# Jean Descloux; Robert Ferro Analysis of Thacker's method for solving the linearized shallow water equations

In: Jaromír Vosmanský and Miloš Zlámal (eds.): Equadiff 6, Proceedings of the International Conference on Differential Equations and Their Applications held in Brno, Czechoslovakia, Aug. 26 - 30, 1985. J. E. Purkyně University, Department of Mathematics, Brno, 1986. pp. [295]--302.

Persistent URL: http://dml.cz/dmlcz/700183

# Terms of use:

© Masaryk University, 1986

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://project.dml.cz

# ANALYSIS OF THACKER'S METHOD FOR SOLVING THE LINEARIZED SHALLOW WATER EQUATIONS

## J. DESCLOUX, R. FERRO

EPFL - Department of Mathematics CH 1015 Lausanne, Switzerland

## 1. INTRODUCTION.

In their simplest form, the shallow water equations read

$$\partial_t \vec{U}(x,t) = -b(x)\vec{v}H(x,t) + fR\vec{U}(x,t), x \in \Omega, t \ge 0,$$
 (1.1)

$$\partial_t H(\mathbf{x}, \mathbf{t}) = -\vec{\mathbf{v}} \cdot \vec{\mathbf{U}}(\mathbf{x}, \mathbf{t}), \quad \mathbf{x} \in \Omega, \quad \mathbf{t} \ge \mathbf{0},$$
 (1.2)

$$\vec{U}(x,t)\cdot\vec{n}(x) = 0, \quad x \in \partial\Omega, \quad t \ge 0, \quad (1.3)$$

$$\vec{U}(x,o) = \vec{U}_{0}(x), H(x,o) = H_{0}(x), x \in \Omega.$$
 (1.4)

Here ,  $\Omega \subset \mathbb{R}^2$  is a bounded open domain with  $\mathbb{C}^{\infty}$  boundary  $\partial\Omega$  and closure  $\overline{\Omega}$ .  $\overline{n}$  is the outwards unit normal to  $\partial\Omega$ .  $\overline{U}(x,t) = (\overline{U}_1(x,t), U_2(x,t))$  is a two components vector related to the average horizontal velocity. H(x,t) is the height of the surface of the basin. Up to a constant factor, b(x) is the depth of the basin. We shall assume that  $b(\cdot)$  is a  $\mathbb{C}^{\infty}(\overline{\Omega})$  strictly positive function. f which represents the intensity of the Coriolis forces is taken to be constant.  $\overline{U}_0$  and  $H_0$  are given initial conditions. R is the  $(-\Pi/2)$  rotation operator acting in  $\mathbb{R}^2$ , i.e.  $R(x_1, x_2) = (x_2, -x_1)$ . The tangential vector  $\vec{t}$  at  $\partial\Omega$  is given by  $\vec{t} = -R\vec{n}$ .

Equations (1.1),(1.2) can be easily set in the framework of the theory of semigroups. Let  $\mathscr{H} = (L_2(\Omega))^3$  be the Hilbert space with scalar product

$$((\vec{u},g),(\vec{v},h))_{\mathcal{H}} = \int_{\Omega} \frac{1}{b(x)} \vec{u}(x) \cdot \vec{v}(x) + g(x)h(x)$$
(1.5)

and associate norm  $||\cdot||_{\mathcal{H}}$  where  $\vec{u} = (u_1, u_2)$ ,  $\vec{v} = (v_1, v_2)$ . We define the operator L with domain  $\mathscr{D}(L)$  by the relations:

$$\mathscr{D}(\mathsf{L}) = \{(\vec{\mathfrak{u}}, \mathsf{g}) \in \mathscr{H} \mid \vec{\mathfrak{f}} \cdot \vec{\mathfrak{u}} \in \mathsf{L}^{2}(\Omega), \ \vec{\mathfrak{u}} \cdot \vec{\mathfrak{n}} = \mathsf{o} \text{ on } \Im\Omega, \ \mathsf{g} \in \mathsf{H}^{1}(\Omega)\}$$
(1.6)

$$L(\vec{u},g) = (-b\vec{\nabla}g, -\vec{\nabla}\cdot\vec{u}), \qquad (1.7)$$

where  $H^k(\Omega)$  is the classical Sobolev space of order k. One can verify that L is a skewadjoint operator, i.e. the adjoint of L is -L which implies in particular for  $(\vec{u},g)$  and  $(\vec{v},h)$  in  $\mathcal{D}(L)$  the following relation:

$$(L(\vec{u},g),(\vec{v},h))_{\mathcal{H}} = -((\vec{u},g),L(\vec{v},h))_{\mathcal{H}}.$$
(1.8)

