

On conditions for weak conservativeness of regularized explicit finite-difference schemes for 1D barotropic gas dynamics equations

A. Zlotnik and T. Lomonosov

Abstract We consider explicit two-level three-point in space finite-difference schemes for solving 1D barotropic gas dynamics equations. The schemes are based on special quasi-gasdynamic and quasi-hydrodynamic regularizations of the system. We linearize the schemes on a constant solution and derive the von Neumann type necessary condition and a CFL type criterion (necessary and sufficient condition) for weak conservativeness in L^2 for the corresponding initial-value problem on the whole line. The criterion is essentially narrower than the necessary condition and wider than a sufficient one obtained recently in a particular case; moreover, it corresponds most well to numerical results for the original gas dynamics system.

Keywords: gas dynamics, barotropic quasi-gas dynamics system of equations, explicit finite-difference schemes, stability criterion, weak conservativeness

1 Introduction

The stability theory for finite-difference schemes for model problems in gas dynamics is well presented in the literature [1, 4, 5, 6, 7, 11, 12]. In this paper we consider some finite-difference schemes for solving 1D barotropic gas dynamics equations. The schemes are explicit, two-level in time and use a symmetric three-point stencil in space. Their construction is based on special quasi-gasdynamic and quasi-hydrodynamic [3, 8, 13, 15, 20] regularizations of the original equations (without a regularization, the schemes are unstable). The schemes of this kind were suc-

A. Zlotnik

National Research University Higher School of Economics, Myasnikskaya 20, 101000 Moscow, Russia e-mail: azlotnik@hse.ru

T. Lomonosov

National Research University Higher School of Economics, Myasnikskaya 20, 101000 Moscow, Russia e-mail: tlomonosov@hse.ru

cessfully applied in numerous and various practical applications, in particular, see [2, 9, 10, 21], but their theory is not developed so well.

We linearize the schemes on a constant solution and derive both the von Neumann type necessary condition and a CFL type criterion for weak conservativeness in L^2 for the corresponding initial-value problem on the whole line. The weak conservativeness in L^2 means the uniform in time bound for the norm of scaled solution by the norm of initial data instead of the energy conservation law for the linearized original system, i.e., the acoustics system of equations. Our numerical experience show that validity of the weak conservativeness property is important since it prevents numerical solutions from the well-known possible spurious oscillations. The property guarantees the uniform in time L^2 stability with respect to initial data.

Since in practice necessary conditions are often in use (a derivation of sufficient conditions is much more complicated in general), it is important to know to what extent this is lawful to do. The criterion turns out to be essentially narrower than the necessary condition and at the same time wider than a sufficient condition obtained recently in a particular case in [14]. Moreover, namely the criterion corresponds most well to results of numerical experiments for the original gas dynamics system. Therefore the criterion (but not the necessary condition or sufficient one) is most adequate and useful for practical purposes.

2 Systems of equations, finite-difference schemes and their linearization

The 1D barotropic gas dynamics (Euler) system of equations consists in the mass and momentum balance equations

$$\partial_t \rho + \partial_x(\rho u) = 0, \quad \partial_t(\rho u) + \partial_x p(\rho) = 0, \quad (1)$$

where $\rho > 0$, u and p are the gas density and velocity (the sought functions) and pressure. We assume that $p'(\rho) > 0$ and consider the equations for $x \in \mathbb{R}$ and $t > 0$.

The 1D barotropic quasi-gas dynamics (QGD) system of equations consists in the regularized mass and momentum balance equations

$$\partial_t \rho + \partial_x j = 0, \quad \partial_t(\rho u) + \partial_x(ju + p(\rho) - \Pi) = 0, \quad (2)$$

$$j = \rho(u - w), \quad w = \frac{\tau}{\rho} u \partial_x(\rho u) + \hat{w}, \quad \hat{w} = \frac{\tau}{\rho} [\rho u \partial_x u + p'(\rho)], \quad (3)$$

$$\Pi = \Pi_{NS} + \rho u \hat{w} + \tau p'(\rho) \partial_x(\rho u), \quad \Pi_{NS} = \mu(\rho) \partial_x u. \quad (4)$$

Here j and Π are the regularized mass flux and stress, w and \hat{w} are the regularizing velocities, $\tau = \tau(\rho) > 0$ is a regularization parameter and Π_{NS} is the Navier-Stokes viscous stress with $\mu(\rho) \geq 0$ being proportional to the viscosity coefficient. In the barotropic case, quasi-gasdynamic and quasi-hydrodynamic systems were introduced and investigated (in multidimensional case) in [20, 16, 17].

The QGD system is simplified into the original system (1) for $\tau = \mu = 0$ and the Navier-Stokes system of equations for viscous compressible barotropic gas flow for $\tau = 0$ and $\mu > 0$.

