RAIRO
ANALYSE NUMÉRIQUE

ALEXANDER ŽENÍŠEK

Finite element methods for coupled thermoelasticity and coupled consolidation of clay

RAIRO – Analyse numérique, tome 18, n° 2 (1984), p. 183-205.
<http://www.numdam.org/item?id=M2AN_1984__18_2_183_0>

R.A.I.R.O. Analyse numérique/Numerical Analysis
(vol 18, n° 2, 1984, p 183 a 205)

FINITE ELEMENT METHODS FOR COUPLED THERMOELASTICITY AND COUPLED CONSOLIDATION OF CLAY (*)

by Alexander ŽENÍŠEK ([1])

Communicated by P G CIARLET

Abstract — *Three linear two-dimensional coupled problems are considered dynamical thermoelasticity, quasistatical thermoelasticity and consolidation of clay In the first two cases an equation parabolic with respect to the temperature T is coupled with a system of equations either hyperbolic or elliptic with respect to the displacement vector* **u**, *in the third case an equation elliptic with respect to the pressure T is coupled with a system elliptic with respect to the displacement vector* **u** *The problems are solved approximately using both triangular and curved triangular finite elements in the space discretization and ν-step A-stable difference methods* (ν = 1 *or* 2) *in the time discretization. The effect of numerical integration is also considered The resulting schemes are unconditionally stable*

Résumé — *Nous traitons trois problèmes linéaires en deux dimensions la thermoélasticité dynamique, la thermoélasticité quasi statique et la consolidation de l'argile Dans les deux premiers cas une équation parabolique par rapport à la température T est couplée avec un système d'équations hyperboliques ou bien elliptiques par rapport au vecteur* **u** *des déplacements Dans le troisième cas, une équation elliptique par rapport à la pression T est couplée avec un système elliptique par rapport au vecteur* **u** *des déplacements Ces problèmes sont approchés en utilisant la méthode des éléments finis avec les triangles rectilignes ou curvilignes pour la discrétisation spatiale et les méthodes des différences finis A-stables a* ν *pas* (ν = 1 *ou* 2) *pour la discrétisation en temps Les schémes qui en résultent sont inconditionnellement stables*

1. FORMULATION OF THE PROBLEM

According to [2] the dynamical two-dimensional problem of coupled linear thermoelasticity can be formulated in the following way : Let Ω be a bounded domain in the x_1, x_2-plane with a sufficiently smooth boundary Γ. Find a displacement vector $\mathbf{u}(x_1, x_2, t)$ and a temperature $T(x_1, x_2, t)$ which satisfy the following equations and boundary and initial conditions (for a greater sim-

(*) Received in April 1982, revised in May 1983
([1]) Computing Center of the Technical University, Obránců míru 21, 602 00 Brno, Czechoslovakia

plicity we restrict ourselves to the case of Dirichlet boundary conditions) :

$$T_{,ii} + Q = c_1 \dot{T} + c_2 \dot{u}_{i,i} \quad \text{in} \quad \Omega \times (0, t^*] \tag{1}$$

$$\sigma_{ij,j} + X_i = c_4 u_i \quad (i = 1, 2) \quad \text{in} \quad \Omega \times (0, t^*] \tag{2}$$

$$T(x_1, x_2, t)|_\Gamma = \overline{T}(x_1, x_2), \quad t > 0 \tag{3}$$

$$u_i(x_1, x_2, t)|_\Gamma = \overline{u}_i(x_1, x_2) \quad (i = 1, 2), \quad t > 0 \tag{4}$$

$$c_1 T(x_1, x_2, 0) = c_1 T_0(x_1, x_2), \quad (x_1, x_2) \in \Omega \tag{5}$$

$$u_i(x_1, x_2, 0) = u_{i0}(x_1, x_2), \quad (x_1, x_2) \in \Omega \quad (i = 1, 2) \tag{6}$$

$$c_4 \dot{u}_i(x_1, x_2, 0) = c_4 v_{i0}(x_1, x_2), \quad (x_1, x_2) \in \Omega \quad (i = 1, 2) \tag{7}$$

where

$$\sigma_{ij} = D_{ijkm}[\varepsilon_{km}(\mathbf{u}) - \alpha(T - T_r)\,\delta_{km}] \tag{8}$$

$$D_{ijkm} = D_{jikm} = D_{kmij} \tag{9}$$

$$\varepsilon_{ij}(\mathbf{v}) = (v_{i,j} + v_{j,i})/2 \tag{10}$$

$$D_{ijkm}\,\xi_{ij}\,\xi_{km} \geqslant \mu_0\,\xi_{ij}\,\xi_{ij} \quad \forall \xi_{ij} = \xi_{ji} \in R \tag{11}$$

where $\mu_0 = \text{const.} > 0$. A summation convention over a repeated subscript is adopted. A comma is employed to denote partial differentiation with respect to spatial coordinates and a dot denotes the derivative with respect to time t. Thus equation (1) is the coupled heat equation and equations (2) are Cauchy's equations of equilibrium. The symbol Q denotes a prescribed sufficiently smooth rate of internal heat generation per unit volume. the symbols X_1, X_2 denote prescribed sufficiently smooth components of body forces per unit volume. The symbols c_1, c_2, c_4 are positive constants; c_1 and c_4 depend only on the material of a considered body, $c_2 = \overline{c}_2\, T_r$ where $\overline{c}_2$ is a positive constant depending only on the material and T_r is a positive constant which has the meaning of the temperature for which the material is stress-free. The functions on the right-hand sides of relations (3)-(7) are prescribed sufficiently smooth functions.

