Shock wave instability and the carbuncle phenomenon: same intrinsic origin?

By J.-Ch. ROBINET[1], J. GRESSIER[1], G. CASALIS[1] AND J.-M. MOSCHETTA[2]

[1]Department of Modelling Aerodynamics and Energetics, ONERA-CERT, 2, avenue Edouard Belin, 31055 Toulouse Cedex 4, France

[2]Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, 31400 Toulouse, France

The theoretical linear stability of a shock wave moving in an unlimited homogeneous environment has been widely studied during the last fifty years. Important results have been obtained by Dýakov (1954), Landau & Lifchitz (1959) and then by Swan & Fowles (1975) where the fluctuating quantities are written as normal modes. More recently, numerical studies on upwind finite difference schemes have shown some instabilities in the case of the motion of an inviscid perfect gas in a rectangular channel. The purpose of this paper is first to specify a mathematical formulation for the eigenmodes and to exhibit a new mode which was not found by the previous stability analysis of shock waves. Then, this mode is confirmed by numerical simulations which may lead to a new understanding of the so-called carbuncle phenomenon.

1. Introduction

The stability of a shock wave is of considerable interest from both a fundamental and a practical point of view. L. Landau made the first attempts to determine the stability of shock waves. Small disturbances were introduced on both sides of a steady, non-dissipative, plane shock wave. Landau & Lifchitz (1959) and Xu (1987) obtained the stability criterion $M_0 > 1$, $M_1 < 1$ for small disturbances which are travelling in the direction perpendicular to the shock wave (one-dimensional perturbations case). This stability criterion is simply a consequence of the requirement of the second law of thermodynamics. A fundamental paper dealing with the stability of a shock wave containing two-dimensional small disturbances in an infinite homogeneous environment with an arbitrary equation of state is that of Dýakov (1954). He found that the shock is unstable when

$$j^2 \frac{\partial \overline{\mathscr{V}}}{\partial \overline{P}}\bigg|_H < -1 \quad \text{and} \quad j^2 \frac{\partial \overline{\mathscr{V}}}{\partial \overline{P}}\bigg|_H \geqslant 1 + 2\overline{M}_1, \tag{1.1}$$

where $j^2 = (\overline{P}_1 - \overline{P}_0)/(\overline{\mathscr{V}}_0 - \overline{\mathscr{V}}_1)$ is the slope of the Rayleigh line, $(\mathrm{d}\overline{\mathscr{V}}/\mathrm{d}\overline{P})_H$ the slope of the Hugoniot curve in the pressure–volume $(\overline{P}, \overline{\mathscr{V}})$ plane, $\overline{M}_1$ the downstream Mach number and $\overline{\mathscr{V}} = 1/\overline{\rho}$ the mean specific volume. Swan & Fowles (1975) completed the calculations of Dýakov by giving a physical interpretation of these instabilities. In their analysis, the marginal case $\mathrm{Im}(k_x) = 0$ (no spatial amplification), $\mathrm{Im}(\omega) = 0$ (no temporal amplification) was also considered. This case corresponds to pure acoustic, entropic and vorticity waves which are neither amplified nor damped; it corresponds in fact to a spontaneous emission of sound by a discontinuity. This emission had

already been studied by Dýakov (1957), Kontorovitch (1957) and Fowles & Houwing (1984). A new possible zone of instability has been obtained:

$$\frac{1 - \overline{M}_1^2 - \overline{M}_1^2\,\overline{U}_0/\overline{U}_1}{1 - \overline{M}_1^2 + \overline{M}_1^2\,\overline{U}_0/\overline{U}_1} < j^2\frac{\partial\overline{\mathcal{V}}}{\partial\overline{P}}\bigg|_H < 1 + 2\overline{M}_1. \qquad (1.2)$$

However, for a perfect gas, it can be easily shown that

$$j^2(\mathrm{d}\overline{\mathcal{V}}/\mathrm{d}\overline{P})_H = -1/M_0^2.$$

Thus, for such a gas, the instability criteria (1.1) and (1.2) can never be satisfied; consequently according to these criteria, in a perfect gas, a shock wave is unconditionally stable.

More recently, the linear stability of a shock wave has been used in the framework of self-sustained oscillations of shock waves in a transonic nozzle flow. Many experiments carried out by Sajben, Bogar & Kroutil (1981) have highlighted some critical configurations for which the shock in the nozzle exhibits well-defined oscillations. The previous stability analyses have been generalized for this non-analytical case and interesting results have been obtained in comparison with the experimental ones, see Casalis & Robinet (1997). In this case also (with a perfect gas), the stability analysis leads to the conclusion that the mean flow is stable; the mean shock along with the downstream core region actually plays the role of a noise selecting system.

From the previous studies, it could be inferred that the shock wave is intrinsically stable. This is not true. First, the analysis performed by Dýakov and Swan & Fowles has some deficiencies: one of the two acoustic waves is discarded but the reason why is not clear, and the existence of disturbances in the upstream region is not clearly stated. Moreover none of the previous analysis took into account a strange mode. The shock wave stability must be inspected again.

In a seemingly different scientific area, the so-called *carbuncle phenomenon* has been observed and discussed in the CFD community for many years and so far has been considered as a purely numerical instability by numerical scheme designers. One of the main findings of this paper is that the carbuncle phenomenon may be the numerical symptom of a more fundamental instability mechanism associated with a shock wave in inviscid flow. The carbuncle phenomenon was first observed by Peery & Imlay (1988) for blunt body computations using Roe's method (Roe 1981). It consists of a spurious steady-state solution obtained when computing a blunt-body flow problem at supersonic speeds. The unphysical solution, although converged in time, includes a non-symmetrical recirculation region ahead of the bow shock in the vicinity of the stagnation line. On figure 1, a comparison between a physically acceptable solution and a spurious solution including the carbuncle phenomenon is presented to illustrate the importance of the flaw. The free-stream Mach number is $M_\infty = 10$ and the geometry is a two-dimensional cylinder whose axis is perpendicular to the incoming flow. Both numerical solutions are obtained by solving the Euler equations in time, starting from an initially uniform flow which corresponds to the free-stream conditions. It is remarkable that both numerical solutions shown on figure 1 are steady-state solutions, i.e. converged in time, obtained using exactly the same geometry, the same grid and the same initial and boundary conditions. The only difference between the two lies in the numerical method used to reach the final solution. This is the reason why aerodynamicists have long attributed the carbuncle phenomenon to a purely numerical mechanism with no connection to the physics. For stable and consistent conservative methods, one might expect the numerical
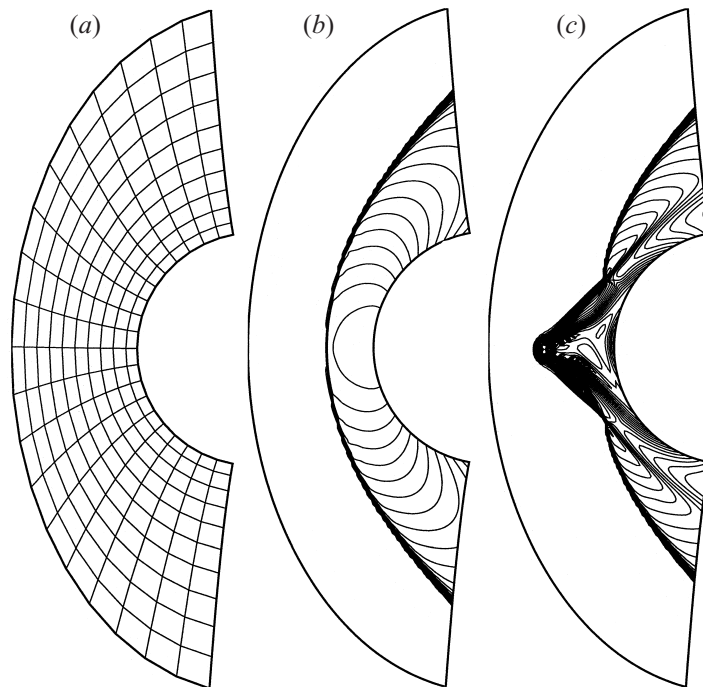
FIGURE 1. Temperature contours for a forward facing cylinder, $M_\infty = 10$, 80×160 computations: (*a*) 10×20 mesh, (*b*) HLLE method, (*c*) Roe's scheme.

procedure to converge toward a unique physical solution according to Lax–Wendroff's theorem. However, one should remember that this convergence theorem applies to a convergence in time and space and is not in contradiction with the occurrence of the carbuncle phenomenon obtained on a grid of finite resolution. By using a grid series with increasing spatial resolution, one should eventually obtain an unstable solution.

The carbuncle phenomenon was first closely studied from a numerical point of view by Quirk (1994) who introduced a simplified test case, dubbed 'Quirk's test', in which a planar moving shock wave is computed as it is propagating down a duct. In Quirk's test case, the symmetry line is slightly perturbed from a uniform grid to initiate the instability. Ever since, many authors, using other upwind schemes, have reported the strong connection between the carbuncle phenomenon and Quirk's test: all numerical schemes which fail Quirk's test also fail the blunt-body problem and vice versa. The main interest of Quirk's test is that it is not as strongly grid-dependent as in the case of the carbuncle phenomenon and that it is easier to study mathematically. In the following, only Quirk's test will be considered but the strong connection between both problems should be kept in mind when considering the relevance of the present study for practical gas dynamics applications.

Amongst the different researchers who have studied the carbuncle phenomenon, many have noticed that the instability is more likely to appear when the bow shock is almost perfectly aligned to grid lines (Quick 1994) and when grid cells are very elongated along a direction normal to the shock (Pandolfi & D'Ambrosio 1998). Consequently, some schemes may produce flawless solutions on one set of grids but can fail on a different grid. In the carbuncle phenomenon, the shock instability is caused by the shock strength itself and all problems vanish when a shock-fitting technique is used as opposed to a shock-capturing technique (Pandolfi &

D'Ambrosio 1998). Some schemes have a reputation for never producing the carbuncle phenomenon. These are the upwind schemes which do not exactly solve the contact waves, such as all the Flux Vector Splitting methods (van Leer 1982; Pullin 1980; Steger & Warming 1981) or some upwind schemes based on the integral approach such as the HLLE method (Harten, Lax & van Leer 1983). These schemes are suitable for Euler problems but are much less attractive for Navier–Stokes applications since they dramatically broaden boundary layers by adding an overwhelming amount of artificial dissipation.

Among the upwind schemes which exactly resolve grid-aligned contact waves, the vast majority, including an exact Riemann solver such as Godunov's method (Godunov 1959), produce the shock instability. Furthermore, when a very small amount of extra numerical dissipation is added to contact waves, all instability problems disappear but that seriously compromises the accuracy of the solution.

Hence, there exists a trade-off between an exact solver for contact waves, which would allow the confident computation of boundary layers, and the addition of a limited amount of dissipation to contact waves where intense shocks waves are present. This trade-off is the basis of some solutions which have been proposed to date, to remove the carbuncle phenomenon from computed solutions. Some cures (Quirk 1994; Wada & Liou 1997) consist of flagging the cell interfaces, which are located in the vicinity of the shock wave, according to an arbitrary test based for instance on the pressure ratio across the cell interface. A dissipative scheme is used to compute fluxes through flagged cell interfaces while a non-dissipative method is used elsewhere. In this family of solutions, all methods differ in the flagging procedure, some involving a tunable parameter, some taking into account the intrinsic multidimensional mechanism associated with the shock instability.

Other *ad hoc* solutions are specified for a given family of schemes such as Roe's method (see Sanders, Morano & Druguet 1998 and Pandolfi & D'Ambrosia 1998), of the HLL family (see Flandrin, Charrier & Dubroca 1994) and can be described as built-in limiters which selectively add some extra numerical dissipation to damp out spurious oscillations near shocks.

