA^{\pm}S^{-1}) j \ U_{j}, \quad \text{with} \quad A^{\pm} = (A \pm |A|)/2, \]

here, the \( A_{j} \)'s are the matrices that contain the eigenvalues of the Jacobians \( A_{j} \)s.

To make a linearized analysis of Eq (2), we consider the Jacobian matrix a constant, and using the similarity transformation \( V_{j} = S^{-1} U_{j} \), we get the system decoupled in three advection eqns, that is

\[ v_{\mu j}^{n+1} = v_{\mu j}^{n} - \frac{\Delta t \lambda_{\mu}}{h} \delta_{\mu} v_{\nu j}^{n}, \quad \text{with} \quad \delta_{\mu} v_{\nu j} = \begin{cases} v_{\mu j}^{n} - v_{\mu j-1}^{n} & \text{if} \ \lambda_{\mu} > 0 \\ v_{\mu j+1}^{n} - v_{\mu j}^{n} & \text{if} \ \lambda_{\mu} < 0 \end{cases} \]

here, \( v_{\mu j} \) stand for the \( \mu \)-th component of \( V_{j} \).

The maximum \( \Delta t \) allowable is set by the condition limit of stability \( CFLN = \Delta t |\lambda_{\mu}|/h \leq 1 \), and considering that \( A = \text{diag}(u, u + c, u - c) \), we get \( \Delta t_{\text{max}} = h/(|u| + c) \). If \( \Delta t_{\text{max}} \) is used, the wave corresponding to the eigenvalue \( (|u| + c) \) moves at a velocity of one grid spacing per time step, while the slowest wave \( \min.(|u|, |u| - c) \) moves at

\[ \min.\{|u|/(|u| + c), (|u| - c)/(|u| + c)\} = \min.\{(M/(M + 1), (|M - 1|/(M + 1)) \}

elements per time step. It follows that when the Mach number \( M = |u|/c | \) is relatively low or it approaches the unity, there will be an extremely slow wave that can be annihilated only by damping. But, as is demonstrated in the following paragraph, if \( N \) is the characteristic number of elements in the mesh, the convergence by damping needs \( O(N) \) times work units as compared with the
convergence by absorption at the boundaries, therefore we have a very slow rate of convergence.

1.2 One-dimensional linear scalar case, rates of convergence:

Consider the advection equation

$$u_t + au_x = 0 \quad \text{on} \quad -\infty < x < \infty$$

where $a > 0$ is the constant transport velocity and $u$ a scalar quantity. We may discretize the Cauchy problem defined in Eq. (3) using one-sided first-order differentiation in space and the forward Euler scheme in time, that is,

$$u_j^{n+1} - u_j^n = -\frac{a\Delta t}{h}(u_j^n - u_{j-1}^n), \quad j = -N + 1, \ldots, 0, \ldots, N - 1, N,$$  (4)

and

$$u_N^{n+1} = u_N^n,$$

here $x = jh$, and $t = n\Delta t$. The adimensional number $\frac{a\Delta t}{h}$ is the Courant number, noted as $C$ hereafter. We imposed periodic boundary conditions in the discrete problem since in the sequel we will make a Fourier analysis. The exact solution of Eq. (4) may be written in terms of Fourier components, that is,

$$u_j^n = \sum_k \hat{u}(k) e^{i(kj - \omega\Delta t)}, \quad k = \frac{lx}{L}, \quad l = -N, \ldots, 0, \ldots, N - 1, N,$$

where $i = \sqrt{-1}$, $k$ is the wave number, $\omega$ the frequency, and $L = Nh$. In the same way, the error of any given solution can be resolved by Fourier analysis in the harmonic components $u(k)e^{i(kx - \omega t)}$, and anyone of these components introduced in eq. (2) leads to

$$u_j^{n+1}(k) = [C(e^{-ik\Delta t}) - 1]u_j^n(k) = Gu_j^n(k),$$  (5)

here, $G$ is the amplification factor, which must verify the condition $|G| \leq 1$ to have stability. The last condition implies $C \leq 1$.

We now turn our attention to the velocity of propagation of the harmonic components, that is,

$$a^* = \frac{\text{Re}(\omega(k))}{k}.$$  

From Eq. (5) we have

$$e^{-i\omega\Delta t} = C(e^{-ik\Delta t} - 1) + 1 = Ae^{-i\varphi},$$  (6)

here,

$$A = \sqrt{(C(\cos(\varphi) - 1) + 1)^2 + (C\sin(\varphi))^2},$$

and

$$\varphi = \arctan\left(\frac{C\sin(\varphi)}{(C(\cos(\varphi) - 1) + 1)}\right).$$
Taking the limit when $kh \to 0$ and $C \to 0$, we get from Eq. (7) $a^* = a$, the expected result to have consistency assured.

To evaluate the velocity of convergence we must use the Group Velocity (GV), that is, the velocity at which a wave packet moves (a wave packet is composed of a number of short wavelength oscillations modulated by a slowly varying envelope). The GV is given by the following formula (see [3,4])

$$GV = \frac{Re(\xi)}{k} = Re\left(\frac{\partial \xi}{\partial k}\right),$$

where the frequency $\xi$ and the wave number $k$ are characteristic values of the modulating envelope, while $\omega$ and $k$ are the corresponding values for the short wavelength oscillations.

From Eq. (6) we obtain

$$\frac{\partial \omega}{\partial t} = C \epsilon^{-ikh} e^{i \omega t},$$

and

$$Re\left(\frac{\partial \omega}{\partial k}\right) = \frac{a}{\Delta t} \{\cos(kh)[C\cos(kh) - 1] + C\sin^2(kh)\}. \tag{8}$$

In Fig. (1) we can see plots of GV vs C with $kh$ constant.

