

A DISCUSSION AND COMPARISON OF NUMERICAL TECHNIQUES USED TO SOLVE THE NAVIER–STOKES AND EULER EQUATIONS

S. D. CONNELL* AND P. STOW†

Theoretical Science Group, Rolls-Royce Ltd., P.O. Box 31, Derby DE2 3BJ, U.K.

SUMMARY

This paper is intended to provide some background to a number of widely used methods for solving the Navier–Stokes and Euler equations. The difference between coupled and uncoupled iterative schemes is discussed together with methods for solving the equations. Methods covered include time marching (both explicit and implicit), pressure correction and a Newton–Raphson technique. The relationship between the methods is illustrated.

1. INTRODUCTION

The Navier–Stokes and Euler equations are highly non-linear and in order to solve them with existing numerical techniques we must first linearize the equations. It is the form of this linearization which determines the form of the numerical technique that is subsequently used.

In Section 3 a number of different linearizations are discussed which give rise to coupled or uncoupled systems of discrete equations. With an uncoupled system the equations may be solved sequentially but with a coupled system the equations must be solved simultaneously.

Three methods of solution of the uncoupled equations are outlined in Section 4. These methods are implicit and explicit time marching and a pressure correction procedure. A stability analysis is also performed on a model equation.

The solution of the coupled equations is described in Section 5. The methods described are implicit time marching and a Newton–Raphson procedure.

It is now apparent that there are a wide variety of methods available for solving the same governing equations. The question of which method is optimal is discussed in Section 6.

2. GOVERNING EQUATIONS

The two-dimensional unsteady continuity, Navier–Stokes and energy equations may be written in strong conservation form as

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{q})}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{q})}{\partial y} = \frac{\partial \mathbf{R}(\mathbf{q})}{\partial x} + \frac{\partial \mathbf{S}(\mathbf{q})}{\partial y}, \quad (1)$$

* Principal Theoretical Scientist; author to whom correspondence should be addressed.

† Chief Aerodynamic Scientist

where

$$\mathbf{q} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ e \end{pmatrix}, \quad (2)$$

and the flux vectors \mathbf{F} , \mathbf{G} , \mathbf{R} and \mathbf{S} are given by

$$\mathbf{F} = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ (e + p)u \end{pmatrix}, \quad (3)$$

$$\mathbf{G} = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ (e + p)v \end{pmatrix}, \quad (4)$$

$$\mathbf{R} = \begin{pmatrix} 0 \\ \tau_{xx} \\ \tau_{xy} \\ R_4 \end{pmatrix}, \quad (5)$$

and

$$\mathbf{S} = \begin{pmatrix} 0 \\ \tau_{xy} \\ \tau_{yy} \\ S_4 \end{pmatrix}. \quad (6)$$

In (5) and (6)

$$\tau_{xx} = (\lambda + 2\mu) \frac{\partial u}{\partial x} + \lambda \frac{\partial v}{\partial y}, \quad (7)$$

$$\tau_{xy} = \mu \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right), \quad (8)$$

$$\tau_{yy} = (\lambda + 2\mu) \frac{\partial v}{\partial y} + \lambda \frac{\partial u}{\partial x}, \quad (9)$$

$$R_4 = u\tau_{xx} + v\tau_{xy} + k \frac{\partial T}{\partial x} \quad (10)$$

and

$$S_4 = u\tau_{xy} + v\tau_{yy} + k \frac{\partial T}{\partial y}, \quad (11)$$

where k is the thermal diffusivity.

For a perfect gas the pressure p is given by the equation of state

$$p = (\gamma - 1) \left[e - \frac{1}{2}(\rho u^2 + \rho v^2) \right], \quad (12)$$

where γ is the ratio of specific heats and the static temperature T is given by

$$T = [e - (\rho u^2 + \rho v^2)/2]/(\rho C_v). \quad (13)$$

Only laminar flow will be considered, but the numerical schemes developed in this paper may be applied to turbulent flow.

