A MacCormack Method for Complete Shallow Water Equations with Source Terms

Eric Ngondiep\textsuperscript{a,b}, Alqahtani Rubayyi\textsuperscript{b}, Jean C. Ntonga\textsuperscript{a}

\textsuperscript{a}Department of Mathematics and Statistics, College of Science, Al-Imam Muhammad Ibn Saud Islamic University (IMSIU), 90950 Riyadh 11632, Kingdom of Saudi Arabia.

\textsuperscript{b}Hydrological Research Centre, Institute for Geological and Mining Research, 4110 Yaounde-Cameroon.

Email addresses: ericngondiep@gmail.com, engondiep@imamu.edu.sa, ratalqahtani@imamu.edu.sa, ntonga jc@yahoo.fr

Abstract. In the last decades, more or less complex physically-based hydrological models, have been developed to solve the shallow water equations or their approximations using various numerical methods. The MacCormack method was developed for simulating overland flow with spatially variable infiltration and microtopography using the hydrodynamic flow equations. The basic MacCormack scheme is enhanced when it uses the method of fractional steps to treat the friction slope or a stiff source term and to upwind the convection term in order to control the numerical oscillations and stability. In this paper we describe, the MacCormack scheme for 1D complete shallow water equations with source terms, analyze the stability condition of the method and we provide the convergence rate of the algorithm. This work improves some well known results deeply studied in the literature which concern the Saint-Venant problem and it represents an extension of the time dependent shallow water equations without source terms. The numerical evidences consider the rate of convergence of the method and compares the numerical solution respect to the analytical one.

Keywords: 1D shallow water equations, source terms, MacCormack scheme, Fourier stability analysis, stability condition, convergence rate.

AMS Subject Classification (MSC). 65M10, 65M05.

1 Introduction

Most open-channel flows of interest in the physical, hydrological, biological, engineering and social sciences are unsteady and can be considered to be one-dimensional (1D). In this paper we are interested in the numerical solutions of one-dimensional complete shallow water equations with source terms introduced in \cite{8} and still widely used in modelling flows in rivers, lakes and coastal areas as well as atmospheric and oceanic flows in certain regimes. In the case of a prismatic channel, the complete shallow water equations with source terms reads as follows

\[
\begin{cases}
\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = r, & \text{for } t \in (0, T_1) \text{ and } x \in \Omega = (0, L), \\
\frac{\partial Q}{\partial t} + g A \frac{\partial A}{\partial x} + \frac{\partial}{\partial x} \left( Q^2 A \right) = gA(S_0 - S_f), & \text{for } t \in (0, T_1) \text{ and } x \in \Omega = (0, L),
\end{cases}
\]

(1)

where the bottom (or bed) slope ($S_0$) and the friction slope ($S_f$) (see \cite{8}) are defined as

\[
S_0 = \frac{\tau_P}{\rho g A} \quad \text{and} \quad S_f = \frac{Q|Q|}{K^2},
\]

(2)
\( r = r(x, t) \) is the lateral inflow per unit length along the channel, \( T_1 \) is the time interval length, \( L \) is the rod
interval length, \( A = A(t, x) \) is the cross section, \( Q = Q(t, x) \) is the discharge, \( g \) is the acceleration of gravity.
\( T = T(x, t) \) is the top width assumed to be constant, \( \tau \) is the average shear stress on the water from the
channel boundary, \( \rho \) is the fluid density, \( P = P(x, y(t, x)) \) is the wetted perimeter, i.e., the length of the
boundary of the cross section that is under water for a given height of water (\( y \)). As in [8], the conveyance
for a compact channel is defined by

\[
K := K(x, y) = \frac{1.49}{n_1}A(x, y)R(x, y)^{2/3},
\]

where \( R = A/P \) is the hydraulic radius and \( n_1 \) is the manning’s roughness coefficient.

By straightforward computations, it is not hard to see that system (1) becomes

\[
\frac{\partial W}{\partial t} + \frac{\partial F}{\partial x} = S,
\]

where \( W = \text{col}(A, Q) \), \( F = \text{col}(Q, g\frac{A^2}{T} + \frac{Q^2}{A}) \), and \( S = \text{col}(r, gA(S_0 - S_f)) \). Equation (4) emphasizes the
conservative character of system (1).

The one-dimensional shallow water equations with source terms (1) are highly nonlinear and therefore
do not have global analytical solutions [8]. When solving the system of balance laws (4) numerically, one
typically faces several difficulties. One difficulty stems from the fact that many physically relevant solutions
of (4) are small perturbations of steady-state solutions. So, using a wrong balance between the flux and
geometric source term in equation (4), the solution may develop spurious waves of a magnitude that can
become larger than the exact solution. Another drawback is when the cross section is very small. In that
case, even small numerical oscillation in the computed solution can result in a very large discharge, which is
not only physically irrelevant, but cause the numerical scheme to break down. To overcome these numerical
challenges, one needs to use a numerical scheme that is both well-balanced and positivity preserving.

A number of well-balanced and positivity preserving numerical methods frequently used in the models
based on the shallow water equations have been proposed in literature [3, 20], or on the boussinesq equations,
which are reduced to shallow water equations, in order to simulate breaking waves [10, 32, 15]. Although the
MacCormack scheme is less accurate than the more recent methods, it is commonly used for engineering
problems due to its greater simplicity. So, we have to approximate the exact solution of problem (4) by
a numerical method based on the MacCormack scheme. This algorithm is a class of higher order finite
difference methods (second order convergent), which provides an effective way of joining spectral method
for accuracy and robust characteristics of finite difference schemes. For example, to compute unsteady flow
specifically in the presence of discontinuity, inherent dissipation and stability, one such widely used method
is MacCormack method [21]. This technique has been used successfully to provide time-accurate solution for
fluid flow and aeroacoustics problems. The applications of this technique to 1D shock tube and 2D acoustic
scattering problems provide good result while comparing with the analytical solution. MacCormack introduced
a simpler variation of Lax-Wendroff scheme which is basically a two-step scheme with second order
Taylor series expansion in time and fourth order in spatial accuracy [21, 15]. This algorithm is computationally
efficient and easy to implement which can be appropriate to obtain reliable results. By using this scheme
with two nodes, the flow field can be simulated for unsteady flows especially for shallow water problem in
the presence of discontinuity and strict gradient conditions. Furthermore, to capture fluid flow in transition
over long periods of time and distance, numerical spatial derivative are required to be determined in few
grid points while error controlled can be accurately computed. The authors [24, 25, 26, 23, 12] extended
MacCormack scheme [21] to an implicit-explicit scheme by coupling the original MacCormack approach and
the Crank-Nicolson method, and to implicit compact differencing scheme by splitting the derivative operator
of a central compact scheme into one-sided forward and backward operators. The one-sided nature of the
MacCormack technique is an essential advantage especially when severe gradients are present.

In [15, 19, 14] the authors compared the Lax-Wendroff scheme to many numerical methods of high
order of accuracy, such as, the linear Central Weighed Essential Non-Oscillatory (CWENO) scheme which is
superior to full nonlinear CWENO method, to high-resolution TVD conservative schemes along with high order Central Schemes for hyperbolic systems of conservative laws [4, 5] and to Central-upwind schemes for the shallow water system [17]. In a search for stable and more accurate shock capturing numerical approach, they observed that the Lax-Wendroff approach is one of the most frequently encountered in the literature related to classical Shock-capturing schemes. However, difficulties have been reported when trying to include source terms in the discretization and to keep the second order accuracy at the same time [29]. The MacCormack approach which is a predictor-corrector version of the Lax-Wendroff algorithm provides a reasonably good result at discontinuities. This method is much easy to apply than the Lax-Wendroff scheme because the Jacobian does not appear. The amplification factor and stability constraint almost are the same as presented for the Lax-Wendroff method (see [1], P. 202-206, case of inviscid burgers equation). It is also important to note that the solutions obtained for the same problem at the same courant number are different from those obtained using the Lax-Wendroff scheme. This is due both to the switched differencing in the predictor and the corrector and the nonlinear nature of the governing PDE. It should be noted that reversing the differencing in the predictor and corrector steps leads to quite different results. Another motivation of this work comes from the fact that, the explicit MacCormack time discretization for the complete nonlinear Burgers equation (which can be served as a model equation for many nonlinear PDEs: Navier-Stokes equations, Stokes-Darcy equations, Parabolized Navier-Stokes equations,...) gives a suitable stability restriction which can be used with an appropriate safety factor (see [1], P. 227-228). For the MacCormack solver, as with other explicit schemes, it requires a time step limitation. In general, the maximum time step (with respect to stability) allowable in the MacCormack scheme applied to linear hyperbolic equations is limited by the CFL condition, as are all explicit finite difference methods. The overland flow equations are nonlinear, however, and a rigorous stability analysis for these equations is exceedingly difficult. The source terms place additional and problem-dependent restrictions on the maximum admissible time step for stability. Therefore, the CFL condition can only be considered as a general guideline here, and the maximum allowable time step for any particular problem will be less than predicted by the CFL condition and determined by numerical experimentation (see [9], page 223).

In a recent work [22], we addressed the problem of mathematical model of complete shallow water equations with source terms in which we provided the stability analysis of the Lax-Wendroff scheme. In this paper we are still interested by the stability analysis, but in the sense of MacCormack. In particular, we consider the case where the channel is prismatic and the interesting result is that the algorithm is at least of second order accuracy in time and space, while the stability limitation does not coincide with the CFL condition widely studied in the literature for hyperbolic partial differential equations (for example: linear advection equation, wave equation, inviscid burgers equations, etc,...). However, while the stability requirement is highly unusual, the result has a potential positive implication since the stability constraint obtained in this work controls the CFL condition. Indeed the nice feature is that, as required in a stability context, we normally find a linear stability condition which can be considered as a necessary condition of stability from a Fourier stability analysis. On the other hand, it comes from this analysis that an instability occurs when $|\Delta t|$ is greater than some $|\Delta t|_{max}$ which can be viewed as $(\Delta t)_{CFL}$. More specifically, the attention is focused in the following three items:

(i1) full description of the MacCormack method for 1D complete shallow water equations with source terms;

(i2) stability restriction of the algorithm: this item together with item (i1) represent our original contributions and they improve the works studied in [22, 5, 28];

(i3) Numerical experiments concerning the convergence of the method, the simulation of the numerical solution obtained by the MacCormack approach along with the analytical solution, and regarding the effectiveness of this method according to the theoretical analysis given in the first two items.

The paper is organized as follows. Section 2 deals with the full description of MacCormack method for 1D complete shallow water equations with source terms. The stability analysis of this scheme is deeply studied in Section 3. In Section 4 some numerical experiments which consider the convergence of the scheme and some simulations are presented and discussed. We draw the general conclusion and present the future direction of works in section 5.
2 Full description of MacCormack method

In this section, we give a detailed description of the MacCormack algorithm for system (1). First, we recall that the MacCormack scheme is a two step explicit method which consists in predictor-corrector steps. The scheme uses the forward difference in predictor step while the corrector step considers the backward difference. Since our aim is to analyze both stability and rate of convergence of the method, without loss of generality we should use a constant time step $\Delta t$ and mesh size $\Delta x$. Let $N$ and $M$ be two positive integers. Noticing $x_j = j\Delta x$, $t^n = n\Delta t$ and let the superscript (resp., the subscript) denoting the time level (resp., space level) of the approximation. We denote by $W^n_j = (A^n_j, Q^n_j)^T$ the approximate solution of equations (1), obtained at time $t^n$ and at point $x_j$, using the MacCormack algorithm and $W(t^n, x_j) = (A(t^n, x_j), Q(t^n, x_j))^T$ the value of the analytical solution of system (1) at discrete time $t^n$ and at discrete point $x_j$. Furthermore, the domain $\Omega = (0, L)$ is subdivided into $M+1$ grid points $\{x_j : j = 0, 1, ..., M\}$, while the time interval $(0, T)$ is subdivided into $N+1$ grid points $\{t^n : n = 0, 1, ..., N\}$. Using this, the full description of MacCormack method for the system (1) reads: Given $W^n_j$, find an approximate solution (the solution can be considered as weak) $w^{n+1}_j$, for $0 \leq n \leq N-1$ and $0 \leq j \leq M$, satisfying

Predictor step: solve equation (5) for predicted value

$$W_j^{n+1} = W^n_j - \frac{\Delta t}{\Delta x} (F_j^{n+1} - F_j^n) + \Delta t S^n_j,$$

(5)

Corrector step: use the predicted value obtained in equation (5) to compute the exact one

$$W_j^{n+1} = \frac{1}{2} \left[ W^n_j + W_j^{n+1} - \frac{\Delta t}{\Delta x} (F_j^{n+1} - F_j^n) + \Delta t S_j^{n+1} \right].$$

(6)

**Proposition 2.1.** Let $n$ and $j$ be two nonnegative integers. Setting $t^n = n\Delta t$ and $x_j = j\Delta x$, where $\Delta t$ and $\Delta x$ are time step and mesh size, respectively. The MacCormack scheme for system (1) is given by

Predictor step: solve equation (7) for predicted values $A_j^{n+1}$ and $Q_j^{n+1}$

$$A_j^{n+1} = A^n_j - \frac{\Delta t}{\Delta x} (Q_{j+1}^n - Q_j^n) + \Delta t r_j^n,$$

(7)

and

$$Q_j^{n+1} = Q^n_j - \frac{\Delta t}{\Delta x} \left\{ \frac{g}{2T} \left[ \left( A_{j+1}^n \right)^2 - \left( A_j^n \right)^2 \right] + \left( Q_{j+1}^n \right)^2 - \left( Q_j^n \right)^2 \right\} + gP\Delta t \left( \frac{\tau}{\rho g} - \frac{n_1^2}{1.49^2} \frac{P}{(A_j^n)^2} \right).$$

(8)

Corrector step: use $A_j^{n+1}$ and $Q_j^{n+1}$ obtained in equations (7) - (8) to compute $A_j^{n+1}$ and $Q_j^{n+1}$

$$A_j^{n+1} = \frac{1}{2} \left\{ A_j^n + A_j^{n+1} - \frac{\Delta t}{\Delta x} (Q_{j+1}^{n+1} - Q_{j-1}^{n+1}) + \Delta t r_j^{n+1} \right\},$$

(9)

and

$$Q_j^{n+1} = \frac{1}{2} \left\{ Q_j^n + Q_j^{n+1} - \frac{\Delta t}{\Delta x} \left[ \frac{g}{2T} \left[ \left( A_{j+1}^{n+1} \right)^2 - \left( A_{j-1}^{n+1} \right)^2 \right] + \left( Q_{j+1}^{n+1} \right)^2 - \left( Q_{j-1}^{n+1} \right)^2 \right] \right\} + gP\Delta t \left( \frac{\tau}{\rho g} - \frac{n_1^2}{1.49^2} \frac{P}{(A_j^{n+1})^2} \right).$$