In relation (8) α is the coefficient of linear thermal expansion, δ_{ij} is the Kronecker delta and D_{ijkm} are constants depending on the material only. We shall consider isotropic materials only; in this case

$$\alpha D_{ijkm}\,\delta_{km} = c_3\,\delta_{ij}, \quad c_3 = \text{const.} > 0\,. \tag{12}$$

If we set $c_1 = c_4 = 0$, replace (6) by

$$u_{i,i}(x_1, x_2, 0) = 0\,, \quad (x_1, x_2) \in \Omega \tag{6*}$$

and define σ_{ij} by

$$\sigma_{ij} = D_{ijkm}\,\varepsilon_{km}(\mathbf{u}) - T\,\delta_{ij} \tag{13}$$

then problem (1)-(4), (6*), (9)-(13) represents the two-dimensional problem of coupled consolidation of clay in the case of incompressible pore water [1, 3, 9]. The symbol T has now the meaning of pore water pressure, the constant c_2 depends on the material only and $Q \equiv 0$. Numerical tests [9] show that the linear model (1)-(4), (6*), (9)-(11), (13) gives satisfactory results. (Let us note that a nonlinear elastoplastic model is studied in [8].)

Now we present a variational formulation of our three problems. Before doing it let us introduce some notation. By $H^m(\Omega)$ we denote the Sobolev space of real functions which together with their generalized derivatives up to order m inclusive are square integrable over Ω. The inner product and the norm are denoted by $(.,.)_{m,\Omega}$ and $\| \,.\, \|_{m,\Omega}$, respectively. $H_0^1(\Omega)$ is the closure in the H^1-norm of the set of infinitely differentiable functions having compact support in Ω. $H^{-1}(\Omega)$ is the space dual to $H_0^1(\Omega)$ (with dual norm). $C^m(H^k(\Omega))$ is the space of continuous functions $f.\ [0, t^*] \to H^k(\Omega)$ which have continuous derivatives up to order m on $[0, t^*]$. $L^2(H^k(\Omega))$ is the space of strongly measurable functions $f : (0, t^*) \to H^k(\Omega)$ such that

$$\int_0^{t^*} \| f(t) \|_{k,\Omega}^2\, dt < \infty .$$

Multiplying equation (1) by $w \in H_0^1(\Omega)$ and using Green's theorem we easily find

$$D(T, w) + c_1(\dot{T}, w)_{0,\Omega} + c_2(\dot{u}_{i,i}, w)_{0,\Omega} = (Q, w)_{0,\Omega} \qquad \forall w \in H_0^1(\Omega),\ t \in (0, t^*] \tag{14}$$

where

$$D(v, w) = \int_\Omega v_{,i}\, w_{,i}\, dx\,, \quad (v, w)_{0,\Omega} = \int_\Omega vw\, dx\,. \tag{15}$$

Multiplying equation (8) or (13) by $\varepsilon_{ij}(\mathbf{v})$, where $\mathbf{v} \in [H_0^1(\Omega)]^2$, integrating over Ω, using relations (2), (9), (10) and Green's theorem we find

$$a(\mathbf{u}, \mathbf{v}) + c_4(\ddot{\mathbf{u}}, \mathbf{v})_{0,\Omega} - c_3(T - T_r, v_{i,i})_{0,\Omega} = = (\mathbf{X}, \mathbf{v})_{0,\Omega} \quad \forall \mathbf{v} \in [H_0^1(\Omega)]^2\,, \quad t \in (0, t^*] \tag{16}$$

where

$$a(\mathbf{v}, \mathbf{w}) = \int_\Omega D_{ijkm}\, v_{i,j}\, w_{k,m}\, dx\,, \quad (\mathbf{v}, \mathbf{w})_{0,\Omega} = \int_\Omega v_i\, w_i\, dx \tag{17}$$

and $c_3 = 1$, $T_r = 0$ in the case of consolidation of clay.

Thus the variational formulation of problem (1)-(11) reads : Find a function T and a vector $\mathbf{u}$ which have the following properties :

a) $T \in L^2(H^1(\Omega))$, $\dot{T} \in L^2(H^{-1}(\Omega))$;

b) $\mathbf{u} \in L^2([H^1(\Omega)]^2)$, $\dot{u}_{i,i} \in L^2(H^{-1}(\Omega))$, $\ddot{\mathbf{u}} \in L^2([H^{-1}(\Omega)]^2)$;

c) the function T satisfies boundary condition (3) and the vector $\mathbf{u}$ satisfies boundary conditions (4) in the sense of traces;

d) relations (14), (16) hold;

e) the function T satisfies initial condition (5) and the vector $\mathbf{u}$ satisfies initial conditions (6), (7).

The remaining two variational formulations can be obtained from *a*)-*e*) by means of the following rules : If $c_4 = 0$ then we do not demand $\ddot{\mathbf{u}} \in L^2([H^{-1}(\Omega)]^2)$ in *b*) and (7) in *e*) and replace (6) by (6*) in *e*). If moreover $c_1 = 0$ then we do not demand $\dot{T} \in L^2(H^{-1}(\Omega))$ in *a*) and (5) in *e*).

The uniqueness of the solutions of all three variational problems can be proved similarly as in [2, pp. 39-40]. As to the existence of the solution see, e.g., [7, 16] where some sufficient conditions are presented.

From the point of view of applications it should be noted that it is permissible in most thermal — stress problems to disregard the effects of both coupling and inertia (see [2, pp. 41-61]). However, in the case of consolidation of clay the coupling effect is not negligible.

2. FINITE ELEMENT SOLUTION

If the domain Ω has not a polygonal boundary we approximate Γ piecewise by arcs of degree $n > 1$ and denote such a changed boundary by Γ_h. The curve Γ_h is the boundary of a domain Ω_h which is the approximation of Ω. (If Ω has a polygonal boundary Γ we set (because of a uniform notation) $\Omega_h = \Omega$ and $\Gamma_h = \Gamma$.) Let us assume that Γ is piecewise of class C^{n+1}.

Let us triangulate the domain Ω_h. If $\Omega_h = \Omega$ then the triangulation τ_h is quite arbitrary; if $\Omega_h \neq \Omega$ then τ_h satisfies the following conditions : Each arc of degree n is the curved side of one boundary triangle; each boundary triangle has only one side lying on Γ_h; the interior triangles have only straight sides.