To evaluate the rate of convergence by absorption at the boundaries, we consider the number of iterations needed for the reduction of one order in the amplitude of the relatively smooth harmonic components, because the high frequency modes are damped out in a few iteration using $C = 1/2$ as we will see later. The formula that gives us the number of iterations is the following

$$N_{\text{abs. bou.}} = \frac{L}{|GV| \Delta t} = \frac{L}{hC} = \frac{N}{C}, \tag{9}$$

where $L$ represents the domain characteristic length that a wave component must travel to reach the boundary. In the above equation we supposed that the wave is not totally absorbed at the boundary, but a 10 per cent of its amplitude is reflected back into the domain.

Besides of being absorbed at the boundary, the harmonic components may be damped out as they travel through the mesh. From equations (4) and (6), we have

$$\frac{|u_j^{n+1}|}{|u_j^n|} = |G| = A = \sqrt{(C\cos(kh) - 1) + 1} + (C\sin(kh))^2, \tag{10}$$

therefore, those harmonic components for which $|G| < 1$, will be damped out as they travel through the mesh. The value of $A^2 = f(kh,C)$ is shown in Fig. (2).

From Eq. (5) we see that $G$ varies linearly with $C$ from (1, i0) for $C = 0$ to $e^{-ikh}$ for $C = 1$, therefore, for any $k$, the distance to the origin $|G|$ will be minimum for $C = 1/2$. Using $C = 1/2$, the max.|G| is the value that limits the rate of
convergence by damping, and this value is reached for $k = \frac{\pi}{2}$ (the case $k = 0$ is not considered).

The rate of convergence by damping is measured as the number of iterations needed to reduce in one order the amplitude of the more slowly decaying by damping harmonic component, that is,

$$\left| G \right|^{N_{\text{damp}}} = \frac{1}{2} (e^{-i2k} + 1)^{N_{\text{damp}}} = 10^{-1},$$

it follows that

$$N_{\text{damp}} = \frac{8N^2 \log 10}{\pi^2}. \quad (11)$$

1.3 Rate of convergence, conclusions:

From the above study we can see that the convergence by damping is extremely slow for smooth errors. For $k < \pi/\sqrt{N}$ the convergence by damping is slower than the convergence by absorption, and the lower the $k$ the slower is the convergence. This means that after all the modes of high frequency have been damped out, the convergence by damping is insignificant and we can reach the exact solution (to reduce the residual several orders) only by advection and absorption at the boundaries. (Note: we can also work out the problem using a multi-grid technique for example.)

From Eq. (9) we see that the number of iterations needed to transport the error component to the boundaries is inversely proportional to the Courant number, therefore, we must use the maximum allowable time step. On the other hand, we know that when a system of eqns, like the Euler eqns, has different eigenvalues, only the eigenmode that has the greatest eigenvalue is integrated with $CFLN = 1$, all the others are integrated with $CFLN < 1$ as was explained (see 1.1).

2. PRECONDITIONING MASS MATRIX

As solution to the aforementioned problem, we propose to use a preconditioning mass matrix which introduced in Eq.(2) give us the following scheme:

$$\frac{1}{(|u|+c)A_j} U_{j}^{n+1} - U_{j}^{n} = -\frac{\Delta t}{h} (F_{j}^{+} - F_{j-1}^{+} + F_{j+1}^{-} - F_{j}^{-}). \quad (12)$$

As before (see 1.1), if we make a linearized analysis of Eq. (12), we obtain

$$v_{p_j}^{n+1} = v_{p_j}^{n} - \frac{\Delta t(|u|+c)}{h} \text{sign}(\lambda_{\mu}) \delta_{p} v_{p_j}^{n}, \quad (13)$$

and now, using $\Delta t_{\text{max}}$ set by the condition $CFLN \leq 1$, all waves move at the same, maximum allowable velocity; that is, one grid space per time step for $CFLN = 1$. 
3. NUMERICAL RESULTS

To compare numerically the rate of convergence of the proposed scheme with that of the original, we ran several test cases, all were run using absorbing boundary conditions and a $CFLN = .90$.

The first is the case of a $Mach = 0.9$ uniform flow perturbed with a pressure peak of high value. The pressure profiles for the original and proposed scheme are shown every two iterations in Figs. (3) and (4), respectively.

For the original scheme, the wave of eigenvalue $u - c = 0.1c$ is 11 times slower than that of the eigenvalue $u + c = 1.1c$, on the other hand, with the proposed scheme both waves have the same velocity. The density profiles have the same pattern of wave velocities, that is, for the original scheme there are two slow components, $u - c = -0.1c$ and $u = 0.1c$, whereas with the proposed scheme all wave velocities are the same.

The next test is to check the velocity of formation of a shock wave. The upstream and downstream boundary conditions correspond with Mach numbers of 1.10 and 0.91, respectively. The initial values of the state variables on the central part of the domain are a linear interpolation between the extreme values. The pressure profiles for the original and proposed scheme are shown every two iterations in Figs. (5) and (6). The resulting profiles talk by themselves.

Another interesting case is the passing or a perturbation through a shock wave. This case was selected mainly because the linearized analysis can not predict the behavior of the scheme in a flow discontinuity. The perturbation is a pressure peak and the shock is from Mach=1.5 to 0.70.

The pressure profiles for the original and proposed scheme are shown every two iterations in Figs. (7) and (8), respectively. We can see that the scheme works equally well in this case, and again the rate of convergence is extremely high as compared with the original scheme.

The last case is the velocity of movement of a shock wave whose extreme Mach numbers (boundary conditions) are not compatible. The pressure profiles for the original and proposed scheme are shown every two iterations in Figs. (9) and (10). The original scheme gives a shock velocity extremely slow when compared with the given by the new scheme.