3. COUPLED AND UNCOUPLED SYSTEMS OF EQUATIONS

In order to solve the equations described in the previous section we must first linearize them. It is the form of this linearization which determines the numerical technique that is subsequently used. The linearization scheme of Beam and Warming¹ will be described in this section, and it is shown how this scheme gives rise to a set of coupled implicit equations. It is then demonstrated that using alternative linearizations the equations may be uncoupled and the resulting sets of equations solved by explicit or implicit techniques.

A generalized time differencing formula may be defined as¹

$$\Delta \mathbf{q}^n = \frac{\theta \Delta t}{1 + \varepsilon} \frac{\partial}{\partial t} (\Delta \mathbf{q}^n) + \frac{\Delta t}{1 + \varepsilon} \frac{\partial}{\partial t} (\mathbf{q}^n) + \frac{\varepsilon}{1 + \varepsilon} \Delta \mathbf{q}^{n-1} + [O[(\theta - \varepsilon - \frac{1}{2})\Delta t^2 + \Delta t^3]], \quad (14)$$

where

$$\Delta \mathbf{q}^n = \mathbf{q}^{n+1} - \mathbf{q}^n. \quad (15)$$

The differencing in (14) is equivalent to many common schemes. The parameters θ and ε determine the type of scheme and its accuracy. For example $\theta = 1$, $\varepsilon = 0$ reduces (14) to the common Euler implicit formula.

$$\mathbf{q}^{n+1} - \mathbf{q}^n = \Delta t \frac{\partial}{\partial t} (\mathbf{q}^{n+1}) + O(\Delta t^2). \quad (16)$$

Substituting (1) into (14) gives

$$\begin{aligned} \Delta \mathbf{q}^n + \frac{\theta \Delta t}{1 + \varepsilon} \left[\frac{\partial}{\partial x} (\Delta \mathbf{F}^n - \Delta \mathbf{R}^n) + \frac{\partial}{\partial y} (\Delta \mathbf{G}^n - \Delta \mathbf{S}^n) \right] \\ = -\frac{\Delta t}{1 + \varepsilon} \left[\frac{\partial}{\partial x} (\mathbf{F}^n - \mathbf{R}^n) + \frac{\partial}{\partial y} (\mathbf{G}^n - \mathbf{S}^n) \right] + \frac{\varepsilon}{1 + \varepsilon} \Delta \mathbf{q}^{n-1}, \end{aligned} \quad (17)$$

neglecting the temporal truncation error. If we now define the Jacobian matrices \mathbf{J}_F , \mathbf{J}_G , \mathbf{J}_R and \mathbf{J}_S as

$$\mathbf{J}_F = \frac{\partial \mathbf{F}}{\partial \mathbf{q}}, \quad (18)$$

$$\mathbf{J}_G = \frac{\partial \mathbf{G}}{\partial \mathbf{q}}, \quad (19)$$

$$\mathbf{J}_R = \frac{\partial \mathbf{R}}{\partial \mathbf{q}} \quad (20)$$

and

$$\mathbf{J}_S = \frac{\partial \mathbf{S}}{\partial \mathbf{q}}, \quad (21)$$

which are of order 4 and easily evaluated from (2)–(6), these definitions imply

$$\Delta \mathbf{F} = \mathbf{J}_F \Delta \mathbf{q}, \quad (22)$$

$$\Delta \mathbf{G} = \mathbf{J}_G \Delta \mathbf{q}, \quad (23)$$

$$\Delta \mathbf{R} = \mathbf{J}_R \Delta \mathbf{q} \quad (24)$$

and

$$\Delta \mathbf{S} = \mathbf{J}_S \Delta \mathbf{q}. \quad (25)$$

Equations (22)–(25) may be used in (17) to obtain

$$\begin{aligned} & \left\{ \mathbf{I} + \frac{\theta \Delta t}{1 + \varepsilon} \left[\frac{\partial}{\partial x} (\mathbf{J}_F^n - \mathbf{J}_R^n) + \frac{\partial}{\partial y} (\mathbf{J}_G^n - \mathbf{J}_S^n) \right] \right\} \Delta \mathbf{q}^n \\ & = - \frac{\Delta t}{1 + \varepsilon} \left[\frac{\partial}{\partial x} (\mathbf{F}^n - \mathbf{R}^n) + \frac{\partial}{\partial y} (\mathbf{G}^n - \mathbf{S}^n) \right] + \frac{\varepsilon}{1 + \varepsilon} \Delta \mathbf{q}^{n-1}. \end{aligned} \quad (26)$$