Let V_h and W_h denote two finite element subspaces of $C^0(\overline{\Omega}_h)$ with the following properties :

1. To every function $v \in H^{n+1}(\tilde{\Omega})$, where $\tilde{\Omega} \supset \Omega \cup \Omega_h$ $(h < \tilde{h})$, there exists a function $v^{\text{int}} \in V_h$, the interpolate of v, such that

$$\| v - v^{\text{int}} \|_{j,\Omega_h} \leqslant \mathrm{C}h^{n+1-j} \| v \|_{n+1,\tilde{\Omega}} \quad (j = 0, 1) \tag{18}$$

where C is an absolute constant and $h = \max h_K$ $(K \in \tau_h)$, h_K being the diameter of the triangle K. (We assume that the smallest angle of all triangles of all triangulations τ_h is bounded away from zero.) Further, it holds

$$\mathbf{v}_1, \mathbf{v}_2 \in \mathrm{V}^1_{hu} \times \mathrm{V}^2_{hu} \Rightarrow \mathbf{v}_1 - \mathbf{v}_2 \in \mathrm{V}_{h0} \times \mathrm{V}_{h0} \tag{19}$$

where the subsets V_{h0}, V^i_{hu} $(i = 1, 2)$ of the space V_h are defined in the following way :

$$V_{h0} = \{ v \in V_h : v = 0 \text{ on } \Gamma_h \} , \tag{20}$$

$$V^i_{hu} = \{ v \in V_h : v = \bar{u}^{\text{int}}_i \text{ on } \Gamma_h \} \quad (i = 1, 2) \tag{21}$$

where $\bar{u}^{\text{int}}_i$ is the finite element interpolation of the function $\bar{u}_i$ appearing in (4). (E.g., if the used finite elements are of the Lagrange type then each curved side of a curved boundary triangle has $n + 1$ common points with Γ. The function $\bar{u}^{\text{int}}_i$ is uniquely determined on the curved side by the function values of $\bar{u}_i$ at these $n + 1$ points.)

2. To every function $w \in H^{p+1}(\tilde{\Omega})$, where $p < n$, there exists a function $w^{\text{int}} \in W_h$, the interpolate of w, such that

$$\| w - w^{\text{int}} \|_{j,\Omega_h} \leqslant Ch^{p+1-j} \| w \|_{p+1,\tilde{\Omega}} \quad (j = 0, 1) \tag{22}$$

with C an absolute constant. It holds

$$w_1, w_2 \in W_{hT} \Rightarrow w_1 - w_2 \in W_{h0} \tag{23}$$

where the subsets W_{h0} and W_{hT} of the space W_h are defined in the following way :

$$W_{h0} = \{ w \in W_h : w = 0 \text{ on } \Gamma_h \} , \tag{24}$$

$$W_{hT} = \{ w \in W_h : w = \bar{T}^{\text{int}} \text{ on } \Gamma_h \} \tag{25}$$

where $\bar{T}^{\text{int}}$ is the finite element interpolation of the function $\bar{T}$ appearing in (3).

As we want to approximate all terms on the right-hand side of (8) (or (13)) with the same accuracy we choose $p = n - 1$. (In applications we usually have $n = 2, p = 1$ — see, e.g., [1, 9].) In the case of a polygonal boundary Γ the construction of the spaces V_h and W_h is straightforward and we can choose the space W_h quite independently on the space V_h. In the case of curved elements the choice of the space V_h determines the choice of the space W_h.

We explain it in detail in the case $n = 3$ and show simultaneously that such constructions are possible. Let us consider finite elements of the Hermite type. In this case the parametric equations of the curved side of a boundary triangle are formed by cubic Hermite interpolation polynomials of the functions which express parametrically the corresponding arc of the exact boundary Γ. (This situation is described in [12, 13, 15] where all details concerning the construction of corresponding finite elements can be found.) The ten parameters uniquely determining the function $v(x_1, x_2)$ on the curved triangle K are

$$D^\alpha v(P_i), |\alpha| \leqslant 1, i = 1, 2, 3; v(P_0) \tag{26}$$

where P_1, P_2, P_3 are the vertices of K and P_0 is the image of the point $R_0(1/3, 1/3)$ in the transformation

$$x_1 = x_1^*(\xi_1, \xi_2), \quad x_2 = x_2^*(\xi_1, \xi_2) \tag{27}$$

which maps one-to-one the boundary triangle K onto the unit triangle K_0 lying in the plane ξ_1, ξ_2 and having the vertices $R_1(0, 0)$, $R_2(1, 0)$, $R_3(0, 1)$. For this element the following interpolation theorem holds : If $u \in H^4(K)$ and

$$D^\alpha v(P_i) = D^\alpha u(P_i), \quad |\alpha| \leqslant 1, \quad i = 1, 2, 3; \quad v(P_0) = u(P_0)$$

then

$$\| v - u \|_{j,K} \leqslant Ch^{4-j} \| u \|_{4,K} \quad (0 \leqslant j \leqslant 4).$$

The same interpolation theorem holds for a polynomial of third degree $v(x_1, x_2)$ which is uniquely determined on a triangle K with straight sides by the parameters (26), P_0 being now the centre of gravity of K. Combining these two types of finite elements we obtain the space V_h with the interpolation property (18).

Let Γ be of class C^1. Then it follows from the construction of curved triangular elements that the subspace $V_{h0} \subset V_h$ consists of those functions $v \in V_h$ for which

$$v(P_j) = v_{,1}(P_j) \cos \alpha_j + v_{,2}(P_j) \sin \alpha_j = 0, \quad P_j \in \Gamma_h$$

where α_j is the angle made by the x-axis and the tangent to the curve Γ at the point P_j. (If Γ is piecewise of class C^1 and P_j a corner of Γ then $v(P_j) = v_{,1}(P_j) = v_{,2}(P_j) = 0$.) The subset $V_{hu}^i (i = 1, 2)$ of V_h consists of those functions $v \in V_h$

for which

$$v(P_j) = \bar{u}_i(P_j), \quad v_{,1}(P_j) \cos \alpha_j + v_{,2}(P_j) \sin \alpha_j = \bar{u}_i'(P_j)$$

where $P_j \in \Gamma_h$ and where $\bar{u}_i'$ is the tangent derivative of the function $\bar{u}_i$ on Γ. It is clear that implication (19) is satisfied.