It follows that L is the infinitesimal generator of a conservative group so that Problem (1.1)-(1.4) possesses a unique solution with the property:

$$\|(\vec{U}(\cdot,t),H(\cdot,t))\|_{\mathcal{H}}^2 = \text{constant}; \tag{1.9}$$

furthermore , from (1.2),(1.3), one deduces immediately the law of mass conservation

$$\int_{\Omega} H(x,t) = \text{constant.}$$
(1.10)

<u>Remarque 1.1</u>: In [4], we give some results concerning the regularity of solutions of Problems (1.1)-(1.4).

The purpose of this paper is to analyse a numerical method proposed by Thacker [1], [2] for solving Problems (1.1)-(1.4); more exactly, we shall consider in fact two variants of Thacker's scheme.

### 2. DISCRETIZATION.

We consider a sequence  $\{\mathscr{D}_h\}$  of standard triangulations of  $\Omega$ , as shown in the figure. h denotes the maximum length of the sides of the triangles of  $\mathscr{D}_h$ . We assume that all angles of all triangles of all triangularizations are bounded from below by some positive constant. For a particular  $\mathscr{D}_h$ , let N be equal to three times the total number of nodes minus the number of nodes belonging to the boundary  $\partial\Omega$ ,  $\Omega_h$  will denote the interior of the union of the triangles of  $\mathscr{D}_h$ .

Let us consider a fixed triangularization  $\mathscr{D}_h$ . For each node  $P_k$ , let  $\Lambda_k$  be the polygon formed by the triangles containing  $P_k$  and let  $\mu_k$  be the measure of  $\Lambda_k$ .  $Z_k$  will denote the set of indices j such that  $P_j \in \partial \Lambda_k$ ; in the figure,  $Z_2 = \{2,3,4,5\}$ ,  $Z_6 = \{7,8,9,10, 11\}$ ; clearly,  $k \in Z_k$  if and only if  $P_k \in \partial \Omega$ . For some node  $P_k$ , let  $P_j \in \partial \Lambda_k$ , i.e.  $j \in Z_k$ ; let  $P_\alpha \in \partial \Lambda_k$  be the node preceding  $P_j$  with respect to the trigonometric orientation and let  $P_B \in \partial \Lambda_k$  be the node following  $P_i$ ; we define the vector:

$$\tilde{q}_{jk} = P_{\alpha}^{\dagger} P_{\beta} ; \qquad (2.1)$$

in the figure, we have, for example,  $\vec{Q}_{8,6} = P_7^{-P}P_5$ ,  $\vec{Q}_{5,2} = P_4^{-P}P_2$ ,  $\vec{Q}_{22} = P_5^{-P}P_3$ . Furthermore, we introduce at  $P_k \in \partial \Omega$  the approximate tangent and normal vectors:

$$\vec{T}_{k} = \frac{1}{|\vec{Q}_{kk}|} \vec{Q}_{kk}$$
,  $N_{k} = \vec{R} T_{k}$ . (2.2)