System (1) can be linearized on a constant solution $\rho_* \equiv \text{const} > 0$ and $u_* = 0$. Substituting the solution in the form $\rho = \rho_* + \Delta\rho$ and $u = u_* + \Delta u$ in the equations and neglecting the terms having the second order of smallness with respect to $\Delta\rho$ and Δu and their derivatives leads us to the following system of equations:

$$\partial_t \Delta\rho + \rho_* \partial_x \Delta u = 0, \quad \rho_* \partial_t \Delta u + p'(\rho_*) \partial_x \Delta\rho = 0. \quad (5)$$

For the dimensionless unknowns $\tilde{\rho} = \frac{\Delta\rho}{\rho_*}$ and $\tilde{u} = \frac{\Delta u}{\sqrt{p'(\rho_*)}}$ we gain the acoustics system of equations:

$$\partial_t \tilde{\rho} + c_* \partial_x \tilde{u} = 0, \quad \partial_t \tilde{u} + c_* \partial_x \tilde{\rho} = 0. \quad (6)$$

Hereafter $c_* = \sqrt{p'(\rho_*)}$ is the background velocity of sound. Given the initial data $\tilde{\rho}|_{t=0} = \tilde{\rho}_0$ and $\tilde{u}|_{t=0} = \tilde{u}_0$ (that one can consider complex-valued), for the solution to the last system the following energy conservation law holds

$$\|\tilde{\rho}(\cdot, t)\|_{L^2(\mathbb{R})}^2 + \|\tilde{u}(\cdot, t)\|_{L^2(\mathbb{R})}^2 = \|\tilde{\rho}_0\|_{L^2(\mathbb{R})}^2 + \|\tilde{u}_0\|_{L^2(\mathbb{R})}^2 \quad \text{for } t \geq 0. \quad (7)$$

Now we pass to discretization. Let ω_h be a uniform mesh on \mathbb{R} with the nodes $x_k = kh$, $k \in \mathbb{Z}$, and step $h = X/N$. Let ω_h^* be an auxiliary mesh with the nodes $x_{k+1/2} = (k+0.5)h$, $k \in \mathbb{Z}$. Define a uniform mesh in t with the nodes $t_m = m\Delta t$, $m \geq 0$, and step $\Delta t > 0$. We define the shift, averaging and difference quotient operators

$$v_{\pm, k} = v_{k\pm 1}, \quad (sv)_{k-1/2} = \frac{v_k + v_{k+1}}{2}, \quad (\delta v)_{k-1/2} = \frac{v_k - v_{k-1}}{h},$$

$$(\delta^* y)_k = \frac{y_{k+1/2} - y_{k-1/2}}{h}, \quad \delta_t v = \frac{v^+ - v}{\Delta t}, \quad v^{+, m} = v^{m+1}.$$

We first consider a standard explicit two-level in time and three-point symmetric in space discretization of the QGD equations (2)-(4):

$$\delta_t \rho + \delta^* j = 0, \quad \delta_t(\rho u) + \delta^*(jsu + p(sp) - \Pi) = 0 \quad \text{on } \omega_h, \quad (8)$$

$$j = (sp)su - (sp)w, \quad (sp)w = (s\tau)[\delta(\rho u)]su + (sp)\hat{w}, \quad (9)$$

$$(sp)\hat{w} = (s\tau)[(sp)(su)\delta u + \delta p(\rho)], \quad (10)$$

$$\Pi = \mu \delta u + (su)(sp)\hat{w} + (s\tau)[p'(sp)]\delta(\rho u). \quad (11)$$

The main unknown functions $\rho > 0$, u and the parameter τ are defined on ω_h whereas j, w, \hat{w}, Π and μ are defined on ω_h^* .

In [18] two non-standard spatial discretizations of the QGD equations (2)-(4) were constructed which are weakly conservative in energy (see their generalization to a multidimensional case in [19]). One of them has the ‘‘enthalpy’’ form

$$\delta_t \rho + \delta^* j = 0, \quad \delta_t(\rho u) + \delta^*(jsu - \Pi) + s^*[(s\rho)\delta h(\rho)] = 0, \quad (12)$$

$$j = s\rho \cdot su - s\rho \cdot w, \quad s\rho \cdot w = [(\tau\partial_x)_h(\rho u)]su + (s\rho)\hat{w}, \quad (13)$$

$$\hat{w} = (s\tau)[(su)\delta u + \delta h(\rho)], \quad \Pi = \mu\delta u + (su)(s\rho)\hat{w} + p'(s\rho)(\tau\partial_x)_h(\rho u), \quad (14)$$

$$(\tau\partial_x)_h(\rho u) = \left(s\frac{\tau}{h'(\rho)}\right)\{[\delta h(\rho)]su + p'(s\rho)\delta u\}, \quad (15)$$

where $h(\rho) = \int_{r_0}^{\rho} \frac{p'(r)}{r} dr$, with some $r_0 > 0$, is the gas *enthalpy* and thus $h'(\rho) = \frac{p'(\rho)}{\rho}$. In the isentropic case $p(\rho) = p_1\rho^\gamma$ with $\gamma > 1$, one can take $r_0 = 0$ and then $h(\rho) = \frac{\gamma}{\gamma-1} \frac{p(\rho)}{\rho}$ and $h'(\rho) = \gamma \frac{p(\rho)}{\rho^2}$. Notice the non-standard $h(\rho)$ -dependent discretizations of $\partial_x p(\rho)$ in (12) and (14) and $\tau\partial_x(\rho u)$ in (13)-(14), see (15).