Liou's (1997) analysis includes an interesting conjecture discussed in Xu (1998) which states that a necessary condition for a numerical flux to develop the carbuncle phenomenon is to have a mass flux which depends on the pressure. Liou makes the observation that not only do all tested numerical fluxes which produce the carbuncle phenomenon have a pressure dependence in the mass component, but that all tested numerical fluxes which do not show the carbuncle phenomenon have mass components which are independent of the pressure. Yet, Liou's conjecture, if true, would imply that one can design a numerical flux function which would not produce the carbuncle phenomenon and still maintain the exact resolution of contact waves. However, this conclusion would be in contradiction with Gressier's theorem which states that strict stability for Quirk's problem and exact resolution of contact waves are incompatible (Gressier & Moschetta 1998b). This theorem only applies to first order schemes which only depend on two neighbouring states.

This brief review of the carbuncle phenomenon and its various possible solutions has emphasized the practical importance of this bizarre instability. It has shown that the carbuncle phenomenon has been long regarded as a purely numerical pathology for which the CFD community would have to face a dilemma between the use of upwind schemes which exactly solve contact waves but are prone to develop the carbuncle phenomenon and the use, either local or global, of stable methods which add too much numerical dissipation to the solution to be applied in shear flow regions.
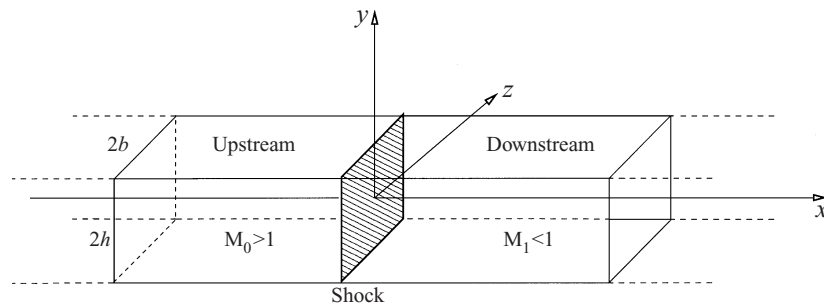
FIGURE 2. The geometry of the flow field considered.

Indeed, there is no doubt that there is a strong influence of the whole numerical procedure (grid stretching, numerical flux functions, higher-order upgrading methods, etc.) on the development of the carbuncle phenomenon. However, this study will demonstrate that the numerical aspects are not the only indication of this phenomenon and that a more fundamental stability mechanism is at stake. Furthermore, this paper aims at highlighting new viewpoint for the CFD community with regards to upwind schemes for the compressible Navier–Stokes equations. It has been common practice for many years to focus on the capability of upwind schemes to exactly solve contact waves and all existing solutions for the carbuncle phenomenon are based on the modification of the linear path in the Riemann problem. This study will suggest that the shock-capturing capability of upwind schemes must be revisited to include the multidimensional interaction between contact waves and shock waves.

To summarize, the general objectives of this paper are first to demonstrate the existence of a new instability mode, and secondly to show that the carbuncle phenomenon seems to have a strong link with this intrinsic instability mode.

The paper is divided into six parts. Following this introduction, the second part is devoted to the theoretical analysis. The linear stability theory of a shock wave is fully analysed, leading to the existence of a new unstable mode (the 'strange mode'). A brief description of Quirk's original problem is given in §3. The proposed methodology is divided in two steps. First, the observed numerical instabilities are shown to follow a linear instability threshold for a given shock-wave Mach number. To the best of our knowledge, this result is not known or not established in the CFD community. These points constitute the fourth part of the report. The second step demonstrates that when a numerical instability occurs it is in agreement with the 'strange mode' highlighted in §2. Therefore the key point is to check the theoretical dispersion relation and the shape of the theoretical eigenfunction. This is achieved in §5. A summary constitutes the last part of the report.

2. The linear stability of plane shock waves

2.1. *Presentation of the problem and assumptions*

The practical configuration corresponds to a planar shock wave propagating at a constant speed $\overline{W}_c$ in a tube, where the mean flow upstream of the shock is at rest. However, in order to simplify the theoretical approach, the Cartesian coordinates system $Oxyz$ is fixed on the shock: the x-coordinate coincides with the axis of channel, and the undisturbed shock front is located at $x = 0$. Both upstream and downstream flows are assumed to be constant and one-dimensional. By convention, the flow moves from region $x < 0$ to region $x > 0$. The flow is considered as a perfect inviscid gas

and the downstream quantities are denoted by subscript 1 and the upstream ones by subscript 0. The geometry and notation are shown in figure 2.

2.2. *Governing equations and boundary conditions*

The general equations of motion for the instantaneous flow are the Euler equations, the energy equation, written for the total enthalpy and the equation of state for a perfect gas:

$$\frac{\partial \rho}{\partial t} + (\boldsymbol{U} \cdot \nabla)\rho = 0, \tag{2.1a}$$

$$\rho \frac{\partial \boldsymbol{U}}{\partial t} + \rho(\boldsymbol{U} \cdot \nabla)\boldsymbol{U} = -\nabla P, \tag{2.1b}$$

$$\rho \frac{\partial h_i}{\partial t} + \rho(\boldsymbol{U} \cdot \nabla)h_i = \frac{\partial P}{\partial t}, \tag{2.1c}$$

$$P = r\rho T, \tag{2.1d}$$

$$h_i = C_p T + \tfrac{1}{2}\boldsymbol{U}^2, \tag{2.1e}$$

where $\boldsymbol{U} = (U, V)^t$ represents both velocity components of the flow. In the channel considered, three types of boundary conditions must be imposed. The first one corresponds to the slip condition on the walls of the channel:

$$V|_{y=\pm h} = 0. \tag{2.2}$$

For each part of the flow, the second boundary condition is imposed at the shock, which is considered as a discontinuity. The instantaneous shock jump relations are the Rankine–Hugoniot relations. The last boundary condition is that there is no fluctuation far from the shock. For all quantities q, the following condition is imposed:

$$\forall y, t \quad \frac{\partial}{\partial t}\left(\lim_{x \to \pm\infty} q(x, y, t)\right) = 0. \tag{2.3}$$

2.3. *Mathematical form of the perturbation*

The present stability theory is based on the classical small-perturbations technique where the instantaneous flow is the superposition of an imposed mean flow and unknown fluctuations. All the physical quantities q (velocity, pressure, etc.) are decomposed into a mean value and a fluctuating one:

$$q(x, y, t) = \bar{q} + q_f(x, y, t). \tag{2.4}$$

The mean value is assumed to be a constant. In fact, this implies that two decompositions are written: one upstream of the shock, the other one downstream. According to the homogeneous form of the boundary conditions as well as the constant form of the mean flow, the perturbation can be described as a normal mode with respect to the different variables x, y, t:

$$q(x, y, t) = \bar{q} + \tilde{q}\mathrm{e}^{\mathrm{i}(k_x x + k_y y - \omega t)} + \text{c.c.}, \tag{2.5}$$

where $\tilde{q}$ is the amplitude of the fluctuations, $\boldsymbol{k} = (k_x, k_y)^t$ is a wavenumber vector with $k_x \in \mathbb{C}$ and $k_y \in \mathbb{R}$, c.c. denoting the complex conjugate. The circular frequency ω is a complex number: its real part ω_r represents the frequency of the perturbation and its imaginary part ω_i a temporal growth rate. The physical meaning of ω_i is in accordance with the classical definition of stability: for $\omega_i > 0$, the mean flow is

unstable whereas for $\omega_i < 0$ the mean flow is stable. Usually, for a stability analysis, either a temporal or a spatial theory is used, depending on the physical nature of the phenomenon. The spatial theory is prefered when the phenomenon takes its origin at a specific place, which corresponds to a complex wavenumber k_x and a real circular frequency ω (Laplace transform with respect to x and Fourier transform with respect to t). The temporal theory is prefered when a temporal origin is defined. In this case, the wavenumber k_x is real and the pulsation ω complex. In our configuration, a space origin in x (the shock) and a temporal origin ($t = 0$ corresponds to the starting position of the shock on the wind tunnel in the computation which will be described in §3) may be introduced. This leads us to consider a space origin in x, in t, and thus to make a Laplace transform in x and in t. In the present case, both k_x and ω are therefore complex numbers.

2.4. *Linearized Euler equations*

The decomposition (2.5) is introduced into equations (2.1*a*) to (2.1*e*). The resulting equations are then simplified, first by taking into account that the mean quantities satisfy the equations and secondly by assuming that the fluctuating quantities are small, so that these equations can be linearized with respect to the disturbance. Finally, the linearized Euler equations become a homogeneous algebraic system:

$$(\boldsymbol{M}_1 - k_x \boldsymbol{M}_2)\boldsymbol{Z} = 0, \tag{2.6}$$

where k_x is the eigenvalue of this problem, $\boldsymbol{Z}$ stands for $(\tilde{T}, \tilde{p}, \tilde{u}, \tilde{v})$ and $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ are (4×4) matrices which depend on the mean flow and the coefficients ω and k_y.

A non-zero solution in (2.6) exists if $\det(\boldsymbol{M}_1 - k_x \boldsymbol{M}_2) = 0$. This condition provides four different wavenumbers:

$$k_x^{(1)} = \frac{-\omega\overline{U} - \bar{a}\Omega}{\bar{a}^2 - \overline{U}^2}, \quad k_x^{(2)} = \frac{-\omega\overline{U} + \bar{a}\Omega}{\bar{a}^2 - \overline{U}^2}, \quad k_x^{(3)} = k_x^{(4)} = \frac{\omega}{\overline{U}}, \tag{2.7}$$

where $\Omega = (\omega^2 - k_y^2(\bar{a}^2 - \overline{U}^2))^{1/2}$, $\bar{a} = (\gamma r \overline{T})^{1/2}$.
The corresponding eigenvectors are

$$\left. \begin{aligned} \boldsymbol{V}_{1,2} &= \left[-\frac{(\overline{U}k_x^{(1,2)} - \omega)}{C_p}, -\frac{\bar{\rho}(\overline{U}k_x^{(1,2)} - \omega)}{\bar{a}^2}, k_x^{(1,2)}, k_y\right]^t, \\ \boldsymbol{V}_3 &= \left[1, -\frac{\bar{\rho}}{\overline{T}}, 0, 0\right]^t, \quad \boldsymbol{V}_4 = \left[0, 0, 1, -\frac{k_x^{(3)}}{k_y}\right]^t. \end{aligned} \right\} \tag{2.8}$$

These eigenvectors can build a base if the determinant of the matrix $\boldsymbol{M}_v$, whose columns are the four eigenvectors $\boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{V}_3, \boldsymbol{V}_4$, is different from zero:

$$\det \boldsymbol{M}_v \neq 0 \Leftrightarrow \omega \neq \pm \mathrm{i}k_y \overline{U} \quad \text{and} \quad \omega \neq \pm k_y(\bar{a}^2 - \overline{U}^2)^{1/2}. \tag{2.9}$$

Physically, if one has $\omega = \pm \mathrm{i}k_y \overline{U}$, then the second acoustic mode has the same phase speed (wavenumber) as the vorticity mode. On the other hand, $\omega = \pm k_y(\bar{a}^2 - \overline{U}^2)^{1/2}$ implies that the both acoustic modes coincide. However, this case is eliminated by the boundary conditions. If the condition (2.9) holds, the general solution of (2.6) can be written

$$\boldsymbol{q}_f = \left(\sum_{j=1}^{4} C_j \boldsymbol{V}_j \mathrm{e}^{\mathrm{i}k_x^{(j)}x}\right) \mathrm{e}^{\mathrm{i}(k_y y - \omega t)}. \tag{2.10}$$

The four coefficients C_j, $j = 1, \dots, 4$, are unknown integration constants which are

respectively related to both acoustic waves ($j = 1, 2$), the entropy wave ($j = 3$) and the vorticity wave ($j = 4$).

2.5. *Introduction of the boundary conditions*

2.5.1. *Wall conditions*

For the fluctuating quantities, the boundary conditions (2.2) become

$$v_f|_{y=\pm h} = 0. \tag{2.11}$$

These relations impose that the wavenumber k_y must have discrete values only if $k_y = n\pi/h$ for $n \in \mathbb{Z}$.