(10)

**Proof.** Since $W = col(A, Q)$, using relations (5) - (8) we get

Predictor step: solve equation (11) - (12) for predicted values $A_j^{n+1}$ and $Q_j^{n+1}$

$$A_j^{n+1} = A^n_j - \frac{\Delta t}{\Delta x} (Q_{j+1}^n - Q_j^n) + \Delta t r_j^n,$$

(11)
\( Q_{j}^{n+1} = Q_{j}^{n} - \frac{\Delta t}{\Delta x} \left\{ \frac{g}{2T} ((A_{j+1}^{n})^{2} - (A_{j}^{n})^{2}) + \frac{(Q_{j+1}^{n})^{2}}{A_{j+1}^{n+1}} - \frac{(Q_{j}^{n})^{2}}{A_{j}^{n+1}} \right\} + g\Delta t A_{j}^{n} ((S_{0})_{j}^{n} - (S_{j})_{j}^{n}) , \) (12)

Corrector step: use \( A_{j}^{n+1} \) and \( Q_{j}^{n+1} \) obtained in equations (13) and (14) into equations (12) and (14) results in

\[
A_{j}^{n+1} = \frac{1}{2} \left\{ A_{j}^{n} + A_{j}^{n+1} - \frac{\Delta t}{\Delta x} (Q_{j}^{n+1} - Q_{j-1}^{n+1}) + \Delta t_{j+1} \right\} ,
\]

(13)

\[
Q_{j}^{n+1} = \frac{1}{2} \left\{ Q_{j}^{n} + Q_{j}^{n+1} - \frac{\Delta t}{\Delta x} \left\{ \frac{g}{2T} \left[ (A_{j+1}^{n+1})^{2} - (A_{j}^{n+1})^{2} \right] + \frac{(Q_{j+1}^{n+1})^{2}}{A_{j+1}^{n+1}} - \frac{(Q_{j}^{n+1})^{2}}{A_{j}^{n+1}} \right\} + g\Delta t A_{j}^{n+1} (S_{j}^{n+1} - (S_{j})_{j}^{n+1}) \right\} .
\]

(14)

Substituting equation (15) into relation (2) yields

\[
S_{j} = \frac{n_{i}^{2}}{1.492} \frac{Q|Q|}{A^{Q}} = \frac{n_{i}^{2}}{1.492} \frac{Q|Q|}{A^{Q}} p^{4/3} .
\]

(15)

In way similar, substituting (2) and (13) into equations (12) and (14) results in

\[
Q_{j}^{n+1} = Q_{j}^{n} - \frac{\Delta t}{\Delta x} \left\{ \frac{g}{2T} ((A_{j+1}^{n})^{2} - (A_{j}^{n})^{2}) + \frac{(Q_{j+1}^{n})^{2}}{A_{j+1}^{n+1}} - \frac{(Q_{j}^{n})^{2}}{A_{j}^{n+1}} \right\} + gP\Delta t \left( \frac{\tau}{\rho g} - \frac{n_{i}^{2}}{1.492} P^{4} Q_{j}^{n+1} (Q_{j}^{n})^{4/3} \right) ,
\]

(16)

and

\[
Q_{j}^{n+1} = \frac{1}{2} \left\{ Q_{j}^{n} + Q_{j}^{n+1} - \frac{\Delta t}{\Delta x} \left\{ \frac{g}{2T} \left[ (A_{j+1}^{n+1})^{2} - (A_{j}^{n+1})^{2} \right] + \frac{(Q_{j+1}^{n+1})^{2}}{A_{j+1}^{n+1}} - \frac{(Q_{j}^{n+1})^{2}}{A_{j}^{n+1}} \right\} + gP\Delta t \left( \frac{\tau}{\rho g} - \frac{n_{i}^{2}}{1.492} P^{4} Q_{j}^{n+1} (Q_{j}^{n})^{4/3} \right) \right\} .
\]

(17)

An assembling of relations (11), (13), (16) and (17) provides the full description of MacCormack scheme which is given by

Predictor step:

\[
A_{j}^{n+1} = A_{j}^{n} - \frac{\Delta t}{\Delta x} (Q_{j+1}^{n} - Q_{j}^{n}) + \Delta t_{j+1}^{n},
\]

and

\[
Q_{j}^{n+1} = Q_{j}^{n} - \frac{\Delta t}{\Delta x} \left\{ \frac{g}{2T} ((A_{j+1}^{n})^{2} - (A_{j}^{n})^{2}) + \frac{(Q_{j+1}^{n})^{2}}{A_{j+1}^{n+1}} - \frac{(Q_{j}^{n})^{2}}{A_{j}^{n+1}} \right\} + gP\Delta t \left( \frac{\tau}{\rho g} - \frac{n_{i}^{2}}{1.492} P^{4} Q_{j}^{n+1} (Q_{j}^{n})^{4/3} \right) .
\]

Corrector step: use \( A_{j}^{n+1} \) and \( Q_{j}^{n+1} \) obtained above to compute \( A_{j}^{n+1} \) and \( Q_{j}^{n+1} \)

\[
A_{j}^{n+1} = \frac{1}{2} \left\{ A_{j}^{n} + A_{j}^{n+1} - \frac{\Delta t}{\Delta x} (Q_{j}^{n+1} - Q_{j-1}^{n+1}) + \Delta t_{j+1}^{n} \right\} ,
\]

and

\[
Q_{j}^{n+1} = \frac{1}{2} \left\{ Q_{j}^{n} + Q_{j}^{n+1} - \frac{\Delta t}{\Delta x} \left\{ \frac{g}{2T} \left[ (A_{j+1}^{n+1})^{2} - (A_{j}^{n+1})^{2} \right] + \frac{(Q_{j+1}^{n+1})^{2}}{A_{j+1}^{n+1}} - \frac{(Q_{j}^{n+1})^{2}}{A_{j}^{n+1}} \right\} + gP\Delta t \left( \frac{\tau}{\rho g} - \frac{n_{i}^{2}}{1.492} P^{4} Q_{j}^{n+1} (Q_{j}^{n})^{4/3} \right) \right\} .
\]
Here, the terms $A^{n+1}$ and $Q^{n+1}$ are "predicted" values of $A$ and $Q$, respectively, at the time level $n + 1$. Assuming further that the superscript $n + 1$ is a time level, it is easy to see that MacCormack algorithm is a tree level method, so the initial data $A^0$ and $Q^0$ are needed to begin the algorithm. However, appropriate initial and boundary conditions must be specified. Further, the presence of cross section in the denominator of several terms disallows zero cross sections, therefore, a finite minimum cross section is assigned to each node that is ponded. It is primarily the $A^{10/3}$ in the denominator of the friction slope term given by relation (15) that limits the magnitude of the minimum cross section and discharge. When the cross sections are very small, the friction slope is very large compared with the other terms in second equation of system (1). As cross sections increase rapidly during the early stages of flow development, the friction slope term magnitude changes much faster than the other terms. This phenomenon renders the second equation in system (1) stiff and severely limits the maximum admissible time step for stability. Indeed, this phenomenon likely forced previous researchers to use very small time steps relative to their mesh size (courant number $\leq 1$) and keep lateral inflows and initial cross sections large $[9]$.

### 3 Stability analysis of MacCormack scheme

This section deals with the stability analysis of the MacCormack numerical scheme for 1D complete shallow water equations with source terms in the case where the channel is prismatic. First, we present a rainfall hydrograph test, based on experimental measurements realized thanks to the SATREPS project METHOD in a flume at the rain simulation facility at Benoué-Garoua (Cameroon). The flume is $1150$ m long with a slope of $4\%$. The simulation duration is $1$s. The rainfall intensity $I(x, t)$ is described by

$$I(x, t) = \begin{cases} 1.18 \times 10^{-5} m/s & \text{if } (t, x) \in [0; 1] \times [0; 1]; \\ 0 & \text{otherwise.} \end{cases}$$

For this test, as there is no rain on the last $150$m, we have a wet/dry transition. The measured output is an hydrograph, that is a plot of the discharge versus time. The mathematical model for this ideal overland flow is the following: we consider a uniform plane catchment whose overall length in the direction of flow is $L$. The surface roughness and slope are assumed to be invariant in space and time. We consider a constant rainfall excess such that

$$r(x, t) = \begin{cases} I & \text{if } t_0 \leq t \leq T_1, \ 0 \leq x \leq L; \\ 0 & \text{otherwise,} \end{cases}$$

where $I$ is the rainfall intensity and $T_1$ is the final time of the rainfall excess. According to relations (18) and (19) we assume in the following that $r$ is more less that $A$ and $Q$, i.e., $r \ll A, Q$. Furthermore, Proposition 3.1 gives the "temporary" stability constraint of the MacCormack algorithm described in section 2.

**Proposition 3.1.** The numerical scheme (11) is stable if estimate (20) holds.

$$\frac{\Delta t^3}{\Delta x^2} \left( 1 + \frac{2\Delta t}{3} \Gamma_0 \mu |A^n|^{-\frac{2}{3}} \right) \leq 3\Gamma_0^{-1} (\mu^n)^{-3} |A^n|^2 |\phi|^2,$$

where $\mu = \frac{|Q|}{|A|}$, $\Gamma_0 = \frac{gn^2}{1.49^2} P^\frac{2}{3}$ and $\phi = k\Delta x$, where $k \neq 0$ is the wave number.

**Proof.** First of all, we give an explicit form of $A^{n+1}$. Combining relations (11) and (18), simple computations provide

$$Q^{n+1}_{j+1} - Q^{n+1}_{j-1} = Q^n_j - Q^n_{j-1} - \frac{\Delta t}{\Delta x} \left\{ \frac{g}{2} \left[ (A^n_{j+1})^2 - 2(A^n_j)^2 + (A^n_{j-1})^2 \right] + \frac{(Q^n_{j+1})^2}{A^n_{j+1}} - 2 \frac{(Q^n_j)^2}{A^n_j} + \frac{(Q^n_{j-1})^2}{A^n_{j-1}} \right\}$$

$$- \Delta t \frac{gn^2}{1.49^2} P^\frac{2}{3} \left( \frac{Q^n_j}{(A^n_j)^{\frac{2}{3}}} - \frac{Q^n_{j-1}Q^n_{j-1}}{(A^n_{j-1})^{\frac{2}{3}}} \right),$$

and

$$A^n_j + A^{n+1}_j = 2A^n_j - \frac{\Delta t \Delta x}{Q^n_j} (Q^n_{j+1} - Q^n_j) + \Delta t n^j.$$
Substituting equations (21) and (22) into (9) results in

\[ A_{j}^{n+1} = A_{j}^{n} - \frac{\Delta t}{2\Delta x} (Q_{j+1}^{n} - Q_{j-1}^{n}) + \frac{1}{2} \left( \frac{\Delta t}{\Delta x} \right)^{2} \left\{ \frac{g}{2T} \left[ (A_{j+1}^{n})^{2} - 2(A_{j}^{n})^{2} + (A_{j-1}^{n})^{2} \right] + \frac{Q_{j+1}^{n}}{A_{j+1}^{n}} \right\} - 2(\frac{Q_{j}^{n}}{A_{j}^{n}}) + \frac{(\frac{Q_{j}^{n}}{A_{j}^{n}})^{2}}{A_{j}^{n}} + \frac{\Delta t^{2} g n_{j}^{2}}{2\Delta x} \cdot P^{\pm} \left( \frac{Q_{j}^{n} |Q_{j}^{n}|}{(A_{j}^{n})^{2}} - \frac{Q_{j-1}^{n} |Q_{j-1}^{n}|}{(A_{j-1}^{n})^{2}} \right) + \frac{1}{2} \Delta t \left( r_{j}^{n} + r_{j}^{n+1} \right) \].

(23)

Neglecting the last term in (23), we obtain

\[ A_{j}^{n+1} = A_{j}^{n} - \frac{\Delta t}{2\Delta x} (Q_{j+1}^{n} - Q_{j-1}^{n}) + \frac{1}{2} \left( \frac{\Delta t}{\Delta x} \right)^{2} \left\{ \frac{g}{2T} \left[ (A_{j+1}^{n})^{2} - 2(A_{j}^{n})^{2} + (A_{j-1}^{n})^{2} \right] + \frac{(Q_{j+1}^{n})^{2}}{A_{j+1}^{n}} \right\} - 2(\frac{(Q_{j}^{n})^{2}}{A_{j}^{n}}) + \frac{(Q_{j}^{n})^{2}}{A_{j}^{n}} \right\} + \frac{\Delta t^{2} g n_{j}^{2}}{2\Delta x} \cdot \Gamma_{0} \left( \frac{Q_{j}^{n} |Q_{j}^{n}|}{(A_{j}^{n})^{2}} - \frac{Q_{j-1}^{n} |Q_{j-1}^{n}|}{(A_{j-1}^{n})^{2}} \right) ,

(24)

where

\[ \Gamma_{0} = \frac{g n_{j}^{2}}{1.492} P^{\pm} . \]

(25)

Indeed, since \(|r(x,t)| < \langle A(x,t) \rangle, \langle Q(x,t) \rangle, \forall (x,t) \in [0,T_{1}] \times [0, L] \), the tracking of the last term in (23) does not compromise the result.

Since the analysis considers the Von Neumann approach, we should put \( \phi = k \Delta x \) and take \( A_{j}^{n} = e^{atn} e^{ikx_{j}} \) and \( Q_{j}^{n} = e^{bx} e^{ikx_{j}} \) (where \( a, b \in \mathbb{C} \) with \( b = b_{1} + ib_{2} \) and \( a = a_{1} + ia_{2} \), where \( a_{j}, b_{j} \in \mathbb{R} \) and \( k \) is the wave number). For the sake of readability, we assume in the following that \( a_{2} = b_{2} \). Replacing this into relation (24) to get

\[ e^{a(t^{n}+\Delta t)} e^{ikx_{j}} = e^{atn} e^{ikx_{j}} - \frac{\Delta t}{2\Delta x} \left( e^{bt} e^{ik(x_{j} + \Delta x)} - e^{bt} e^{ik(x_{j} - \Delta x)} \right) + \frac{1}{2} \left( \frac{\Delta t}{\Delta x} \right)^{2} \left\{ \frac{g}{2T} \left[ 2e^{2at} e^{2ik(x_{j} + \Delta x)} - 2e^{2at} e^{2ik(x_{j})} + e^{2at} e^{2ik(x_{j} - \Delta x)} \right] + e^{2(b-a)t} \right\} - 2e^{2(b-a)t^{n}} + e^{2(b-a)t} e^{2it^{n}} + \frac{\Delta t^{2} g n_{j}^{2}}{2\Delta x} \cdot \Gamma_{0} \left( e^{bt} e^{ikx_{j}} e^{bt} - e^{bt} e^{ikx_{j}} e^{bt} \right) \right] \cdot (1 - e^{2it^{n}}) \].