With a few examples we have shown numerically the fast rate of convergence given by the proposed modification to the explicit time-stepping scheme. Using multigrid algorithms, we add to the intrinsic high rate of convergence of the algorithm, the aforementioned property of equal velocity of propagation for all the wave components.
4. TWO DIMENSIONAL CASE

4.1 Preconditioning Mass Matrices (PMM) - Diagonalizable systems:

Let's consider a system of PDEs of hyperbolic type written in conservation form

\[ U_{xt} + A_{x} U_{xt} = 0, \quad \forall (x_{i}, t) \in \Omega \times (0, \infty), \quad (14) \]

if the system is diagonalizable there exists a similarity transformation that simultaneously diagonalizes all the Jacobians, that is, if \( T = \frac{\partial U}{\partial \mathbf{V}} \), then \( U_{\mathbf{x}} = TV_{\mathbf{x}} \), and introducing this transformation in Eq. 14 we get

\[ V_{\mathbf{x}t} + B_{\mathbf{x}} V_{\mathbf{x}t} = 0, \quad (15) \]

where

\[ B_{\mathbf{x}} = T^{-1} A_{\mathbf{x}} T, \]

and now the \( B_{\mathbf{x}} \)s are diagonal matrices.

Making a linearized analysis of Eq. 15, we can see that the new system is a system of decoupled, scalar, advection equations.

Now if the diagonal matrix \( C = (B_{\mathbf{x}} B_{\mathbf{x}})^{1/2} \) has elements (its eigenvalues) of different magnitude, physically this means that the transport velocity of the decoupled scalar eqns are different, the max. time step for the explicit integration schemes is imposed by the maximum diagonal element, while the number of iterations needed to reach convergence is imposed by the minimum diagonal element. Therefore, the velocity of convergence to the steady solutions is inversely proportional to the condition number of \( C \), that is, the greater the quotient \( (\text{max} C_{ii}/\text{min} C_{ii}) \) the slower the rate of convergence. The natural solution in this case is to use \( C \) as a preconditioning mass matrix, then Eq. 15 is modified as follows

\[ CV_{\mathbf{x}t} + B_{\mathbf{x}} V_{\mathbf{x}t} = 0, \quad (16) \]

and we can see that only the temporal part is modified while the condition number of the new system is one. The PMM to be used with the original scheme is

\[ M = T(B_{\mathbf{x}} B_{\mathbf{x}})^{1/2} T^{-1} = (A_{\mathbf{x}} A_{\mathbf{x}})^{1/2} \]

4.2 (PMM) - Nondiagonalizable systems of PDEs:

Design conditions for the PMM:

1) If the system is of Hyperbolic type, then the change in the temporal part must not change the hyperbolic condition of the system of PDEs.

2) The number of Dirichlet boundary conditions to be specified on the boundaries must not be changed by the use of the PMM.

The first condition is related with stability. A system of PDEs of order one must be hyperbolic (real eigenvalues), otherwise the system is unconditionally unstable.
For the case of the Euler eqns, it is well known that there exists a similarity transformation that symmetrizes simultaneously all the Jacobians (using the so called entropy variables [5-7]). Writing the Euler system in this base, we obtain

$$
\hat{M}\hat{p} + \hat{A}_{x_j} \hat{U}_{x_j} = 0,
$$

and introducing in Eq. 17 a typical solution $e^{i(k_j x_j - \omega t)}$, we get

$$
i[-\omega \hat{M} + k_j \hat{A}_{x_j}]e^{i(k_j x_j - \omega t)} = 0.
$$

Considering that any linear combination of $\hat{A}_{x_j}$s is symmetric, then the eigenvalues $\omega$ of the eigenvalue problem

$$
|k_j \hat{A}_{x_j} - \omega \hat{M}| = 0,
$$

are real on the condition that $\hat{M}$ be a real, symmetric, definite matrix.

The second condition to be fulfilled by the PMM has to be with the number of ingoing characteristics at the boundary. This number is the number of negative eigenvalues of the matrix $n_i A_{x_i}$, where the $n_i$s are the components of the unit outward normal to the boundary. We will show that this condition is fulfilled by positive definite, symmetric $M$s.

Using the change of variables defined by $\hat{M}^{-1/2} = \frac{\hat{e}}{\hat{e}^*}$, then Eq. 17 can be written as

$$
\hat{\psi}_x + \hat{B}_{x_i} \hat{\psi}_{x_i} = 0,
$$

where

$$
\hat{B}_{x_i} = \hat{M}^{-1/2} \hat{A}_{x_i} \hat{M}^{-1/2},
$$

and from Eq. 19 it can be demonstrated using the eigenvalue separation property, and the theory of determinants, that for the case of simultaneously symmetrizable systems, the number of negative eigenvalues of $M^{-1/2}(n_i A_{x_i})M^{-1/2}$ is the same as that one of $n_i A_{x_i}$ with the only condition that $M$ be positive definite with real eigenvalues.

Condition (1) imposes the symmetry and definiteness of $M$ in the base where the $\hat{A}_{x_j}$s are symmetric. Condition (2) imposes only the positive definiteness because being $M$ symmetric it is diagonalizable and has real eigenvalues.

5. GROUP VELOCITIES

5.1 Continuous problem:

The analysis of the group velocities for the continuous problem is very important because for small wave numbers the consistent discretizations approach the differential operators, and on the other hand, only the convergence by absorption at the boundaries works because this is in the nature of the hyperbolic systems.

All the consistent schemes of discretization in space reduce to the continuous differential operators when the size of the discretization tends to zero. With regard to the discretization in time, we can use any nonconsistent scheme for driving the
explicit algorithm. The only conditions that these schemes have to verify are those related with stability and convergence.