2.5.2. *Conditions at infinity*

For the fluctuating quantities, boundary condition (2.3) becomes

$$\lim_{x \to \pm\infty} q_f(x, y, t) = 0, \quad \forall y, t. \tag{2.12}$$

The amplitudes of the disturbed physical quantities should vanish at infinity. That leads to studying the intrinsic stability of the shock itself and not an external excitation. According to condition (2.12), the following inequalities must be satisfied:

$$\mathrm{Im}\left(k_x^{(i)}\right) > 0 \quad (\text{resp.} < 0) \quad \text{for} \quad x > 0 \quad (\text{resp.} \ x < 0) \quad \text{for} \quad i = 1, 2, 3, \tag{2.13}$$

where $\mathrm{Im}(z)$ is the imaginary part of z. In the present analysis, when the shock wave is temporally unstable ($\mathrm{Im}(\omega) > 0$), equations (2.4) lead to

$$\left.\begin{array}{l} \mathrm{Im}\left(k_x^{(1)}\right) < 0, \quad \mathrm{Im}\left(k_x^{(2)}\right) > 0, \quad \mathrm{Im}\left(k_x^{(3)}\right) > 0 \quad \text{for} \quad x > 0 \\ \mathrm{Im}\left(k_x^{(1)}\right) > 0, \quad \mathrm{Im}\left(k_x^{(2)}\right) > 0, \quad \mathrm{Im}\left(k_x^{(3)}\right) > 0 \quad \text{for} \quad x < 0. \end{array}\right\} \tag{2.14}$$

Thus, the constants C_1, C_2, C_3, C_4 of the upstream flow must be equal to zero. Therefore no unstable fluctuation may exist in the upstream region. On the other hand, downstream of the shock, only the constant C_1 which corresponds to an acoustic wave must be equal to zero. Finally, the fluctuating quantities are written as

$$\left.\begin{array}{l} T_f = \left(-C_2 \dfrac{\overline{U}k_x^{(2)} - \omega}{k_y C_p} \mathrm{e}^{ik_x^{(2)}x} + C_3 \mathrm{e}^{ik_x^{(3)}x}\right) E, \\[3mm] \rho_f = \left(-C_2 \dfrac{\bar{\rho}(\overline{U}k_x^{(2)} - \omega)}{k_y \bar{a}^2} \mathrm{e}^{ik_x^{(2)}x} - \dfrac{\bar{\rho}}{\overline{T}}C_3 \mathrm{e}^{ik_x^{(3)}x}\right) E, \\[3mm] u_f = \left(\dfrac{k_x^{(2)}}{k_y}C_2 \mathrm{e}^{ik_x^{(2)}x} + C_4 \mathrm{e}^{ik_x^{(3)}x}\right) E, \\[3mm] v_f = \left(C_2 \mathrm{e}^{ik_x^{(2)}x} - \dfrac{\omega}{k_y \overline{U}}C_4 \mathrm{e}^{ik_x^{(3)}x}\right) E, \end{array}\right\} \tag{2.15}$$

with $E = \mathrm{e}^{i(k_y y - \omega t)}$.

This analysis clearly shows the reason why Dýakov (1954) removed one of the acoustic waves in his analysis. It also explains and justifies why Dýakov (1954) and Swan & Fowles (1975) did not consider any fluctuation upstream of the shock.

2.5.3. *Linearized Rankine–Hugoniot relations*

The same small perturbation technique (2.5) is used for the shock equations. The perturbed position of the shock is written as

$$x = f(y, t) = \bar{x}_c + X\mathrm{e}^{i(k_y y - \omega t)} + \text{c.c.}, \tag{2.16}$$

where $\bar{x}_c$ is the mean shock position (here $\bar{x}_c = 0$ due to the choice of the coordinate system) and X represents the shock oscillation amplitude. The latter is assumed to be a small (complex) quantity. The expressions for the normal vector $\boldsymbol{n}$ and the tangential vector $\boldsymbol{\tau}$ can be deduced from equation (2.16). At first order:

$$\left.\begin{aligned}
\boldsymbol{n} &= (1, -\partial f/\partial y)^t = (1, -ik_y XE)^t \\
\boldsymbol{\tau} &= (\partial f/\partial y, 1)^t = (ik_y XE, 1)^t.
\end{aligned}\right\} \tag{2.17}$$

The Rankine–Hugoniot relations are then linearized by performing a first-order Taylor expansion with respect to the fluctuating quantities. For example, q_1 is the value of the quantity q, which is itself the sum of the mean and the fluctuating quantities. Both are evaluated just downstream of the perturbed shock position. q_1 is expressed as

$$q_1(\bar{x}_c + XE, y, t) = \bar{q}_1(\bar{x}_c + XE) + q_{1_f}(\bar{x}_c + XE, y, t). \tag{2.18}$$

As the coordinate system is such that $\bar{x}_c = 0$, q_1 and q_0 are expanded into

$$q_i(XE, y, t) = \bar{q}_i(0) + \breve{q}_i E, \quad \text{for} \quad i = 0, 1, \tag{2.19}$$

where $\breve{q}$ is the amplitude of the fluctuation at the mean shock position. After some calculation, the linearized shock relations lead to an algebraic system of equations:

$$\boldsymbol{A}_1 \boldsymbol{Z}_1(0) = \boldsymbol{\xi} X + \boldsymbol{A}_0 \boldsymbol{Z}_0(0), \tag{2.20}$$

where $\boldsymbol{Z}_i(0)$ (for $i = 0, 1$) is the fluctuating-amplitude vector calculated at the mean shock position. $\boldsymbol{\xi}$ is a complex vector and $\boldsymbol{A}_0$, $\boldsymbol{A}_1$ are fourth-order complex matrices. As explained above, the stability analysis of the Euler equations showed that no intrinsic fluctuation can exist upstream of the shock ($\boldsymbol{Z}_0 \equiv 0$). Thus, for a one-dimensional constant flow on both sides of the shock, the general linearized equations of Rankine–Hugoniot (2.20) can be simplified into

$$\left.\begin{aligned}
&\overline{U}_1 \rho_f + \bar{\rho}_1 u_f = i\omega(\bar{\rho}_0 - \bar{\rho}_1)XE, \\
&(r\overline{T}_1 + \overline{U}_1^2)\rho_f + r\bar{\rho}_1 T_f + 2\bar{\rho}_1 \overline{U}_1 u_f = 0, \\
&C_p T_f + \overline{U}_1 u_f = i\omega(\overline{U}_0 - \overline{U}_1)XE, \\
&v_f = ik_y(\overline{U}_0 - \overline{U}_1)XE.
\end{aligned}\right\} \tag{2.21}$$

2.6. Eigenvalue problem

Substituting the expressions for T_f, ρ_f, u_f, v_f (2.15) at $x = 0$ into the shock relations (2.21) leads to an algebraic system:

$$\boldsymbol{G}\boldsymbol{\chi} = 0, \tag{2.22}$$

where $\boldsymbol{\chi} = (C_2, C_3, C_4, X)^t$ is the unknown vector and $\boldsymbol{G}$ is a fourth-order matrix, which depends on ω, k_y and the mean flow values:

$$\boldsymbol{G} = \begin{pmatrix}
\dfrac{\bar{\rho}_1}{k_y}\left((1 - \overline{M}_1^2)k_x^{(2)} + \dfrac{\omega \overline{M}_1}{\bar{a}_1}\right) & \dfrac{\bar{\rho}_1 \overline{U}_1}{\overline{T}_1} & \bar{\rho}_1 & -i\omega\,[\bar{\rho}] \\[2ex]
\dfrac{\bar{\rho}_1}{k_y}(2\overline{U}_1 k_x^{(2)} - (1 + \overline{M}_1^2)(\overline{U}_1 k_x^{(2)} - \omega)) & \dfrac{\bar{\rho}_1 \overline{U}_1^2}{\overline{T}_1} & 2\bar{\rho}_1 \overline{U}_1 & 0 \\[2ex]
\dfrac{\omega}{k_y} & C_p & \overline{U}_1 & -i\omega[\overline{U}] \\[2ex]
1 & 0 & -\dfrac{\omega}{k_y \overline{U}_1} & -ik_y[\overline{U}]
\end{pmatrix}. \tag{2.23}$$

A non-zero solution can exist if the rank of this system of four relations is less than four. Hence, that the determinant must be zero: det $\mathbf{G} = 0$. This condition yields a dispersion relation.

After tedious calculations, this dispersion relation has been analytically obtained as

$$2\frac{\overline{U}_1}{\overline{U}_0}\omega\left(k_y^2 + \frac{\omega^2}{\overline{U}_1^2}\right) - \left(\frac{\omega^2}{\overline{U}_1\overline{U}_0} + k_y^2\right)(\omega - \overline{U}_1 k_x^{(2)})\left(\frac{\overline{M}_0^2 - 1}{\overline{M}_0^2}\right) = 0. \qquad (2.24)$$

One can immediately note that this relation is identical to the one obtained by Dýakov and by Fowles & Houwing, in the case of a perfect gas.

2.7. *Explicit resolution of the dispersion equation*

Let us solve (2.24) in more detail. Substituting $k_x^{(2)}$ into the dispersion relation (2.24) leads to an algebraic equation in ω. Then, defining $\Theta = \omega/(\overline{a}_1 k_y)$, the dispersion relation (2.24) becomes

$$2\frac{\overline{U}_1}{\overline{U}_0}(1 - \overline{M}_1^2)\Theta(\Theta^2 + \overline{M}_1^2) - \left(\frac{\overline{U}_1}{\overline{U}_0}\Theta^2 + \overline{M}_1^2\right)(\Theta - \overline{M}_1(\Theta^2 + \overline{M}_1^2 - 1)^{1/2})\phi = 0, \quad (2.25)$$

with $\phi = (\overline{M}_0^2 - 1)/\overline{M}_0^2$. After some calculations, the dispersion relation (2.25) becomes

$$(\Theta^2 + \overline{M}_1^2)(f_1\Theta^4 + f_2\Theta^2 + f_3) = 0, \qquad (2.26)$$

with

$$f_1 = \frac{\overline{U}_1^2}{\overline{U}_0^2}(\phi^2 - 4\phi + 4(1 - \overline{M}_1^2)),$$

$$f_2 = 2\frac{\overline{U}_1}{\overline{U}_0}\left(\phi^2 - 2\phi + 2\frac{\overline{U}_1}{\overline{U}_0}(1 - \overline{M}_1^2)\right),$$

$$f_3 = (\overline{M}_1^2\phi)^2.$$

The dispersion relation (2.26) has a solution given by $\Theta = \pm i\overline{M}_1$, where only the positive root is the solution of the dispersion relation (2.25). It corresponds to

$$\omega_r = 0 \quad \text{and} \quad \omega_i = k_y\overline{U}_1. \qquad (2.27)$$

This mode corresponds to a value of ω which has been excluded, see (2.9), and a special analysis is necessary, see §2.8. The other roots of (2.26) lead to real values of ω.

However, although these roots are mathematical solutions of (2.25), they are not physically acceptable. Indeed, the conditions (2.14) are not satisfied. This result is in agreement with those of Kontorovitch (1957). He showed that the existence of the marginal mode, corresponding to real and undamped sound waves and entropic waves moving away from a discontinuity, i.e. to the spontaneous emission of sound by the shock wave, is possible for real gas only.

2.8. *The unstable mode*

The dispersion equation (2.24) has been obtained for all complex values of ω, with some exceptions, (2.9), for which the form of the disturbance (2.10) is no longer valid. Surprisingly, this is the case for the mode considered (2.27). To sum up, a stability analysis has been performed assuming that the four eigenvectors in (2.8) are linearly independent (four values of ω are thereby eliminated). A unique mode has

been obtained, but it corresponds to one of the four excluded values. The previous analysis must be therefore performed again from the beginning (§ 2.4) by writing the disturbance in an adapted base. In principle, this must be done for each of the four eliminated values. In fact only $\omega = i\overline{U}_1 k_y$ leads to a non-zero solution compatible with the boundary conditions.