For a function  $\phi$  defined on  $\Omega$ , let  $\phi_0$  be the continuous, piecewise linear (with respect to  $\mathcal{D}_h$ ) function defined on  $\Omega_h$  and equal to  $\phi$  at the nodes; for a vector function  $\vec{\psi}$ , we define, in the same way, componentwise its interpolant  $\vec{\psi}_0$ . By Greens's formula, we have the identities:

$$\frac{1}{\mu_{k}} \int_{\Lambda_{k}} \vec{\nabla}_{\phi_{0}} = \frac{1}{2\mu_{k}} \sum_{j \in \mathbb{Z}_{k}} \phi(P_{j})(R\vec{Q}_{jk}), \qquad (2.3)$$

$$\frac{1}{\mu_k} \int_{\Lambda_k} \vec{\nabla} \cdot \vec{\phi}_0 = \frac{1}{2\mu_k} \sum_{j \in Z_k} \vec{\phi}(P_j) \cdot (R\vec{Q}_{jk}); \qquad (2.4)$$

clearly, the right members of (2.3),(2.4) define a natural approximation of  $\vec{\nabla}\phi(P_k)$  and  $\vec{\nabla}\cdot\vec{\sigma}(P_k)$  respectively.

With the help of (2.3),(2.4), we now define a space semi-discretization of Problem (1.1)-(1.4). For all nodes  $P_k$ ,  $H_k(t)$  is an approximation of  $H(P_k,t)$ ; for interior nodes  $P_k \in \Omega$ ,  $\vec{U}_k(t) = (U_{k1}(t), U_{k2}(t))$  is an approximation of  $\vec{U}(P_k,t)$ ; for boundary nodes  $P_k \in \partial\Omega$ ,  $U_{Tk}(t)$  is an approximation of the tangential component of  $\vec{U}(P_k,t)$ , i.e. of  $\vec{U}(P_k,t) \cdot \hat{t}(P_k)$ . Method IS is then defined by the relations:

$$\frac{\mu_k}{b_k} \vec{\hat{U}}_k(t) = -\frac{1}{2} \sum_{j \in Z_k} H_j(t) (R\vec{Q}_{jk}) + \frac{\mu_k}{b_k} fR\vec{U}_k(t), \quad P_k \in \Omega; \quad (2.5)$$

$$\frac{\mu_k}{b_k} \dot{U}_{Tk}(t) = -\frac{1}{2} \sum_{j \in Z_k} H_j(t) (R \dot{Q}_{jk}) \cdot \dot{T}_k , \qquad P_k \in \partial \Omega ; \qquad (2.6)$$

$$\mu_{k} \dot{H}_{k}(t) = -\frac{1}{2} \sum_{j \in \mathbb{Z}_{k}} \vec{U}_{j}(t) \cdot (R\vec{Q}_{jk}), \qquad P_{k} \in \bar{\Omega} ; \qquad (2.7)$$

$$\vec{v}_{k}(o) = \vec{v}_{o}(P_{k}), P_{k} \in \Omega; U_{Tk}(o) = \vec{v}_{o}(P_{k}) \cdot \vec{t}(P_{k}), P_{k} \in \partial\Omega; \qquad (2.8)$$

$$H_{k}(o) = H_{o}(P_{k}),$$
  $P_{k} \in \bar{a};$  (2.9)

here the "dot" represents the time derivative,  $b_k = b(P_k)$  and in (2.7),

$$\vec{U}_{j}(t) = U_{Tj}(t)\vec{T}_{j}$$
 if  $P_{j} \in \partial \Omega$ .

By choosing any fixed order, all the unknown function  $U_{k1}(t)$ ,  $U_{k2}(t)$ ,  $U_{Tk}(t)$  and  $H_k(t)$  can be set in a single vector w(t) of dimension N. Then Problem (2.5)-(2.9) can be written in the compact form

$$D\dot{w}(t) = Aw(t); w(o) = w_{o};$$
 (2.10)

where D is a diagonal matrix with diagonal elements of the form  $\mu_k$  or  $\mu_k/b_k$ . Because of property (1.8) for L, one could expect that A is an antisymmetric matrix. Due to difficulties at the boundary, which seem inherent to the problems and impossible to overcome in a natural way, A is only "almost" antisymmetric. By inspection of the figure, one can show:



<u>Lemma 2.1</u>: Let  $P_k$  and  $P_j$  be two different nodes belonging to a same triangle. We suppose that at most one of them belong to the boundary  $\partial\Omega$ . Then

$$\vec{\mathsf{q}}_{\mathbf{j}\mathbf{k}} + \vec{\mathsf{q}}_{\mathbf{k}\mathbf{j}} = \vec{\mathsf{o}}. \tag{2.11}$$