The equations have been linearized in this step by evaluation of the Jacobian matrices at time level n . When (26) is integrated over suitable control volumes using central differences $\Delta \mathbf{q}$ now represents a vector containing values of the variables at all the discrete points in the domain. (26) may now be written in matrix form as

$$\mathbf{A}^n \Delta \mathbf{q}^n = \frac{\Delta t \mathbf{E}^n}{1 + \varepsilon} + \frac{\varepsilon}{1 + \varepsilon} \Delta \mathbf{q}^{n+1}, \quad (27)$$

where \mathbf{E} is a vector containing the flux balance residuals over each of the control volumes.

It is the ordering of $\Delta \mathbf{q}$ which determines the form of \mathbf{A} . Usually this vector is ordered as follows:

$$\Delta \mathbf{q} = \begin{pmatrix} \Delta q_1 \\ \Delta q_2 \\ \vdots \\ \Delta q_k \end{pmatrix}, \quad (28)$$

where

$$\Delta \mathbf{q}_l = \begin{pmatrix} \rho_l \\ (\rho u)_l \\ (\rho v)_l \\ e_l \end{pmatrix}, \quad (29)$$

k is the number of grid points and $1 \leq l \leq k$. \mathbf{A} is then a block penta-diagonal matrix (each block of order 4) and may be factorized into two block tri-diagonal matrices and inverted using ADI techniques. This factorization is performed without reduction in the order of the temporal or spatial truncation error.¹⁻⁴

To illustrate the coupling between the equations we may adopt an alternative ordering of $\Delta \mathbf{q}$, thus:

$$\Delta \mathbf{q} = \begin{pmatrix} \Delta \rho \\ \Delta(\rho u) \\ \Delta(\rho v) \\ \Delta e \end{pmatrix}. \quad (30)$$

The coefficient matrix A then takes the following form:

$$A = \begin{pmatrix} I & C_1 & C_2 & 0 \\ X_1 & I + X_2 & X_3 & X_4 \\ Y_1 & Y_2 & I + Y_3 & Y_4 \\ E_1 & E_2 & E_3 & I + E_4 \end{pmatrix}. \quad (31)$$

Each of the submatrices in A is penta-diagonal and of order k . It can now be clearly seen that (31) represents a coupled system of equations, i.e. we cannot solve for $\Delta\rho$, $\Delta\rho u$, $\Delta\rho v$ and $\Delta\varepsilon$ individually but require to invert A .

If submatrices in A are neglected so that A becomes block lower triangular, i.e.

$$A = \begin{pmatrix} I & 0 & 0 & 0 \\ X_1 & I + X_2 & 0 & 0 \\ Y_1 & Y_2 & I + Y_3 & 0 \\ E_1 & E_2 & E_3 & I + E_4 \end{pmatrix}, \quad (32)$$

then the equations become partially uncoupled. We may now solve the equations sequentially and only require to invert a submatrix for each equation. We are also free to choose the order in which the equations are solved so long as the block triangular nature of A is maintained.

Inverting A in these various forms corresponds to alternative linearizations to that defined by (31) which can only affect the time accuracy, stability and rate of convergence of the scheme; the steady-state solution remains unchanged.

It is possible to go one stage further and fully uncouple the equations by adopting an alternative linearization such that A becomes block diagonal, i.e.

$$A = \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & I + X_2 & 0 & 0 \\ 0 & 0 & I + X_3 & 0 \\ 0 & 0 & 0 & I + E_4 \end{pmatrix}. \quad (33)$$

Again this linearization cannot affect the steady-state solution. The equations may now be solved in any order.

With the partially and fully uncoupled systems we also have the choice of how to calculate the coefficient matrices. We could use the most up to date value of each variable or use all variables at the previous time level. These alternative schemes correspond to different linearizations which do not affect the steady-state solution.