(26)

Dividing side by side relation (26) by \( e^{atn} e^{ikx_{j}} \) results in

\[ e^{a\Delta t} = 1 - \frac{\Delta t}{2\Delta x} \left( e^{i\phi} - e^{-i\phi} \right) e^{(b-a)t^{n}} + \frac{1}{2} \left( \frac{\Delta t}{\Delta x} \right)^{2} \left\{ e^{atn} e^{ikx_{j}} e^{2i\phi} - 2e^{atn} e^{ikx_{j}} \right\} \]

\[ + e^{atn} e^{ikx_{j}} e^{-2i\phi} + \frac{\Delta t^{2} g n_{j}^{2}}{2\Delta x} \cdot \Gamma_{0} \left( e^{bt} e^{ikx_{j}} e^{bt} - e^{bt} e^{ikx_{j}} e^{bt} \right) \cdot (1 - e^{2it^{n}}). \]

(27)

Using the identities: \( e^{i\phi} - e^{-i\phi} = 2i \sin(\phi) \), \( e^{2i\phi} - e^{-2i\phi} = -4 \sin^{2}(\phi) \) and \( 1 - e^{2it^{n}} = 2 \sin^{2}(\frac{t^{n}}{2}) + i \sin(\frac{t^{n}}{2}) \), equation (27) becomes

\[ e^{a\Delta t} = 1 - i \frac{\Delta t}{\Delta x} \sin(\phi) e^{(b_{1}-a_{1})t^{n}} + \frac{\Delta t^{2} g n_{j}^{2}}{2T} \sin^{2}(\phi) e^{a_{1}t^{n}} [\cos(a_{2}t^{n} + kx_{j}) + i \sin(a_{2}t^{n} + kx_{j})] \]

\[ - i \frac{\Delta t^{2} g n_{j}^{2}}{2T} \sin(\phi) \cos(\phi) e^{a_{1}t^{n}} [\cos(\phi) e^{(b_{1}-a_{1})t^{n}}] \left( \cos \left( \frac{7}{3} [a_{2}t^{n} + kx_{j}] \right) - i \sin \left( \frac{7}{3} [a_{2}t^{n} + kx_{j}] \right) \right) \]

\[ - 2 \left( \frac{\Delta t}{\Delta x} \right)^{2} \sin(\frac{1}{2} \phi) e^{(b_{1}-a_{1})t^{n}} \]

\[ = 1 - i \frac{\Delta t}{\Delta x} \mu \sin(\phi) \left( \frac{\Delta t}{\Delta x} \right)^{2} \frac{g}{2T} \sin^{2}(\phi) e^{a_{1}t^{n}} [\cos(a_{2}t^{n} + kx_{j}) + i \sin(a_{2}t^{n} + kx_{j})] \]

\[ - i \frac{\Delta t^{2} g n_{j}^{2}}{2T} \cos(\phi) e^{a_{1}t^{n}} [\cos(\phi) e^{(b_{1}-a_{1})t^{n}}] \left( \cos \left( \frac{7}{3} [a_{2}t^{n} + kx_{j}] \right) - i \sin \left( \frac{7}{3} [a_{2}t^{n} + kx_{j}] \right) \right) \]

\[ - 2 \left( \frac{\Delta t}{\Delta x} \right)^{2} \mu^{2} \sin^{2}(\frac{1}{2} \phi) , \]

(28)
where
\[ \mu^n = \frac{|Q^n|}{|A^n|} = e^{\mu_1 t^n} = e^{(b_1 - a_1) t^n}. \] (29)

Putting
\[ \alpha_2 = \frac{7}{3}(a_2 t^n + k x_j); \quad \gamma_2 = \Gamma_0 (\mu^n)^2 e^{-\frac{4}{9} a_1 t^n} e^{-i \alpha_2}; \quad \alpha_3 = a_2 t^n + k x_j \quad \gamma_1 = \mu^n; \quad \gamma_4 = \gamma_1^2; \quad \gamma_3 = \frac{q}{T} e^{a_1 t^n} e^{i \alpha_3}. \]

Utilizing this, equation (28) yields
\[
e^{a \Delta t} = 1 - i \frac{\Delta t}{\Delta x} \gamma_1 \sin(\phi) - \left( \frac{\Delta t}{\Delta x} \right)^2 \gamma_2 \sin^2(\phi) - i \frac{\Delta t^2}{\Delta x} \gamma_2 \sin(\frac{2}{3} \phi) \cos(\frac{2}{3} \phi) - 2 \left( \frac{\Delta t}{\Delta x} \right)^2 \gamma_3 \sin^2(\frac{1}{2} \phi) = 1 - \frac{\Delta t^2}{\Delta x} |\gamma_2| \sin \alpha_2 \sin(\frac{2}{3} \phi) \cos(\frac{2}{3} \phi) - \left( \frac{\Delta t}{\Delta x} \right)^2 \left[ |\gamma_3| \cos \alpha_3 \sin^2(\phi) + 2 |\gamma_4| \sin^2(\frac{1}{2} \phi) \right] - i \left\{ \frac{\Delta t}{\Delta x} |\gamma_1| \sin(\phi) + \frac{\Delta t^2}{\Delta x} |\gamma_2| \cos \alpha_2 \sin(\frac{2}{3} \phi) \cos(\frac{2}{3} \phi) + \left( \frac{\Delta t}{\Delta x} \right)^2 |\gamma_3| \sin \alpha_3 \sin^2(\phi) \right\}. \] (31)

Of course the aim of this report is to give the general picture of necessary condition of stability. Since the formulae can become quite heavy, for the sake of simplicity, we consider in all the proofs the following remark which plays a crucial role in our study. However, the obtained result is a linear stability condition which can be considered as a necessary condition of stability.

**Remark 3.1.** Since the considered problem is a nonlinear partial differential equations, the solution may contain discontinuity even if the initial conditions are smooth enough. To overcome this numerical challenge, we should assume that the phase angle \( \phi = k \Delta x \) satisfies \( |\phi| \ll \pi \). In fact, the method could be generally stabilized by adding additional dissipation to the scheme without affecting the order of accuracy.

Now, we must analyze some extreme cases. The extreme cases are obtained when \( |\phi| = \pi \), on the one hand, and when \( \phi \) equals zero on the other.

- **Case** \( |\phi| = \pi \). For \( |\phi| = \pi \), it comes from equation (31) that the amplification factor becomes
\[
e^{a \Delta t} = 1 + \frac{\sqrt{3} \Delta t^2}{4 \Delta x} |\gamma_2| \sin \alpha_2 - 2 \left( \frac{\Delta t}{\Delta x} \right)^2 |\gamma_4| - i \frac{\sqrt{3} \Delta t^2}{4 \Delta x} |\gamma_2| \cos \alpha_2. \]

The squared modulus of the amplification factor equals
\[
|e^{a \Delta t}|^2 = 1 + \frac{3 \Delta t^4}{16 \Delta x^2} |\gamma_2|^2 + 4 \left( \frac{\Delta t}{\Delta x} \right)^4 |\gamma_4|^2 + 2 \left( \frac{\sqrt{3} \Delta t^2}{4 \Delta x} |\gamma_2|(1 + \Delta t^2 |\gamma_4|) \sin \alpha_2 - 2 \left( \frac{\Delta t}{\Delta x} \right)^2 |\gamma_4| \right) .
\]

But there exists values of \( \alpha_2 \) for which \( \sin \alpha_2 = 1 \). So \( |e^{a \Delta t}|^2 > 1 \). Thus the scheme is unconditionally unstable.

- **Case** \( \phi = 0 \). In that case, the amplification factor given by equation (31) provides \( e^{a \Delta t} = 1 \). Then the modulus is \( |e^{a \Delta t}| = 1 \). Then, the numerical scheme is neutrally stable. Thus, the MacCormack method is not dissipative in the sense of Kreiss [15] and when applied to complete uni-dimensional shallow water equations with source terms [1]. That is, the computations should become unstable in certain circumstances. This instability is entirely due to the non-linearity of the equations, since the same scheme applied to linear shallow water equations without source terms does not diverge, although strong oscillations are generated (see for example, [11] [21] [22]).

- **Case where** \( 0 < |\phi| \ll \pi \). Using the Taylor expansion around \( \phi = 0 \), and neglecting high-order terms, the squared modulus of the amplification factor given by (31) is approximated as
\[
|e^{a \Delta t}|^2 = \left( 1 - \frac{2 \Delta t^2}{3 \Delta x} |\gamma_2| \phi \sin \alpha_2 \right)^2 + \left( \frac{\Delta t}{\Delta x} |\gamma_1| + \frac{2 \Delta t^2}{3 \Delta x} |\gamma_2| \cos \alpha_2 \right)^2 \phi^2 . \] (32)
Proposition 3.2. MacCormack scheme (9) is stable if the following estimate is satisfied

\[ \Delta t \left( 3 P_n W_1(\Delta t, \Delta x) + \frac{1}{\Delta x} \max \{ W_2(\Delta t, \Delta x); W_3(\Delta t, \Delta x) \} \right) \leq \max \left\{ 1 + \sqrt{1-r^*}; \sqrt{r^*} \right\}, \]

where \( r^* \in (0; 1) \) and

\[ W_1(\Delta t, \Delta x) = \frac{1}{2} + \left[ 1 + 4 \left( \Delta t + \Delta t^2 + \frac{\Delta t}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \left( 1 + \frac{\Delta t^2}{\Delta x} \right)^2 \right) \right] \max \left\{ P_n, 2 P_n^2 \frac{1}{2} R_n |\phi|, \mu^n |\phi|, \right\}, \]

\[ W_2(\Delta t, \Delta x) = (P_n + \frac{1}{2} R_n) |\phi| + \mu^n \left[ 1 + 4 \left( \Delta t + \Delta t^2 + \frac{\Delta t}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \left( 1 + \frac{\Delta t^2}{\Delta x} \right)^2 \right) \right] \left[ 1 + (4 + R_n (\mu^n)^{-1} + R_n^{-1} + N_n R_n (\mu^n)^{-1}) \left( 1 + \frac{\Delta t}{\Delta x} \mu^n \right) |\phi| \right], \]

\[ W_3(\Delta t, \Delta x) = \left( \frac{\Delta t}{\Delta x} \right)^2 \left( 1 + 4 \left( \Delta t + \Delta t^2 + \frac{\Delta t}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \left( 1 + \frac{\Delta t^2}{\Delta x} \right)^2 \right) \right) \left[ 1 + (4 + R_n (\mu^n)^{-1} + R_n^{-1} + N_n R_n (\mu^n)^{-1}) \left( 1 + \frac{\Delta t}{\Delta x} \mu^n \right) |\phi| \right]. \]
\[
\max \left\{ P_n, N_n, R_n P_n, N_n P_n, N_n P_n, R_n \mu^n, P_n^2, N_n^2, P_n^2 \mu^n, R_n^2 \mu^n, R_n^2 (\mu^n)^2, R_n P_n \mu^n, R_n N_n \mu^n \right\};
\]

\[
W_3(\Delta t, \Delta x) = (P_n + \frac{1}{2} R_n) |\phi| + \frac{3}{4} \mu^n \left\{ |\phi| + 2 \left[ \Delta t + \Delta t^2 + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^3}{\Delta x} + \left( \frac{\Delta t}{\Delta x} \right)^2 \right] \right\}
\]

\[
\left[ 1 + N_n R_n (\mu^n)^{-1} + \left( 4 + 2 P_n + R_n^{-1} + R_n^{-1} \mu^n + 4 N_n R_n^{-1} (\mu^n)^{-1} + \frac{\Delta t}{\Delta x} \mu^n \right) |\phi| \right]
\]

\[
\max \left\{ P_n, N_n, N_n P_n, R_n \mu^n, P_n^2, N_n^2, R_n^2 \mu^n, P_n R_n \mu^n, R_n^2 (\mu^n)^2, R_n N_n \mu^n \right\};
\]

with\[
\mu^n = \frac{|Q_j^n|}{|A_j^n|} = e^{(b_1-a_1)^n}; \quad R_n = \frac{g}{T} |A_j^n| |\mu^n|^{-1} + \mu^n; \quad P_n = \frac{P_T}{\rho} |Q_j^n|^{-1} + \Gamma_0 \mu^n |A_j^n|^\frac{3}{2}; \quad N_n = \frac{g}{2T} |Q_j^n||\mu^n|^{-1}.
\]

The proof of Proposition 3.2 requires some intermediate results which play a crucial role in the study of the amplification factor associated with \([10]\).