Now we want to construct the finite element space W_h on the same triangulation τ_h in the case $p = 2$. To this end, on the interior triangles which have no common point with the boundary we choose quadratic polynomials uniquely determined by function values prescribed at the vertices and at the mid-points of the sides. On the boundary triangles, which have a cubic curved side, we choose functions which are uniquely determined by parameters

$$D^\alpha w(P_1), |\alpha| \leqslant 1, \quad w(P_2), \quad w(P_3), \quad w(Q_{23}) \tag{28}$$

where P_1, P_2, P_3 is a local notation of the vertices of a boundary triangle K chosen in such a way that P_1, P_3 lie on Γ; Q_{23} is the mid-point of the segment $P_2 P_3$. If K_1, K_2 are two boundary triangles with a common vertex and the local notation of the vertices of K_1 has been chosen then we must choose the local notation of the vertices of K_2 in such a way that the common vertex of K_1 and K_2 is denoted by the same symbol in both local notations. (This implies a restriction on triangulations τ_h : the number of boundary triangles must be even.)

The function $w(x_1, x_2)$ uniquely determined on K by parameters (28) is defined in the following way : The function

$$w^*(\xi_1, \xi_2) = w(x_1^*(\xi_1, \xi_2), x_2^*(\xi_1, \xi_2)), \tag{29}$$

where $x_1^*(\xi_1, \xi_2)$, $x_2^*(\xi_1, \xi_2)$ are the same functions as in (27), is a quadratic polynomial uniquely determined by the parameters

$$D^\alpha w^*(R_1), \quad |\alpha| \leqslant 1, \quad w^*(R_2), w^*(R_3), w^*(S_{23}) \tag{30}$$

where $S_{23} = (1/2, 1/2)$. Parameters (30) are linear combinations of parameters (28) and can be computed by means of (29) and the rule of differentiation of a composite function.

It remains to define finite elements on the interior triangles with one vertex lying on Γ. If this vertex is the vertex P_1 of a boundary triangle then we choose a quadratic polynomial uniquely determined by parameters (28), where P_1, P_2, P_3 denote now vertices of the interior triangle and P_1 lies on Γ. If the vertex lying on Γ is the vertex P_3 of a boundary triangle then we choose a quadratic polynomial uniquely determined by function values prescribed at the vertices and the mid-points of the sides of the interior triangle.

Combining triangular finite elements just described we can construct the finite dimensional space W_h. It is easy to see that W_h has the interpolation property (22). The construction of the subsets W_{h0}, W_{hT} with property (23) is similar as in the case of the space V_h.

The constructions of the spaces V_h, W_h in the case of finite elements of the Lagrange type are simpler than in the preceding case. Thus we do not introduce them.

In order to define the approximate solution of the variational problem *a)-e)* let us introduce the bilinear forms

$$\tilde{D}_h(v, w) = \int_{\Omega_h} v_{,i}\, w_{,i}\, dx\,, \quad (v, w)_{0,\Omega_h} = \int_{\Omega_h} vw\, dx\,, \tag{31}$$

$$\tilde{a}_h(\mathbf{v}, \mathbf{w}) = \int_{\Omega_h} D_{ijkm}\, v_{i,j}\, w_{k,m}\, dx, (\mathbf{v}, \mathbf{w})_{0,\Omega_h} = \int_{\Omega_h} v_i\, w_i\, dx\,. \tag{32}$$

In the case of polygonal boundary $\tilde{a}_h(\mathbf{v}, \mathbf{w}) = a(\mathbf{v}, \mathbf{w})$, etc.

In order to get numerical results in the case of a curved boundary Γ we approximate the integrals appearing in (31), (32) by quadrature formulas with integration points lying in Ω in the same way as in [5] or [10, 15]. Doing it we obtain forms $D_h(v, w)$, $(v, w)_h$, $a_h(\mathbf{v}, \mathbf{w})$ and $(\mathbf{v}, \mathbf{w})_h$.

Let us choose an integer M and set

$$\Delta t = t^*/M\,, \quad t_m = m\,\Delta t \quad (m = 0, 1, ..., M)\,. \tag{33}$$

If $f = f(x_1, x_2, t)$ then the symbol f^m will denote a function in two variables x_1, x_2 defined by the relation

$$f^m \equiv f^m(x_1, x_2) = f(x_1, x_2, m\,\Delta t)\,. \tag{34}$$

Finally, we denote

$$\Delta f^m = f^{m+1} - f^m, \Delta^2 f^m = \Delta f^{m+1} - \Delta f^m\,. \tag{35}$$

Now we can define the discrete problem for approximate solving our variational problem *a)-e)* :

Let $\nu = 1$ or 2 in the case $c_4 = 0$ and $\nu = 2$ in the case $c_4 > 0$. For each $m = 0, 1, ..., M - \nu$ find a vector $\mathbf{u}_h^{m+\nu} \in V_{hu}^1 \times V_{hu}^2$ and a function $T_h^{m+\nu} \in W_{hT}$ such that

$$\Delta t D_h\left(\sum_{j=0}^{\nu} \beta_j\, T_h^{m+j}, w\right) + c_1\left(\sum_{j=0}^{\nu} \alpha_j\, T_h^{m+j}, w\right)_h^{**} + \\ + c_2\left(\sum_{j=0}^{\nu} \alpha_j\, u_{hi,i}^{m+j}, w\right)_h = \Delta t\left(\sum_{j=0}^{\nu} \beta_j\, Q^{m+j}, w\right)_h^{*} \quad \forall w \in W_{h0}\,, \tag{36}$$