For $\omega = ik_y\overline{U}_1$, the eigenvalues of (2.6) are given by

$$k_x^{(1)} = -ik_y\frac{1+\overline{M}_1^2}{1-\overline{M}_1^2} \quad \text{and} \quad k_x^{(2)} = k_x^{(3)} = k_x^{(4)} = ik_y \quad \text{with} \quad k_y \in \mathbb{N}. \tag{2.28}$$

We will now focus on this mode alone. The eigenspace $E_{k_x^{(1)}} = \mathrm{Ker}\left(\boldsymbol{M}_1 - k_x^{(1)}\boldsymbol{M}_2\right)$ associated with the eigenvalue $k_x^{(1)}$ is of dimension 1 and is generated by

$$\hat{V}_1 = \left(1, \frac{\overline{\rho}_1}{(\gamma-1)\overline{T}_1}, -\frac{\overline{a}_1^2+\overline{U}_1^2}{2(\gamma-1)\overline{U}_1\overline{T}_1}, -i\frac{\overline{a}_1^2-\overline{U}_1^2}{2(\gamma-1)\overline{U}_1\overline{T}_1}\right)^t.$$

The eigenspace $E_{k_x^{(2)}} = \mathrm{Ker}\left(\boldsymbol{M}_1 - k_x^{(2)}\boldsymbol{M}_2\right)^3$ associated with the eigenvalue $k_x^{(2)}$ is of dimension 2 only, and is hence generated by two vectors $\hat{V}_2$, $\hat{V}_3$. Thus there is no base in which the matrix $\boldsymbol{M}_2^{-1}\boldsymbol{M}_1$ is diagonal. The last vector $\hat{V}_4$ is sought such that the matrix $\boldsymbol{M}_2^{-1}\boldsymbol{M}_1$ is in the Jordan form, i.e.

$$\left(\boldsymbol{M}_1 - k_x^{(2)}\boldsymbol{M}_2\right)\hat{V}_4 = \boldsymbol{M}_2\hat{V}_3.$$

Finally

$$\hat{V}_2 = \left(1, -\frac{\overline{\rho}_1}{\overline{T}_1}, 0, 0\right)^t, \quad \hat{V}_3 = (0, 0, 1, -i)^t, \quad \hat{V}_4 = \left(\frac{i\overline{U}_1}{rk_y}, 0, \frac{1}{k_y}, -\frac{1+i}{k_y}\right)^t.$$

In this base, the matrix $\boldsymbol{M}_2^{-1}\boldsymbol{M}_1$ is given by

$$\boldsymbol{J} = \begin{pmatrix} -ik_y\dfrac{1+\overline{M}_1^2}{1-\overline{M}_1^2} & 0 & 0 & 0 \\ 0 & ik_y & 0 & 0 \\ 0 & 0 & ik_y & 1 \\ 0 & 0 & 0 & ik_y \end{pmatrix}. \tag{2.29}$$

The primitive fluctuating quantities, $(\tilde{T}, \tilde{p}, \tilde{u}, \tilde{v})$ can be found by

$$\boldsymbol{Z} = \boldsymbol{P}\mathrm{e}^{\mathrm{i}x\boldsymbol{J}}\hat{C},$$

where $\boldsymbol{P}$ is the base transformation matrix, the columns of which are the eigenvectors of matrix $\boldsymbol{M}_2^{-1}\boldsymbol{M}_1$ and $\hat{C}$ is the vector $(\hat{C}_1, \hat{C}_2, \hat{C}_3, \hat{C}_4)^t$ formed with the integration constants. Ultimately, the general solution of (2.6) for the mode $\omega = ik_y\overline{U}_1$ is

$$\mathbf{q}_f = [\hat{C}_1\hat{V}_1\mathrm{e}^{ik_x^{(1)}x} + \hat{C}_2\hat{V}_2\mathrm{e}^{ik_x^{(2)}x} + (\hat{C}_3 + ix\hat{C}_4)\hat{V}_3\mathrm{e}^{ik_x^{(3)}x} + \hat{C}_4\hat{V}_4\mathrm{e}^{ik_x^{(3)}x}]\mathrm{e}^{i(k_yy-\omega t)}. \tag{2.30}$$

For the same reasons as before, $\hat{C}_1$ is equal to zero. Expression (2.30) where $x = 0$ is then introduced into the linearized Rankine–Hugoniot relations (2.21). The following algebraic system is obtained:

$$\hat{\boldsymbol{G}}\hat{\chi} = 0, \tag{2.31}$$

where $\hat{\chi} = (\hat{C}_2, \hat{C}_3, \hat{C}_4, X)^t$ is the unknown vector and $\hat{\boldsymbol{G}}$ is a fourth-order matrix, which depends on k_y and on the mean flow values. The mode $\omega = ik_y\overline{U}_1$ only exists

if the determinant of $\hat{\mathbf{G}}$ is equal to zero. It can be easily shown that the determinant of $\hat{\mathbf{G}}$ is only a function of the upstream Mach number $\overline{M}_0$ and γ:

$$\det \hat{\mathbf{G}} \propto (\gamma^2 - 3\gamma)\overline{M}_0^4 + (\gamma^2 + 6\gamma - 3)\overline{M}_0^2 + \gamma + 5. \tag{2.32}$$

A first consequence of this result is that the wavenumber k_y does not act upon the condition of instability. The roots of this biquadratric equation are

$$\overline{M}_0^{(1)} = i(\gamma)^{-1/2}, \quad \overline{M}_0^{(2)} = -i(\gamma)^{-1/2}, \quad \overline{M}_0^{(3)} = \left(\frac{5+\gamma}{3-\gamma}\right)^{1/2}, \quad \overline{M}_0^{(4)} = -\left(\frac{5+\gamma}{3-\gamma}\right)^{1/2}. \tag{2.33}$$

The only acceptable solution is:

$$\boxed{\overline{M}_{0_c} = \left(\frac{5+\gamma}{3-\gamma}\right)^{1/2}.} \tag{2.34}$$

This instability is very surprising because it exists only for one value of the upstream Mach number; for example: $\overline{M}_{0_c} = 2$ for $\gamma = 1.4$. Using the critical upstream Mach number $\overline{M}_{0_c}$ as a function of γ, one can calculate the associated critical downstream Mach number $\overline{M}_{1_c}$ with the Rankine–Hugoniot relations:

$$\overline{M}_1^2 = \frac{2 + (\gamma - 1)\overline{M}_{0_c}^2}{2\gamma\overline{M}_{0_c}^2 - (\gamma - 1)} = \frac{1}{3},$$

which shows a surprising independence an γ.

To determine the shape of the eigenfunctions, the coefficients $\hat{C}_j$ are calculated according to the amplitude of the shock displacement X. The resolution of the system $\hat{\mathbf{G}}\hat{\chi} = 0$ gives

$$\left.\begin{aligned}
\hat{C}_2 &= \frac{\overline{U}_1[\overline{U}]}{C_p} \frac{2(2\gamma - 1)}{(\gamma - 1)((2\gamma - 1)\overline{M}_1^2 + 3)} k_y X, \\
\hat{C}_3 &= -[\overline{U}] \frac{2(i+1)}{((2\gamma - 1)\overline{M}_1^2 + 3)} k_y X, \\
\hat{C}_4 &= \frac{2ik_y^2[\overline{U}]}{((2\gamma - 1)\overline{M}_1^2 + 3)} X.
\end{aligned}\right\} \tag{2.35}$$

The general solution (2.30) becomes

$$\left.\begin{aligned}
p_f &= \kappa_p e^{-k_y(x - \overline{U}_1 t)} e^{ik_y y}, \\
u_f &= \kappa_u \left(1 + k_y x\right) e^{-k_y(x - \overline{U}_1 t)} e^{ik_y y}, \\
v_f &= \kappa_v k_y x e^{-k_y(x - \overline{U}_1 t)} e^{ik_y y},
\end{aligned}\right\} \tag{2.36}$$

with: $\kappa_p = -\kappa_1 \overline{\rho_1 U_1}$, $\kappa_u = -\kappa_1$, $\kappa_v = i\kappa_1$ and $\kappa_1 = 2[\overline{U}]k_y X / [(2\gamma - 1)\overline{M}_1^2 + 3]$.

It is remarkable that the transverse fluctuation speed (and only this quantity) is continuous through the shock. To calculate these eigenfunctions, non-dimensional quantities have been defined:

$$x^* = \frac{x}{h}, \quad t^* = \frac{\overline{a}_0 t}{h} \quad \text{and} \quad v_f^* = \frac{v_f}{\overline{a}_0}, \quad u_f^* = \frac{u_f}{\overline{a}_0}, \quad p_f^* = \frac{p_f}{\overline{\rho}_0 \overline{a}_0^2}.$$

Figure 3 shows the evolution of the amplitude of u_f^*, v_f^* and p_f^* for $n = 1$ ($k_y = n\pi/h$) and the same X.
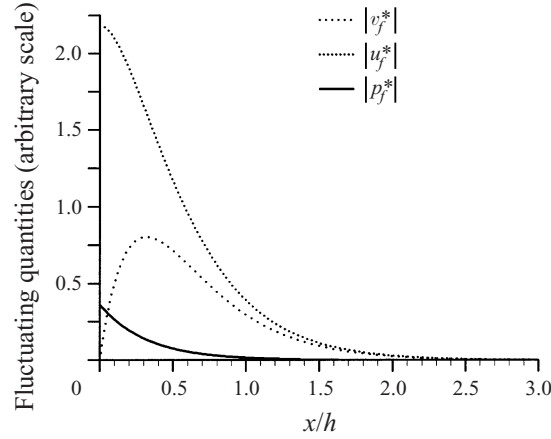
FIGURE 3. Eigenfunctions of the mode $\omega = \mathrm{i}k_y \overline{U}_1$.

This stability analysis of a shock wave has highlighted an unknown unstable mode, which exists for one, and only one, value of the upstream Mach number, and furthermore characterizes the above eigenfunctions. The existence of this intrinsic instability in the continuous Euler equations has significant consequences for the numerical resolution of these equations. The following sections aim to show that the intrinsic instability of the continuous equations allows a different explanation of the phenomenon, up until now regarded as numerical pathology: Quirk's problem.

3. The computational instability

First, a general description of the numerical method is given. Some features of the various schemes used in this study are described. Then, a review of the available knowledge concerning the carbuncle phenomenon and the odd–even decoupling is detailed.

3.1. *Numerical background*

Let us consider the Euler equations in the conservative form

$$\frac{\partial \mathscr{U}}{\partial t} + \nabla \cdot \mathscr{H} = 0, \tag{3.1}$$

where $\mathscr{U}$ is the vector of conservative variables $(\rho, \rho U, \rho e)^t$ where ρe denotes the total energy and $\mathscr{H}$ the tensor of convective terms $(\rho U, \rho U \cdot U^t + P, \rho U h_i)^t$. Partial differential Euler equations are written in integral form, by integrating over a volume. According to the finite volume method, cells are

$$\int_{\Omega_{i,j}} \frac{\partial \mathscr{U}}{\partial t} \mathrm{d}\Omega + \oint_{\partial \Omega_{i,j}} \mathscr{H} \cdot n \mathrm{d}S = 0. \tag{3.2}$$

If $\mathscr{U}_{i,j}$ denotes the average value of $\mathscr{U}$ in $\Omega_{i,j}$, conservative explicit methods on a structured mesh can be expressed in the following form:

$$\mathscr{U}_{i,j}^{n+1} - \mathscr{U}_{i,j}^{n} + \frac{\Delta t}{\Omega_{i,j}} \left(\left[SF_n^\star \right]_{i+1/2,j} - \left[SF_n^\star \right]_{i-1/2,j} + \left[SF_n^\star \right]_{i,j+1/2} - \left[SF_n^\star \right]_{i,j-1/2} \right) = 0, \tag{3.3}$$

where $S_{i+1/2,j}$ is a measure of the interface between $\Omega_{i,j}$ and $\Omega_{i+1,j}$ and $(F_n^\star)_{i+1/2,j}$ the numerical flux evaluated on the same interface with the associated normal vector $n_{i+1/2,j}$. The numerical flux is a function of two states and completely depends on the scheme used. In first-order schemes, both these states are the average states of cells separated by the interface considered. The numerical flux must satisfy the consistency condition

$$F_n^\star(\mathcal{U}, \mathcal{U}) = \mathcal{H}(\mathcal{U}) \cdot n = \begin{pmatrix} \rho u_n \\ \rho U u_n + p n \\ \rho u_n h_i \end{pmatrix}, \qquad (3.4)$$

where $u_n = U \cdot n$.