**Lemma 3.1.** Let \(n\) and \(j\) be nonnegative integers. Then the terms \(\frac{1}{2} \left( Q_j^n + Q_j^{n+1} \right)\) and \(\frac{g}{4T} \left[ (A_j^{n+1})^2 - (A_j^{n+1})^2 \right]\) can be approximated as
\[
\frac{1}{2} \left( Q_j^n + Q_j^{n+1} \right) = Q_j^n \left\{ 1 + \frac{\Delta t}{2 \Delta x} C_{11} + \frac{1}{2} \Delta t C_{12} + i \left( \frac{\Delta t}{2 \Delta x} \overline{C}_{11} \phi + \frac{1}{2} \Delta t \overline{C}_{12} \right) \right\} + O(\phi^2),
\]

and
\[
\frac{g}{4T} \left[ (A_j^{n+1})^2 - (A_j^{n+1})^2 \right] = Q_j^n \left( C_{21} + i \overline{C}_{21} \right) \phi + O(\phi^2),
\]

where
\[
C_{11} = \frac{g}{T} |A^n| |\mu^n|^{-1} \sin \alpha_3; \quad \overline{C}_{11} = -\frac{g}{T} |A^n| |\mu^n|^{-1} \cos \alpha_3 - \mu^n; \quad C_{12} = \frac{P_T}{\rho} |Q_j^n|^{-1} \cos \alpha_3 - \Gamma_0 \mu^n |A_j^n|^{-\frac{3}{2}} \cos \alpha_2;
\]
\[
\overline{C}_{12} = \frac{P_T}{\rho} |Q_j^n|^{-1} \sin \alpha_3 - \Gamma_0 \mu^n |A_j^n|^{-\frac{3}{2}} \sin \alpha_2; \quad C_{21} = -\frac{g}{2T} |Q_j^n||\mu^n|^{-1} \sin \alpha_3; \quad \overline{C}_{21} = \frac{g}{2T} |Q_j^n||\mu^n|^{-1} \cos \alpha_3;
\]

where
\[
\alpha_3 = a_2 t^n + k x_j, \quad \text{and} \quad \alpha_2 = \frac{7}{3} \alpha_3.
\]

**Proof.** First, we recall that \(A_j^n = e^{a_1 \phi} e^{i(k x_j)} = e^{(a_1 + i a_2) t^n e^{i k x_j}}\) and \(Q_j^n = e^{b_1 t^n} e^{i k x_j} = e^{(b_1 + i a_2) t^n e^{i k x_j}}\). Expanding the Taylor series around \(\phi\) and neglecting the terms of high-order to obtain \(e^{i \phi} - 1 = i \phi + O(\phi^2)\) and \(e^{2i \phi} - 1 = 2i \phi + O(\phi^2)\) (this is true according to Remark 3.1). Utilizing this, we get
\[
\frac{1}{2} \left( Q_j^n + Q_j^{n+1} \right) = Q_j^n - \frac{\Delta t}{2 \Delta x} \left\{ \frac{g}{2T} \left[ (A_j^{n+1})^2 - (A_j^n)^2 \right] + \frac{(Q_j^{n+1})^2}{A_j^{n+1}} - \frac{(Q_j^n)^2}{A_j^n} + \frac{1}{2} \Delta t \left( \frac{P_T}{\rho} - \Gamma_0 \frac{Q_j^n}{(A_j^n)^2} \right) \right\}
\]

\[
= Q_j^n - \frac{\Delta t}{2 \Delta x} \left\{ \frac{g}{2T} \left[ e^{2a_1 \phi} e^{i(k x_j)} e^{i(k x_j)} - e^{2a_1 \phi} e^{i(k x_j)} \right] + \frac{e^{2b_1 t^n} e^{2i(k x_j)}}{e^{a_1 t^n} e^{i(k x_j)}} - \frac{e^{2b_1 t^n} e^{2i(k x_j)}}{e^{a_1 t^n} e^{i(k x_j)}} + \frac{1}{2} \Delta t \left( \frac{P_T}{\rho} - \frac{e^{\phi}}{e^{a_1 t^n} e^{i(k x_j)}} \right) \right\}
\]

\[
- \Gamma_0 \left[ \frac{|e^{b_1 t^n}|}{e^{a_1 t^n} e^{i(k x_j)}} \right] = Q_j^n \left\{ 1 - i \frac{\Delta t}{2 \Delta x} \phi \left( \frac{g}{T} 2(a_2 - b_1) t^n e^{i k x_j} + e^{(b_2 - a_2) t^n} \right) + \frac{1}{2} \Delta t \left( \frac{P_T}{\rho} (Q_j^n)^{-1} \right)^{-1}
\]

\[
\Gamma_0 \left[ \frac{|Q_j^n||A_j^n|^\frac{3}{2}}{e^{a_1 t^n} e^{i(k x_j)}} \right] + O(\phi^2) = Q_j^n \left\{ 1 - i \frac{\Delta t}{2 \Delta x} \phi \left( \frac{g}{T} 2(a_2 - b_1) t^n e^{i(k x_j)} + e^{(b_2 - a_2) t^n} \right) + \frac{1}{2} \Delta t \left( \frac{P_T}{\rho} (Q_j^n)^{-1} \right)^{-1}
\]

10
\[
\Gamma_0 [\mu^n] A_j^n \left( \frac{-4}{e^{\nu (\alpha_2 + k x_j)}} \right) \right) + O(\phi^2) = Q_j^n \left\{ 1 - i \frac{\Delta t}{2\Delta x} \phi \left( \frac{g}{T} A^n (\mu^n)^{-1} e^{i\alpha x} + \mu^n \right) + \frac{1}{2} \Delta t \left( \frac{\Gamma \rho}{\rho} (Q_j^n)^{-1} - 3 \phi \right) \right\} + O(\phi^2),
\]

On the other hand, using equation (7) and applying the Taylor expansion around \(0\), we get

\[
\frac{g}{4T} \left[ (A_j^{n+1})^2 - (A_j^{n+1})^2 \right] = \frac{g}{4T} \left( A_j^{n+1} - A_j^{n+1} \right) \left( A_j^{n+1} + A_j^{n+1} \right) = \frac{g}{4T} \left( A_j^n - A_j^n \right)
\]

\[
- \frac{\Delta t}{\Delta x} \left( Q_j^{n+1} - 2Q_j^n + Q_j^{n-1} + \Delta t(r_j^n - r_{j-1}^n) \right) \left( A_j^n + A_j^n - \frac{\Delta \phi}{\Delta x} (Q_j^{n+1} - Q_j^{n-1}) + \Delta t(r_j^n + r_{j-1}^n) \right)
\]

\[
\approx \frac{g}{4T} \left( e^{a^n x} e^{ikx_j} - e^{a^n x} e^{ikx_{j-1}} - \frac{\Delta t}{\Delta x} \left( e^{b^n x} e^{ikx_{j+1}} + 2e^{b^n x} e^{ikx_{j-1}} + e^{b^n x} e^{ikx_{j-1}} \right) \right) \left( e^{b^n x} e^{ikx_j} + e^{a^n x} e^{ikx_{j-1}} - \frac{\Delta t}{\Delta x} \left( e^{a^n x} e^{ikx_{j+1}} + 2e^{b^n x} e^{ikx_{j-1}} + e^{b^n x} e^{ikx_{j-1}} \right) \right)
\]

\[
\frac{\left( e^{a^n (a+b) x} (1 - e^{-i\phi}) - \frac{\Delta \phi}{\Delta x} (e^{i\phi} - e^{-i\phi}) \right) x e^{c^n (a+b) x} (1 + e^{-i\phi}) - \frac{\Delta \phi}{\Delta x} (e^{i\phi} - e^{-i\phi}) \right) \}
\]

\[
\frac{\left( e^{a^n (a+b) x} (1 - e^{-i\phi}) - \frac{\Delta \phi}{\Delta x} (e^{i\phi} - e^{-i\phi}) \right) x e^{c^n (a+b) x} (1 + e^{-i\phi}) - \frac{\Delta \phi}{\Delta x} (e^{i\phi} - e^{-i\phi}) \right) \}
\]

\[
\frac{\left( e^{a^n (a+b) x} (1 - e^{-i\phi}) - \frac{\Delta \phi}{\Delta x} (e^{i\phi} - e^{-i\phi}) \right) x e^{c^n (a+b) x} (1 + e^{-i\phi}) - \frac{\Delta \phi}{\Delta x} (e^{i\phi} - e^{-i\phi}) \right) \}
\]

\[
\frac{\left( e^{a^n (a+b) x} (1 - e^{-i\phi}) - \frac{\Delta \phi}{\Delta x} (e^{i\phi} - e^{-i\phi}) \right) x e^{c^n (a+b) x} (1 + e^{-i\phi}) - \frac{\Delta \phi}{\Delta x} (e^{i\phi} - e^{-i\phi}) \right) \}
\]

where \(C_{11}, C_{12}, C_{12} \) and \(C_{12} \) are defined by relation (39).

\[\square\]

**Lemma 3.2.** Let consider \(n\) and \(j\) be two nonnegative integers. Then the term \(\frac{(Q_j^{n+1})^2}{A_j^{n+1}} - \frac{(Q_j^{n+1})^2}{A_j^{n+1}}\) can be approximated as

\[
\frac{(Q_j^{n+1})^2}{A_j^{n+1}} - \frac{(Q_j^{n+1})^2}{A_j^{n+1}} = Q_j^n n \left\{ H_1 - K_1^2 + K_2^2 + 2K_1 K_2 (1 - \frac{\Delta t}{\Delta x} \mu^n) \phi + i (H_2 - 2K_1 K_2 +
\]

\[
(K_2^2 - K_1^2) (1 - \frac{\Delta t}{\Delta x} \mu^n) \phi \right\} + O(\phi^2),
\]

where the functions \(H_1, H_2, K_1^2, K_2^2\) and \(K_1 K_2\) are given by

\[
H_1 = 1 + 2\Delta t C_{12} + \Delta t^2 (C_{12}^2 - C_{12}^2) + 2 \frac{\Delta t^2}{\Delta x} \left[ C_{11} C_{12} - C_{11} C_{12} - \mu^n (C_{12} + \Delta t C_{12} C_{12}) \right] \phi;
\]

\[
H_2 = \Delta t C_{12} + \Delta t^2 C_{12} C_{12} + \frac{\Delta t}{\Delta x} \left[ C_{11} + \mu^n + \Delta t (C_{11} C_{12} + C_{11} C_{12} + \mu^n (2C_{12} + \Delta t (C_{12}^2 - C_{12}^2))) \right] \phi;
\]

\[
K_1^2 = 1 + \left( \frac{\Delta t}{2\Delta x} \right)^2 \left[ C_{21} + 4C_{21} C_{21} \phi + \Delta t \left[ C_{22}^2 - 2C_{22}^2 C_{22} \phi - \frac{\Delta t}{\Delta x} (C_{22} + 2C_{21} C_{22} \phi) + \Delta t (C_{12} - 2C_{22} \phi) - \frac{\Delta t^2}{\Delta x} [C_{12} C_{21} - (C_{21} C_{22} - 2C_{21} C_{22}) \phi];
\]

\[
K_2 = \left( \frac{\Delta t}{2\Delta x} \right)^2 (\mu^n)^2 \left[ C_{11}^2 + 2C_{11} (1 - C_{11} - \mu^n) \phi + \Delta t^2 [C_{12}^2 - 2C_{12} C_{22} \phi + \Delta t^2 \mu^n C_{11} \phi +
\]

\[
2 \Delta t C_{12} \phi + \frac{\Delta t^2}{\Delta x} \mu^n [C_{11} C_{12} - C_{11} C_{22} \phi + C_{12} (1 - C_{11} - \mu^n) \phi];
\]

11
\[
K_1K_2 = - \left\{ \phi - \left( \frac{\Delta t}{2\Delta x} \right)^2 \mu^n [C_{11}C_2 + C_{21}(1 - C_{11} - \mu^n) + 2C_{11}C_{21}\phi] + \Delta t [C_{12} - (C_{12}C_{22} - C_{12})\phi] \\
+ \frac{\Delta t}{2\Delta x} \mu^n [C_{11} + (1 - C_{11} - \mu^n - C_{21}(\mu^n)^{-1})\phi] + \frac{\Delta t^2}{2\Delta x} \mu^n [C_{11}C_{12} - C_{12}C_{21}(\mu^n)^{-1} + [C_{12}(1 - C_{11} - \mu^n) - C_{11}C_{22} + C_{21}C_{22}(\mu^n)^{-1} - 2C_{12}C_{22}(\mu^n)^{-1}]\phi] + \Delta t^2 [C_{12}C_{12} - (C_{12}C_{22} + C_{12}C_{22})\phi] \right\};
\]

(46)

\( C_{rs}, \overline{C}_{rs}, r, s = 1, 2, \) are given by (39) and \( \alpha_2, \alpha_3 \) come from relation (20).
Furthermore
\[
C_{22} = \frac{4}{3} \Gamma_0 \mu^n |A_j^n|^{-\frac{3}{2}} \sin \alpha_2 \quad \text{and} \quad \overline{C}_{22} = \frac{4}{3} \Gamma_0 \mu^n |A_j^n|^{-\frac{3}{2}} \cos \alpha_2.
\]

(47)

Proof. The following identity holds
\[
Q_j^{n+1} = 2 \left( \frac{1}{2} (Q_j^{n+1} + Q_j^n) \right) - Q_j^n.
\]

This fact, along with equation (37) result in
\[
Q_j^{n+1} = Q_j^n \left\{ 1 + \frac{\Delta t}{\Delta x} C_{11} \phi + \Delta t C_{12} + i \left( \frac{\Delta t}{\Delta x} C_{11} \phi + \Delta t \overline{C}_{12} \right) \right\} + O(\phi^2),
\]

(48)

where \( C_{1j} \) and \( \overline{C}_{1j}, j = 1, 2, \) are given by relations (39) and (40).

Taking the square of \( Q_j^{n+1} \) and neglecting the high-order terms in \( \phi \), it comes from equation (45) that
\[
(Q_j^{n+1})^2 = (Q_j^n)^2 \left\{ 1 + \frac{\Delta t}{\Delta x} C_{11} \phi + \Delta t C_{12} + i \left( \frac{\Delta t}{\Delta x} C_{11} \phi + \Delta t \overline{C}_{12} \right) \right\}^2 + O(\phi^2) =
\]

\[
(Q_j^n)^2 \left\{ \left( 1 + \frac{\Delta t}{\Delta x} C_{11} \phi + \Delta t C_{12} \right)^2 + \frac{\Delta t^2}{\Delta x} \left( C_{11} \phi + \Delta t \overline{C}_{12} \right)^2 + 2i \left( 1 + \frac{\Delta t}{\Delta x} C_{11} \phi + \Delta t C_{12} \right) \times \left( \frac{\Delta t}{\Delta x} C_{11} \phi + \Delta t \overline{C}_{12} \right) \right\} + O(\phi^2) = (Q_j^n)^2 \left\{ 1 + 2\Delta t C_{12} + \Delta t^2 (C_{12}^2 - \overline{C}_{12}) + 2 \frac{\Delta t^2}{\Delta x} (C_{11}C_{12} - \overline{C}_{11} \overline{C}_{12}) \phi + 2i \left( \Delta t \overline{C}_{12} + \Delta t \overline{C}_{12} C_{12} + \frac{\Delta t}{\Delta x} \left( C_{11} + \Delta t (C_{11} \overline{C}_{12} + \overline{C}_{11} C_{12}) \right) \phi \right) \right\} \right\} + O(\phi^2).
\]

(49)

Utilizing equation (7) together with Remark 3.1 simple calculations give
\[
A_j^{n+1} = e^{at^n} e^{ikx_j} - \frac{\Delta t}{\Delta x} e^{b^n} e^{ikx_j} (e^{i\phi} - 1) = Q_j^n \left\{ e^{(a-b)t^n} - \frac{\Delta t}{\Delta x} (e^{i\phi} - 1) \right\} = Q_j^n (\mu^n)^{-1} \left\{ 1 - i \frac{\Delta t}{\Delta x} \mu^n \phi \right\} + O(\phi^2),
\]

(50)

and
\[
A_{j-1}^{n+1} = e^{at^n} e^{ikx_{j-1}} - \frac{\Delta t}{\Delta x} e^{b^n} e^{ikx_{j-1}} (1 - e^{-i\phi}) = Q_j^n \left\{ e^{(a-b)t^n} e^{-i\phi} - \frac{\Delta t}{\Delta x} (1 - e^{-i\phi}) \right\} = Q_j^n \left\{ e^{(a_1-b_1)t^n} e^{-i\phi} - \frac{\Delta t}{\Delta x} (1 - e^{-i\phi}) \right\} = Q_j^n (\mu^n)^{-1} \left\{ 1 + i \left( -1 + \frac{\Delta t}{\Delta x} \mu^n \phi \right) \right\} + O(\phi^2).
\]