$$\Delta t^2\, a_h\left(\sum_{j=0}^{\nu} \beta_j\, \mathbf{u}_h^{m+j}, \mathbf{v}\right) - c_3\, \Delta t^2\left(\sum_{j=0}^{\nu} \beta_j\, T_h^{m+j} - T_r, v_{i,i}\right)_h + \\ + c_4(\Delta^2 \mathbf{u}_h^m, \mathbf{v})_h = \Delta t^2\left(\sum_{j=0}^{\nu} \beta_j \mathbf{X}^{m+j}, \mathbf{v}\right)_h \quad \forall \mathbf{v} \in [V_{h0}]^2 \quad (37)$$

with initial conditions (38), (39) in the case $c_1 > 0$, $c_4 > 0$:

$$\mathbf{u}_h^0 = \mathbf{u}_0^{\text{apr}} \in V_{hu}^1 \times V_{hu}^2, \quad T_h^0 = T_0^{\text{apr}} \in W_{hT}, \tag{38}$$

$$\mathbf{u}_h^1 = \mathbf{z}^{\text{apr}} \in V_{hu}^1 \times V_{hu}^2, \quad T_h^1 = Y^{\text{apr}} \in W_{hT}, \tag{39}$$

where the meaning of the symbols on the right-hand sides of (38), (39) is defined in Remark 3. The coefficients α_j, β_j are defined in Remark 2. In the case of consolidation of clay initial conditions reduce to

$$u_{hi,i}^0 = 0 . \tag{40}$$

This restricts a choice of finite difference formulas.

Remark 1 : In (36) the symbols $(v, w)_h^*$ and $(v, w)_h{}^{**}$ denote two approximations of $(v, w)_{0,\Omega_h}$ which are, in general, different from $(v, w)_h$ — see Theorem 1.

Remark 2 : For $\nu = 1$ we have

$$\alpha_0 = -1, \quad \alpha_1 = 1; \quad \beta_0 = \theta, \quad \beta_1 = 1 - \theta \quad (\theta \leqslant 1/2) \tag{41}$$

and for $\nu = 2$

$$\alpha_0 = -1 + \theta, \quad \alpha_1 = 1 - 2\theta, \quad \alpha_2 = \theta, \quad \beta_0 = 1/2 - 1/2\,\theta + \delta, \\ \beta_1 = 1/2 - 2\delta, \quad \beta_2 = 1/2\,\theta + \delta \quad (\theta \geqslant 1/2) \quad (42)$$

where $\delta > 0$ if $c_4 = 0$ and $\delta \geqslant 0$ if $c_4 > 0$. In the case $c_4 = 0$ relations (41) and (42) define the coefficients of ν-step A-stable methods (see [10, 14]). In the case $c_4 > 0$ relations (35_2) and (42) define the coefficients of the general Newmark method. In this case the β's are written usually in the form :

$$\beta_0 = 1/2 + \beta_2 - \theta, \quad \beta_1 = 1/2 - 2\beta_2 + \theta, \quad \beta_2 \geqslant 1/2\,\theta . \tag{43}$$

If we set $\theta = 1/2$ we obtain the special form of the Newmark method used, e.g., in [4] and [6].

In the case of consolidation of clay we shall use only two schemes : for $\nu = 1$ the Euler backward method (the special case of (41) with $\theta = 0$)

$$\alpha_0 = -1, \quad \alpha_1 = 1, \quad \beta_0 = 0, \quad \beta_1 = 1 \tag{44}$$

and for $\nu = 2$ the special case of (42) with $\theta = 3/2, \delta = 1/4$:

$$\alpha_0 = 1/2\,, \quad \alpha_1 = -2, \alpha_2 = 3/2\,, \quad \beta_0 = \beta_1 = 0 \quad \beta_2 = 1\,. \tag{45}$$

If we use the two-step method (45) we must compute $\mathbf{u}_h^1$ by means of the one-step method (44). Let us note that in the case $\nu = 2$ we could use all schemes (42) satisfying $\beta_0 = 0$.

Remark 3. The symbol $\mathbf{u}_0^{\text{apr}}$ denotes a vector whose components approximates the right-hand sides u_{i0} of (6). The function T_0^{apr} is an approximation of the right-hand side of (5). If $T_0 = \overline{T}$ on Γ then we usually define $T_0^{\text{apr}} = T_0^{\text{int}}$ where T_0^{int} is the interpolate of T_0 in W_h. In the case $c_4 > 0$ the function $Y(x_1, x_2)$ is defined by

$$Y(x_1, x_2) = T(x_1, x_2, 0) + \Delta t \dot{T}(x_1, x_2, 0) \tag{46}$$

where $\dot{T}(x_1, x_2, 0)$ can be computed from equation (1) by means of initial conditions (5), (7). Similarly we define

$$\mathbf{z}(x_1, x_2) = \mathbf{u}(x_1, x_2, 0) + \Delta t \dot{\mathbf{u}}(x_1, x_2, 0) \tag{47}$$

or

$$\mathbf{z}(x_1, x_2) = \mathbf{u}(x_1, x_2, 0) + \Delta t\, \dot{\mathbf{u}}(x_1, x_2, 0) + 1/2\, \Delta t^2\, \ddot{\mathbf{u}}(x_1, x_2, 0)\,. \tag{48}$$

(Definitions (47) and (48) correspond to the cases $q = 1$ and $q = 2$ from Theorem 1, respectively.) The first two members on the right-hand sides of (47), (48) are given by initial conditions (6), (7). The third member in (48) can be computed from relations (2), (8), (10), (12) by means of initial conditions (5), (6). We define $\mathbf{u}_0^{\text{apr}}$ and $\mathbf{z}^{\text{apr}}$ to be the discrete Ritz approximations of $\mathbf{u}_0$ and $\mathbf{z}$, respectively. (In detail see Section 3.) This definition is a modification of the definition of starting values from [6]. T_0^{apr} and Y^{apr} can be defined similarly.

If we do not use the numerical integration then we define $\mathbf{u}_0^{\text{apr}}$ and $\mathbf{z}^{\text{apr}}$ to be the Ritz approximations of $\mathbf{u}_0$ and $\mathbf{z}$, respectively.