We rewrite the Euler eqns as follows

$$\hat{M}\dot{\bar{U}} + \hat{A}_x \bar{U}_x + \hat{A}_y \bar{U}_y = 0,$$  \hspace{1cm} (20)

here, \(\hat{M}\) represents the general run of preconditioning mass matrices that generally give the nonconsistent time-marching schemes we were referring to.

To study the group velocities for the continuous case, we must know the numerical frequency as a function of the vector wave number \(k\). We can obtain \(\omega = \omega(k)\) by introducing in Eq. 20 a typical Fourier component \(\bar{U} e^{i(k_j x - \omega t)}\), that is

$$[-i\omega \hat{M} e^{i(k_j x - \omega t)} + (i k_1 \hat{A}_x + i k_2 \hat{A}_y) e^{i(k_j x - \omega t)}] \bar{U} = 0,$$

therefore, the frequencies can be obtained from the following eigenvalue problem

$$|k_1 \hat{A}_x + k_2 \hat{A}_y| = \omega \hat{M} = 0,$$  \hspace{1cm} (21)

and the group velocities, that is, the velocities of wave packets representing part of the solution error, are [3-4]:

$$GV_{x_j} = \frac{\partial \omega_{\mu}}{\partial x_j},$$  \hspace{1cm} (22)

here, we have the group velocity of the \(\mu\)-th eigenmode in the \(x_j\)-th cartesian direction. It is evident from Eq. 21 that a scalar multiple in \(k\) modifies \(\omega\) as follows

$$\omega_{\mu}(\lambda k) = \lambda \omega_{\mu}(k),$$  \hspace{1cm} (23)

therefore the \(\omega_{\mu}\)s are homogeneous functions of degree one in \(k\). Taking derivatives of Eq. 23 with respect to \(k_j\), we obtain

$$\frac{\partial}{\partial k_j} (\omega_{\mu}(\lambda k)) = \lambda \frac{\partial}{\partial k_j} (\omega_{\mu}(k)),$$

$$\lambda \frac{\partial}{\partial k_j} (\omega_{\mu}(\hat{k})) = \lambda GV_{\mu}(k),$$

where \(\hat{k} = \lambda k\). Therefore \(GV_{\mu}(\hat{k}) = GV_{\mu}(k)\), and this means that the GV's are homogeneous functions of degree zero in \(k\).

The locus of points that describe in the \((GV_x - GV_y)\) plane the group velocity for the \(\mu\)-th branch of eigenmodes is obtained varying the \(k\)-versor on the unit circle. For each \(\mu\) we obtain a closed curve that is characteristic of the system we are dealing with. For instance, the curves for the pressure eigenmodes of the unmodified Euler eqns \((M = 1)\) are two circles with center at \(V\) and radius \(c\). The remaining two branches of eigenmodes, which represent the vorticity and entropy transport, reduce to a point, that is the velocity \(V\) in the plane of group velocities.

The above description will be more clear after paragraph 5.1.1.

If we consider the CFL condition, we presume, and the real numerical stability analysis for the discretized eqns confirms it, that the maximum allowable time step has to be less than \(h / |GV|_{\text{max}}\). The fastest eigenmode \((\mu, k)_{\text{max}}\) travels at a speed of one element per time step. Other eigenmodes travel at a speed of
elements per time step. The slowest eigenmode travels at a speed of \(|GV^\infty(k_{min})|/|GV^\infty(k_{max})|\) elements per time step. We define

\[ \kappa = |GV^\infty(k_{max})|/|GV^\infty(k_{min})|, \]

as the GV-condition number for the continuous problem. From this definition, we see that the lower \(\kappa\), the greater the rate of convergence.

5.1.1 Group velocities with the identity matrix as the PMM:

We can rewrite the eigenvalue problem Eq. 21 by using the similarity transformation \(\hat{A}_s = \Psi \hat{A}_s \Psi^{-1}\), where \(\hat{A}_s\) is the diagonal matrix containing the eigenvalues of \(\hat{A}_s\), and \(\Psi\) is the matrix containing the right eigenvectors of \(\hat{A}_s\).

\[ |(k_1 \hat{A}_s + k_2 \Psi^{-1} \hat{\phi} \Psi) - \omega I| = 0. \]

The solution to this eigenvalue problem is

\[ \omega_1,2 = V k_1, \quad \text{and} \quad \omega_3,4 = V k_1 \pm c|k|. \]

Now we can obtain the group velocities by taking the derivatives of the \(\omega_s\) with respect to the \(k_s\), that is

\[ GV^{(1,2)} = (V, 0), \]

and the pressure eigenmodes have group velocities

\[ GV^{(3,4)} = (V \pm c \frac{k_x}{|k|}, \pm c \frac{k_y}{|k|}). \]

We can see that in the \((GV_x - GV_y)\) plane, the pressure eigenmodes describe circles of radius \(c\) with center in \((V, 0)\), and the eigenmodes of vorticity and entropy transport are represented by a single point of coordinates \((V, 0)\).

For \(k = (1, 0)\), the group velocities are

\[ GV^{(1,2)} = (V, 0), \quad GV^{(3,4)} = (V \pm c, 0), \]

and for \(k = (0, 1)\), the group velocities are

\[ GV^{(1,2)} = (V, 0), \quad GV^{(3,4)} = (V, \pm c). \]

It can be demonstrated, and for this case is evident from the curves, that the versor \(\mathbf{k}\) is perpendicular to the locus of points that represent the different eigenmodes in the \((GV_x - GV_y)\) plane.

For small wave numbers, all the consistent differential operators represent the continuous case. Having real eigenvalues, the hyperbolic systems have no damping, therefore for small wave numbers there is no damping and the only mechanism of convergence is the absorption at the boundaries.