We linearize scheme (8)-(11) on a constant solution $\rho_* \equiv \text{const} > 0$ and $u_* = 0$. To do that, we write its solution in the form $\rho = \rho_* + \Delta\rho$ and $u = u_* + \Delta u$, neglect terms having the second order of smallness with respect to $\Delta\rho$ and Δu and obtain

$$\delta_t \Delta\rho + \rho_* \delta^* s \Delta u - \tau(\rho_*) p'(\rho_*) \delta^* \delta \Delta\rho = 0,$$

$$\rho_* \delta_t \Delta u + p'(\rho_*) \delta^* s \Delta\rho - [\mu(\rho_*) + \tau(\rho_*) \rho_* p'(\rho_*)] \delta^* \delta \Delta u = 0.$$

For the dimensionless unknowns $\tilde{\rho} = \frac{\Delta\rho}{\rho_*}$ and $\tilde{u} = \frac{\Delta u}{c_*}$ we get equations

$$\delta_t \tilde{\rho} + c_* \delta^* s \tilde{u} - \tau(\rho_*) c_*^2 \delta^* \delta \tilde{\rho} = 0, \quad (16)$$

$$\delta_t \tilde{u} + c_* \delta^* s \tilde{\rho} - \left[\frac{\mu(\rho_*)}{\rho_*} + \tau(\rho_*) c_*^2\right] \delta^* \delta \tilde{u} = 0 \quad (17)$$

(cf. systems (5) and (6)). The linearization of scheme (12)-(15) is the same.

Notice that since $u_* = 0$, the linearization result remains the same if it would be $w = \hat{w}$ and the terms dependent on u were omitted in the definition of these variables, i.e., for example, $w = \hat{w} = \frac{\tau}{s\rho} \delta p(\rho)$ instead of formulas in (9)-(10).

We assume that the regularization parameter and viscosity coefficient are given by usual QGD-formulas

$$\tau(\rho) = \frac{\alpha h}{\sqrt{p'(\rho)}}, \quad \mu(\rho) = \alpha_s \tau(\rho) \rho p'(\rho),$$

where $\alpha > 0$ and $\alpha_s \geq 0$ are parameters. Then omitting tildes above ρ and u , equations (16)-(17) can be rewritten in the following recurrent form

$$\rho^+ = \rho - \frac{\beta}{2}(u_+ - u_-) + \alpha\beta(\rho_- - 2\rho + \rho_-), \quad (18)$$

$$u^+ = u - \frac{\beta}{2}(\rho_+ - \rho_-) + \varkappa\alpha\beta(u_+ - 2u + u_-) \quad (19)$$

with three parameters α , $\beta := c_* \frac{\Delta t}{h}$ and $\varkappa := \alpha_s + 1 \geq 1$. The functions ρ^0 and u^0 are given, i.e., we consider the initial-value problem for the scheme. Below it is convenient to consider ρ and u as complex-valued mesh functions.

3 Weak conservativeness analysis

Let $\mathbf{y}^m = (\rho^m \ u^m)^T$, $m \geq 0$, be a column-vector function on ω_h and the linearized difference scheme (18)-(19) be rewritten in a matrix form

$$\mathbf{y}^+ = \begin{pmatrix} \alpha\beta & \frac{\beta}{2} \\ \frac{\beta}{2} & \varkappa\alpha\beta \end{pmatrix} \mathbf{y}_- + \begin{pmatrix} 1-2\alpha\beta & 0 \\ 0 & 1-2\varkappa\alpha\beta \end{pmatrix} \mathbf{y}^+ + \begin{pmatrix} \alpha\beta & -\frac{\beta}{2} \\ -\frac{\beta}{2} & \varkappa\alpha\beta \end{pmatrix} \mathbf{y}_+. \quad (20)$$

Let H be a Hilbert space of complex valued square-summable on ω_h vector functions, i.e. having a finite norm

$$\|\mathbf{y}\|_H = \left(h \sum_{k=-\infty}^{\infty} |\mathbf{y}_k|^2 \right)^{1/2}.$$

For $\mathbf{y}^0 = (\rho^0 \ u^0)^T \in H$ we have that $\mathbf{y}^m \in H$ for all $m \geq 1$. We define a *weak conservativeness* of scheme (20) as validity of the bound

$$\sup_{m \geq 0} \|\mathbf{y}^m\|_H \leq \|\mathbf{y}^0\|_H \quad \forall \mathbf{y}^0 \in H. \quad (21)$$

This definition is motivated by the energy conservation law (7) for the acoustics system of equation (6). It is essential to notice that for the linearized QGD-system (2)-(4) namely the corresponding inequality holds in place of equality (7) so that it is natural to study the bound for schemes based on such a system. Of course, estimate (21) guarantees the uniform in time stability in H with respect to initial data.