Remark 4 : In [4] the approximation of coupled linear thermoelasticity by the finite element method is also studied. However, the authors restrict themselves to the case $p = n = 1$; they do not analyze the effect of numerical integration and consider the Newmark method only with $\theta = 1/2$.

3. ERROR ESTIMATES

In this section we prove the existence and uniqueness of the approximate solution and establish the maximum rate of convergence. We shall start with some definitions and lemmas.

The symbols $\tilde{T}$ and $\tilde{\mathbf{u}}$ will denote the Calderon extensions of the exact solution T and $\mathbf{u}$, respectively. The function $\eta \in W_{hT}$ satisfying

$$\tilde{D}_h(\tilde{T} - \eta, w) = 0 \quad \forall w \in W_{h0} \tag{49}$$

is called the Ritz approximation of the function $\tilde{T}$. The function $\eta_d \in W_{hT}$ satisfying

$$D_h(\eta_d, w) = -(\tilde{T}_{,ii}, w)_h \quad \forall w \in W_{h0} \tag{50}$$

is called the discrete Ritz approximation of the function $\tilde{T}$. The vector $\mathbf{r} \in V_{hu}^1 \times V_{hu}^2$ satisfying

$$\tilde{a}_h(\tilde{\mathbf{u}} - \mathbf{r}, \mathbf{v}) = 0 \quad \forall \mathbf{v} \in [V_{h0}]^2 \tag{51}$$

is called the Ritz approximation of the vector $\tilde{\mathbf{u}}$. The vector $\mathbf{r}_d \in V_{hu}^1 \times V_{hu}^2$ satisfying

$$a_h(\mathbf{r}_d, \mathbf{v}) = -(D_{ijkm}\, \tilde{u}_{k,mj}, v_i)_h \quad \forall \mathbf{v} \in [V_{h0}]^2 \tag{52}$$

is called the discrete Ritz approximation of the vector $\tilde{\mathbf{u}}$.

LEMMA 1 : *Let the boundary Γ of the domain Ω be piecewise of class C^{p+1}. Let $T(x, t) \in H^{p+3}(\Omega)$, $t \in [0, t^*]$. Then*

$$\| \tilde{T} - \eta \|_{j,\Omega_h} \leqslant C h^{p+1-j} \| T \|_{p+3,\Omega} \quad (j = 0, 1), \tag{53}$$

C being a constant independent on T, h and t. In addition, let quadrature formulas on the unit triangle K_0 for calculation of the forms $D_h(v, w)$ and $(v, w)_h$ appearing in (50) *be of degree of precision $d = \max(1, 2p - 2)$. Then*

$$\| \tilde{T} - \eta_d \|_{j,\Omega_h} \leqslant C h^{p+1-j} \| T \|_{p+3,\Omega} \quad (j = 0, 1). \tag{54}$$

LEMMA 2 : *Let the boundary Γ of the domain Ω be piecewise of class C^{n+1}. Let $\mathbf{u}(x, t) \in [H^{n+1}(\Omega)]^2$, $t \in [0, t^*]$. Then*

$$\| \tilde{\mathbf{u}} - \mathbf{r} \|_{1,\Omega_h} \leqslant C h^n \| \mathbf{u} \|_{n+1,\Omega} \tag{55}$$

C being a constant independent on $\mathbf{u}$, h and t. In addition, let $\mathbf{u}(x, t) \in [H^{n+2}(\Omega)]^2$, $t \in [0, t^*]$ and let quadrature formulas on the unit triangle K_0 for calculation of the forms $a_h(\mathbf{v}, \mathbf{w})$ and $(\mathbf{v}, \mathbf{w})_h$ appearing in (52) *be of degree of precision $d = 2n - 2$. Then*

$$\| \tilde{\mathbf{u}} - \mathbf{r}_d \|_{1,\Omega_h} \leqslant C h^n \| \mathbf{u} \|_{n+2,\Omega} . \tag{56}$$

Lemma 1 is proved in [10, 11] in the case $W_{hT} = W_{h0}$. The generalization to the case $W_{hT} \neq W_{h0}$ is not difficult. Lemma 2 is an immediate consequence of the interpolation theorem (18), inequalities (67) and standard devices used in the analysis of the effect of numerical integration [5, 10, 11].

THEOREM 1 : *Let $c_1 > 0$, $c_4 > 0$, $p = n - 1$ and the boundary Γ of the domain Ω be piecewise of class C^{n+1}. Let*

$$T \in C^{q+1}(H^{n+2}(\Omega)) \tag{57}$$

$$u_i \in C^{q+3}(H^{n+2}(\Omega)) \quad (i = 1, 2) \tag{58}$$

Let quadrature formulas on the unit triangle K_0 for calculation of the forms $D_h(v, w)$ and $a_h(\mathbf{v}, \mathbf{w})$ be of degree of precision $2p - 2$ and $2n - 2$, respectively, let quadrature formulas on K_0 for calculation of the forms $(v, w)_h$ and $(\mathbf{v}, \mathbf{w})_h$ be of degree of precision $2p = 2n - 2$, let the quadrature formula on K_0 for calculation of the form $(v, w)_h^*$ be of degree of precision $2p - 1$ and let the quadrature formula on K_0 for calculation of the form $(v, w)_h^{**}$ be of degree of precision $2p - 1$ and such that the first inequality (68) holds. Let all weights of quadrature formulas used for calculation of the form $(\mathbf{v}, \mathbf{w})_h$ be positive. Then for sufficiently small h problem (36)-(39) has one and only one solution $\mathbf{u}_h^m$, T_h^m $(m = 2, \ldots, M)$ and in the case $\delta > 0$ (see (42)) the following estimate holds :

$$\| \tilde{T}^m - T_h^m \|_{0,\Omega_h} + \| \tilde{\mathbf{u}}^m - \mathbf{u}_h^m \|_{1,\Omega_h} \leqslant C \Big\{ \Delta t^q + h^n + \\ + \sum_{j=0}^{1} (\| \varepsilon^j \|_{0,\Omega_h} + \| \mathbf{e}^j \|_{1,\Omega_h}) + \Delta t^{-1} \| \Delta \mathbf{e}^0 \|_{1,\Omega_h} \Big\} \tag{59}$$

where $q = 2$ for $\theta = 1/2$ and $q = 1$ for $\theta > 1/2$ and where the constant C does not depend on Δt and h. The symbols $\mathbf{e}^m$ and ε^m are defined by the relations

$$\mathbf{e}^m = \mathbf{r}_d^m - \mathbf{u}_h^m, \quad \varepsilon^m = \eta_d^m - T_h^m. \tag{60}$$