From Eqns 26 and 27 we can see that for the case \(V \to c\), \(|GV^\infty|_{\min}\) appears for \(k = (1, 0)\) and \(\mu = 4\), that is \(|GV^\infty|_{\min} < |V - c|\). On the other hand, \(|GV^\infty|_{\max}\)
appears for \( k = (1, 0) \) and \( \mu = 3 \), that is \( |GV^\mu|_{\text{min}} = |V + c| \). Therefore the GV-condition number for this case is

\[
\kappa = \frac{|V + c|}{|V - c|} = \frac{|M + 1|}{|M - 1|}. \tag{28}
\]

The number of iterations that the slowest eigenmode needs to pass one element is \( N_{\text{damp}} = \kappa / \text{CFLN} \), where CFLN makes reference to the value obtained in the stability analysis for wave numbers \( (k_x, k_y) \in [(-\pi, x), (-\pi, x)] \). For example, for the case \( M=0.95 \), CFLN=0.76, and then \( N_{\text{damp}} = (1+0.95)/(1-0.95)/0.76 = \approx 51 \). This value can be verified in Fig. (13) curve 1.

5.1.2 Group velocities with

\[
(1 |\hat{A}_x| + |\hat{A}_y|)/(c + V)
\]

as the PMM:

The PMM is defined as

\[
\left( \frac{1}{c + V} \right) (|\hat{A}_x| + |\hat{A}_y|) = \left( \frac{1}{c + V} \right) \Phi (|\hat{A}_x| + \Phi^{-1} |\hat{A}_y|) \Phi^{-1}. \tag{29}
\]

In this case, the eigenvalue problem to be solved is

\[
([k_1 \hat{A}_x + k_2 \Phi^{-1} \hat{A}_y] - \omega \left( \frac{1}{c + V} \right) (|\hat{A}_x| + \Phi^{-1} |\hat{A}_y|) \Phi^{-1}. \tag{30}
\]

In Fig. (11) we can see in the \( (GV_x - GV_y) \) plane plots representing the group velocities for the different eigenmodes, for Mach numbers ranging from 0.50 till 0.95 in Fig. (11)-a, and for Mach numbers ranging from 1.05 till 2.00 in Fig. (11)-b.

Looking at the locus of group velocities for \( M=0.86 \), we can see that the characteristic circles of the two pressure wave branches have been distorted in the loop ABCB'. The vorticity transport branch is the triangular shaped figure DED', and the entropy transport branch is the point C.

For \( V \to c \), \( k_x = 1 \), and \( k_y = 0 \), the group velocities are

\[
GV^{(1,2)} = (V , 0), \quad GV^{(3)} = (V/2 , 0), \quad GV^{(4)} = (3/4(M^2 - 1)c , 0) \tag{31}
\]

and for \( V \to c \), \( k_x = 0 \), and \( k_y = 1 \), the group velocities are

\[
GV^{(1,3)} = (V , 0), \quad GV^{(3)} = (V/4 , c/\sqrt{2}), \quad \text{and} \quad GV^{(4)} = (V/4 , -c/\sqrt{2}) \tag{32}\]

From Eqns 31 and 32 we can see that for the case \( V \to c \), \( |GV^\mu|_{\text{min}} \) appears for \( k = (1, 0) \) and \( \mu = 4 \), that is \( |GV^\mu|_{\text{min}} = \frac{3}{4}|M^2 - 1|c \). On the other hand, \( |GV^\mu|_{\text{max}} \) appears for \( \mu = (1, 2) \) and for all \( k \), its value is \( |GV^\mu|_{\text{max}} = V \). Therefore the GV-condition number for this case is

\[
\kappa = \frac{4}{3|M^2 - 1|}. \tag{33}
\]

The number of iterations that the slowest eigenmode needs to pass one element is \( N_{\text{damp}} = \kappa \). For example, for the case \( M=0.95 \), CFLN=1.00, and then \( N_{\text{damp}} \approx 10 \). This value can be verified in Fig. (13) curve 3.
5.1.3 Group velocities with $|\hat{A}_z|/(c + V)$ as the PMM:

The PMM is defined as

$$|\hat{A}_z|/(c + V) = \left(-\frac{1}{c + V}\right) \Phi |\hat{A}_z|\Phi^{-1},$$

where $\hat{A}_z$ is the diagonal matrix containing the eigenvalues of $\hat{A}_z$, and the columns of $\Phi$ are the right eigenvectors of $\hat{A}_z$. In this case, the eigenvalue problem to be solved is

$$|(k_1\hat{A}_z + k_2(-\frac{1}{c + V}) \Phi^{-1}\hat{A}_z\Phi) - \omega(-\frac{1}{c + V})|\hat{A}_z| | = 0. \quad (35)$$

In Fig. (12) we can see in the $(GV_1 - GV_2)$ plane plots representing the group velocities for the different eigenmodes and for Mach numbers ranging from 0.50 till 0.95 in Fig. (12)-a, and for Mach numbers ranging from 1.05 till 2.00 in Fig. (12)-b.

The solution to the eigenvalue problem when $M \to 1$ is discontinuous. First we consider the case $M < 1$, the group velocities for this case are

$$GV^{(1,2)} = (V, 0), \quad GV^{(3,4)} = \pm \frac{c}{\sqrt{k_x^2 + k_y^2 \frac{M^2}{1 - M^2}}} (k_x, k_y \left(\frac{M^2}{|1 - M^2|}\right)) \quad (36)$$

and for the case $M > 1$ with $M \to 1$, the group velocities are

$$GV^{(1,2)} = (V, 0), \quad GV^{(3,4)} = (c, \pm \frac{V}{\sqrt{|1 - M^2|}}) \quad (37)$$

From Eqns 36 and 37 we can see that for the case $V \to c$, $|GV|^\mu_{\min}$ appears for $k = (1, 0)$, with value $|GV|^\mu_{\min} = |V|$. On the other hand, $|GV|^\mu_{\max}$ appears for $k = (0, 1)$, and its value is $|GV|^\mu_{\max} = \frac{V}{\sqrt{|1 - M^2|}}$. Therefore the GV-condition number for this case is

$$\kappa = \frac{1}{\sqrt{|1 - M^2|}}. \quad (38)$$

Comparing this GV-condition number with that of $M = 1$, we see that there exist a great improvement if we limit the analysis to the continuous case.