(51)

In way similar, one easily shows that
\[
Q_{j-1}^{n+1} = Q_j^n \left\{ e^{-i\phi} - \frac{\Delta t}{\Delta x} \left( \frac{g(t^n)}{2\Delta x} |A_j^n| e^{i\phi_s} (1 - e^{-2i\phi}) + \mu^n (1 - e^{-i\phi}) \right) + \Delta t \left( \frac{P_T}{\rho} |Q_j^n|^{-1} e^{-i\phi_s} - \Gamma_0 \mu^n |A_j^n|^{-\frac{3}{2}} e^{-i\alpha_2} e^{i\phi} \right) \right\} = Q_j^n (K_1 + iK_1) + O(\phi^2),
\]

(52)
where
\[ K_1 = 1 - \frac{\Delta t}{2\Delta x}(C_{21} + 2C_{21}\phi) + \Delta t(C_{12} - C_{22}\phi), \]  
(53)
and
\[ K_2 = -\phi - \frac{\Delta t}{2\Delta x}\mu^n[C_{11} + (1 - C_{11} - \mu^n)\phi] - \Delta t(C_{12} - C_{22}\phi), \]  
(54)
where \(C_{rs}, \overline{C}_{rs}, r, s = 1, 2,\) are given by (39) and \(\alpha_2, \alpha_3\) come from relation (40). Furthermore
\[ C_{22} = \frac{4}{3}\Gamma_0\mu^n|A_j^n|^{\frac{3}{2}}\sin\alpha_2 \]  
and \(\overline{C}_{22} = \frac{4}{3}\Gamma_0\mu^n|A_j^n|^{\frac{3}{2}}\cos\alpha_2.\)

From equation (52), simple computations give
\[ (Q_j^{n+1})^2 = (Q_j^n)^2(K_1^2 - K_2^2 + 2iK_1K_2) + O(\phi^2). \]  
(55)
Applying Remark 3.3, we obtain the following approximation
\[ K_1^2 = 1 + \left(\frac{\Delta t}{2\Delta x}\right)^2\left[\overline{C}_{21}^2 + 4C_{21}\overline{C}_{21}\phi\right] + \Delta t^2\left[C_{12}^2 - 2C_{12}\overline{C}_{22}\phi\right] - \frac{\Delta t^2}{\Delta x}(C_{21} + 2C_{21}\phi) + \]  
\[ \Delta t(C_{12} - 2C_{22}\phi) - \frac{\Delta t^2}{\Delta x}\left[C_{12}\overline{C}_{21} - (C_{21}C_{22} - 2C_{12}C_{21})\phi\right] + O(\phi^2), \]
\[ K_2^2 = \left(\frac{\Delta t}{2\Delta x}\right)^2(\mu^n)^2\left[C_{11}^2 + 2C_{11}(1 - C_{11} - \mu^n)\phi\right] + \Delta t^2\left[C_{11}^2 - 2C_{11}\overline{C}_{22}\phi\right] + \frac{\Delta t^2}{\Delta x}\mu^nC_{11}\phi + \]  
\[ 2\Delta tC_{12}\phi + \frac{\Delta t^2}{\Delta x}\mu^n\left[C_{11}\overline{C}_{12} - C_{11}\overline{C}_{22}\phi + C_{12}(1 - C_{11} - \mu^n)\phi\right] + O(\phi^2), \]

and
\[ K_1K_2 = -\left\{\phi - \left(\frac{\Delta t}{2\Delta x}\right)^2\mu^n\left[C_{11}\overline{C}_{21} + C_{21}(1 - C_{11} - \mu^n) + 2C_{11}C_{21}\phi\right] + \Delta t\left[C_{12} - (C_{12}\overline{C}_{22} - C_{12})\phi\right] + \]  
\[ + \frac{\Delta t}{2\Delta x}\mu^n\left[C_{11} + (1 - C_{11} - \mu^n - C_{21}(\mu^n)^{-1})\phi\right] + \frac{\Delta t^2}{2\Delta x}\mu^n\left[C_{11}C_{12} - C_{12}\overline{C}_{21}(\mu^n)^{-1} + C_{12}(1 - C_{11} - \mu^n)\right] - \]  
\[ C_{11}C_{22} + C_{21}C_{22}(\mu^n)^{-1} - 2C_{21}\overline{C}_{12}(\mu^n)^{-1}\phi\]  
\[ + \Delta t^2\left[C_{12}\overline{C}_{12} - (C_{12}\overline{C}_{22} + C_{12}C_{22})\phi\right]\} + O(\phi^2). \]

Combining approximations (49) and (50) on the one hand, (53) and (51) on the other hand, straightforward computations yield
\[ \frac{(Q_j^{n+1})^2}{A_j^{n+1}} = Q_j^n\mu^n\{H_1 + iH_2\} + O(\phi^2), \]  
(56)
and
\[ \frac{(Q_j^{n+1})^2}{A_j^{n-1}} = Q_j^n\mu^n\left\{K_1^2 - K_2^2 - 2K_1K_2\left(1 - \frac{\Delta t}{\Delta x}\mu^n\right)\phi + i\left[2K_1K_2 + \left(1 - \frac{\Delta t}{\Delta x}\mu^n\right)(K_1^2 - K_2^2)\phi\right]\right\} + O(\phi^2), \]  
(57)
where \(H_1, H_2, K_1^2, K_2^2\) and \(K_1K_2\) are given by equations (12), (13), (14), (15) and (16), respectively. Subtracting equation (57) from approximation (56) completes the proof.

**Lemma 3.3.** For \(n\) and \(j\) be nonnegative integers, the term \(\frac{Q_j^{n+1}|Q_j^{n+1}|^2}{(A_j^{n+1})^2}\) can be approximated as
\[ \frac{Q_j^{n+1}|Q_j^{n+1}|^2}{(A_j^{n+1})^2} = Q^n|Q^n|^{\frac{3}{2}}(\mu^n)^{\frac{3}{2}}\left\{1 + \left(\frac{\Delta t}{\Delta x}\right)^2(\mu^n)^2\phi^2\right\}^{\frac{3}{2}}\left\{1 + \frac{\Delta t}{\Delta x}C_{11}\phi + 2\Delta tC_{12}\right\}^2 + \]
\[
\left( \frac{\Delta t}{\Delta x} \bar{C}_{11}\phi + 2\Delta t \bar{C}_{12} \right) \left\{ 1 + \frac{\Delta t}{\Delta x} C_{11}\phi + 2\Delta t \bar{C}_{12} + i \left( \frac{\Delta t}{\Delta x} \bar{C}_{11}\phi + 2\Delta t \bar{C}_{12} \right) \right\} \times
\left( C_{31} + \frac{\Delta t}{\Delta x} \bar{C}_{32}\phi + i \left[ -\bar{C}_{31} + \frac{\Delta t}{\Delta x} C_{32}\phi \right] \right) + O(\phi^2).
\]
\end{equation}

where the functions \( C_{11} \) and \( \bar{C}_{11} \), \( l = 1, 2 \), are defined by relations (39) and (40), \( C_{31} \) and \( \bar{C}_{31} \), \( l = 1, 2 \), are given by
\[
C_{31} = \cos \alpha_2; \quad \bar{C}_{31} = -\sin \alpha_2; \quad C_{32} = \mu^n \cos \alpha_2; \quad \bar{C}_{32} = \mu^n \sin \alpha_2.
\]

Proof. It comes from approximation (48) that
\[
Q_n^{n+1} = Q_n \left\{ 1 + \frac{\Delta t}{\Delta x} C_{11}\phi + 2\Delta t \bar{C}_{12} + i \left( \frac{\Delta t}{\Delta x} \bar{C}_{11}\phi + 2\Delta t \bar{C}_{12} \right) \right\} + O(\phi^2),
\]
where \( C_{1j} \) and \( \bar{C}_{1j} \), \( j = 1, 2 \), are given by relations (39) and (40). So, the modulus of the discharge at the temporary time level \( n + 1 \) and at the position \( x_j \) is approximated as
\[
|Q_n^{n+1}| = |Q_n| \left\{ 1 + \frac{\Delta t}{\Delta x} C_{11}\phi + 2\Delta t \bar{C}_{12} \right\} + O(\phi^2),
\]

Similarly, from relation (60) we have that
\[
A_n^{n+1} = Q_n \left\{ 1 - \frac{\Delta t}{\Delta x} \mu^n \phi \right\} + O(\phi^2).
\]

So, the modulus of the cross section at predicted time \( t^{n+1} \) and at position \( x_j \) is approximated by
\[
|A_n^{n+1}| = |Q_n^{n+1}| \left( \mu^n \right)^{-1} \left\{ 1 + \left( \frac{\Delta t}{\Delta x} \right)^2 (\mu^n)^2 \phi^2 \right\} + O(\phi^2).
\]

On the other hand, the following equality holds
\[
\frac{Q_n^{n+1}}{|(A_n^{n+1})^2|} = |Q_n^{n+1}||Q_n^{n+1}| \left( A_n^{n+1} \right)^{-2},
\]
where \( \overline{A_n^{n+1}} \) designates the conjugate of \( A_n^{n+1} \).
Now, combining relations (60), (61), (60), and (62), by straightforward computations, an approximate formula of (63), is given by
\[
\frac{Q_n^{n+1}}{|(A_n^{n+1})^2|} = |Q_n| \left\{ 1 + \frac{\Delta t}{\Delta x} C_{11}\phi + 2\Delta t \bar{C}_{12} \right\} + O(\phi^2),
\]

\[
\left\{ \left( 1 + \frac{\Delta t}{\Delta x} C_{11}\phi + 2\Delta t \bar{C}_{12} \right)^2 + \left( \frac{\Delta t}{\Delta x} \bar{C}_{11}\phi + 2\Delta t \bar{C}_{12} \right)^2 \right\} + O(\phi^2).
\]

Since \( Q_n^{n+1} = e^{(b_1+i\alpha_2)\eta}e^{i\kappa x_j} = e^{b_1\mu^n}\phi^{i\kappa x_j} = e^{b_1\mu^n}e^{i\kappa x_j} \), and the conjugate of \( Q_n \) is given by \( \overline{Q_n^{n+1}} = e^{b_1\mu^n}e^{-i\kappa x_j} = |Q_n^{n+1}|e^{-i\kappa x_j} \). This fact together with equation (61) yield
\[
\frac{Q_n^{n+1}}{|(A_n^{n+1})^2|} = |Q_n| \left( \mu^n \right)^{-\frac{1}{2}} \left\{ 1 + \left( \frac{\Delta t}{\Delta x} \right)^2 (\mu^n)^2 \phi^2 \right\} + O(\phi^2).
\]
Lemma 3.4. The proof is obvious according to Lemmas 3, 2 and 3.3. Armed with Lemmas 3, 2, and 3.3, we are ready to describe the amplification factor of the numerical scheme \([10]\).

Lemma 3.4. The amplification factor of the numerical scheme \([10]\) is approximated by

\[
e^{ih\Delta t} = 1 + \Delta t \left( \frac{1}{2} C_{12} + C_{33} - \Gamma_0 |Q_j^n|^{-\frac{1}{2}} (\mu^n)^\frac{1}{2} \left[ 1 + \left( \frac{\Delta t}{\Delta x} \right)^2 (\mu^n) \frac{1}{2} \right] \right) + O(\phi^2) + \left( 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11}\phi \right)^2 \left[ C_{31} + \frac{\Delta t}{\Delta x} C_{32}\phi \right] \left[ (C_{31} + \frac{\Delta t}{\Delta x} C_{32}\phi)^2 + (C_{31} - \frac{\Delta t}{\Delta x} C_{32}\phi)^2 \right] \cos \left( \theta_1 + \frac{7}{3} \theta_2 \right) - \frac{\Delta t}{\Delta x} \left[ (C_{21} - \frac{1}{2} C_{11})\phi + \frac{1}{2} \mu^n \left[ H_1 - K_1^2 + K_2^2 + 2 K_1 K_2 \left( 1 - \frac{\Delta t}{\Delta x} \mu^n \right) \right] \right] + i \left[ \left( \frac{1}{2} C_{12} - C_{33} - \Gamma_0 |Q_j^n|^{-\frac{1}{2}} (\mu^n)^\frac{1}{2} \left[ 1 + \left( \frac{\Delta t}{\Delta x} \right)^2 (\mu^n) \right] \right) \left[ (1 + 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11}\phi)^2 \right] + \left( 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11}\phi \right)^2 \left[ (C_{31} + \frac{\Delta t}{\Delta x} C_{32}\phi)^2 + (C_{31} - \frac{\Delta t}{\Delta x} C_{32}\phi)^2 \right] \sin \left( \theta_1 + \frac{7}{3} \theta_2 \right) \right) - \frac{\Delta t}{\Delta x} \left[ (C_{21} - \frac{1}{2} C_{11})\phi + \frac{1}{2} \mu^n \left[ H_2 - 2 K_1 K_2 + (K_2^2 - K_1^2) \left( 1 - \frac{\Delta t}{\Delta x} \mu^n \right) \right] \right] \right) \right) + O(\phi^2). \tag{65}
\]

where

\[
C_{33} = \frac{P}{\rho} |Q_j^n|^{-1} \cos \alpha_3, \quad C_{33} = \frac{P}{\rho} |Q_j^n|^{-1} \sin \alpha_3,
\]

\(H_1, H_2, K_1^2, K_2^2\) and \(K_1 K_2\) are given by equations \([12], [13], [14], [15] and [16]\), respectively; \(C_{ij}, C_{ij}\), \(j, i = 1, 2\); come from \([9]\); \(\alpha_2\) and \(\alpha_3\) follow from equation \([10]\) and \(C_{3i}, C_{3i}, i = 1, 2\), are defined by relation \([69]\). The functions \(\theta_1\) and \(\theta_2\) are given implicitly by relations

\[
e^{ih_1} = \frac{1 + 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11}\phi + i \left( 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11}\phi \right)}{\sqrt{\left( 1 + 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11}\phi \right)^2 + \left( 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11}\phi \right)^2}}
\]

and

\[
e^{ih_2} = \frac{C_{31} + \frac{\Delta t}{\Delta x} C_{32}\phi + i \left( C_{31} + \frac{\Delta t}{\Delta x} C_{32}\phi \right)}{\sqrt{(C_{31} + \frac{\Delta t}{\Delta x} C_{32}\phi)^2 + (C_{31} - \frac{\Delta t}{\Delta x} C_{32}\phi)^2}}
\]