We first substitute a partial solution in the form $\mathbf{y}_k^m = e^{ik\xi} \mathbf{v}^m(\xi)$, $k \in \mathbb{Z}$, $m \geq 0$, where \mathbf{i} is the imaginary unit and $0 \leq \xi \leq 2\pi$ is a parameter, into (20) and obtain

$$\mathbf{v}^+(\xi) = G(\xi) \mathbf{v}(\xi), \quad G(\xi) = \begin{pmatrix} 1 - \omega_1 & -\mathbf{i}\omega_2 \\ -\mathbf{i}\omega_2 & 1 - \varkappa\omega_1 \end{pmatrix}, \quad (22)$$

where we denote $\omega_1 = 4\alpha\beta\theta$, $\theta = \sin^2 \frac{\xi}{2} \in [0, 1]$ and $\omega_2 = \beta \sin \xi$ for brevity. Below it is important that $\omega_2^2 = 4\beta^2\theta(1-\theta)$.

It is known (see similar formulas in [6]) that if $\mathbf{y}^0 = (\rho^0 \ u^0)^T \in H$, then there exists a function $\mathbf{v}^0 \in L^2(0, 2\pi)$ such that

$$\mathbf{v}^0(\xi) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \mathbf{y}_k^0 e^{-ik\xi},$$

and we can write the solution to scheme (20) in an integral form

$$\mathbf{y}_k^m = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \mathbf{v}^m(\xi) e^{ik\xi} d\xi, \quad k \in \mathbb{Z},$$

where $\mathbf{v}^m \in L^2(0, 2\pi)$ due to (22). The following Parseval identity also holds

$$\|\mathbf{y}^m\|_H = \sqrt{h} \|\mathbf{v}^m\|_{L^2(0, 2\pi)}, \quad m \geq 0. \quad (23)$$

The von Neumann type spectral condition

$$\max_{0 \leq \xi \leq 2\pi} \max_l |\lambda_l(G(\xi))| \leq 1 \quad (24)$$

is known to be a *necessary condition* for property (21) to hold (see similar result in [6]). Hereafter $\lambda_l(A)$ are eigenvalues of a matrix A .

Let us determine the spectral form of property (21).

Lemma 1. *Validity of the spectral bound*

$$\max_{0 \leq \xi \leq 2\pi} \max_l \lambda_l((G^*G)(\xi)) \leq 1 \quad (25)$$

is necessary and sufficient for the weak conservativeness property (21) to hold.

Proof. Due to the Parseval identity (23) and formula (22) we have

$$h^{-1} \|\hat{\mathbf{y}}\|_H^2 = \|\hat{\mathbf{v}}\|_{L^2(0, 2\pi)}^2 = \|\mathbf{G}\mathbf{v}\|_{L^2(0, 2\pi)}^2 = (\mathbf{G}^*\mathbf{G}\mathbf{v}, \mathbf{v})_{L^2(0, 2\pi)}.$$

Since $(\mathbf{G}^*\mathbf{G})(\xi) \geq 0$ is a Hermitian matrix, it has a spectral decomposition $(\mathbf{G}^*\mathbf{G})(\xi) = U^*(\xi)\Lambda(\xi)U(\xi)$, where $U(\xi)$ is a unitary matrix, and $\Lambda(\xi)$ is a diagonal matrix with numbers $\lambda_l((\mathbf{G}^*\mathbf{G})(\xi)) \geq 0$ forming its diagonal. Hence for $z(\xi) := U(\xi)\mathbf{v}(\xi)$ we have

$$(\mathbf{G}^*\mathbf{G}\mathbf{v}, \mathbf{v})_{L^2(0, 2\pi)} = (\Lambda\mathbf{z}, \mathbf{z})_{L^2(0, 2\pi)} = \|\Lambda^{1/2}\mathbf{z}\|_{L^2(0, 2\pi)}^2.$$

Thus $\|\mathbf{y}^m\|_H^2 = h\|\Lambda^{m/2}\mathbf{z}^0\|_{L^2(0, 2\pi)}^2$ for $m \geq 0$, and bound (21) is equivalent to the following one

$$\sup_{m \geq 0} \|\Lambda^{m/2}\mathbf{z}^0\|_{L^2(0, 2\pi)}^2 \leq \|\mathbf{z}^0\|_{L^2(0, 2\pi)}^2 \quad \forall \mathbf{z}^0 \in L^2(0, 2\pi).$$

It holds if and only if the spectral bound (25) holds.

Remark 1. Under validity of the spectral bound (25), the norm $\|\mathbf{y}^m\|_H$ is actually non-increasing in $m \geq 0$ that serves as a stronger property than (21).

In the proof of this lemma, the specific form and dimension of the matrix G are clearly inessential, and actually it holds in general case.

In our case the matrix G^*G has the form

$$G^*G = \begin{pmatrix} (1 - \omega_1)^2 + \omega_2^2 & -\mathbf{i}(1 - \varkappa)\omega_1\omega_2 \\ \mathbf{i}(1 - \varkappa)\omega_1\omega_2 & (1 - \varkappa\omega_1)^2 + \omega_2^2 \end{pmatrix}.$$

Note that $G^*G = [(1 - \omega_1)^2 + \omega_2^2]I$ in the simplest case $\varkappa = 1$, where I is a unit matrix.