4. METHODS OF SOLVING THE UNCOUPLED EQUATIONS

4.1. Uncoupled implicit time marching

This method corresponds to the inversion of (32) or (33). The solution of the individual equations could be achieved by the use of an A.D.I. method. To the authors' knowledge there are no reports on the use of this method in the literature.

4.2. Pressure correction methods

These methods^{5,6} solve the uncoupled steady-state equations using primitive variables. The continuity equation is expressed in terms of pressure as opposed to density. An A.D.I. method

is then employed to solve the uncoupled equations sequentially. These methods bear an obvious resemblance to the uncoupled implicit time marching method described in Section 4.1. It has been shown by Connell⁷ that the SIMPLE procedure of Caretto *et al.*⁶ is equivalent to solving the unsteady momentum equations with a spatially varying time step.

4.3. Explicit time marching

Explicit time marching may be applied in a form which corresponds to a further simplification of \mathbf{A} in (33), reducing it to the identity matrix.

In this form the scheme is similar to that of Denton⁸ and may be recognized as one iteration of a point Jacobi procedure. An alternative scheme may be proposed where the most up to date values of the variables are used as we sweep through the matrix. \mathbf{A} then becomes lower triangular and the scheme corresponds to one Gauss–Siedel iteration.

There are a number of other explicit schemes which are the multi-step in nature.^{9,10}

4.4. Stability restrictions

A general linearized analysis for the implicit methods described previously indicates that they are unconditionally stable. However in reality, owing to the non-linearities and the method of applying the boundary conditions, there will be some limit on the forward time step.

In order for an explicit scheme to be stable a limit has to be put on the forward time step Δt . An approximation to this limit can be obtained by analysing the stability of simple model problems. For example consider the linearized Burgers' equation:

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} - v \frac{\partial^2 u}{\partial x^2} = 0, \quad (34)$$

where a is the speed of sound and $v = \mu/\rho$ (both assumed constant). The discrete approximation to (34) for an internal control volume using central differences for the space terms may be written as

$$u_p^{n+1} = u_p^n + \frac{\Delta t v}{\Delta x^2} (u_w^n - 2u_p^n + u_e^n) + \frac{a \Delta t}{2 \Delta x} (u_e^n - u_w^n). \quad (35)$$

In order to establish a stability criterion we neglect any effect the boundary conditions may have and look for a solution to (35) of the form

$$u_p^n = e^{\alpha n \Delta t} e^{i \beta x}, \quad (36)$$

giving

$$e^{\alpha \Delta t} = 1 + \frac{\Delta t v}{\Delta x^2} (e^{-i \beta \Delta x} - 2 + e^{i \beta \Delta x}) + \frac{a \Delta t}{2 \Delta x} (e^{i \beta \Delta x} - e^{-i \beta \Delta x}), \quad (37)$$

which may be expressed as

$$e^{\alpha \Delta t} = 1 - \frac{4v \Delta t}{\Delta x^2} \sin^2(\beta \Delta x/2) + \frac{a \Delta t}{\Delta x} i \sin(\beta \Delta x) \quad (38)$$

For stability we require $|e^{\alpha \Delta t}| \leq 1$, which leads to

$$\Delta t \left(\frac{4v^2}{\Delta x^2} \sin^2(\beta \Delta x/2) + a^2 \cos^2(\beta \Delta x/2) \right) \leq 2v. \quad (39)$$

If $v \ll a \Delta x$ then

$$\Delta t \leq 2v/a^2, \quad (40)$$

which is very restrictive.

If upwind differencing is used on the $a \partial u / \partial x$ term then, by a similar method to that used above, we find that if

$$\Delta t \leq 1/(a/\Delta x + 2v/\Delta x^2), \quad (41)$$

then the scheme is stable. This condition is much less restrictive than (40) at the expense of a larger spatial truncation error. If the fluid under consideration is compressible and inviscid, then the stability criterion may be derived by considering the one-dimensional momentum and continuity equations under the assumption of constant entropy. Richardson¹¹ performs the analysis and derives the well known Courant–Friedricks–Lewy (CFL) condition of

$$\Delta t \leq \Delta x/(u \pm a), \quad (42)$$

where the first derivatives have been upwind differenced. This condition is also derived by MacCormack⁹ where the equations of momentum, continuity and energy are considered.