Several schemes are used in this study. Most of them share the same origin: the numerical flux involves a more or less sophisticated solution of the local Riemann problem. This includes the classical Roe (1981) and Osher & Chakravarthy (1983) schemes and the more recent HLLC scheme first proposed by Toro, Spruce & Speares (1994) and modified by Batten *et al.* (1997). All of them share the property of exact resolution of a stationary contact discontinuity. In other words, the numerical dissipation vanishes transversely to a contact discontinuity. In order to investigate numerical dependences, another scheme which does not share this property is used in the numerical study: the EFM kinetic scheme has been proposed by Pullin (1980) and yields both robustness and a great deal of numerical dissipation on contact waves. It belongs to the class of flux vector splitting schemes and generates similar results to those obtained by van Leer (1982) scheme or the two-wave variants of the HLL family, see Harten *et al.* (1983) and Einfeldt *et al.* (1991).

All these methods are originally put forward with first-order accuracy. Some higher-order computations are presented using the classical MUSCL extension of van Leer (1979).

Note that the time step is computed from a classical CFL-like condition. It is checked that it is ruled by the longitudinal flow and does not depend on the crosswise size of cells for all computations presented.

3.2. *Description of Quirk's original problem*

This test has been proposed by Quirk (1994). It consists of an unsteady computation of the propagation of a planar shock in a duct, where the flow is initially at rest. Although initial conditions and the expected solution are one-dimensional, the computation involves the two-dimensional Euler equations on a two-dimensional structured grid. The physical problem only depends on the shock-wave Mach number M_s. It is defined as u_s/a_0 where u_s is the speed of the travelling shock wave and a_0 is the speed of sound of the upstream flow, i.e. the flow at rest. M_s was set to 6 in the original test case.

Although the computation is unsteady, one can include the basis of the theoretical analysis by choosing a shock-related coordinate system. The upstream Mach number $\overline{M}_0$ is then the shock-wave Mach number M_s.

The computational mesh is initially a uniform Cartesian grid of 800×20 cells for a 40×1 length unit duct. However, the centreline of the mesh has been slightly perturbed as

$$y_{i,j_{mid}} = y_{j_{mid}} + (-1)^i \times 10^{-6}. \qquad (3.5)$$

A sketch of the grid where the perturbation of the centreline has been exaggerated to make it visible is presented in figure 4. Initial conditions are for a flow at rest. As a left-hand boundary condition, the flow is set to the right inflow state computed from
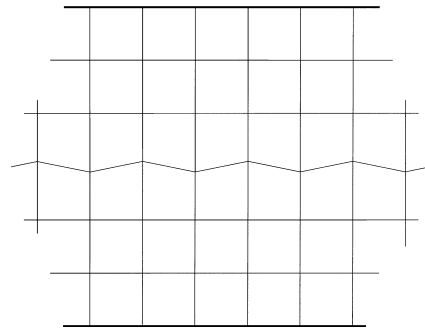
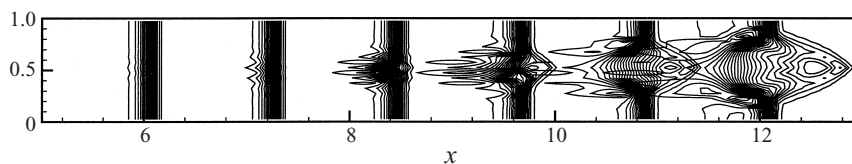FIGURE 4. Sketch of the grid in Quirk's problem.

FIGURE 5. Temporal evolution of the unstable shock wave in Quirk's problem: density contours.

the Rankine–Hugoniot equations so that the shock propagates with the right speed. The right-hand boundary condition is a simple extrapolation technique. The upper and lower bounds are treated as symmetry lines to simulate the wall condition for inviscid flows.

This test is known to result in an unexpected disturbance of the shock shape. The expected solution is a discrete representation of a sharp shock with a constant velocity, which is one obvious solution of the continuous equations. Quirk reported this insidious failing and linked it to high-resolution computations of planar shock waves. This instability has been named *odd–even decoupling* and is the unexpected growth of perturbations along planar shock which are aligned with the mesh.

A classical example of this instability follows: in figure 5, the time evolution is represented through six successive snapshots which have been superimposed on the same duct. As the shock propagates downstream, perturbations appear at the intersection of the shock and the centreline. They grow in the transverse direction and dramatically perturb the shock shape the velocity of which increases slightly until it completely breaks down.

In the same paper, Quirk (1994) proposed an analysis of Roe's scheme and the HLLE Einfeldt *et al.* (1991) scheme. A more detailed analysis has been proposed by Gressier & Moschetta (1998*a*) and links this pathological behaviour to the marginal or neutral stability of the method in a simplified form of the discrete conservation equation (3.3). The same results have been confirmed for many schemes by Pandolfi & D'Ambrosia (1998). Later, all schemes which yield vanishing numerical dissipation on contact waves were proved to be neutrally stable and then exposed to highlight the instability (see Gressier & Moschetta 1999).

However, in the following section, numerical computations are performed in order to demonstrate the connection with the analysis presented in §2. The numerical schemes will be analysed not for their discrete stability properties but for their capability of reproducing the right dynamics of the continuous equations which have been discretized.
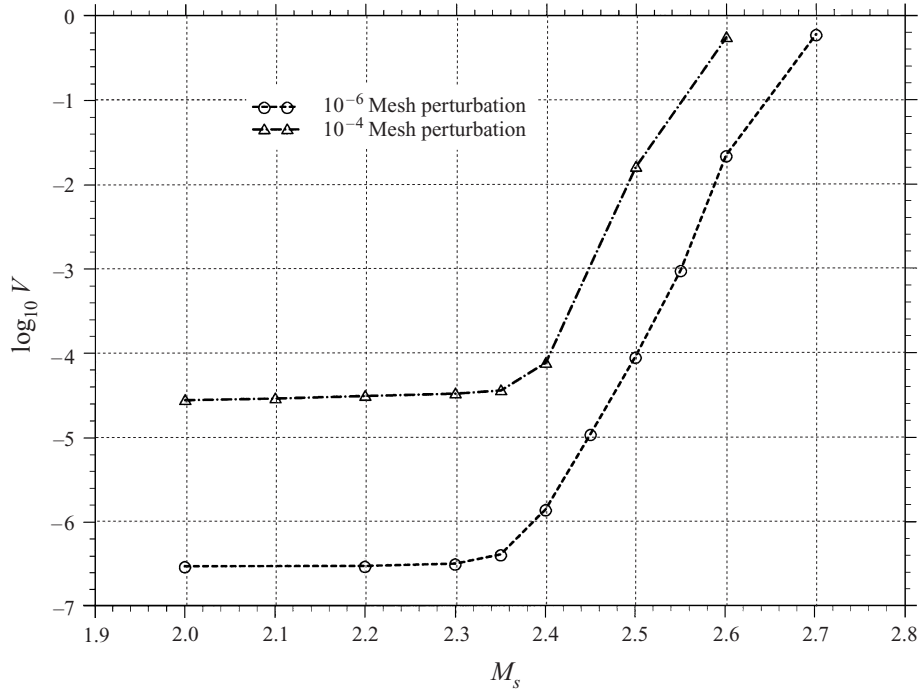
FIGURE 6. Instability threshold for two values of the mesh perturbation
(Roe's scheme, CFL = 0.45).

4. Numerical results and dependences

The aim of this section is to compare the shock instability presented by the theoretical analysis (§2) with the numerical disturbance of the shock profile observed in Quirk's problem. Although the time evolution observed in figure 5 is numerical evidence of this instability, that behaviour appears in the nonlinear regime and cannot be used to study the linear stability of the flow.

Perturbations must be extracted when their magnitudes are still small enough compared to the average flow quantities. But perturbations around the shock are extremely difficult to extract: the computation is not stationary and the shock is thick. Since the variations in the shock thickness are not known, one cannot extract perturbations without introducing errors which could be larger than the perturbations themselves.

On the other hand, since the propagation of a stable shock is one-dimensional, the transverse velocity is expected to remain zero. Then, every non-zero value of the transverse velocity directly stands for the perturbation of this quantity. In the following, the state of the perturbed flow has therefore been represented by the maximum value of the transverse velocity in the flow.

The instability of the shock is produced using Quirk's problem with different shock-wave Mach numbers M_s.

4.1. *Instability threshold*

The classical test case is performed with a shock-wave Mach number of 6. When computing the same problem with different shock-wave Mach numbers M_s, a threshold is found: the shock instability does not appear when M_s is below a threshold M_s^c.

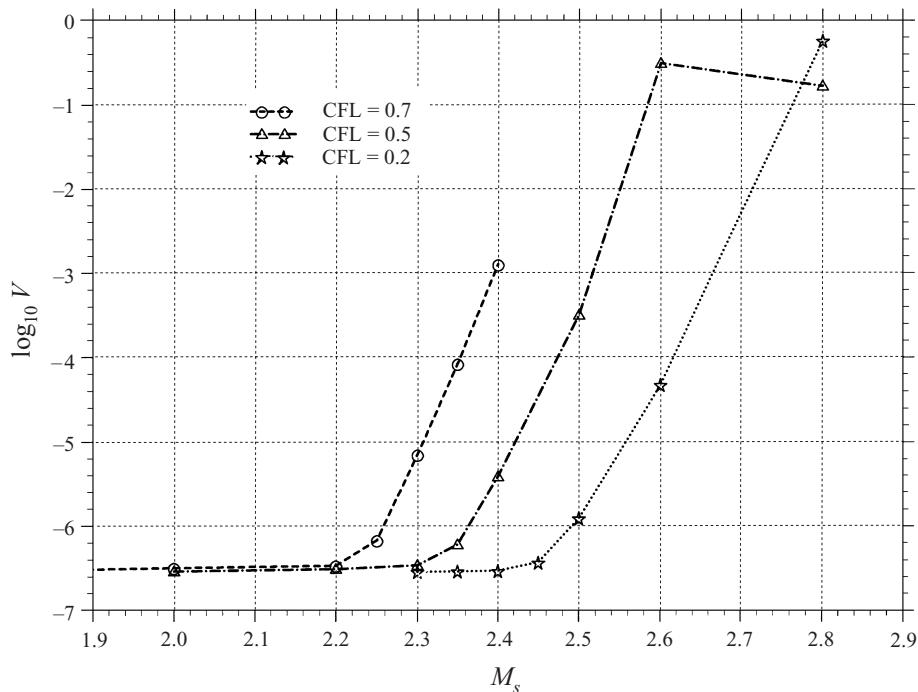In figure 6, the magnitude of the transverse velocity when the shock arrives near

FIGURE 7. Instability threshold dependence on the CFL number (Roe's scheme).

the end of the duct is plotted versus M_s, the shock-wave Mach number. When M_s is below approximately 2.3, perturbations are about the mesh perturbation magnitude (10^{-6} or 10^{-4}) at the end of the computation. This magnitude can be observed even for very dissipative schemes which do not make the shock instability appear: it is just a consequence of a mesh perturbation as a forced mode. Indeed, the mesh perturbation produces small perturbations of the physical quantities but they remain at a low level and are restricted to around the centreline. In this case, the instability is not expected to appear even for longer ducts. On the other hand, when the shock Mach number is above the threshold, at the end of the duct, perturbations have been amplified and are expected to carry on growing. Using the same numerical choices, the threshold is about $M_s^c = 2.35$ and is thus almost independent of the mesh perturbation. This result confirms a linear instability mechanism for which the initial amplitudes do not play any role.

However, the numerical threshold is somewhat larger than the theoretical one, $\overline{M}_{0_c} = 2$. Moreover, it should also be pointed out that the theory predicts an unstable mode only for this singular Mach number while numerical computations demonstrate a threshold. However, to the authors' knowledge, this intrinsic stability threshold has never been mentioned.