Theorem 1. *The necessary spectral condition (24) holds if and only if*

$$\beta \leq \min \left\{ (\varkappa + 1)\alpha, \frac{1}{2\varkappa\alpha} \right\}. \quad (26)$$

Proof. The characteristic polynomial for the matrix G has the following form

$$q_1(\lambda) = \lambda^2 - (\operatorname{tr} G)\lambda + \det G = \lambda^2 + [(\varkappa + 1)\omega_1 - 2]\lambda + [\varkappa\omega_1^2 + \omega_2^2 + 1 - (\varkappa + 1)\omega_1]. \quad (27)$$

We set $a_0 := q_1(1) = 1 - (\operatorname{tr} G) + \det G = \varkappa\omega_1^2 + \omega_2^2 \geq 0$ and notice that $a_0 = q_1(1) = 0$ if and only if $G = I$. We transform the unit circle $\{|\lambda| \leq 1\}$ with a punctured point $(1, 0)$ on \mathbb{C} into the closed left half-plane $\{\operatorname{Re} z \leq 0\}$ and put

$$\hat{q}_1(z) := (z - 1)^2 q_1\left(\frac{z+1}{z-1}\right) = a_0 z^2 + 2a_1 z + a_2,$$

where $a_1 = 1 - \det G$, $a_2 = 1 + \operatorname{tr} G + \det G$. It is well known that for $a_0 > 0$ the roots $\hat{q}_1(z)$ lie in $\{\operatorname{Re} z \leq 0\}$ under the conditions $a_1 \geq 0$ and $a_2 \geq 0$, i.e.

$$\varkappa\omega_1^2 + \omega_2^2 - (\varkappa + 1)\omega_1 \leq 0, \quad \varkappa\omega_1^2 + \omega_2^2 - 2(\varkappa + 1)\omega_1 + 4 \geq 0.$$

We rewrite these conditions as

$$\beta(4\varkappa\alpha^2\theta + 1 - \theta) - (\varkappa + 1)\alpha \leq 0 \quad \text{for } 0 \leq \theta \leq 1, \quad (28)$$

$$r(\theta) := \beta^2(4\varkappa\alpha^2 - 1)\theta^2 - \beta(2(\varkappa + 1)\alpha - \beta)\theta + 1 \geq 0 \quad \text{for } 0 \leq \theta \leq 1. \quad (29)$$

The left-hand side of (28) is linear in θ , thus it suffices to test it for $\theta = 0, 1$ that leads us to the condition

$$\beta \leq \min \left\{ (\varkappa + 1)\alpha, \frac{\varkappa + 1}{4\varkappa\alpha} \right\}. \quad (30)$$

Next we analyze condition (29). Notice that $r(0) = 1$ and due to (30) we have $2(\varkappa + 1)\alpha - \beta > 0$. For $a = 4\varkappa\alpha^2 - 1 \neq 0$ the vertex of the parabola $r(\theta)$ is given by

$$\theta_v = \frac{2(\varkappa + 1)\alpha - \beta}{2\beta(4\varkappa\alpha^2 - 1)}.$$

For $4\varkappa\alpha^2 - 1 > 0$ the property $\theta_v > 1$ means that $\beta < \frac{(\varkappa + 1)\alpha}{4\varkappa\alpha^2 - 0.5}$, and it holds due to (30). Hence condition (29) reduces to $r(1) \geq 0$, i.e.

$$4\kappa\alpha^2\beta^2 - 2(\kappa+1)\alpha\beta + 1 = 4\kappa\alpha^2\left(\beta - \frac{1}{2\kappa\alpha}\right)\left(\beta - \frac{1}{2\alpha}\right) \geq 0. \quad (31)$$

For $4\kappa\alpha^2 - 1 < 0$ we have $\theta_v < 0$, so that condition (29) reduces again to $r(1) \geq 0$. For $4\kappa\alpha^2 - 1 = 0$ the condition also reduces to $r(1) \geq 0$ (since $r(0) = 1$).

Since $\kappa \geq 1$, inequality (31) means that either of the conditions

$$\beta \leq \frac{1}{2\kappa\alpha}, \quad \beta \geq \frac{1}{2\alpha} \quad (32)$$

holds. Combining them with (30), we obtain (26).

Now we turn to the spectral criterion (25).