To extend the analysis of Richardson and MacCormack to include the effects of viscosity is complicated. However if the viscous effects are small enough for the assumption of constant entropy to be valid an extension of Richardson's analysis can be shown to give

$$\Delta t \leq 1/[(u \pm a)/\Delta x + \nu/\Delta x^2], \quad (43)$$

indicating that the viscous effects tend to reduce the maximum permissible time step.

5. METHODS OF SOLUTION OF THE COUPLED EQUATIONS

In addition to the implicit time marching procedure described in Section 3 a Newton–Raphson procedure may be used to solve the coupled equations. It is this technique and its relationship to implicit time marching that will be described in this section. With this approach the time derivatives are omitted from the governing equations and we solve (iteratively) for the steady-state solution.¹² As with the implicit time marching method described in Section 2 the steady-state governing equations may be expressed as

$$\frac{\partial \mathbf{F}(\mathbf{q})}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{q})}{\partial y} = \frac{\partial \mathbf{R}(\mathbf{q})}{\partial x} + \frac{\partial \mathbf{S}(\mathbf{q})}{\partial y}, \quad (44)$$

where \mathbf{q} , \mathbf{F} , \mathbf{G} , \mathbf{R} and \mathbf{S} are defined by (2)–(6). It is assumed that we have an approximate solution \mathbf{q}^n to (44), where n refers to iteration number. To \mathbf{q}^n we add $\Delta \mathbf{q}^n$ such that we have a better approximation to the solution, i.e.

$$\frac{\partial \mathbf{F}(\mathbf{q}^n + \Delta \mathbf{q}^n)}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{q}^n + \Delta \mathbf{q}^n)}{\partial y} = \frac{\partial \mathbf{R}(\mathbf{q}^n + \Delta \mathbf{q}^n)}{\partial x} + \frac{\partial \mathbf{S}(\mathbf{q}^n + \Delta \mathbf{q}^n)}{\partial y}. \quad (45)$$

Adopting a Taylor expansion with some rearrangement gives

$$\frac{\partial}{\partial x} \{(\mathbf{J}_F^n - \mathbf{J}_R^n) \Delta \mathbf{q}^n\} + \frac{\partial}{\partial y} \{(\mathbf{J}_G^n - \mathbf{J}_S^n) \Delta \mathbf{q}^n\} = \frac{\partial}{\partial x} (\mathbf{F}^n - \mathbf{R}^n) + \frac{\partial}{\partial y} (\mathbf{G}^n - \mathbf{S}^n), \quad (46)$$

where terms of order $(\Delta \mathbf{q})^2$ have been neglected and \mathbf{J}_F , \mathbf{J}_G , \mathbf{J}_R and \mathbf{J}_S are the Jacobian matrices defined in (18)–(21).

Equation (46) is now discretized in a suitable manner and solved for $\Delta \mathbf{q}^n$. At convergence the R.H.S. of (46) will be zero, and hence $\Delta \mathbf{q}$ will be zero.

The solution of the discrete approximation to (46) can be achieved by use of a direct or iterative method. Baldwin¹² uses a preconditioned bi-conjugate gradient method.

The Newton–Raphson scheme bears an obvious resemblance to the coupled implicit time marching method described in Section 3. It will be seen that (46) is identical to (26) with $\varepsilon = 0$ and $\theta = 1$ (the Euler implicit time discretization) and $\Delta t \rightarrow \infty$.

The stability of coupled systems is extremely difficult to analyse. However, the conclusions drawn in Section 4.4 are applicable. One may expect that owing to the linearization and consequent strong coupling between the equations, the coupled methods described above will be faster to converge and more robust than uncoupled systems.

6. SUMMARY AND DISCUSSION

Linearization techniques for the iterative solution of the Navier–Stokes equations have been discussed that give rise to coupled or uncoupled systems of equations.