Similar curves are plotted in figure 7 using different CFL numbers. The threshold M_s^c is shown to be dependent on the CFL number. This is one of the direct dependences on numerical parameters, which obviously cannot be predicted by the theoretical analysis of continuous equations. However, all the numerical thresholds M_s^c are greater than the theoretical one $\overline{M}_{0_c}$, given by (2.34). Furthermore, the higher the CFL number, the closer to $\overline{M}_{0_c}$ the numerical threshold M_s^c is. For numerical stability reasons, one cannot use CFL numbers higher than 0.7 in order to avoid numerical oscillations in the flow.
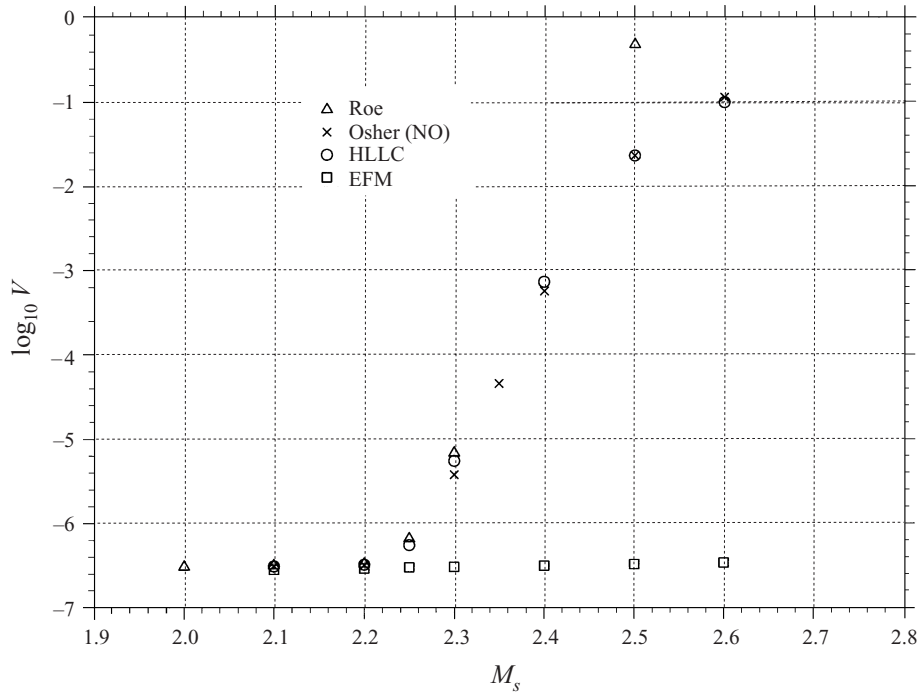
FIGURE 8. Relative independence of instability threshold of schemes (CFL = 0.7).

This dependence is not surprising since the CFL number is intrinsically involved in the numerical dissipation of a given scheme.

Some remarkable results are presented in figure 8: three of the schemes presented yield the same numerical threshold M_s^c. These three schemes are Godunov-type methods, which yield the exact resolution of contact discontinuities and are known to make the odd–even decoupling appear, see Gressier & Moschetta (1998b) and Pandolfi & D'Ambrosio (1998). The linear amplification is the same for the three schemes. Only the nonlinear response differs.

As expected, the EFM kinetic scheme is very robust and does not make the instability appear. Only the forced response can be observed: it does not depend on the shock Mach number M_s and remains at the same level (3×10^{-7}).

These results tend to raise an intrinsic instability in numerical schemes if they are not too dissipative. In other words, the better a scheme is able to solve Euler equations, the more it could suffer from the shock instability.

The final results (figure 9) are aimed at proving that second-order methods do not avoid this instability, see Gressier & Moschetta (1999) for additional results. Moreover, using the same CFL number, which must be low in second-order computations, the shock instability is shown to appear more easily with the second-order scheme since the threshold has decreased. The second-order computation has been performed with a classical MUSCL extension of Roe's scheme (van Leer 1979).

4.2. *Temporal amplification*

In this section, the aim is to determine a numerical amplification factor for comparison with the theoretical one $\omega_i = k_y \overline{U}_1$. This test is severe because high numerical dependences are expected. Nevertheless, the aim is to show similar behaviours of
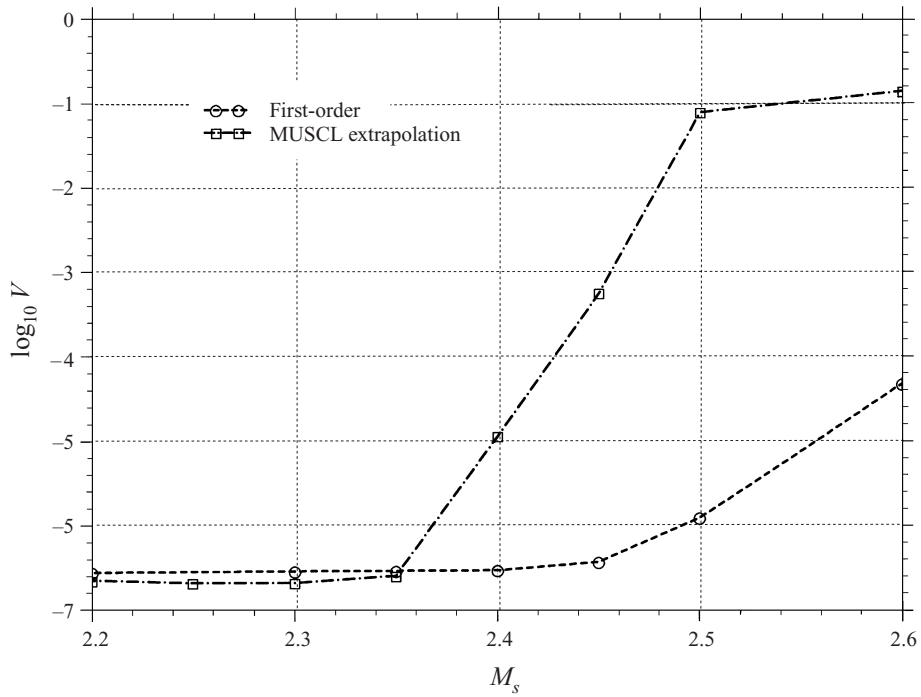
FIGURE 9. Instability threshold for first- and second-order schemes (CFL = 0.2).
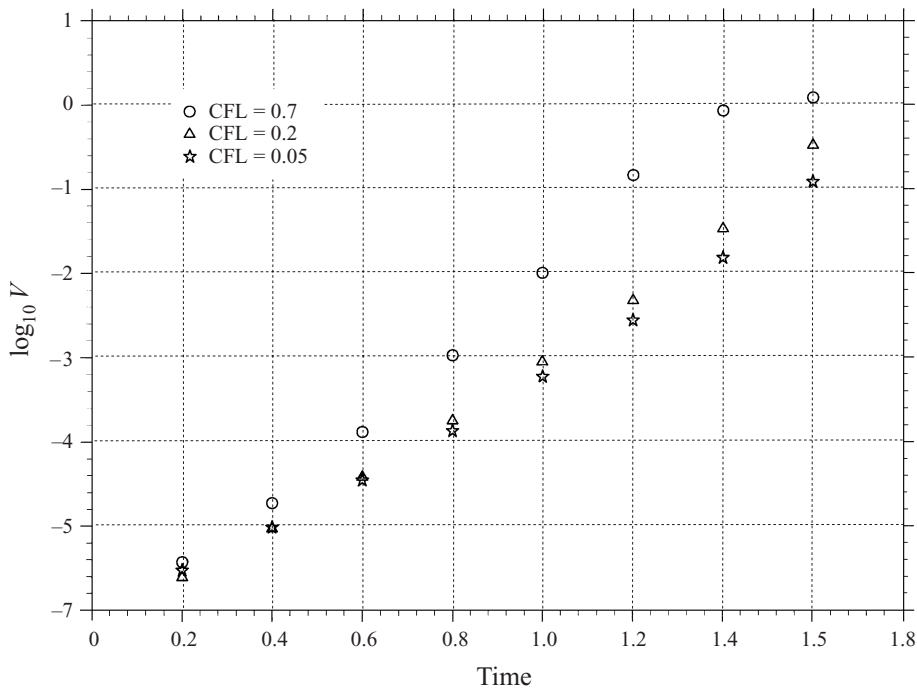

FIGURE 10. Temporal amplification dependence on the CFL number (Roe's scheme, $M_s = 6$).
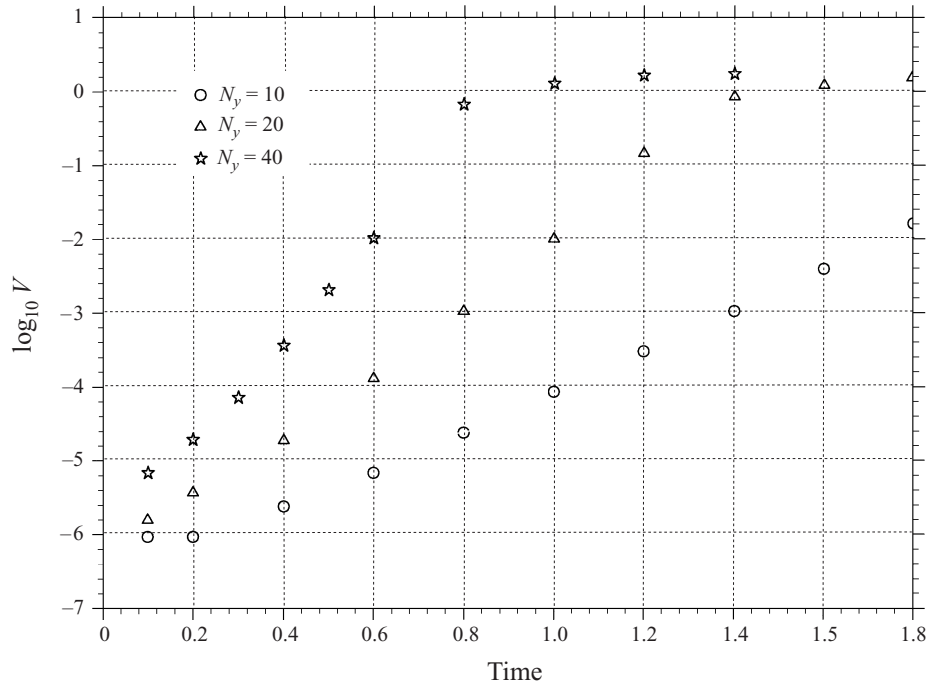
FIGURE 11. Temporal amplification dependence on the number of cells
(Roe's scheme, $M_s = 6$, CFL = 0.7).

N_y	10	20	40
ω_i	6.34	10.8	16.6

TABLE 1. Temporal amplification coefficients.

numerical and theoretical amplification factors. Several computations are performed. The maximum transverse velocity perturbation is plotted at successive time intervals.

Figure 10 shows a marked dependence on the CFL number. This dependence prevents any attempt at precisely recovering the theoretical one. But, as expected, points are more or less aligned for small amplitudes. This indicates, therefore, an exponential growth.

The transverse wavenumber k_y is inversely proportional to the wavelength. Hence, the theory predicts that the smallest wavelengths are the most amplified. The minimum wavelength is the length of two cells. It is verified that oscillations are sawtooth-like. Since ω_i is proportional to k_y, it is predicted that the amplification factor should be related to the number of cells N_y (for a fixed width of the duct). Using successive levels of perturbations, one can compute numerical amplifications factors. Points are expected to be aligned in order to represent an exponential growth. Values between 10^{-5} and 10^{-2} are used to compute the slope in order to avoid the influence of forced responses for low perturbation magnitudes and of the nonlinear response for high perturbation magnitudes.

Figure 11 presents three computations with successively halved sizes of crosswise computing cells. They have been performed using the same shock-wave Mach number

M_s and the same CFL-like condition, i.e. the same time step. While the amplification factors can be expected to double with successive refinement, the computational factors are underestimated and the linear dependence on N_y is not proved. However, the expected trend of the evolution of the amplification factors is clearly confirmed.