Proof. The proof is obvious according to Lemmas 3.1, 3.2 and 3.3. Now, let us prove Proposition 3.2
Proof. (of Proposition 5.2)
Considering relation (64) and applying Remark 5.1 the squared modulus of the amplification factor of the numerical scheme (10) is approximated as

\[ |e^{\Delta t}|^2 = \left\{ 1 + \Delta t \left[ \frac{1}{2} C_{12} + C_{33} - \Gamma_0 |Q_j^n|^{-\frac{2}{3}} (\mu^n)^{\frac{2}{3}} \left[ 1 + \left( \frac{\Delta t}{\Delta x} \right)^2 (\mu^n)^2 \phi \right]^{\frac{2}{3}} \left[ 1 + 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi \right]^2 + \left( 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi \right)^2 \right] \left[ \left( C_{31} + \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 + \left( \overline{C}_{31} - \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 \right]^{\frac{2}{3}} \cos \left( \theta_1 + \frac{7}{3} \theta_2 \right) \right\} - \frac{\Delta t}{\Delta x} \left\{ (C_{21} - \frac{1}{2} C_{11}) \phi + \frac{1}{2} \mu^n \left[ H_1 - K_1^2 + K_2^2 + 2 K_1 K_2 \left( 1 - \frac{\Delta t}{\Delta x} \mu^n \right) \phi \right] \right\} \right\}^2 + \frac{\Delta t}{\Delta x} \left\{ \left( C_{21} - \frac{1}{2} C_{11} \phi + \frac{1}{2} \mu^n \left[ H_2 - 2 K_1 K_2 \left( K_2^2 - K_1^2 \right) \left( 1 - \frac{\Delta t}{\Delta x} \mu^n \right) \phi \right] \right\} \right\}^2 + O(\phi^4). \] (67)

Of course, the aim of this paper is to give the general picture of a necessary stability condition. Since we are working under the assumptions 0 < \Delta t < \Delta x < 1, 0 < |k| \Delta x = |\phi| << \pi, and the notations can become quite heavy, for the sake of readability we neglect the O(\phi^4). Furthermore, applying Remark 5.1 \[ 1 + (\frac{\Delta t}{\Delta x})^2 (\mu^n)^2 \phi^2 \leq 1 + O(\phi^2), \] neglecting the term O(\phi^4), \[ 1 + (\frac{\Delta t}{\Delta x})^2 (\mu^n)^2 \phi^2 \leq 1 \] is approximated by 1. However, the tracking of the infinitesimal terms does not compromise the result. Using this, estimate \[ |e^{\Delta t}| \leq 1, \] means that there exists a nonnegative number \( r^* \) between 0 and 1 such that

\[ \left\{ 1 + \Delta t \left[ \frac{1}{2} C_{12} + C_{33} - \Gamma_0 |Q_j^n|^{-\frac{2}{3}} (\mu^n)^{\frac{2}{3}} \left[ 1 + \left( \frac{\Delta t}{\Delta x} \right)^2 (\mu^n)^2 \phi \right]^{\frac{2}{3}} \left[ 1 + 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi \right]^2 + \left( 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi \right)^2 \right] \left[ \left( C_{31} + \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 + \left( \overline{C}_{31} - \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 \right]^{\frac{2}{3}} \cos \left( \theta_1 + \frac{7}{3} \theta_2 \right) \right\} - \frac{\Delta t}{\Delta x} \left\{ (C_{21} - \frac{1}{2} C_{11}) \phi + \frac{1}{2} \mu^n \left[ H_1 - K_1^2 + K_2^2 + 2 K_1 K_2 \left( 1 - \frac{\Delta t}{\Delta x} \mu^n \right) \phi \right] \right\} \right\} \leq \sqrt{1-r^*} \] (68)

and

\[ \left\{ 1 + \Delta t \left[ \frac{1}{2} C_{12} + C_{33} - \Gamma_0 |Q_j^n|^{-\frac{2}{3}} (\mu^n)^{\frac{2}{3}} \left[ 1 + \left( \frac{\Delta t}{\Delta x} \right)^2 (\mu^n)^2 \phi \right]^{\frac{2}{3}} \left[ 1 + 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi \right]^2 + \left( 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi \right)^2 \right] \left[ \left( C_{31} + \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 + \left( \overline{C}_{31} - \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 \right]^{\frac{2}{3}} \sin \left( \theta_1 + \frac{7}{3} \theta_2 \right) \right\} - \frac{\Delta t}{\Delta x} \left\{ (C_{21} - \frac{1}{2} C_{11}) \phi + \frac{1}{2} \mu^n \left[ H_2 - 2 K_1 K_2 \left( K_2^2 - K_1^2 \right) \left( 1 - \frac{\Delta t}{\Delta x} \mu^n \right) \phi \right] \right\} \right\} \leq \sqrt{r^*}, \] (69)

which are equivalent to

\[ 1 - \sqrt{1-r^*} \leq -\Delta t \left[ \frac{1}{2} C_{12} + C_{33} - \Gamma_0 |Q_j^n|^{-\frac{2}{3}} (\mu^n)^{\frac{2}{3}} \left[ 1 + \left( \frac{\Delta t}{\Delta x} \right)^2 (\mu^n)^2 \phi \right]^{\frac{2}{3}} \left[ 1 + 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi \right]^2 + \left( 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi \right)^2 \right] \left[ \left( C_{31} + \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 + \left( \overline{C}_{31} - \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 \right]^{\frac{2}{3}} \cos \left( \theta_1 + \frac{7}{3} \theta_2 \right) \right\} - \frac{\Delta t}{\Delta x} \left\{ (C_{21} - \frac{1}{2} C_{11}) \phi + \frac{1}{2} \mu^n \left[ H_1 - K_1^2 + K_2^2 + 2 K_1 K_2 \left( 1 - \frac{\Delta t}{\Delta x} \mu^n \right) \phi \right] \right\} \right\} \leq \sqrt{r^*}, \]
\[
\left[ \left( C_{31} + \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 + \left( C_{31} - \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 \right]^\frac{2}{3} \cos \left( \theta_1 + \frac{7}{3} \theta_2 \right) \bigg] + \frac{\Delta t}{\Delta x} \left\{ (C_{21} - \frac{1}{2} C_{11}) \phi + \frac{1}{2} \mu^n \left[ H_1 - K_2^2 + K_3^2 + 2K_1K_2 \left( 1 - \frac{\Delta t}{\Delta x} \mu^n \right) \phi \right] \right\} \leq 1 + \sqrt{1 - r^2},
\]
(70)
and
\[
\left| \frac{\Delta t}{\Delta x} \left\{ \frac{1}{2} C_{12} - C_{33} - \Gamma_0 |Q_j^\mu|^{-\frac{4}{3}} (\mu^n)^{\frac{2}{3}} \left[ (1 + 2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi)^2 + (2\Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi)^2 \right] \right. \right.
\left. \left. \left[ \left( C_{31} + \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 + \left( C_{31} - \frac{\Delta t}{\Delta x} C_{32} \phi \right)^2 \right]^\frac{2}{3} \sin \left( \theta_1 + \frac{7}{3} \theta_2 \right) \right\} - \frac{\Delta t}{\Delta x} \left\{ (C_{21} - \frac{1}{2} C_{11}) \phi + \frac{1}{2} \mu^n \left[ H_2 - 2K_1K_2 + (K_2^2 - K_3^2) \left( 1 - \frac{\Delta t}{\Delta x} \mu^n \right) \phi \right] \right\} \right| \leq \sqrt{r^2}.
\]
(71)
Since we are interested in a linear stability condition, we should find a restriction satisfies by \( \Delta t \) and \( \Delta x \) for which inequalities (70) and (71) hold. First, we must bound each term in estimates (70) and (71). Using relations (31), (39), (69), (66) and (47), we get
\[
\gamma_4 = (\mu^n)^2; \quad |C_{11}|, |C_{11}| \leq \frac{g}{2} |A_j^\mu| |\phi||^2 + \mu^n := R_n; \quad |C_{12}|, |C_{12}| \leq \frac{P_\eta}{\rho} |Q_j^\mu|^{-1} + \Gamma_0 |A_j^\mu|^{-\frac{4}{3}} := P_n;
\]
\[
|C_{31}|, |C_{31}| \leq 1; \quad |C_{32}|, |C_{32}| \leq \mu^n; \quad |C_{33}|, |C_{33}| \leq \frac{P_\eta}{\rho} |Q_j^\mu|^{-1} \leq P_n;
\]
\[
|C_{21}|, |C_{21}| \leq \frac{g}{2} |Q_j^\mu| |\phi||^2 := N_n; \quad |C_{22}|, |C_{22}| \leq \frac{4}{3} \Gamma_0 |A_j^\mu|^{-\frac{4}{3}} \leq 2P_n.
\]
(72)
From this and by simple calculations, it comes from equations (12), (13), (44), (45) and (46) that the quantities \( H_1, H_2, K_1^2, K_2^2 \) and \( K_1K_2 \) can be bounded as
\[
H_1 \leq 1 + 2\left[ 1 + (2 + R_n^{-1} \mu^n) |\phi|| \right] \max \left\{ P_n, P_nR_n, P_n^2, P_n^2\mu^n \right\} \left[ \Delta t + \Delta t^2 + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^3}{\Delta x} \right];
\]
(73)
\[
H_2 \leq 2 \max \left\{ P_n, P_nR_n, P_n^2, P_n(R_n + \mu^n) |\phi||, \mu^n P_n^2 |\phi|| \right\} \left[ \Delta t + \Delta t^2 + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^3}{\Delta x} \right];
\]
(74)
\[
K_1^2 \leq 1 + (1 + 4 |\phi||) \max \left\{ P_n, N_n, P_nN_n, P_n^2, N_n^2 \right\} \left[ \Delta t + \Delta t^2 + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^3}{\Delta x} + \left( \frac{\Delta t}{\Delta x} \right)^2 \right];
\]
(75)
\[
K_2^2 \leq 1 + (1 + 4R_n^{-1} + R_n^{-1}(\mu^n)^{-1}) |\phi|| \max \left\{ P_n, P_nR_n, P_n^2, R_n^2\mu^n, R_n^2(\mu^n)^2 \right\} \left[ \Delta t + \Delta t^2 + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^3}{\Delta x} + \left( \frac{\Delta t}{\Delta x} \right)^2 \right];
\]
(76)
\[
|K_1K_2| \leq |\phi|| + 1 + (4 + N_nR_n^{-1} + (4 + R_n^{-1} + 2P_n + R_n^{-1}\mu^n + 4N_nR_n^{-1}\mu^n)^{-1}) |\phi|| \max \left\{ P_n, P_nR_n, P_n^2, R_n, N_n, N_n^2 \right\} \left[ \Delta t + \Delta t^2 + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^3}{\Delta x} + \left( \frac{\Delta t}{\Delta x} \right)^2 \right];
\]
(77)
A combination of inequalities (73), (74), (75), (76) and (77), results in
\[
|H_1| + K_1^2 + K_2^2 + 2|K_1K_2| - \frac{\Delta t}{\Delta x} \mu^n |\phi|| \leq 2 + 8 \left[ \Delta t + \Delta t^2 + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^3}{\Delta x} + \left( \frac{\Delta t}{\Delta x} \right)^2 \right]
\leq 1 + (4 + R_n(\mu^n)^{-1} + R_n^{-1} + N_nR_n(\mu^n)^{-1}) \left( 1 + \frac{\Delta t}{\Delta x} \mu^n \right) |\phi||
\]
(77)
\[
\max \{ P_n, N_n, R_n P_n, N_n P_n, R_n \mu^n, P_n^2, N_n^2, P_n^2 \mu^n, R_n^2, (\mu^n)^2, R_n P_n \mu^n, R_n N_n \mu^n \} ; \tag{78}
\]
\[
|H_2| + 2|K_1 K_2| + (K_1^2 + K_2^2) (1 - \frac{\Delta t}{\Delta x} \mu^n) |\phi| \leq \left( 3 + \frac{\Delta t}{\Delta x} \mu^n \right) |\phi| + 6 \left[ \Delta t + \Delta t^2 + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^3}{\Delta x} \right] \leq 3 \left( 1 + 3 \frac{\Delta t}{\Delta x} \mu^n |\phi| \right). \tag{79}
\]

But
\[
\Gamma_0 |Q^n|^{-\frac{1}{2}} (\mu^n)^{\frac{1}{2}} = \Gamma_0 |Q^n| A^n|^{-\frac{1}{2}} \leq P_n ; \tag{80}
\]
applying Remark 33 and neglecting the terms of high order in \(\phi\), it is easy to see that
\[
\left( 1 + 2 \Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi \right)^2 + \left( 2 \Delta t C_{12} + \frac{\Delta t}{\Delta x} C_{11} \phi \right)^2 \leq 1 + 4 \Delta t P_n + 8 \Delta t^2 P_n^2 + 2 \frac{\Delta t}{\Delta x} R_n |\phi| + 8 \frac{\Delta t}{\Delta x} P_n R_n |\phi| ; \tag{81}
\]
\[
\left[ (C_3 + \frac{\Delta t}{\Delta x} C_{32} \phi)^2 + (C_3 - \frac{\Delta t}{\Delta x} C_{32} \phi)^2 \right]^{\frac{1}{2}} \leq 2 \left( 1 + \frac{\Delta t}{\Delta x} \mu^n |\phi| \right)^2 \leq 2 \sqrt{2} \left( 1 + \frac{\Delta t}{\Delta x} \mu^n |\phi| \right) \leq 3 \left( 1 + 3 \frac{\Delta t}{\Delta x} \mu^n |\phi| \right). \tag{82}
\]

We recall that the aim of this work is to find a linear (or necessary) stability condition of the numerical scheme (10). Plugging estimates (78)-(82), by straightforward computations, a necessary condition to obtain inequalities (10) and (11), is given by
\[
3P_n \Delta t \left\{ \frac{1}{2} + \left[ 1 + 4 \left( \Delta t + \Delta t^2 + \frac{\Delta t}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \left( \frac{\Delta t}{\Delta x} \right)^2 \right) \right] \max \{ P_n, 2P_n^2, \frac{1}{2} R_n |\phi|, \mu^n |\phi|, P_n \mu^n |\phi|, \right\} \leq 1 + \sqrt{1 - r^2} ; \tag{83}
\]
and
\[
3P_n \Delta t \left\{ \frac{1}{2} + \left[ 1 + 4 \left( \Delta t + \Delta t^2 + \frac{\Delta t}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \left( \frac{\Delta t}{\Delta x} \right)^2 \right) \right] \max \{ P_n, 2P_n^2, \frac{1}{2} R_n |\phi|, \mu^n |\phi|, P_n \mu^n |\phi|, \right\} \leq \sqrt{1 - r^2} . \tag{84}
\]