For the uncoupled systems of equations two solution algorithms have been described, time marching and pressure correction. The time marching method may be subdivided into implicit and explicit time marching. Both these methods solve the governing equations for a transient-like solution and obtain the steady-state solution after a sufficient number of time steps.

With a pressure correction method the time terms are omitted and the steady equations are solved iteratively. However this method is similar to uncoupled implicit time marching.

For the coupled systems of equations two possible methods are described, implicit time marching and a Newton–Raphson procedure. Both methods require the inversion of large matrices of coefficients. With the time marching method this is achieved by the use of an efficient block ADI procedure. With the Newton–Raphson method a direct or iterative method could be used. The similarity between implicit time marching and the Newton–Raphson method is also demonstrated.

The obvious question which arises from this summary is ‘is it best to use an implicit or explicit method?’ There is no easy answer to this question. However if we are interested in minimizing the CPU time to obtain the steady-state solution to a given problem, then it can be seen that although an explicit method is cheap in terms of CPU time per time step, there is a restriction on the forward time step. Hence the method may take many time steps to reach the steady-state solution. In contrast, with implicit methods a larger time step may be used and fewer time steps will be required to reach a converged solution. However implicit methods use more CPU time per time step. It can therefore be seen that there may be some balance in terms of overall CPU time (to reach the steady-state solution) between implicit and explicit methods. One advantage of fully coupled systems is that they may suffer less from convergence and stability problems associated with an uncoupled system. It is unlikely that much can be done analytically and the question of which procedure is more efficient will only be answered with numerical experimentation.

ACKNOWLEDGEMENTS

The first author wishes to extend his thanks to the U.K. Science and Engineering Research Council for a Cooperative Award in Science and Engineering and to Rolls-Royce Ltd., (Derby) for their additional financial support.

REFERENCES

1. R. M. Beam and R. F. Warming, 'An implicit factored scheme for the compressible Navier–Stokes equations', *AIAA Journal*, **16**, 393 (1978).
2. W. N. Dawes, 'Computation of viscous compressible flow in blade cascades using an iterative replacement algorithm', in *Computational Methods in Turbomachinery*, Inst. of Mechanical Engineers, Birmingham University England, April 1984.
3. J. L. Steger, T. H. Pulliam and R. V. Chima, 'An implicit finite difference code for inviscid and viscous cascade flow', *AIAA Paper 80-1427* (1980).
4. W. T. Thompkins, S. S. Tong, R. H. Bush, W. J. Usab and R. J. G. Norton, 'Solution procedures for accurate numerical simulations of flows in Turbomachinery cascades', *AIAA Paper 83-0257* (1983).
5. J. Moore and J. G. Moore, 'A calculation procedure for three-dimensional viscous, compressible duct flow. Part 1 inviscid considerations', *Journal of Fluids Engineering*, **101**, 415–422 (1979).
6. L. S. Carretto, A. D. Gosman, S. V. Patankar and D. B. Spalding, 'Two calculation procedures for steady three dimensional flows with recirculation', *Proceedings of the Third International Conference on Numerical methods in Fluid Dynamics*, Springer, Berlin, Heidelberg, New York, 1972.
7. S. D. Connell, 'Numerical solution of the equations of viscous flow', *Ph.D. Thesis*, University of Nottingham, Nottingham, England, 1983.
8. J. D. Denton, 'A time marching method for two and three dimensional blade to blade flows', *ARC & M No 3775* (1974).
9. R. W. McCormack, 'The effect of viscosity in hypervelocity impact cratering', *AIAA Paper 69-354* (1969).
10. A. Jameson, W. Schmidt and E. Turkel, 'Numerical solutions of the Euler equations by finite volume methods using Runge Kutta time stepping schemes', *AIAA Paper 81-1259* (1981).
11. S. M. Richardson, 'Stability-Theory of time marching methods', Cambridge University Engineering Department, *Report CUED/A-Turbo/TR81* (1976).
12. D. I. Baldwin, 'The numerical solution of the two-dimensional Navier–Stokes equations', *M.Sc. Thesis*, Hertford College, Oxford, 1981.