6. DISCRETIZATION OF THE EULER EQUATIONS

To assess the velocity of convergence using the preconditioning mass matrix, we have to choose a given scheme to discretize the problem both in space and time. An explicit, first-order accurate time-discretization is the forward Euler scheme, that is

$$\frac{U^{n+1} - U^n}{\Delta t} + F_{i,j}(U^n) = 0.$$
With regard to the space discretization, we choose a FVM that uses the Steger-Warming flux-splitting. Let the computational domain $\Omega$ be decomposed in disjunctive quadrangular elements (subdomains). A restructuring of the quadrangulation is done by constructing a new mesh in which each element contains exactly one node of the quadrangulation. This is accomplished by joining the centroids of the quadrangles having a node as a common corner and the midpoints of the sides passing through that node. In Fig. (14) the following notations are used: $m$ is the number of quadrangles that have to $N_i$ as a common node, $G_{ij}$ is the centroid of the $j$-th quadrangle, $I_{ij}$ the midpoint of the side $N_iN_{ij}$, and $Ar(N_i)$ the shaded area that corresponds to the node $N_i$.

The approximation test function space $V_h$ consists of piecewise constant functions, constant in each cell. The approximation problem is to find $U_{h}^{n+1} \in V_h$ such that

$$\int_{\Omega} \left( \frac{U_{h}^{n+1} - U_{h}^{n}}{\Delta t} \right) d\Omega + \int_{\Gamma} F_{ij}(U_\Omega^{n}) d\Gamma = 0,$$

Using Green's formula, we get

$$\frac{U_{h}^{n+1} - U_{h}^{n}}{\Delta t} Ar(N_i) + \int_{\Gamma} F_{ij}(U_\Omega^{n}) n_j d\Gamma = 0,$$  \hspace{1cm} (39)

where

$$U_{h}^{n} = (U_\Omega^{n})_{cell(N_i)}.$$

Equation 39 reduces to

$$\frac{U_{h}^{n+1} - U_{h}^{n}}{\Delta t} Ar(N_i) + \sum_{j=1}^{m} H_{ij}^{n} = 0,$$

where

$$H_{ij}^{n} = (A^+)_{ij}^{n} U_{h}^{n} + (A^-)_{ij}^{n} U_{h}^{n},$$

here

$$A_{\pm} = (\eta_{x,y} A_x + \eta_{y,x} A_y)_{\pm},$$

and

$$\eta_{ij} = n_{ij}^{1} \frac{|G_{ij} - I_{ij}|}{|G_{ij} - I_{ij}|} + n_{ij}^{2} \frac{|G_{ij} - I_{ij}|}{|G_{ij} - I_{ij}|}.$$

7. FOURIER STABILITY ANALYSIS

Both the stability analysis and the velocity of convergence are evaluated by considering a constant flow perturbed with an error which written in Fourier components has the form

$$U = \sum_{k_1,k_2} \hat{U}(k_1,k_2)e^{i(k_1x_1 + k_2x_2)},$$  \hspace{1cm} (40)

where, $k_1$, $k_2$ are the wave numbers in the $x_1$, $x_2$ directions, and $V^2$ is the velocity of propagation of the harmonic component $(k_1,k_2)$.

The fundamental tool for the stability analysis will be the Von Neumann method, which ignores boundary conditions, yet still, usually yields the best results [8-9,10,12].
The domain with periodic b.c. in the square \([-L,L],[-L,L]\), and the wave number vector are all the combinations
\[(k_1, k_2) = \left(\frac{n_1 \pi}{L}, \frac{n_2 \pi}{L}\right), \quad n_{1,2} = -N + 1, ..., -1, 0, 1, ..., N,\]
where \(N = L/h\), and \(h\) is the element size.
The \((I,J)\)-th node of the domain \(\Omega = \{(x,y)/(x,y) \in \mathbb{R}^2 \in \mathbb{R}^2\}\), has coordinates \((Ih, Jh)\). We can rewrite the FV formulation already described as follows:

\[
\left(\frac{\Delta t}{c}\right) (U^{n+1}_{(I,J)} - U^n_{(I,J)}) = -\frac{\Delta t}{h} \left[ A_x^+(U^n_{(I,J)} - U^n_{(I-1,J)}) + A_x^-(U^n_{(I,J-1)} - U^n_{(I,J)}) + A_y^-(U^n_{(I,J+1)} - U^n_{(I,J)}) \right],
\]

where
\[A_{\pm i} = \frac{A_{xi} \pm |A_{xi}|}{2}, \quad \text{with } |A_{xi}| = \Phi|A_{xi}|\Phi^{-1}.\]

Equation 41 can be put in the following form

\[
U^{n+1}_{(I,J)} = U^n_{(I,J)} - \frac{\Delta t}{h} \left[ A_x \left( e^{i\Delta k_1 h} - e^{-i\Delta k_1 h} \right) + A_y \left( e^{i\Delta k_2 h} - e^{-i\Delta k_2 h} \right) \right],
\]

and substituting a typical Fourier component of wave number \((k_1, k_2)\) in the above Eq., we get

\[
U^{n+1}_{(I,J)} = \left\{ I - \frac{\Delta t}{h} M^{-1} \left[ A_x \left( e^{i\Delta k_1 h} - e^{-i\Delta k_1 h} \right) + A_y \left( e^{i\Delta k_2 h} - e^{-i\Delta k_2 h} \right) \right] \right\} U^n_{(I,J)},
\]