Now, setting
\[
W_1(\Delta t, \Delta x) = \frac{1}{2} + \left[ 1 + 4 \left( \Delta t + \Delta t^2 + \frac{\Delta t}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \left( \frac{\Delta t}{\Delta x} \right)^2 \right) \right] \max \{ P_n, 2P_n^2, \frac{1}{2} R_n |\phi|, \mu^n |\phi|, \right\} ; \tag{85}
\]
\[
P_n \mu^n |\phi|, 6P_n \mu^n |\phi|, 2R_n P_n |\phi| ; \tag{85}
\]
18
This fact, together with inequalities (83) and (84) provide

\[ W_2(\Delta t, \Delta x) = (P_n + \frac{1}{2} R_n)\phi + \mu^n \left\{ 1 + 4 \left[ \frac{\Delta t}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^3}{\Delta x} + \left( \frac{\Delta t}{\Delta x} \right)^2 \right] \right\} \]

and

\[ W_3(\Delta t, \Delta x) = (P_n + \frac{1}{2} R_n)\phi + \frac{3}{2} \mu^n \left\{ \phi \right\} \left\{ 1 + \frac{\Delta t}{\Delta x} + \frac{\Delta t^2}{\Delta x} + \frac{\Delta t^3}{\Delta x} + \left( \frac{\Delta t}{\Delta x} \right)^2 \right\} \]

\[ \max \left\{ P_n, N_n, R_n R_n, P_n R_n, P_n R_n, N_n R_n, R_n R_n, R_n R_n, R_n R_n, R_n R_n \right\} \]

This fact, together with inequalities (83) and (84) provide

\[ 3P_n \Delta t W_1(\Delta t, \Delta x) + \frac{\Delta t}{\Delta x} W_2(\Delta t, \Delta x) \leq 1 + \sqrt{1-r^*} \] and \[ 3P_n \Delta t W_1(\Delta t, \Delta x) + \frac{\Delta t}{\Delta x} W_3(\Delta t, \Delta x) \leq \sqrt{r^*}, \]

which are equivalent to

\[ \Delta t \left( 3P_n W_1(\Delta t, \Delta x) + \frac{1}{\Delta x} \max \left\{ W_2(\Delta t, \Delta x); W_3(\Delta t, \Delta x) \right\} \right) \leq \max \left\{ 1 + \sqrt{1-r^*}; \sqrt{r^*} \right\}, \]

where \( r^* \in (0; 1) \), \( W_1(\Delta t, \Delta x) \), \( W_2(\Delta t, \Delta x) \), and \( W_3(\Delta t, \Delta x) \), are given by relations (85), (86) and (87), respectively. Estimate (88) comes from the inequality: \( \max\{a + dx, a + dy\} = a + d\max\{x, y\} \), whenever the numbers \( a, d, x \) and \( y \) are nonnegative. Furthermore, Estimate (88) represents a necessary condition of stability for the numerical scheme (10). This ends the proof of Proposition 3.2 \( \Box \)

Using the above results (namely, Propositions 3.1 and 3.2) we are ready to give a necessary stability constraint of the MacCormack method (7)-(10) and to compare it with what is available in the literature (for example, Courant-Friedrich-Lewy condition for linear hyperbolic partial differential equations).

**Theorem 3.1.** The MacCormack scheme for 1D complete shallow water equations with source terms (1) is stable if

\[ \frac{\Delta t^4}{\Delta x^2} \left( 3P_n W_1(\Delta t, \Delta x) + \frac{1}{\Delta x} \max \left\{ W_2(\Delta t, \Delta x); W_3(\Delta t, \Delta x) \right\} \right) \left( 1 + \frac{2\Delta t}{3} \Gamma_0 \mu^n |A^n|^{-\frac{3}{2}} \right) \leq \]

\[ 3 \max \left\{ 1 + \sqrt{1-r^*}; \sqrt{r^*} \right\} \Gamma_0^{-1} (\mu^n)^{-3} |A^n|^\frac{3}{2} |\phi|^2, \]

with the requirement: \( |\phi| = |k\Delta x| << \pi \). In relations (89): \( e^{\alpha t} = |A|, e^{\beta t} = |Q|, \mu = \left| \frac{Q}{P} \right|, r^* \in (0; 1), \Gamma_0 = \frac{an^2}{4\nu^2} P^\frac{4}{\nu}, W_1(\Delta t, \Delta x), W_2(\Delta t, \Delta x) \) and \( W_3(\Delta t, \Delta x) \), are given by relations (85), (86) and (87), respectively.

**Proof.** The proof of Theorem 3.1 is obvious according to Propositions 3.1 and 3.2 \( \Box \)

The Von Neumann stability approach, based on a Fourier analysis in the space domain has been developed for nonlinear one-dimensional complete shallow water equations with source terms. Although the stability condition has not be derived analytically, we have analyzed the properties of amplification factor numerically (by use of Taylor series expansion), which contain information on the dispersion and diffusion errors of the considered numerical scheme. It is worth noticing that we used a local, linearized stability analysis to obtain estimate (89), which must be considered as a necessary condition of stability for the numerical scheme (1)-(10).
Some important remarks on stability analysis

This section considers some useful remarks on the stability restrictions obtained in this note and compares it with what is known in the literature (Courant-Friedrichs-Lewy condition).

1. The stability restrictions (89) suggests that a small space step $\Delta x$ forces the time step $\Delta t$ to be more potentially small. This makes the MacCormack scheme extremely slow. However, because consistency requires that $\frac{\Delta t}{\Delta x^n}$ $(n \geq 1)$ approached zero as $\Delta t$ and $\Delta x$ approach zero, a much smaller time step than allowed by the stability condition (89) is implied. For this reason, the MacCormack method seems better suitable for the calculation of steady solutions (where time accuracy is unimportant) than for the unsteady solutions (for example, see [27] for analysis of numerical solutions for time dependent PDEs).

2. The MacCormack approach (7)-(10) for 1D complete surface water equations has a stability limitation (89) that limits the maximum time step. This stability requirement does not coincide with the Courant-Friedrichs-Lewy (CFL) condition obtained for linear hyperbolic partial differential equations (for example: linear advection equation, wave equation, linearized burgers equations, etc...) because the MacCormack method is applied to complex time dependent partial differential equations. As discussion on the stability restrictions one can refer to the stability analysis of the two-step Lax-Wendroff method and the MacCormack scheme applied to complete burgers equations (for example, see [1], P. 245-247). The linear stability condition (89) is highly unusual. Since we normally find this condition from a Fourier stability analysis, it follows from inequalities (89) that an instability occurs when $|\Delta t|$ is greater than some $|\Delta t|_{\text{max}}$ which can be viewed as $(\Delta t)_{CFL}$. As observed in proving Proposition 3, it was extremely difficult to obtain the stability criterion for our numerical method. However, it comes from condition given by relation (89) that the empirical formula

$$\frac{\Delta t^4}{\Delta x^2} \left( 3|A^n| \frac{\phi}{\Delta t} + 2\Delta t \Gamma_0 \mu^n \right) \leq 9 \max \left\{ 1 + \sqrt{1 - r^2}; \sqrt{r^2} \right\} \Gamma_0^{-1}(\mu^n)^{-3} |A^n| \frac{\phi}{\Delta t} |\phi|^{-2},$$

can be used with an appropriate safety factor. It should be remembered that the "heuristic" stability analysis, i.e., estimates (89) can only provide a necessary condition for stability. Thus, for some finite difference algorithms, only partial information about the complete stability bound is obtained and for others (such as algorithms for the heat equation, wave equation and linearized Burgers equations) a more complete theory must be employed.

4 Numerical experiments

This section simulates the MacCormack scheme described in section 3 for 1D complete shallow water equations with source terms. We focus on a practical application of a shallow water flow based on the Benoué river. This river is a 7000 m long reach of the upstream part (altitude=174.22 m) and it is located in Cameroon. Being a mountain river, it is characterized by strong irregularities in the cross section, by a rather steep part in the first kilometers and by a low base discharge (708 m³/s) which, altogether, produce a high velocity basic flow, transcritical in some parts. More specifically, we consider the problem of floods observed in this river in 2012 because it is a classical example of time dependent nonlinear flow with shocks to expect floods and to test conservation in numerical schemes. Furthermore, we assume that this problem is generated by the 1D complete shallow water equations with source terms for the ideal case of a flat and frictionless channel with prismatic cross section, i.e., constants top width ($T = 348$ m) and wetted perimeter ($P = 366, 4$ m). We also use the initial data given by relation (93).

Before describing the analytical solution considered in this note, we first approximate the $L^2$-norm of the space $L^2(0,T_1; L^2(0,L))$ by a full discrete norm which plays a crucial role in the analysis of the error
estimates together with the convergence rate of our method. Let \( w \in L^2(0,T_1;L^2(0,L)) \), we have that

\[
\left[ \int_0^{T_1} \int_0^L |w(t,x)|^2 \, dx \, dt \right]^{1/2} \approx \left[ \sum_{n=0}^{N} \int_{t^n}^{t^{n+1}} \int_0^L |w(t,x)|^2 \, dx \, dt \right]^{1/2}
\]

\[
\approx \left[ \Delta t \sum_{n=0}^{N} \int_0^L |w(t^n,x)|^2 \, dx \right]^{1/2}
\]

\[
\approx \left[ \Delta t \cdot \Delta x \sum_{n=0}^{N} \sum_{j=0}^{M} |w(t^n,x_j)|^2 \right]^{1/2}
\]

Using this, we introduce the following fully discrete norm

\[
\|w\|_{L^2(0,T_1;L^2(0,L))} = \left[ \Delta t \cdot \Delta x \sum_{n=0}^{N} \sum_{j=0}^{M} |w(t^n,x_j)|^2 \right]^{1/2}
\]

Denoting by \( w^n_j = (A^n_j, Q^n_j) \) the value of the approximate solution at time \( t^n \) and point \( x_j \) obtained with the MacCormack scheme and by \( w(t^n,x_j) = (A(t^n,x_j), Q(t^n,x_j)) \) the value of the analytical solution at \( (t^n,x_j) \), the exact error at time \( t^n \) and point \( x_j \) is defined by \( e(t^n,x_j) = w(t^n,x_j) - w^n_j = (A(t^n,x_j) - A^n_j, Q(t^n,x_j) - Q^n_j) \). Thus the errors-norm are given by

\[
\|e_A\|_{L^2(0,T_1;L^2(0,L))} = \left[ \Delta t \cdot \Delta x \sum_{n=0}^{N} \sum_{j=0}^{M} |A(t^n,x_j) - A^n_j|^2 \right]^{1/2}
\]

(90)

and

\[
\|e_Q\|_{L^2(0,T_1;L^2(0,L))} = \left[ \Delta t \cdot \Delta x \sum_{n=0}^{N} \sum_{j=0}^{M} |Q(t^n,x_j) - Q^n_j|^2 \right]^{1/2}
\]

(91)

The exact solution considered in this paper is due to Dressler’s dam break with friction [7]. In the literature different approaches are presented and deeply studied for this case. Dressler’s analyzed Chézy friction law and has used a perturbation scheme in the Ritter’s method, i.e., both velocity \( (u) \) and height \( (h) \) of the water are expanded as power series in the friction coefficient \( C_f = 1/C^2 \). We consider the initial conditions defined as

\[
h(0,x) = h^0(x) = \begin{cases} 
    h_1 > 0, & \text{for } 0 \leq x \leq x_0; \\
    0, & \text{for } x_0 < x < L,
\end{cases}
\]

\[
u(0,x) = u^0(x) = \begin{cases} 
    10^{-1}, & \text{for } 0 \leq x \leq x_0; \\
    0, & \text{for } x_0 < x < L.
\end{cases}
\]

(93)

We assume that \( C = 40m^{1/2}/s \) (Chézy coefficient), \( h_1 = 5 \times 10^{-3}m, x_0 = L/2, T_1 = 1s, \) and \( L = 1m \). Dressler’s first-order developments for the flow resistance give the following corrected height and velocity

\[
h_c(t,x) = \frac{1}{9} \left( \frac{2}{3} \sqrt{gh_1} \left( -\frac{x-x_0}{3} + \frac{4}{3} \alpha_1 t \right) \right)^2,
\]

\[
u_c(t,x) = \frac{2}{3} \sqrt{gh_1} + \frac{2(x-x_0)}{3t} + \frac{2}{3} \alpha_2 t,
\]

(94)

where

\[
\alpha_1 = \frac{6}{5} \left( \frac{2-x}{t\sqrt{gh_1}} \right) - \frac{2}{3} + \frac{4\sqrt{3}}{135} \left( 2 - \frac{x-x_0}{t\sqrt{gh_1}} \right)^{3/2},
\]

(95)
where

$$\alpha_2 = \frac{12}{2 - \frac{x-x_0}{t^\sqrt{gh_t}}} - \frac{8}{3} + \frac{8\sqrt{3}}{180} \left( \frac{2 - \frac{x-x_0}{t^\sqrt{gh_t}}}{\sqrt{3}} \right)^{3/2} - \frac{108}{7} \left( \frac{2 - \frac{x-x_0}{t^\sqrt{gh_t}}}{\sqrt{3}} \right)^2. \quad (96)$$

Following the Dressler’s approach we consider four regions: from upstream to downstream (a steady state region \((h_1, 10^{-1})\) for \(x \leq x_1(t)\)); a corrected region \(([h_c, u_c])\) for \(x_1(t) \leq x \leq x_2(t)\); the tip region (for \(x_2(t) \leq x \leq x_3(t)\)) and the dry region \((0, 0)\) for \(x_3(t) \leq x \leq L\). In the tip region, friction term is preponderant thus \((94)\) is no more valid. In the corrected region, the velocity increases with \(x\). Dressler assumed that at \(x_2(t)\) the velocity reaches the maximum of \(u_c\) and that the velocity is constant in space in the tip region

$$u_{tip}(t) = \max_{x \in [x_2(t), x_3(t)]} u_c(t, x). \quad (97)$$

Utilizing these assumptions together with relations \((94)-(97)\), the analytic solution of 1D complete shallow water equations with friction terms is then given by

$$h(t, x) = \begin{cases} h_1, & \text{for } 0 \leq x \leq x_1(t) \text{ and } t \in (0, T_1], \\ \frac{1}{g} \left( \frac{4}{3} \sqrt{gh_1} - \frac{x-x_0}{3t} + \frac{2^2}{3^2} \alpha_1 t \right)^2, & \text{for } x_1(t) \leq x \leq x_3(t) \text{ and } t \in (0, T_1], \\ 0, & \text{for } x_3(t) \leq x \leq L \text{ and } t \in (0, T_1], \end{cases} \quad (98)$$

and

$$u(t, x) = \begin{cases} 0, & \text{for } 0 \leq x \leq x_1(t) \text{ and } t \in (0, T_1], \\ \frac{2}{g} \sqrt{gh_1} + \frac{2(x-x_0)}{3t} + \frac{2^2}{3^2} \alpha_2 t, & \text{for } x_1(t) \leq x \leq x_2(t) \text{ and } t \in (0, T_1], \\ \max_{x \in [x_2(t), x_3(t)]} u_c(t, x), & \text{for } x_2(t) \leq x \leq x_3(t) \text{ and } t \in (0, T_1], \\ 0, & \text{for } x_3(t) \leq x \leq L \text{ and } t \in (0, T_1], \end{cases} \quad (99)$$

where \(\alpha_1\) and \(\alpha_2\) are given by equations \((95)\) and \((96)\), respectively, \(x_1(t) = x_0 - t \sqrt{gh_t}\), \(x_3(t) = x_0 + 2t \sqrt{gh_t}\) and \(x_2(t) \in [x_1(t), x_3(t)]\) is the point where the velocity \(u_c(t, x)\) attains its maximum.