Theorem 2. *The spectral criterion (25) holds if and only if*

$$\beta \leq \min\left\{2\alpha, \frac{1}{2\kappa\alpha}\right\}. \quad (33)$$

Proof. The characteristic polynomial of G^*G has the following form

$$q_2(\lambda) = \lambda^2 - \text{tr}(G^*G)\lambda + (\det G)^2.$$

Since $\lambda_l(G^*G) \geq 0$, the property $|\lambda_l(G^*G)| \leq 1$ means validity of the conditions

$$\frac{1}{2}\text{tr}(G^*G) \leq 1, \quad q_2(1) = 1 - \text{tr}(G^*G) + (\det G)^2 \geq 0. \quad (34)$$

The first of them has the form

$$\frac{\kappa^2 + 1}{2}\omega_1^2 + \omega_2^2 - (\kappa + 1)\omega_1 \leq 0$$

and can be specified as

$$8\alpha^2\beta^2(\kappa^2 + 1)\theta^2 + 4\beta^2\theta(1 - \theta) - 4\alpha\beta(\kappa + 1)\theta \leq 0 \quad \text{for } 0 \leq \theta \leq 1. \quad (35)$$

After dividing by $4\beta\theta$ we get that it suffices to confine ourselves with the values $\theta = 0, 1$ that leads to the condition

$$\beta \leq \min\left\{(\kappa + 1)\alpha, \frac{\kappa + 1}{2(\kappa^2 + 1)\alpha}\right\}. \quad (36)$$

In order to transform the second condition (34) we notice that

$$\text{tr}(G^*G) = 2(b + 1) + (\kappa - 1)^2\omega_1^2, \quad (\det G)^2 = (b + 1)^2,$$

where $b := \kappa\omega_1^2 + \omega_2^2 - (\kappa + 1)\omega_1$, see (27). Hence the following factorization holds

$$q_2(1) = b^2 - (\kappa - 1)^2\omega_1^2 = (b - (\kappa - 1)\omega_1)(b + (\kappa - 1)\omega_1),$$

that is decisive for the simplicity of our analysis. Since $\varkappa \geq 1$, the condition $q_2(1) \geq 0$ is equivalent to validity of either of the conditions

$$\varkappa\omega_1^2 + \omega_2^2 - 2\omega_1 \leq 0, \quad \varkappa\omega_1^2 + \omega_2^2 - 2\varkappa\omega_1 \geq 0,$$

i.e., more specifically, to validity of either of the conditions

$$\begin{aligned} \beta(4\varkappa\alpha^2\theta + 1 - \theta) - 2\alpha &\leq 0 \quad \text{for } 0 \leq \theta \leq 1, \\ \beta(4\varkappa\alpha^2\theta + 1 - \theta) - 2\varkappa\alpha &\geq 0 \quad \text{for } 0 \leq \theta \leq 1. \end{aligned}$$

As above they respectively mean that the inequalities

$$\beta \leq \min \left\{ 2\alpha, \frac{1}{2\varkappa\alpha} \right\}, \quad \beta \geq \max \left\{ 2\varkappa\alpha, \frac{1}{2\alpha} \right\} \quad (37)$$

hold. Combining them with (36) we complete the proof.

It is essential that the function on the right-hand side of condition (33) reaches its maximal value at $\alpha = \alpha_* := \frac{1}{2\sqrt{\varkappa}} \leq \frac{1}{2}$ and the maximal value equals $\frac{1}{\sqrt{\varkappa}} \leq 1$. Hence the criterion coincides with the standard CFL stability condition $\beta \leq 1$ if and only if $\alpha = \alpha_*$ and $\varkappa = 1$. The criterion gives an important information on the optimal choice of α since in practice for the original non-linear problem α is normally sought experimentally. Note also that criterion (33) and the necessary condition (26) coincide only in the case $\alpha \geq \alpha_*$.

We call attention to a paradoxical moment: criterion (33) becomes *stronger* as the coefficient of “effective viscosity” \varkappa increases (it is harder to say that about the necessary condition (26)). Therefore the best choice in the present bounds is $\alpha_s = 0$, i.e. $\varkappa = 1$. But this conclusion is not universal in practice and it is known that in some situations $\alpha_s > 0$ has to be taken (see, for example, [21]).

In Fig. 1 we compare the necessary condition, the criterion of stability and the sufficient condition as well as the results of numerical experiments for the original system (1) for $p(\rho) = \rho^2$ (the scaled case of the shallow water equations) and $\varkappa = \frac{7}{3}$. The sufficient condition was obtained in [14] only for these $p(\rho)$ and \varkappa by the energy method and has the form

$$\beta \leq \min \left\{ \frac{2\alpha}{1 + 6\alpha + 4\alpha^2}, \frac{4\alpha}{1 + 6\alpha + 16\alpha^2} \right\}.$$

Notice that the first fraction in it is less than the second one for $0 < \alpha < \frac{3+\sqrt{17}}{8} \approx 0.890$. The corresponding graph is almost flat for $0.3 \leq \alpha \leq 0.9$ in contrast to the cases of necessary condition and criterion. The computations are accomplished for $0 \leq t \leq 0.5$ for the Riemann problem with the discontinuous initial data

$$\rho_0(x) = \begin{cases} 1, & x < 0 \\ 0.1, & x > 0 \end{cases}, \quad u_0(x) = \begin{cases} 0.1, & x < 0, \\ 0, & x > 0 \end{cases}$$

for both schemes (8)-(11) and (12)-(15) with $h = 1/125$.