In addition, let us denote

$$f^{m-1/2} = (f^m + f^{m-1})/2. \tag{61}$$

Then in the case $\delta = 0$, $\theta = 1/2$ the expression

$$\| \tilde{T}^m - T_h^m \|_{0,\Omega_h} + \| \tilde{\mathbf{u}}^{m-1/2} - \mathbf{u}_h^{m-1/2} \|_{1,\Omega_h}, \tag{62}$$

where $m = 2, \ldots, M$, is bounded by the right-hand side of (59).

Remark 5 : The following quadrature formulas of degree of precision $2p-1$ guarantee the validity of the first inequality (68) : for $p = 1$ the formula

$$\int_{K_0} v^*(\xi_1, \xi_2)\, d\xi = \frac{1}{6}(v^*(R_1) + v^*(R_2) + v^*(R_3))$$

and for $p = 2$ the formula (4.1.18) from [5].

Remark 6 : The norms appearing on the left-hand side of (59) are natural norms in thermoelasticity : 1) in applications we need to know the values of T, u_i and $u_{i,j}$; 2) T and $u_{i,j}$ should be computed with the same accuracy.

THEOREM 2 : *Let $c_1 = c_4 = 0, p = n - 1$ and the boundary Γ of the domain Ω be piecewise of class C^{n+1}. Let*

$$T \in C^1(H^{n+2}(\Omega)), \quad u_i \in C^{\nu+1}(H^{n+2}(\Omega)) \quad (i = 1, 2). \tag{63}$$

Let $Q \equiv 0$ and let the assumptions concerning the forms $D_h(v, w)$, $a_h(\mathbf{v}, \mathbf{w})$, $(v, w)_h$, $(\mathbf{v}, \mathbf{w})_h$ be the same as in Theorem 1. Then for sufficiently small h problem (36), (37), (40) *has one and only one solution $\mathbf{u}_h^m$, $T_h^m (m = 1, ..., M)$ and the following estimate holds* :

$$\| \tilde{T} - T_h \|_{l_2} + \| \tilde{\mathbf{u}}^m - \mathbf{u}_h^m \|_{1,\Omega_h} \leqslant C\{ \Delta t^{(\nu+1)/2} + h^n + \| \mathbf{u}^0 \|_{n+2,\Omega} h^n \Delta t^{-1/2} \} \tag{64}$$

where the constant C does not depend on Δt and h and the norm $\| . \|_{l_2}$ is defined by

$$\| f \|_{l_2}^2 = \Delta t \sum_{m=1}^{M} \| f^m \|_{0,\Omega_h}^2 . \tag{65}$$

If $\nu = 1$ then we use Euler's backward formula (44); *if $\nu = 2$ then we use two-step backward formula* (45) *for $m = 2, ..., M$ and one step formula* (44) *in the first step.*

Proof of theorems 1 *and* 2 : The assumptions of Theorems 1 and 2 and Friedrichs' and Korn's inequalities imply for sufficiently small h

$$C_1 \| w \|_{1,\Omega_h}^2 \leqslant D_h(w, w) \leqslant C_2 \| w \|_{1,\Omega_h}^2 \quad \forall w \in W_{h0} \tag{66}$$

$$C_1 \| \mathbf{v} \|_{1,\Omega_h}^2 \leqslant a_h(\mathbf{v}, \mathbf{v}) \leqslant C_2 \| \mathbf{v} \|_{1,\Omega_h}^2 \quad \forall \mathbf{v} \in [V_{h0}]^2 \tag{67}$$

$$C_1 \| w \|_{0,\Omega_h}^2 \leqslant (w, w)_h^{**} \leqslant C_2 \| w \|_{0,\Omega_h}^2 \quad \forall w \in W_h \tag{68}$$

$$0 \leqslant (\mathbf{v}, \mathbf{v})_h \leqslant C_2 \| \mathbf{v} \|_{1,\Omega_h}^2 \quad \forall \mathbf{v} \in [V_h]^2 \tag{69}$$

where C_1 and C_2 are positive constants independent on h.

a) As relations (36), (37) represent for each m a system of linear algebraic equations it is sufficient to prove the uniqueness of the approximate solution. Owing to implications (19) and (23) it suffices to prove that the following homogeneous problem has only a trivial solution : Find a vector $\mathbf{u}_h^{m+\nu} \in [V_{h0}]^2$ and a function $T_h^{m+\nu} \in W_{h0}$ such that

$$\Delta t\, \beta_\nu\, D_h(T_h^{m+\nu}, w) + c_1\, \alpha_\nu (T_h^{m+\nu}, w)_h^{**} + c_2\, \alpha_\nu (u_{hi,i}^{m+\nu}, w)_h = 0 \quad \forall w \in W_{h0}, \tag{70}$$

$$\Delta t^2\, \beta_\nu\, a_h(\mathbf{u}_h^{m+\nu}, \mathbf{v}) + c_4 (\mathbf{u}_h^{m+\nu}, \mathbf{v})_h - c_3\, \Delta t^2\, \beta_\nu (T_h^{m+\nu}, v_{i,i})_h = 0 \quad \forall \mathbf{v} \in [V_{h0}]^2. \tag{71}$$