With this approach, we should remark that the water height is constant in the tip zone. This is a limit of Dressler’s approach. Thus the authors \[1\] coded the second order interpolation used in \[30, 31\] and recommended by Valerio Caleffi. Even if the authors \[7\] have no information concerning the shape of the wave tip, this case shows if the scheme is able to locate and treat correctly the wet/dry transition.

Armed with the above tools, we are ready to provide the exact solution of the system of PDEs \[1\]. Our analysis consider the case where the channel is prismatic with a constant top width \((T)\) and the average velocity \((u)\) is defined as \(u(t, x) = Q(t, x)/A(t, x)\). Using this, the following formulas hold

$$A(t, x) = Th(t, x) \quad \text{and} \quad Q(t, x) = Th(t, x)u(t, x). \quad (100)$$

Since the water height is constant in the tip region it comes from relation \((100)\) that the cross section \((A)\) and discharge \((Q)\) are not modified in that region. A combination of relations \((98), \ (99)\) and \((100)\) gives the explicit formulae of cross section and discharge

$$A(t, x) = Th(t, x) = \begin{cases} Th_1, & \text{for } 0 \leq x \leq x_1(t) \text{ and } t \in (0, T_1], \\ \frac{T}{g} \left( \frac{2}{3} \sqrt{gh_1} - \frac{x-x_0}{3t} + \frac{2^2}{3^2} \alpha_1 t \right)^2, & \text{for } x_1(t) \leq x \leq x_3(t) \text{ and } t \in (0, T_1], \\ 0, & \text{for } x_3(t) \leq x \leq L \text{ and } t \in (0, T_1], \end{cases} \quad (101)$$
and

\[
Q(t, x) = Th(t, x)u(t, x) = \begin{cases} 
0, & \text{for } 0 \leq x \leq x_1(t) \text{ and } t \in (0, T_1], \\
Th_c(t, x)u_c(t, x), & \text{for } x_1(t) \leq x \leq x_2(t) \text{ and } t \in (0, T_1], \\
Th_c(t, x)u_{tip}(t, x), & \text{for } x_2(t) \leq x \leq x_3(t) \text{ and } t \in (0, T_1], \\
0, & \text{for } x_3(t) \leq x \leq L \text{ and } t \in (0, T_1].
\end{cases}
\]  

(102)

The following values are considered for simulations: shear stress \( \tau = 1.329 N/m^2 \); Top width \( T = 348 m \); wetter perimeter \( P = 366, 4m \); wavelength \( K_\lambda = 2\pi \approx 6.28 m \); manning’s number \( n_1 = 0.025 s/m^{1/3} \); the acceleration of gravity \( g = 10 m/s^2 \); the rainfall intensity defined by relation (103) – (104), \( I(t, x) \) is the rainfall intensity and \( r(t, x) \) is the constant rainfall excess.

\[
I(t, x) = \begin{cases} 
1.18 \times 10^{-5} m/s, & \text{if } (t, x) \in [0, T_1] \times [0, L]; \\
0, & \text{otherwise}.
\end{cases}
\]  

(103)

The mathematical model for this ideal overland flow is the following: we consider a uniform plane catchment whose overall length in the direction of flow is \( L = 1 m \). The surface roughness and shear stress are assumed invariant in space and time. It comes from equation (103) the constant rainfall excess is defined as

\[
r(t, x) = \begin{cases} 
I(t, x), & \text{for } (t, x) \in [t_0, T_1] \times [0, L]; \\
0, & \text{otherwise}.
\end{cases}
\]  

(104)

The mesh size \( \Delta x \) takes values: \( 2^{-4}, 2^{-5}, 2^{-6} \) and \( 2^{-7} \), while the time step \( \Delta t \) varies in range: \( 2^{-7}, 2^{-8}, 2^{-9} \) and \( 2^{-10} \). \( I \) is the rainfall intensity defined by relation (103). \( t_0 = 0 s \) and \( T_1 = 1 s \) are initial and final time, respectively, of the rainfall excess computed above, and \( L = 1 m \) is the rod interval length. The approximate solutions given by numerical schemes (7-10) obtained from 1 to 20 iterations, respectively, are displayed in Figures 1 and 2. Different values of \( k = \Delta t = 2^{-7}, 2^{-8}, 2^{-9}, 2^{-10} \), numbers obtained from the stability restriction (89) as the steady flow cases and space step \( h = \Delta x = 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7} \), in the mesh are used. Before 3 iterations are encountered, the discharge wave propagates with almost a perfectly constant value at different positions (see Figures 1 and 2). Furthermore, after 3 iterations, the discharge wave also tends to zero at different times (Figures 1 and 2). So, the graphs show that the solution of the difference equations cannot grow with time and so must still satisfy the Von Neumann necessary condition. We obtain similar observations for the cross section. In addition, Table 1 suggests that the errors associated with the cross section is of second order accurate while the approximate solution corresponding to the discharge coincides with the exact one. This suggests that for the considered analytical solution the MacCormack approach (7-10) converges with second order accuracy. Furthermore, the graphs indicate that the numerical solutions start to destroy after a fixed time. Specifically, combining the different values of \( \Delta x \) and \( \Delta t \), we observe from the graphs that that good approximate solutions are obtained for small mesh sizes \( \Delta x \) and time steps \( \Delta t \) satisfying the stability restriction (89). Thus, physical insight must be used when the stability limitation (89) of the MacCormack method is investigated. Finally, both Table 1 and Figures 1 and 2 show that the numerical solutions do not increase with time and converge to the analytical one. More specifically, they indicate that stability for the MacCormack scheme is subtle. It is not unconditionally unstable, but stability depends on the parameters \( \Delta x \) and \( \Delta t \).

Table 1. Analyzing of convergence rate \( O(h^\theta + \Delta t^\beta) \) for MacCormack scheme by \( r(\cdot) \), with varying spacing \( h = \Delta x \) and time step \( k = \Delta t \).

| \((k,h)\) | \(\|A-A_1\|_{L^2}\) | \(\|Q-Q_1\|_{L^2}\) | \(r(A)\) | \(r(Q)\) |
|------|------|------|------|------|
| \((2^{-9},2^{-4})\) | 0.0384 | 0 | 0 | 0 |
| \((2^{-8},2^{-5})\) | 0.0192 | 0 | 2.0000 | - |
| \((2^{-7},2^{-6})\) | 0.0093 | 0 | 2.0645 | - |
| \((2^{-6},2^{-7})\) | 0.0047 | 0 | 1.9787 | - |
5 General conclusions and future works

In this paper we have presented a full description of the MacCormack approach for complete shallow water equations with source terms, have studied in details the stability analysis of the numerical scheme and we have provided the convergence rate of the method (which is computed numerically). The graphs (Figures 1 and 2) show that the considered method is both stable and convergent while Table 1 indicates the rate of convergence (second order) of the algorithm. After a few number of iterations, the figures suggest that the numerical solutions strongly converge to the analytical one. This is not a surprise since the exact solutions are discontinuity and tend to zero whenever the time \( t \) is different from zero. From Figures 1 and 2 one should observe that the only case where the exact solutions are not close to zero corresponds to the initial condition. This comes from assumptions made by Dressler when he constructs the analytical solutions. Our future investigations will consist to find a good analytical solution for the proposed problem and to extend the analysis to open channel flows.

Acknowledgment. The authors thank the anonymous referees for detailed and valuable comments which have helped to greatly improve the quality of this paper.

References

[1] F. A. Anderson, R. H. Pletcher, J. C. Tannehill. "Computational fluid mechanics and Heat Transfer". Second Edition, Taylor and Francis, New York, 1997.
[2] A.J.C. Barré de Saint Venant. Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l’introduction des marées dans leur lit. Compte Rendu de l’Académie des Sciences, Paris, 73:147-154, (1871).
[3] D.R. Basco. Computation of rapidly varied unsteady free surface flow. US Geological Sur6ey, Water Resources Investigations Report 83-4284, (1987).
[4] F. Bianco, G. Puppo and G. Russo. High order central schemes for hyperbolic systems of conservation laws. SIAM J. Sci. Comput. Vol. 21, No. 1, PP. 294-322, (1999).
[5] P. Brufau, J. Burguete, P. García-Navarro, and J. Murillo. The shallow water equations: An example of hyperbolic system. Monografías de la Real Academia de Ciencias de Zaragoza 31, 89-119, (2008).
[6] G. Cannata, L. Lasaponara and F. Gallerano. Nonlinear shallow water equations numerical integration on curvilinear boundary-conforming grids. WSEAS Transactions on Fluids Mechanics, 10, P. 13-25, (2015).
[7] O. Delestre, C. Lucas, P.-A. Ksinant, F. Darboux, C. Laguerre, T.N.T. Vo, F. James, S. Cordier. SWASHES: a compilation of shallow water analytic solutions for hydraulic and environmental Studies. Int. J. Numer. Meth. Fluids, V. 72 (3), pp. 269-300, 2013).
[8] D. D. Franz, C. S. Melching. Full equations (FEQ) model for the solution of the full, dynamic equations of motion for one-dimensional unsteady flow in open channels and through control structures. U.S. Geological Survey Water-Resources Investigations report 96-4240, (1997).
[9] Fritz R. Fiedler, Jorge A. Ramirez. A numerical method for simulating discontinuous shallow flow over an infiltrating surface. Int. J. Numer. Meth. Fluids, 32 : 219-240, (2000).
[10] F. Gallerano, G. Cannata and L. Lasaponara. A new numerical model for simulations of wave transformation, breaking and long-shore, currents in complex coastal regions. International Journal of Numerical Methods in Fluids, 80, P. 571-613, (2016a).
[11] R. Garcia and R. A. Kahawaita. Numerical solution of the Saint-Venant equations with MacCormack finite-difference scheme. J. Numer. Meth. Fluids, 6(5), 259-274, (1986).
[12] R.Hixon, E. Turkel. Compact implicit MacCormack type schemes with high accuracy. J. Comp. Phys., 158 pp. 51-70, (2000).
[13] G. S. Jiang, D. Levy, C. T. Lin, S. Osher and E. Tadmor. High-Resolution Non-Oscillatory Central Schemes with Non-Staggered Grids for Hyperbolic Conservation Laws. SINUM 35 P. 2147-2168, (1998).

[14] G. S. Jiang and C. W. Shu. Efficient Implementation of Weighted ENO Schemes. JCP 126 P. 202-228, (1996).

[15] H. O. Kreiss. On difference approximations of the dissipative type for hyperbolic differential equations. Comm. Pure Appl. Math., Vol. 17, P. 335-353, MR 29#4210, (1964).

[16] A. Kurganov and G. Petrova. A second-order well-balanced positivity preserving central-upwind schemes for the Saint-Venant system. Comm. Math. Sc., Vol. 5, 133-160, (2007).

[17] A. Kurganov and D. Levy. Central-upwind schemes for the Saint-Venant system. M2AN, Vol. 36, No3, P. 397-429, (2002).

[18] P.D. Lax, B. Wendroff, Systems of conservation laws. J. Commun. Pure. Appl. Math, 13, pp. 217-237 (1960).

[19] D. Levy, G. Puppo, and G. Russo. Central WENO schemes for hyperbolic systems of conservation laws. M2AN, Vol. 33, No3, P. 547-571, (1999).

[20] G. Li, V. Caleffi and J. Gao. High-order well-balanced central WENO scheme for pre-balanced shallow water equations. Computers & Fluids, 99, P. 182-189, (2014).

[21] R.W. Maccormack. The effect of viscosity in hypervelocity impact cratering. AIAA, 69-354, (1969).

[22] F. T. Namio, E. Ngondiep, R. Ntchantcho, and J. C. Ntonga. "Mathematical models of complete shallow water equations with source terms, stability analysis of Lax-Wendroff scheme", J. Theor. Comput. Sci., Vol. 2(132), (2015).

[23] E. Ngondiep, "Stability analysis of MacCormack rapid solver method for evolutionary Stokes-Darcy problem", J. Comput. Appl. Math. 345(2019), 269-285, 17 pages.

[24] E. Ngondiep, "Long time stability and convergence rate of MacCormack rapid solver method for non-stationary Stokes-Darcy problem", Comput. Math. Appl., Vol 75, (2018), 3663-3684, 22 pages.

[25] E. Ngondiep, "Long time unconditional stability of a two-level hybrid method for nonstationary incompressible Navier-Stokes equations", J. Comput. Appl. Math. 345(2019), 501-514, 14 pages.

[26] E. Ngondiep, Error estimates and rate of convergence of MacCormack rapid solver method for nonstationary incompressible Navier-Stokes equations, submitted.

[27] R. Peyret and H. Viviand. Computation of viscous compressible flows based on the Navier-Stokes equations. AGARD-AG-212, (1975).

[28] Md. Saiduzzaman and Sobuj Kumar Ray, Comparison of numerical schemes for shallow water equation. Global Journal of Science Frontier Research Mathematics and Decision Sciences, Volume 13, Issue 4, Version 1.0, (2013).

[29] R. Sanders and A. Weiser, A High Resolution Staggered Mesh Approach for Nonlinear Hyperbolic Systems of Conservation Laws. JCP 1010, P. 314-329, (1992).

[30] A. Valiani, V. Caleffi and A. Zanni, Finite volume scheme for 2D shallow water equations. Application to the Malpasset dam-break. In Proceedings of the 4th Concerted Action on Dambreak Modelling Workshop, pp. 63-94, Zaragoza, Spain, November (1999).

[31] A. Valiani, V. Caleffi and A. Zanni, Case study: Malpasset dam-break simulation using a two-dimensional finite volume method. Journal of Hydraulic Engineering, 128(5), 460-472, (2002).

[32] Q. Zhou, J. Zhan and Y. Li, High-order finite-volume WENO schemes for Boussinesq modelling of nearshore wave processes. Journal of Hydraulic Research, 54(6), 646-662, (2016).
Graphs of cross section and discharge for shallow water flow.

Figure 1: Stability analysis and convergence rate of MacCormack for shallow water equations with source terms.
Figure 2: Stability analysis and convergence rate of MacCormack for shallow water equations with source terms.