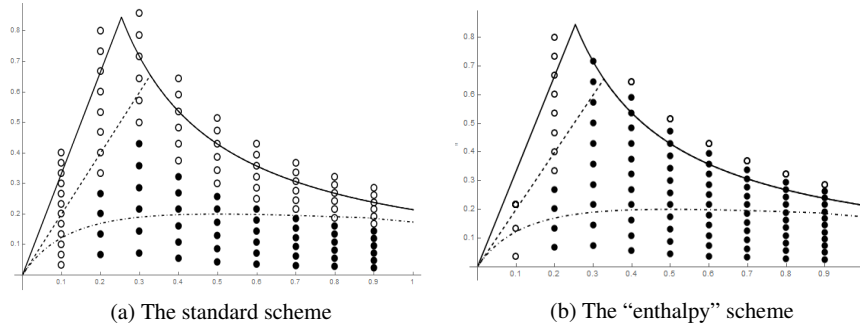


Fig. 1 The weak conservativeness analysis: the necessary condition (solid line), the criterion (dash line), the sufficient condition (dotdash line) together with conservative (painted balls) and non-conservative (unpainted balls) computations for the Riemann problem in dependence with α

We observe a good correspondence of the obtained criterion with the experimental results, and that the sufficient condition underestimates the criterion up to several times in the most interesting region $\alpha \approx \alpha_*$. Also the results for the “enthalpy” scheme (12)-(15) are clearly different from and better than for the standard one (8)-(11) though the above linearized analysis gives the same results for them.

We have identified non-conservative computations by noticeable well-known oscillations of the numerical solutions (some of computations have not even been completed due to overflow). In Fig. 2 we give an example of conservative and non-conservative solutions ρ and u for the “enthalpy” scheme (at time $t = 0.5$) for $\alpha = 0.4$ and two neighboring values of β from Fig. 1 (b).

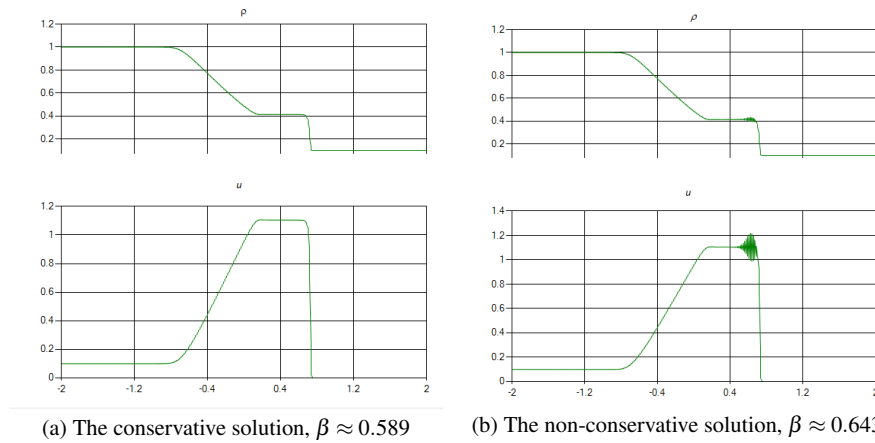


Fig. 2 The examples of conservative and non-conservative solutions for the “enthalpy” scheme for $\alpha = 0.4$ (at time $t = 0.5$)

4 The case of the schemes based on a simplified regularization

We also consider a simplified (quasi-hydrodynamic [8, 13, 15]) regularization (2)-(4), where the terms with $\partial_x(\rho u)$ are omitted, in particular, it becomes $w = \hat{w}$. Correspondingly in schemes (8)-(11) and (12)-(15) we have to omit both terms with respectively $\delta(\rho u)$ and $(\tau \partial_x)_h(\rho u)$. In the linearized scheme the term $\tau(\rho_*)c_*^2$ disappears from equation (17), hence now $\varkappa = \alpha_s$. Notice that usually $0 < \alpha_s \leq 1$ though in specific cases $\alpha_s > 1$ can be also taken.

Theorem 3. *For the simplified scheme based on the quasi-hydrodynamic regularization the following results are valid:*

(1) *in the case $0 \leq \alpha_s \leq 1$ the necessary condition (24) and criterion (25) hold if and only if respectively*

$$\beta \leq \min \left\{ (\alpha_s + 1)\alpha, \frac{1}{2\alpha} \right\}, \quad (38)$$

$$\beta \leq \min \left\{ 2\alpha_s\alpha, \frac{1}{2\alpha} \right\}; \quad (39)$$

(2) *in the case $\alpha_s \geq 1$ the results of Theorems 1 and 2 remain valid with $\varkappa = \alpha_s$.*

Proof. The above given analysis holds true except for some changes in the case $0 \leq \varkappa = \alpha_s \leq 1$. In this case, inequalities (32) are replaced by the following ones

$$\beta \leq \frac{1}{2\alpha}, \quad \beta \geq \frac{1}{2\varkappa\alpha};$$

they being combined with (30) lead to (38). Also inequalities (37) are replaced by the following ones

$$\beta \leq \min \left\{ 2\varkappa\alpha, \frac{1}{2\alpha} \right\}, \quad \beta \geq \max \left\{ 2\alpha, \frac{1}{2\varkappa\alpha} \right\},$$

they being combined with (36) lead to (39).