5. Theoretical and numerical agreements

5.1. *Space–time behaviour*

The aim of this section is to prove that the numerical instability originates in the shock instability detailed in the first section. In the linear stability context, the space–time behaviour of the instability is ruled by the dispersion equation. It links the temporal amplification ω to the spatial behaviour, which is represented by k_x and k_y. Hence, the following relations strongly characterize the instability:

$$\omega = ik_y\overline{U}_1 \quad \text{and} \quad k_x = ik_y. \tag{5.1}$$

In order to check these relations, the spatial behaviour of the fluctuating quantities in the numerical computations are compared with theoretical predictions. The method used is to measure the temporal amplification of the perturbations from the numerical results. Then, one can derive the wavenumbers k_x and k_y and compare the spatial behaviour of the fluctuating quantities in the numerical computation to the behaviour which is theoretically predicted from the amplification factor via the dispersion relation.

Applying this general framework to a given computation, which has a shock-wave Mach number of 6 and a CFL number of 0.7, the eigenfunctions will be compared to the theoretical predictions at the time $t = 1$ s. Only the evolution of the fluctuating transverse velocity is presented: extracting the fluctuations of the other quantities from the instantaneous field generates errors that are too significant to be properly compared. Moreover, comparisons are performed far away from the centreline where the mesh perturbation is likely to perturb the accuracy of the comparisons.

Figure 12 shows the temporal evolution of the maximum value of the transversal velocity in the computation. The comparison should be performed in the linear zone: it must not have too low an amplitude, otherwise it would be disturbed by the forced regime which is caused by the mesh perturbation (3×10^{-7}), and have small enough amplitude to avoid saturation where nonlinear effects cannot be neglected any more (see figure 12). Between these two zones, one can observe the exponential growth of the perturbation: this corresponds to the linear zone. The evolution of the slope of the transverse fluctuation velocity in this zone allows evaluation of the numerical amplification factor $\omega_i = 10.8$. Note that this amplification factor does not depend on the magnitude of the mesh perturbation (see figure 12). Using the temporal amplification factor ω_i, one can write the spatial evolution of the fluctuating transverse velocity eigenfunction as

$$v_f^* = \mathrm{Re}\left[Ax^* e^{ik_x x^*}\right] \stackrel{\text{Dispersion relation}}{\Rightarrow} v_f^* = \mathrm{Re}\left[Ax^* e^{-(\omega_i/\overline{U}_1)x^*}\right],$$

where A is an arbitrary complex amplitude.

Figure 13 shows the comparison between the numerical spatial evolution and the prediction through the dispersion relation. Several evolutions have been plotted using different heights y. Since they are extracted from the same computation, the amplification factor ω_i is the same. The four spatial evolutions are then predicted to be proportional to each other. Since the theoretical perturbation is defined within a
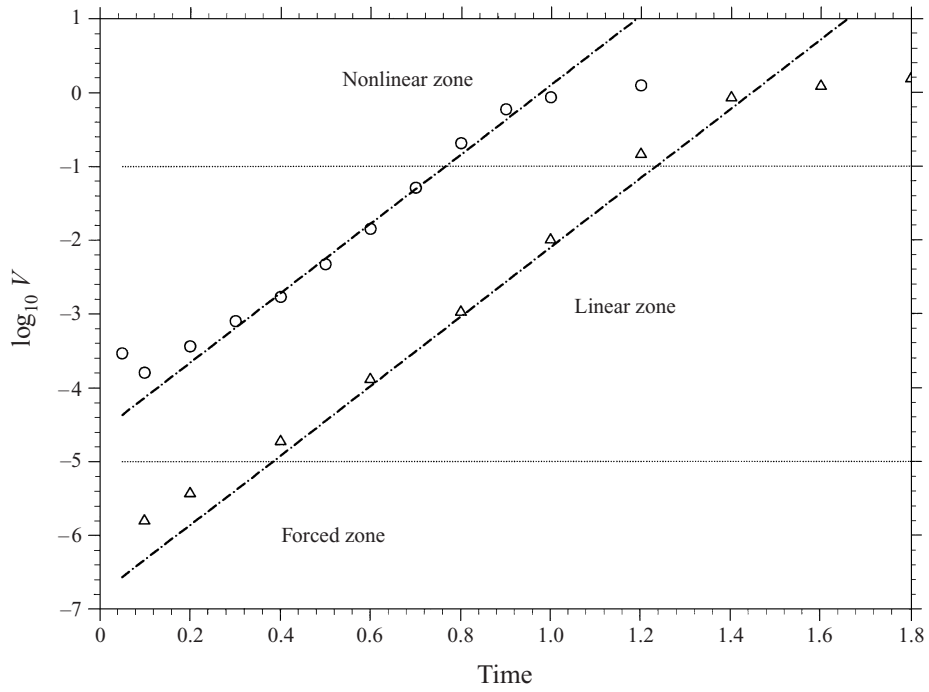
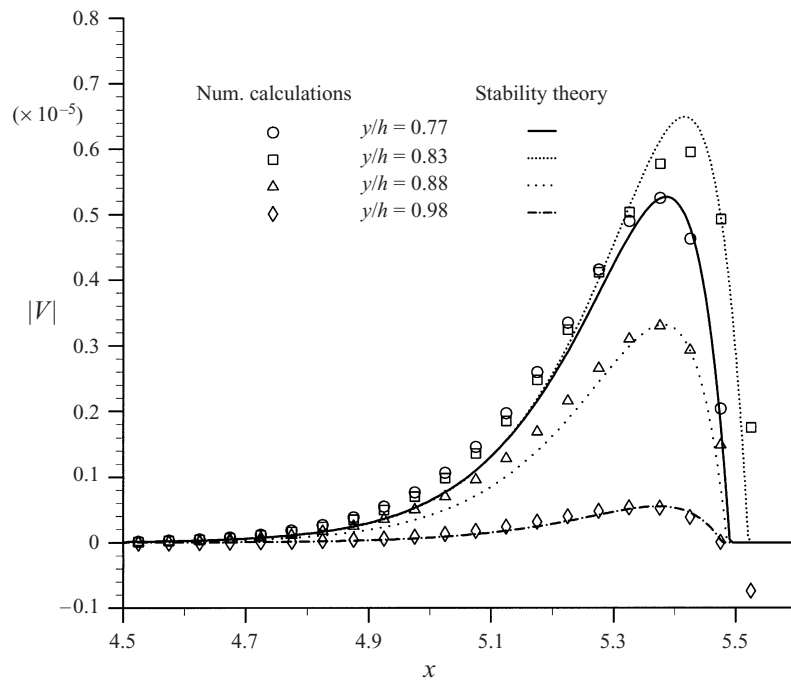FIGURE 12. Temporal amplification ($N_y = 20$, CFL $= 0.7$).
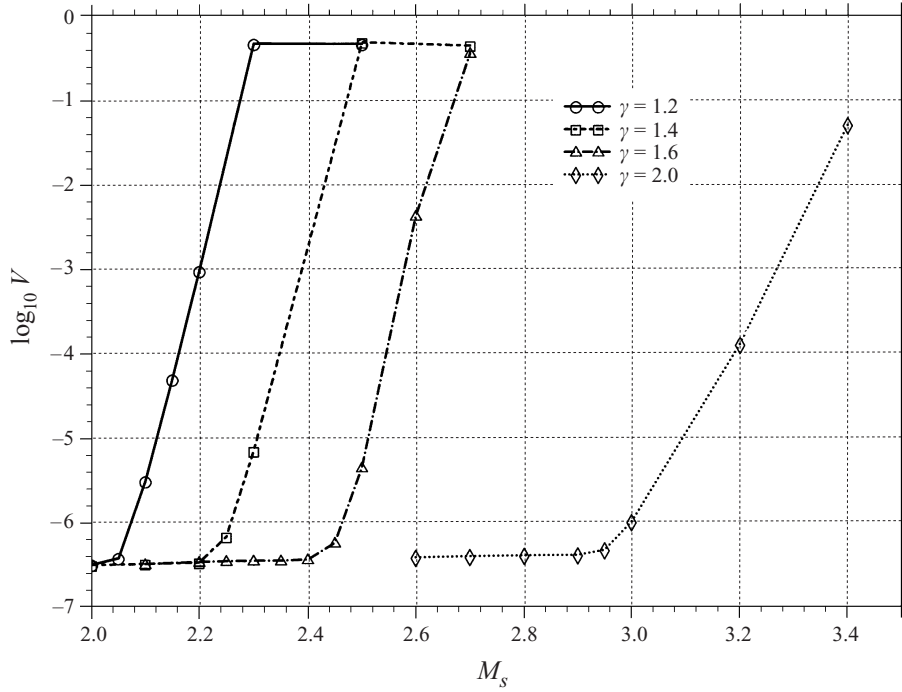

FIGURE 13. Transversal velocity comparison ($t = 1\,\mathrm{s}$).

FIGURE 14. Instability threshold dependence on γ (Roe's scheme, CFL $= 0.7$).

constant of proportionality, the amplitude A of the theoretical eigenfunction has been tuned to fit the numerical evolution in x; this amplitude has been determined for each y independently. Results in figure 13 show good agreement for each section in the channel and so establish the strong link between the theoretical dispersion relation and the computation.

The principal features of the transverse fluctuation velocity are retrieved. The fluctuating velocity is continuous across the shock (see (2.36)), it yields a maximum before decreasing far from the shock and the location of its maximum is correct. Theoretically, the location of this maximum, x_m, is proportional to $\overline{U}_1/\omega_i$.

The validation of the theoretical dispersion relation which strongly features the instability proves that both numerical and theoretical phenomena are related. Indeed, this numerical pathology would be an intrinsic instability of the continuous Euler equations.

Even if absolute values cannot be predicted by the theory, mainly because of unavoidable numerical dependences, the link between both temporal and spatial behaviour has been confirmed.

5.2. *Gamma dependence*

Several computations have been performed using Roe's scheme, two different CFL numbers (0.45 and 0.7), a large range of shock-wave Mach numbers M_s, and several values of the ratio of specific heats γ. Some of these results are presented in figure 14.

Then, one can determine some numerical thresholds, which are plotted in figure 15(*a*) and which are compared to the theoretical prediction. There is striking agreement between theoretical and numerical results in the physical range of 1.0 to 2.0 for γ. For
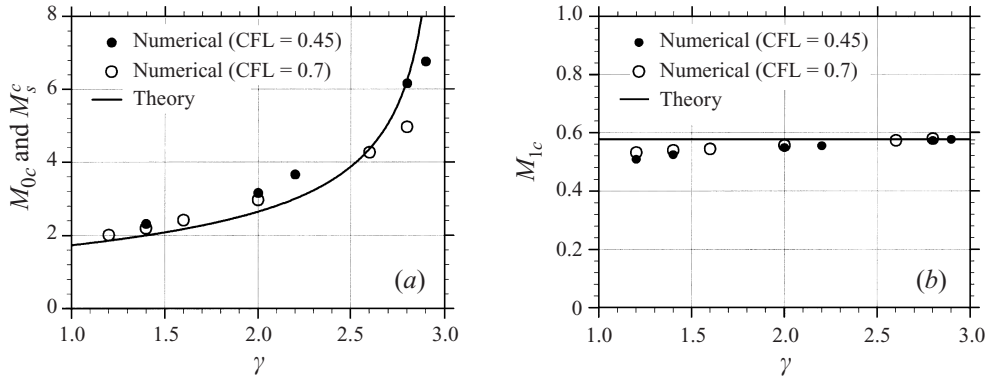
FIGURE 15. (*a*) Upstream and (*b*) downstream instability threshold versus γ (Roe's scheme).

larger values of γ, the numerical thresholds are more sensitive to the CFL numbers used and do not follow the theoretical value as closely.