As easily seen from (2.5)-(2.7), A would be exactly antisymmetric if (2.11) would hold when if both  $P_k$  and  $P_j$  belong to  $\partial \Omega$ . There are several ways to modify Scheme (2.5)-(2.7) for obtaining the desired property of antisymmetry. One of them consists in remplacing in (2.5),(2.6)  $\vec{q}_{jk}$  by  $-\vec{q}_{kj}$ ; by Lemma 2.1, (2.5) is not modified whereas (2.6) becomes

$$\frac{\mu_{k}}{b_{k}}\dot{U}_{Tk}(t) = \frac{1}{2} \sum_{j \in Z_{k}} H_{j}(t)(R\vec{Q}_{jk})\cdot\vec{T}_{k}, P_{k} \in \partial\Omega.$$
(2.12)

<u>Method IIS</u> is then defined by Relations (2.5), (2.12), (2.7), (2.8), (2.9) and can be written as

$$D\dot{w}(t) = Bw(t), w(o) = w_0$$
 (2.13)

where B is an antisymmetric matrix of order N.

For a vector v  $\in \mathbb{R}^{N}$  and a matrix G of order N, let

$$\|v\| = \left(\sum_{j=1}^{N} |v_{j}|^{2}\right)^{1/2}, \|G\| = \sup_{v \in \mathbb{R}^{N}} \frac{\|Gv\|}{\|v\|}.$$

270

By using the smoothness of  $\Im \Omega$  and the angle property of the sequence  $\{\mathcal{D}_h\},$  one can verify:

Lemma 2.2: There exist three constants  $c_1, c_2, c_3$ , independent of h such that

a)  $\|D^{-1/2}(A-B)D^{-1/2}\| \le c_1$ , b)  $c_2 h^{-1} \le \|D^{-1/2}BD^{-1/2}\| \le c_3 h^{-1}$ .

By using Lemma 2.2a, the antisymmetry of B, Relations (2.7),(2.9) in connection with the definition (2.2) of  $\vec{T}_k$ , one can deduce the following properties of Methods IS and IIS.

Proposition 2.1:

a) If w is solution of (2.13), then  $\|D^{1/2}w(t)\|^2 = \|D^{1/2}w_0\|^2$ .

b) There is a constant c. independent of t and of h, such that if w is solution of (2.10), then  $\||D^{1/2}w(t)\|^2 \leq e^{Ct} \|D^{1/2}w_{A}\|^2$ .

c) If w is solution of (2.10) or of (2.13), then  $\sum_{P_k \in \overline{\Omega}} \mu_k H_k(t) = \sum_{P_k \in \overline{\Omega}} \mu_k H_o(P_k)$ .

<u>Remark 2.1</u>: Since the sum is all  $\mu_k$ 's is equal to three times the area of  $\Omega_k$ , Proposition 2.1 appears to be, up to a factor 3, the discrete counterpart of Properties (1.9),(1.10) of the exact solution.

We now turn to the time discretisation. Let  $\tau > o$  be the time increment and set  $t_n = n\tau$ . We shall apply to (2.10) and (2.13) the two-step method, sometimes called "leap-frog" scheme. <u>Method I</u> is then defined by the relations:

$$w_{1/2} = w_0 + \frac{\tau}{2} D^{-1} A w_0$$
(2.14)

$$w_1 = w_0 + \tau D^{-1} A w_{1/2}$$
 (2.15)

$$w_{n+1} = w_{n-1} + 2\tau D^{-1}Aw_n$$
,  $n = 1, 2, 3, ...$  (2.16)

Method II is defined by replacing in (2.14)-(2.16). A by B.

The stability analysis of Method II is trivial since  $D^{-1}B$  or  $D^{-1/2}BD^{-1/2}$  have a pure imaginary spectrum. By using Strang's Lemma [3] in connection with Lemma 2.2a, we easily deduce the stability of Method I from that of Method II.