Equation 43 is rewritten for convenience as

\[
\hat{U}^{n+1}_{(I,J)} = \hat{U} e^{i\Delta k_1 h + \Delta k_2 h - \omega(n+1)\Delta t} = e^{-i\omega\Delta t} \hat{U}^n_{(I,J)} = (I - zB) \hat{U}^n_{(I,J)}, \quad (44)
\]

here,
\[B = B(\rho, u_1, u_2, Mach, c_p/c_s, k_1, k_2).\]

The spectral decomposition of \(B\), that is, \(\Phi A \Phi^{-1}\), is such that in general both the eigenvectors and eigenvalues are complex. Using this decomposition of \(B\), we rewrite Eq. 44 as
\[ \phi^{n+1} = e^{-i\omega \Delta t} \phi^n = (I - \epsilon A)\phi^n = G\phi^n, \]  

and to have conditional stability, the amplification matrix \( G \) must have all its eigenvalues within the unit complex circle, that is

\[
(1 - \epsilon \Re(\lambda_i))^2 + (\epsilon \Im(\lambda_i))^2 \leq 1, \quad i = 1, \ldots, 4,
\]

therefore,

\[
\epsilon = \min_{i=1,\ldots,4} \frac{2\Re(\lambda_i)}{|\lambda_i|^2}.
\]

By making sweepings through a discrete number of angles of the flow with the mesh (\( \alpha \in [0, 45] \) degrees for the domain \(( -L, L ), (-L, L ) \)), for all the combinations of wave numbers

\[
(k_1, k_2) = \left( \frac{n_1 \pi}{L}, \frac{n_2 \pi}{L} \right), \quad n_{1,2} = -N + 1, \ldots, -1, 0, 1, \ldots, N,
\]

where \( N = L/h \), and for a discrete number of flow conditions (Mach numbers within the range we are interested in), we obtain for each combination by using Eq. 46 the max. \( \bar{\epsilon} \) admissible, then the min. (max. \( \bar{\epsilon} \)) will give us the Courant limit \( C(M) \) as a function of the Mach number.

7.1 Convergence by absorption at the boundaries, Group velocities:

The group velocity \((GV)\) of each eigenmode for each \((k_1, k_2)\), can be evaluated rewriting Eq. 45 as

\[-i \omega_l \Delta t = \log(1 - \bar{\epsilon} \lambda_l),\]

where the subscript \( l \) represents the \( l \)-th eigenmode, and differentiating the above equation with respect to \( k_j \), we obtain the group velocity in the cartesian direction \( x_j \) for the \( l \)-th eigenmode [3-4]

\[
GV_j = \frac{\partial \omega_l}{\partial k_j} = \Re \left( \frac{-1}{i \bar{\epsilon}} \frac{(-\bar{\epsilon})}{(1 - \bar{\epsilon} \lambda_l)} \frac{\partial \lambda_l}{\partial k_j} \right) = \left( \frac{c}{h} \right) \Im \left( \frac{1}{(1 - \bar{\epsilon} \lambda_l)} \frac{\partial \lambda_l}{\partial k_j} \right),
\]

where \( \bar{\epsilon} \) represents the value \( \min. (\max. (\bar{\epsilon})) \) obtained from equations 45-46 for all the combinations \((k_1, k_2, \alpha)\).

To measure the convergence velocity in terms of the number of iterations needed for the reduction of one order in the amplitude of the harmonic component by absorption at the boundary, we define

\[
N_{\text{group}} = \frac{L_{\text{mesh}}}{|GV_{\text{min}}(\Delta t)_{\text{max}}|} = \frac{L_{\text{mesh}}}{\epsilon(\frac{1}{\bar{\epsilon}})|GV_{\text{min}}|},
\]

where \( L_{\text{mesh}} \) represents the characteristic length of the domain that a wave component has to travel to reach the boundary,

\[
|GV_{\text{min}}| = \min_{i=1,2,3} \sqrt{|GV_i|^2 + |GV_j|^2}.
\]
and \( \tilde{c} \) is the max. value allowed by the stability condition, Eq. 46. In Eq. 48 we supposed that the wave is not totally absorbed at the boundary, but a 10 per cent of its amplitude is reflected back as a result of the obliqueness of the incidence direction with the normal to the boundary, that is, \( (k_1, k_2) \) isn't parallel to \( (n_1, n_2) \). For the calculations \( L_{\text{mesh}} = h = 1 \), and \( N_{\text{group}} \) refers to the number of iterations that the eigenmode needs to pass through a single element.

7.2 Convergence by Damping:

From eq. (12) we can observe that when

\[
|\lambda_{B_i}| = \sqrt{(1 - \Re(\lambda_i))^2 + (\Im(\lambda_i))^2} < 1,
\]

the harmonic component of the error we are dealing with will be damped out at a velocity given by

\[
N_{\text{damp}} = \frac{-\log(10)}{\log(|\lambda_{B_i}|)},
\]

where \( N_{\text{damp}} \) is the number of iterations needed for the reduction of one order in its amplitude.