The maximal value of the function on the right-hand side of criterion (39) is reached at $\alpha = \alpha_* := \frac{1}{2\sqrt{\alpha_s}} \geq \frac{1}{2}$ and equals $\sqrt{\alpha_s} \leq 1$. We notice that the necessary condition (38) is especially rough compared to criterion (39) for $\alpha_s \approx 0$, including the case $\alpha_s = 0$ when actually the stability is absent at all.

Acknowledgements The study was partially supported by the RFBR, project nos. 16-01-00048 and 18-01-00587.

References

1. Bakhvalov, N.S., Zhidkov, N.P., Kobelkov, G.M.: Numerical methods. Binom, Moscow, (2011) [in Russian]

2. Balashov, V., Zlotnik, A., Savenkov, E.: Analysis of a regularized model for the isothermal two-component mixture with the diffuse interface. *Russ. J. Numer. Anal. Math. Model.* **32** (6) (2017), 347–358. See also: *Keldysh Inst. Appl. Math.* Preprint **89** 1–26 (2016). <http://library.keldysh.ru/preprint.asp?id=2016-89> [in Russian]
3. Chetverushkin, B.N.: Kinetic schemes and quasi-gas dynamic system of equations, CIMNE, Barcelona (2008)
4. Coulombel, J.-F.: Stability of finite difference schemes for hyperbolic initial boundary value problems. In: HCDTE Lecture Notes. Part I. Dispersive and Transport Equations, G. Alberti, F. Ancona, S. Bianchini et al., eds, Amer. Inst. Math. Sci., Springfield, 97–226 (2013)
5. Ganzha, V.G., Vorozhtsov, E.V.: Computer-aided analysis of difference schemes for partial differential equations, Wiley, New York (1996)
6. Godunov, S.K., Ryabenkii, V.S.: Difference schemes. Studies in Mathematics and its Applications **19**, North Holland, Amsterdam (1987)
7. Gustafsson, B., Kreiss, H.-O., Olinger, J.: Time dependent problems and difference methods. Wiley, New York (1995)
8. Elizarova, T.G.: Quasi-gas dynamic equations. Springer, Dordrecht (2009)
9. Elizarova, T.G., Bulatov, O.V.: Regularized shallow water equations and a new method of simulation of the open channel flows. *Comput. Fluids* **46** 206–211 (2011)
10. Elizarova, T.G., Zlotnik, A.A., Istomina, M.A.: Hydrodynamic aspects of spiral-vortex structure formation in rotating gas discs. *Astron. Reports* **62** (1) (2018), 9–18. See also: *Keldysh Inst. Appl. Math.* Preprint **1** 1–30 (2017). <http://library.keldysh.ru/preprint.asp?id=2017-1> [in Russian]
11. LeVeque, R.J.: Finite volume methods for hyperbolic problems. Cambridge University Press, Cambridge (2004)
12. Richtmyer, R.D., Morton, K.W.: Difference methods for initial-value problems. 2nd ed., Wiley-Interscience (1967)
13. Sheretov, Yu.V.: Continuum dynamics under spatiotemporal averaging. RKhD, Moscow-Izhevsk (2009) [in Russian]
14. Suhomozgii, A.A., Sheretov, Yu.V.: Stability analysis of a finite-difference scheme for solving the Saint-Venant equations in the shallow water theory. In: *Appl. of functional anal. in approx. theory*, Tver State Univ. (2013) 48–60 [in Russian]
15. Zlotnik, A.A.: Parabolicity of a quasihydrodynamic system of equations and the stability of its small perturbations. *Math. Notes* **83** (5) 610–623 (2008)
16. Zlotnik, A.A.: Energy equalities and estimates for barotropic quasi-gasdynamic and quasi-hydrodynamic systems of equations. *Comput. Math. Math. Phys.* **50** (2) 310–321 (2010)
17. Zlotnik, A.A.: On construction of quasi-gasdynamic systems of equations and the barotropic system with the potential body force. *Math. Modeling* **24** (4) 65–79 (2012) [in Russian]
18. Zlotnik, A.A.: Spatial discretization of the one-dimensional barotropic quasi-gasdynamic system of equations and the energy balance equation, *Math. Modeling* **24** (10) 51–64 (2012) [in Russian].
19. Zlotnik, A.A.: On conservative spatial discretizations of the barotropic quasi-gasdynamic system of equations with a potential body force. *Comput. Math. Math. Phys.* **56** (2) 303–319 (2016)
20. Zlotnik, A.A., Chetverushkin, B.N.: Parabolicity of the quasi-gasdynamic system of equations, its hyperbolic second-order modification, and the stability of small perturbations for them. *Comput. Math. Math. Phys.* **48** (3) 420–446 (3) (2008)
21. Zlotnik, A., Gavrilin, V.: On a conservative finite-difference method for 1D shallow water flows based on regularized equations. In: *Math. Problems in Meteorological Model. Math. in Industry*, A. Batkai, P. Csomos, A. Horanyi, etc., eds. **24** 16–31 (2016)