<u>Proposition 2.2</u>: There exist a function  $\alpha$ :  $(o,1) \rightarrow \mathbb{R}$  and a constant c, both independent of h,  $\tau$  and n such that if  $\tau < 1/\|D^{-1/2} BD^{-1/2}\|$ , we have:

a) 
$$\||D^{-1/2}w_n\||^2 \leq \alpha(\tau \||D^{-1/2}BD^{-1/2})\||)\||D^{1/2}w_0\|$$
 for Method II,

b) 
$$\|D^{1/2}w_n\|^2 \leq \alpha(\tau \|0^{-1/2}BD^{-1/2}\|)e^{Ct_n}\|D^{1/2}w_0\|$$
 for Method I.

<u>Remark 2.2</u>: Methods I and II satisfy a law of mass conservation as Methods IS and IIS (see Proposition 2.1c).

<u>Remark 2.3</u>: Proposition 2.2a,b prove the stability of both Method I and II; however Method II appears to be "more" stable than Method I.

<u>Remark 2.4</u>:  $w_n$  can be written as a vector of order N with components  $U_{kn1}$ ,  $U_{kn2}$ ,  $U_{Tkn}$ and  $H_{kn}$ . If the Coriolis term f = 0, then for Methods I and II, it is possible to compute  $U_{kn1}$ ,  $U_{kn2}$ ,  $U_{Tkn}$  only at even values of n and  $H_{kn}$  at odd values of n which reduces the computer time and the storage requirements by a factor 2; real Thacker's scheme, which is somewhat more difficult to analyse, keeps this property even for f  $\neq$  0. In fact, if f = 0, Method I is identical to Thacker's scheme (10),(11),(11') in [2] p.683.

<u>Remark 2.5</u>: Thacker [2] has remarked that his scheme can be considered, to some extend, as a lumped version of a Galerkin method. In [4], we briefly analyse the effect of "lumping" on stability.

<u>Remark 2.6</u>: The stability condition  $\tau < 1/\|D^{-1/2}BD^{-1/2}\|$  in Proposition 2.2, implies, by Lemma 2.2b that  $\tau = 0(h)$ .

#### 3. ERROR ESTIMATES

Our estimates will be based on the standard consistency+stability argument. Stability has already been analyzed in Section 2.

We begin with a classical study of consistency by assuming that the components of the solution of Problem (1.1)-(1.4) belong to  $C^{O}([o,T]; C^{3}(\bar{\Omega}))$ . We first associate to this solution a vector u(t)  $\in \mathbb{R}^{N}$  in the following way; let w(t) the solution of the semi-discritized problem (2.10); we set:  $u_{i}(t) = U_{g}(P_{k},t)$  if  $w_{i}(t) = U_{k\ell}$ ,  $\ell = 1,2$ ;

$$u_i(t) = \vec{U}(P_k,t) \cdot \vec{t}(P_k)$$
 if  $w_i(t) = U_{Tk}(t)$ ;  $u_i(t) = H(P_k,t)$  if  $w_i(t) = H_k(t)$ .

Clearly, the time discretization which is of order two, will induce errors of size  $O(\tau^2)$ , which, by Remark 2.6, can be written as  $O(h^2)$ . Let us define for T > o:

$$\begin{split} & \varepsilon_{I}(T) = \max_{\substack{0 \leq t \leq T \\ 0 \leq t \leq T}} \|D^{1/2}(\dot{u}(t) - D^{-1}Au(t))\|, \ \varepsilon_{II}(T) = \max_{\substack{0 \leq t \leq T \\ 0 \leq t \leq T}} \|D^{1/2}(\dot{u}(t) - D^{-1}Bu(t))\|, \\ & R_{I(II)}(T) = \max_{\substack{0 \leq t \\ n \leq T}} \|D^{1/2}(u(t_{n}) - w_{n})\| \text{ if } w_{n} \text{ is obtained by Method I(II).} \end{split}$$

For i = I,II,  $R_i(T)$  is the error of Method i, whereas, by using (1.1),(1.2),  $\epsilon_i(T)$  is the space consistency error.