However, the slight discrepancy observed in the vicinity of $\gamma = 3$ has no bad consequence for practical gas dynamics applications. It should be noticed that $\gamma = 3$ corresponds to a pure one-dimensional gas even though the instability studied is two-dimensional. Furthermore, it was expected that the stiff asymptotic behaviour in the vicinity of $\gamma = 3$ would be difficult to reproduce, see (2.34). The corresponding downstream Mach numbers $\overline{M}_{1_c}$ have been plotted in figure 15(*b*) and are compared to the theoretical prediction which is $1/\sqrt{3}$. In the vicinity of $\gamma = 3$, the downstream Mach number is more important to retrieve since the upstream Mach number increase to infinite values and the downstream flow becomes insensitive to the upstream Mach number. Since the numerical thresholds $\overline{M}_{0_c}$ are minimum values, $\overline{M}_{1_c}$ are maximum values. There is a striking agreement between theoretical and numerical results.

Initially, the existence of a threshold for this numerical behaviour was not known. Moreover, the dependence on γ shows a remarkable agreement with the theoretical prediction of the shock stability analysis. This tends to prove that the two phenomena are closely related.

6. Conclusions

The analytical methodology of the present work is based on three successive steps: identification of a 'strange mode' of the continuous inviscid equations; proof that the numerical carbuncle phenomenon is triggered by an instability mechanism; and demonstration that the two instabilities coincide. Let us summarize in turn the major points of each step.

1. General solutions of linearized perturbed Euler equations are obtained upstream and downstream of the shock respectively. This study has allowed clarification of the mathematical formalism used by Dýakov (1954) and Swan & Fowles (1975). In each zone, the disturbance is written as the sum of four waves, the magnitudes of which are unknown constants. These constants are determined by boundary conditions. These are the linearized Rankine–Hugoniot relations, the slip conditions at the sidewalls and the damping condition for fluctuation far away from the shock. Considering a temporally amplified mode, it is proved that all four unknown constants which feature in the perturbations upstream of the shock should be zero, i.e. there is no amplified mode upstream. Downstream of the shock, one of two acoustic-like perturbations

should vanish (i.e. the corresponding constant should be zero). Finally, the remaining three unknown constants and an additional one which represents the amplitude of the shock displacement, are determined by the shock relations. A dispersion relation is then obtained. After analysing this dispersion relation, one 'curious' unstable mode is shown to satisfy it, although this has never been found by previous analysis. This mode is curious because it does not have the classical form of a normal mode, it comes in fact from a Jordan decomposition of the stability matrix. Moreover, this instability is very surprising because it only exists for one value of the upstream Mach number, $\overline{M}_{0c}$. With the exception of this value this instability does not exist. From a theoretical point of view, the linear stability of a shock wave in a constant and uniform mean flow is now completely solved.

2. The first consequence of the present study affects the field of numerical calculation of shock waves. For many years, a pathological phenomenon, the so-called *carbuncle phenomenon*, has been encountered when computing shock waves. In the CFD community, this behaviour has been usually considered as a purely numerical instability. The carbuncle phenomenon is one example of the numerous situations in which multiple solutions to the Euler equations can be obtained starting from the same initial conditions (Ivanov, Gimelshein & Beylich 1995; Li & Ben-Dor 1997). Even the presence of viscosity does not guarantee the uniqueness of a solution to the Navier–Stokes equation at high Reynolds numbers (Hafez & Guo 1999). In this present paper, a numerical study of the pathology has been performed through a simplified test case, namely Quirk's problem which is likely to be represented by the following analysis. The development of the instability has been proven to be ruled by a linear instability mechanism. Indeed, the temporal growth of the perturbation fits an exponential curve and does not depend on the magnitude of the mesh perturbation. An intrinsic numerical threshold has been determined. It yields a relative independence from the numerical schemes used. This result is totally new, although it was known that the carbuncle phenomenon would be more likely to appear when the Mach number is high (Quirk 1994).

3. Concerning the link between the numerical simulations and the 'curious mode' the following results have been obtained. First, the numerical computations, if unstable, exhibit an instability threshold $\overline{M}_s^c$ in terms of the shock-wave Mach number $\overline{M}_s$ which is independent of the forced perturbation. The value of $\overline{M}_s^c$ obtained is somewhat larger than the theoretical critical value $\overline{M}_{0c}$. Moreover, in the computations, $\overline{M}_s^c$ appears as a threshold whereas $\overline{M}_{0c}$ is a theoretically unique value. These aspects may be due to unavoidable numerical dissipation and to shock thickness. It can be guessed that the instability occurs when the numerical local Mach number crosses the theoretical $\overline{M}_{0c}$ in the shock thickness. We can propose the following explanation. The computed shock has a small but non-zero thickness (there are typically 3 or 4 cells in x in the shock thickness). When the upstream Mach number $\overline{M}_0$ is higher than the critical Mach number $\overline{M}_{0_c}$, there necessarily exists a cell in the shock thickness for which the Mach number that enters is higher than the Mach number $\overline{M}_{0_c}$ and that leaving is lower. One can think that in this cell, the local system to solve is more or less singular; in any case the determinant of the linearized system (for the continuous case) passes through zero. It is the crossing of this singular value which is likely to start the mechanism of instability. Since the numerical shock is not a pure discontinuity, unlike the case of the continuous Euler equations, the critical Mach number appears for higher value than the mechanism of instability. Additional effort must be made in terms of discrete dynamics in order to understand precisely the flow structure in the 3 or 4 cells present in the shock thickness. The key result however is

that this numerical instability is already present in the continuous equations. The two instability mechanisms coincide, as demonstrated by the space–time structure of the perturbation. If unstable, the numerical results exhibit an exponential growth. Due to unavoidable numerical dependences, absolute values are difficult to retrieve. However it is possible for a given computation to extract a theoretical growth rate ω_i. Then the theory predicts the spatial wavenumbers k_x and k_y (from the dispersion relation) and consequently the shape of the eigenfunction. As the latter is in perfect agreement with the numerical results, it can be concluded that the numerical instability coincides with the theoretical mode. Moreover, concerning the link between the threshold $\overline{M}_s^c$ and the coefficient γ, the behaviour is well represented by the theoretical expression.

For many years, it was tacitly assumed in the CFD community that the carbuncle phenomenon was a purely numerical problem. The present work demonstrates that this is not true. The findings of this paper point the way to further analysis which would include a particular form of the numerical flux in order to account for the observed numerical dependences. Since the pathology is intrinsic to the Euler equations, the numerical scheme should be designed in the framework of the Navier–Stokes equations.

REFERENCES

BATTEN, P., CLARKE, N., LAMBERT, C. & CAUSON, D. M. 1997 On the choice of waves speeds for the HLLC Riemann solver. *SIAM J. Sci. Comput.* **18**, 1553–1570.

CASALIS, G. & ROBINET, J.-Ch. 1997 Linear stability analysis in transonic diffuser flows. *Aerospace Sci. Technol.* **1**, 37–47.

DÝAKOV, S. P. 1954 On the stability of shock waves. *Sov. Phys. JETP* **27**, 288–295.

DÝAKOV, S. P. 1957 The interaction of shock waves with small perturbations. *Sov. Phys. JETP* **33**, 948–974.

EINFELDT, B., MUNZ, C. D., ROE, P. L. & SJÖGREEN, B. 1991 On Godunov-type methods near low densities. *SIAM J. Numer. Anal.* **92**, 273–295.

FLANDRIN, L., CHARRIER, P. & DUBROCA, B. 1994 A robust finite volume method for computations on two-dimensional unstructured hybrid meshes. In *Computational Fluid Dynamics*. John Wiley & Sons.

FOWLES, G. R. & HOUWING, A. F. P. 1984 Instabilities of shock and detonation waves. *Phys. Fluids* **27**, 1982–1990.

GODUNOV, S. K. 1959 A difference scheme for numerical computation of discontinuous solutions of hydrodynamics equations. *Math. Sbornik* **47**, 271–306.

GRESSIER, J. & MOSCHETTA, J.-M. 1998a On the marginal stability of upwind schemes. In *16th ICNMFD*, pp. 385–390. Springer.

GRESSIER, J. & MOSCHETTA, J.-M. 1998b On the pathological behavior of upwind schemes. *AIAA Paper* 98-0110.

GRESSIER, J. & MOSCHETTA, J.-M. 1999 Robustness versus accuracy in shock wave computations. *Intl J. Numer. Meth. Fluids* (accepted).

HAFEZ, M. M. & GUO, W. H. 1999 Some anomalies of numerical simulation of shock waves. Part I: inviscid flows. *Comput. Fluids* **28**, 701–719.

HARTEN, A., LAX, P. D. & LEER, B. VAN 1983 On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Rev.* **25**, 35–61.

IVANOV, M. M., GIMELSHEIN, S. F. & BEYLICH, A. E. 1995 Hysteresis effect in stationary reflection of shock waves. *Phys. Fluids* **7**, 685–687.

KONTOROVITCH, V. M. 1957 Reflection and refraction of sound by shock waves. *Sov. Phys. JETP* **33**, 1527.

LANDAU, L. D. & LIFCHITZ, E. M. 1959 *Fluid Mechanics*. Addison-Wesley.

LEER, B. VAN 1979 Towards the ultimate conservative difference scheme V. A second-order sequel to Godunov's method. *J. Comput. Phys.* **32**, 101–136.

LEER, B. VAN 1982 Flux vector splitting for the Euler equations. In *8th ICNMFD*, Lecture Notes in Physics, vol. 170, pp. 505–512. Springer.

LI, H. & BEN-DOR, G. 1997 A parametric study of Mach reflection in steady flows. *J. Fluid. Mech.* **341**, 101–125.

LIOU, M. S. 1997 Probing numerical fluxes: mass flux, positivity and entropy satisfying property. *AIAA Paper* 97-2035.

OSHER, S. & CHAKRAVARTHY, S. 1983 Upwind schemes and boundary conditions with applications to Euler equations in general geometries. *J. Comput. Phys.* **50**, 447–481.

PANDOLFI, M. & D'AMBROSIO, D. 1998 Upwind methods and carbuncle phenomenon. *4th ECCOMAS, Computational Fluid Dynamics 98*, vol. 1, pp. 126–131.

PEERY, K. M. & IMLAY, S. T. 1988 Blunt body flow simulations. *AIAA Paper* 88-2924.

PULLIN, D. I. 1980 Direct simulation methods for compressible inviscid ideal gas flow. *J. Comput. Phys.* **34**, 231–244.

QUIRK, J. J. 1994 A contribution to the great Riemann solver debate. *Intl J. Numer. Meth. Fluids* **18**, 555–574.

ROBINET, J.-CH. & CASALIS, G. 1998 Self-sustained oscillations in diffuser: modelling and predictions. *AIAA Paper* 98-3419.

ROE, P. L. 1981 Approximate Riemann solvers, parameters vectors, and difference schemes. *J. Comput. Phys.* **43**, 357–372.

SAJBEN, M., BOGAR, T. J. & KROUTIL, J. C. 1981 Characteristic frequency and length scales in transonic diffuser flow oscillations. *AIAA Paper* 81-1291.

SANDERS, R., MORANO, E. & DRUGUET, M.-C. 1998 Multidimensional dissipation for upwind schemes: stability and applications to gas dynamics. *J. Comput. Phys.* **145**, 511–537.

STEGER, J. L. & WARMING, R. F. 1981 Flux vector splitting of the inviscid gasdynamics equations with application to finite-difference methods. *J. Comput. Phys.* **40**, 263–293.

SWAN, G. W. & FOWLES, G. R. 1975 Shock wave stability. *Phys. Fluids* **18**, 28–35.

TORO, E. F., SPRUCE, M. & SPEARES, W. 1994 Restoration of the contact surface in the HLL-Riemann solver. *Shock Waves* **4**, 25–34.

WADA, Y. & LIOU, M. S. 1997 An accurate and robust flux splitting scheme for shock and contact discontinuities. *SIAM J. Sci. Comput.* **18**, 633–657.

XU, F. 1987 Interaction of a gasdynamic shock with small disturbance. *Acta Mechanica Sinica* (**3**), 113–122.

XU, K. 1998 Gas kinetic schemes for unsteady compressible flow simulations. *VKI Lecture Series* 1998–03.