7.3 Rate of Convergence, evaluation:

The two mechanisms of convergence already described work together, therefore to assess the convergence rate we can use the value

\[
\text{Num. of Iter.} = \min(N_{\text{group}}, N_{\text{damp}}).
\]

For a given Mach number, the evaluation of the rate of convergence is done by evaluating

\[
\text{Num. of Iter.} = \max_{(k_1, k_2, m)} (\min(N_{\text{group}}, N_{\text{damp}})),
\]

where the \( m \)s are a number of flow-mesh angles \( \in [0, 45] \) and the \( k_i \)s stand for all the possible combinations of values

\[
k = \frac{l\pi}{Nh}, \quad l \in [-N + 1, ..., -1, 0, 1, ..., N],
\]

here \( h \) is the mesh spacing usually taken as unity, and a good value for \( N \) may be 16.

7.4 Rate of Convergence using PMMs:

The calculation of GVs for the continuous case was intended to give some insight into the working of the PMMs. But now we know that the value of \( N_{\text{group}} \) must be considered instead of the single value of GV. From Eq. 48 we see that \( N_{\text{group}} \) is inversely proportional to the product \( (|GV_{\min}|\Delta t) \), therefore for a given PMM
\[|GV_{\text{min}}|\] could be very high, while \(\Delta t_{\text{max}}\) may be so small as to give an exceedingly high \(N_{\text{group}}\); and taking into account that \(N_{\text{damp}}\) is also very high for a very low \(\Delta t_{\text{max}}\), the resulting rate of convergence will be very slow. This is the case for \(M = \frac{|\hat{A}_z|}{(c + V)}\). However, for \(M = \frac{(|\hat{A}_x| + |\hat{A}_y|)}{(c + V)}\) the rate of convergence is much higher than that of the identity matrix.

In what follows we give the numerical results that compare the rate of convergence of the original scheme, \(M = I\), with that of \(M = \frac{(|\hat{A}_x| + |\hat{A}_y|)}{(c + V)}\). In Fig. (15) is plotted the Num. of Iter. vs the wave number for \(k_y = 0, \alpha = 0\) and 45 degrees. From this we can see that the max Num. of Iter. needed is for \(\alpha = 0\) and wave numbers \(k_y \approx 0.60\) with a gain of 250 per cent when the PMM is used. Fig. (16) shows the case \(k_z = k_y\) with \(\alpha = 0\) and 45 degrees. In this case again the max Num. of Iter. appears for \(k \approx 0.60\) and for the wave number vector aligned with the flow direction, that is 45 degrees. For this case the gain is also 250 per cent.

It was seen from the numerical studies that the max Num. of Iter. always appears for wave numbers aligned with the flow velocity and that for small wave numbers the convergence is mainly by absorption at the boundaries, while for medium and high wave numbers the mechanism of convergence is damping.

8. NUMERICAL RESULTS:

A channel of length 7 and height 3 units contains a circular arc profile of chord length unity as a part of the bottom solid wall. The circular arc is centered on the bottom at a distance of 2.5 units from the inlet. The thickness of the arc is equal to 4.2\% of the chord [11].

The computations were started imposing a uniform flowfield with Mach 0.85 over the entire mesh. At solid walls, tangential velocity is imposed as the boundary condition [13,16]. At the inlet and outlet far field boundary, the number of boundary conditions to be imposed depends on the normal Mach number [16]. For this case, both velocity and density were specified on the inflow boundary and pressure on the outflow boundary.

The convergence rate on a coarse 14 \times 42 mesh of quadrilateral elements is depicted in Fig. (17), where the improvement due to the preconditioning mass matrix is shown with the RMS of \(\Delta p/\Delta t\) vs. the Num. of Iterations. It took 600 iterations to arrive at a residue of order \(10^{-5.0}\) when the PMM was used, while without the PMM the residue for the same number of iterations is of order \(10^{-2.3}\).

In Fig. (18) we can see the iso-Mach lines while in Fig. (19) we see the pressure distribution.

9. CONCLUDING REMARKS:

The numerical scheme presented in this paper is:

(i) very efficient in terms of computing time,
(ii) simple to program,
(iii) so robust as the original scheme,
(iv) the steady solutions are captured in many less iterations as compared with the same space discretization but without using the PMM.
10. ACKNOWLEDGEMENTS

The authors wish to express their gratitude to CONICET for its financial support. They also wish to thank Nestor Aguilera and Norberta Nigro for helpful comments, and to Karl for his outstanding job in typing the manuscript.

REFERENCES


Fig. 1. Group Velocity vs Courant Number - One-dimen. scalar case.
Fig. 2. Damping vs Courant Number - One-dimen. scalar case.
Fig. 3. Pressure profile of the perturbation - Original scheme.

Fig. 4. Pressure profile of the perturbation - Proposed scheme.
Fig. 5. Formation of a shock wave - Original scheme.

Fig. 6. Formation of a shock wave - Proposed scheme.
Fig. 7. Passing of a perturbation through a shock - Orig. scheme.

Fig. 8. Passing of a perturbation through a shock - Prop. scheme.
Fig. 9. Movement of an unstable shock wave - Original scheme.

Fig. 10. Movement of an unstable shock wave - Proposed scheme.
Fig. 11. Group velocities with \((|\hat{A}_x| + |\hat{A}_y|)/(c + V)\) as the PMM.
Fig. 12. Group velocities with $|\hat{A}_z|/(c + V)$ as the PMM.
Fig. 13. GV-condition number vs Mach number for different PMMs.
Fig. 14. Finite Volume formulation terminology.
Fig. 15. Convergence Velocity for different PMMs with mesh \( \parallel V \).
Fig. 16. Convergence Velocity for different PMMs with mesh \( \ell V = 45 \text{deg.} \)
Fig. 17. Residual for flow past a profile in a channel at $M_\infty = 0.85$.
(a) Without the PMM. (b) Using the PMM.
Fig. 18. Flow lines of constant Mach number for flow past a profile in a channel at $M_{\infty} = 0.85$. Sonic line.
Fig. 10. Flow lines of constant pressure for flow past a profile in a channel at $Ma = 0.85$. 