In the following, we shall say that for a node  $P_k$ ,  $\Lambda_k$  is <u>symmetric</u>, if for each  $P_j \in \partial \Omega_k$ , there exists  $P_\ell \in \partial \Lambda_k$  which is symmetric to  $P_j$  with respect to  $P_k$ . Clearly if  $P_k \in \partial \Omega$ ,  $\Lambda_k$  cannot be symmetric. The basic difference schemes defined by (2.3), (2.4) are of order 2, with respect to h, if  $\Lambda_k$  is symmetric; otherwise there are only of order 1. We shall say that the sequence of triangularizations  $\{\mathcal{D}_h\}$  possesses Property G if there exists a constant c, independent of h such that for all  $P_k \in \partial \Omega$  one has  $|\hat{T}_k - \hat{t}(P_k)| \leq ch^2$ ; Property G implies a certain regularity in the distribution of the nodes on the boundary. Elementary but tidious calculations allow to establish:

<u>Lemma 3.1</u>: For any fixed T > o and i = I,II, we have: a)  $\varepsilon_i(T) = O(h^{1/2})$ ; b)  $\varepsilon_i(T) = O(h)$  if Property G is satisfied; c)  $\varepsilon_i(T) = O(h^{3/2})$  if  $\Lambda_k$  is symmetric for all  $P_k \in \Omega$  and if Property G is satisfied.

From Lemma 3.1 follows immediately.

<u>Proposition 3.1</u>: Let T > o and  $\xi \in (o,1)$  be fixed numbers. For each triangularization  $\mathscr{D}_{h}$ ,  $\tau$  is chosen in such a way that  $o < \tau < \xi/||D^{-1/2} BD^{-1/2}||$ . Then for i = I, II: a)  $R_{i}(T) = O(h^{1/2})$ ; b)  $R_{i}(T) = O(h)$  if Property G is satisfied; c)  $R_{i}(T) = O(h^{3/2})$  if  $\Lambda_{k}$  is symmetric for all  $P_{k} \in \Omega$  and if Property G is satisfied.

<u>Remark 3.1</u>: Suppose, that, instead of (2.2), we set  $\vec{\uparrow}_k = \vec{t}(P_k)$  (exact tangent vector). Then: a) We loss the exact mass conservation property for both Methods I and II /see Remark 2.2); b) Proposition 3.1 remains valid for Method I; Proposition 3.1a remains valid for Method II.

We now turn to an error analysis under weaker regularity assumptions. We shall suppose that the components of the solution of Problem (1.1)-(1.4) belong to  $C^{0}([o,T]; H^{2}(\Omega))$ . For simplicity, we shall furthermore assume that  $\Omega$  is convex so that  $\Omega_{h^{c}} \Omega$ . The difficulty here comes from the fact that the time derivative of the solution is not a continuous function of the space variable. Let  $\vec{U}_{kn} = (U_{kn1}, U_{kn2}), H_{kn}$  be the approximate solution obtained by Method I or II corresponding to the exact solution  $\vec{U}(P_{k}, t_{n})$ ,  $H(P_{k}, t_{n})$ ; here we set  $\vec{U}_{kn} = U_{Tkn}\vec{T}_{k}$  if  $P_{k} \in \partial\Omega$ . Let  $V_{h}$  be the space of continuous piecewise linear functions on  $\Omega_{h}$  corresponding to  $\mathscr{D}_{h}$ . We define the function  $\vec{U}_{h}(x,t) = (U_{h1}(x,t), U_{h2}(x,t)), H_{h}(x,t)$ , for  $x \in \Omega_{h}$  and  $t = t_{n}$ , in the following way:  $U_{h\beta}(\cdot, t_{n}) \in V_{h}, \beta = 1, 2, H_{h}(\cdot, t_{n}) \in V_{h}; \vec{U}_{h}(P_{k}, t_{n}) = \vec{U}_{kn}, H_{h}(P_{k}, t_{n}) = H_{kn}$  for all  $P_{k} \in \bar{\Omega}$ .

The main trick will consist in introducing functions  $Y_h(x,t) = (Y_{h_1}(x,t),Y_{h_2}(x,t))$ ,  $Z_h(x,t)$  belonging to  $V_h$  for fixed t and which are, for fixed t, Clément's approximations of the exact solution's components; for the notion of Clément's approximation, see [7]. Corresponding to the exact equation (1.2), we have for all  $P_k \in \overline{\Omega}$  the following identity:

$$= \{ \mu_{k} Z_{h}(P_{k}, t_{n+1}) - \mu_{k} Z_{h}(P_{k}, t_{n-1}) - 2\tau \int_{\Lambda_{k}} \dot{H}(x, t_{n}) \}$$
(3.2)

+ 
$$\tau \{ \sum_{j \in \mathbb{Z}_{k}} \overline{\Psi}(P_{j}, t_{n}) \cdot (R\overline{d}_{jk}) - 2 \int_{k} \overline{\Psi} \cdot \overline{U}(x, t_{n}) \}.$$
 (3.3)

The "time" error term (3.2) can be easily estimated by using the fact that the operations of time derivative and Clément's approximation commute. The "space" error term (3.3) can be handled by remarking the following identity which is a direct consequence of (2.4):

$$\sum_{j \in Z_k} \overline{\tilde{Y}}(P_j, t_n) \cdot (R \overline{Q}_{jk}) - 2 \int_{\Lambda_k} \overline{\tilde{V}} \cdot \overline{\tilde{Y}}(x, t_n) = 0.$$

Similarly to (3.1)-(3.3), we can write an equation corresponding to (1.1); it is slightly more complicated to handle because of the boundary condition (1.3) and of the presence of the function b(x). With Proposition 2.2, this allows to get error estimates between Clément's approximation and the solution of Method I or II. The final result is contained in the following proposition:

Proposition 3.2: Let T > o and  $\xi \in (o,1)$  be fixed numbers. We suppose: a) each component of the exact solution belong to  $C^{O}([0,T]; H^{2}(\Omega))$ ; b) Property G is satisfied; c) for each triangularization,  $\tau$  is chosen such that  $o < \tau < \xi / |D^{-1/2} BD^{-1/2}||$ . Then for both Methods I and II we have:

$$\max_{0 \le t_n \le T} \{ \prod_{\Omega_h} \left( \frac{1}{b(x)} | \vec{U}(x, t_n) - \vec{U}_h(x, t_n) |^2 + | H(x, t_n) - H_h(x, t_n) |^2 ) \}^{1/2} = O(h).$$

Remark 3.2: In [4], we give some numerical results.

Remark 3.3: In order to compute the spectrum of the operator L defined in Section 1, one could think of using the same space discretization as in Method I or II; however this generates spurious eigenvalues. For a proper treatment of this problem, see [5], [6].

#### References

- [1] W.C.THACKER, Irregular Grid Finite-Difference Techniques: Simulations of Oscillations in Shallow Circular Basins, Journal of Oceanography, Vol. 7, 1977, 284-292.
- [2] W.C.THACKER, Comparison of Finite-Element and Finite-Differences Schemes, Part I: One-Dimension-al Cravity Wave Motion, Part II: Two-Dimensional Gravity Wave Motion, Journal of Oceanography, vol. 8, 1978, 676-689.
- [3] G.STRANG, Accurate Partial Difference Methods, Numerische Mathematik, 6, 1964, 37-46.
   [4] J.DESCLOUX, R.FERRO, On Thacker's scheme for solving the linearized shallow water equations,
- Report. Departement de Mathématiques. Ecole Polytechnique Fédérale de Lausanne, 1985 [5] M.LUSKIN, Convergence of a Finite Element Method for the Approximation of Normal Modes of the
- M.LUSKIN, Convergence of a Printe Element method for the Approximation of Normal Inducts of Oceans, Math. Comp. 33, 1979, 493-519.
   J.DESLOUX, M.LUSKIN, J. RAPPAZ, Approximation of the Spectrum of Closed Operators: The Determination of Normal Modes of a Rotating Basin, Math. Comp. 36, 1981, 137-154.
- [7] Ph.CLEMENT, Approximation by Finite Element Functions using Local Regularizations, RAIRO 9, 1975, 77-84.