Let us set $w = T_h^{m+\nu}$ in (70) and $\mathbf{v} = \mathbf{u}_h^{m+\nu}$ in (71), multiply equation (70) by $c_3\, \Delta t^2\, \beta_\nu$, equation (71) by $c_2\, \alpha_\nu$, add up the obtained equations and use the first inequalities (68), (69). We get

$$k_1\, D_h(T_h^{m+\nu}, T_h^{m+\nu}) + k_2\, a_h(\mathbf{u}_h^{m+\nu}, \mathbf{u}_h^{m+\nu}) \leqslant 0 \tag{72}$$

where $k_1 > 0$, $k_2 > 0$. The first inequalities (66), (67) together with (72) imply $T_h^{m+\nu} \equiv 0$, $\mathbf{u}_h^{m+\nu} \equiv \mathbf{0}$.

b) Let us set

$$\mathbf{s} = \tilde{\mathbf{u}} - \mathbf{r}_d, \qquad \xi = \tilde{T} - \eta_d. \tag{73}$$

Then, according to (60) and (73),

$$\| \tilde{T}^m - T_h^m \|_0 + \| \tilde{\mathbf{u}}^m - \mathbf{u}_h^m \|_1 \leqslant \| \xi^m \|_0 + \| \mathbf{s}^m \|_1 + \| \varepsilon^m \|_0 + \| \mathbf{e}^m \|_1 \tag{74}$$

in the case of Theorem 1 and

$$\| \tilde{T} - T_h \|_{l_2} + \| \tilde{\mathbf{u}}^m - \mathbf{u}_h^m \|_1 \leqslant \| \xi \|_{l_2} + \| \mathbf{s}^m \|_1 + \| \varepsilon \|_{l_2} + \| \mathbf{e}^m \|_1 \tag{75}$$

in the case of Theorem 2. In (74), (75) and in what follows we write for simplicity $\| \,.\, \|_k$ instead of $\| \,.\, \|_{k,\Omega_h}$ and $(.,.)_0$ instead of $(.,.)_{0,\Omega_h}$. Lemmas 1 and 2 imply

$$\| \xi^m \|_0 + \| \mathbf{s}^m \|_1 \leqslant Ch^n, \quad \| \xi \|_{l_2} + \| \mathbf{s}^m \|_1 \leqslant Ch^n. \tag{76}$$

Here and in what follows C denotes a constant not depending on h and Δt and not necessarily the same in any two places. It remains to estimate the last two terms on the right-hand sides of relations (74) and (75).

Let $\tilde{T}$, $\tilde{\dot{T}}$ and $\tilde{\dot{u}}_i$ be the Calderon extensions of the functions T, $\dot{T}$ and $\dot{u}_i$, respectively. Let us define a function $\tilde{Q}$, an extension of Q, by

$$\tilde{T}_{,ii} - c_1\, \tilde{\dot{T}} - c_2\, \tilde{\dot{u}}_{i,i} + \tilde{Q} = 0 \quad (x_1, x_2) \in \Omega. \tag{77}$$

Let us multiply relation (36) by -1 and to the both sides let us add the expression

$$\Delta t\, D_h\left(\sum_{j=0}^{\nu} \beta_j\, \eta_d^{m+j}, w\right) + c_1\left(\sum_{j=0}^{\nu} \alpha_j\, \eta_d^{m+j}, w\right)_h^{**} + c_2\left(\sum_{j=0}^{\nu} \alpha_j\, r_{di,i}^{m+j}, w\right)_h . \quad (78)$$

On the right-hand side of the obtained relation let us express the first term of (78) by means of (50) and (77); in the last two terms of (78) let us express η_d^{m+j} and $r_{di,i}^{m+j}$ by means of relations (73). As to the left-hand side of the obtained relation let us simplify it by means of (60). Multiplying the resulting relation by c_3 we obtain :

$$\begin{aligned} c_3\, \Delta t\, D_h\left(\sum_{j=0}^{\nu} \beta_j\, \varepsilon^{m+j}, w\right) &+ c_1\, c_3\left(\sum_{j=0}^{\nu} \alpha_j\, \varepsilon^{m+j}, w\right)_h^{**} + \\ &+ c_2\, c_3\left(\sum_{j=0}^{\nu} \alpha_j\, e_{i,i}^{m+j}, w\right)_h = c_1\, c_3(\pi^m - \omega^m, w)_h \\ &+ c_3\, \Delta t\left[\left(\sum_{j=0}^{\nu} \beta_j\, Q^{m+j}, w\right)_h - \left(\sum_{j=0}^{\nu} \beta_j\, Q^{m+j}, w\right)_h^{*}\right] \\ &+ c_1\, c_3\left[\left(\sum_{j=0}^{\nu} \alpha_j\, \eta_d^{m+j}, w\right)_h^{**} - \left(\sum_{j=0}^{\nu} \alpha_j\, \eta_d^{m+j}, w\right)_h\right] \\ &+ c_2\, c_3(\lambda_{i,i}^m - \rho_{i,i}^m, w)_h \quad \forall w \in W_{h0} \end{aligned} \quad (79)$$

where

$$\pi^m = \sum_{j=0}^{\nu} (\alpha_j\, \tilde{T}^{m+j} - \Delta t\, \beta_j\, \tilde{\tilde{T}}^{m+j}), \quad \omega^m = \sum_{j=0}^{\nu} \alpha_j\, \xi^{m+j}, \quad (80)$$

$$\lambda_i^m = \sum_{j=0}^{\nu} (\alpha_j\, \tilde{u}_i^{m+j} - \Delta t\, \beta_j\, \tilde{\tilde{u}}_i^{m+j}), \quad \rho_i^m = \sum_{j=0}^{\nu} \alpha_j\, s_i^{m+j}. \quad (81)$$

Let us define extensions $\tilde{X}_i$ $(i = 1, 2)$ of X_i by

$$D_{ijkm}\, \tilde{u}_{k,mj} - c_3\, \tilde{T}_{,i} + \tilde{X}_i = c_4\, \tilde{\ddot{u}}_i \quad (x_1, x_2) \in \tilde{\Omega} \quad (82)$$

where $\tilde{u}_i$, $\tilde{\ddot{u}}_i$ and $\tilde{T}$ denote the Calderon extensions of u_i, $\ddot{u}_i$ and T, respectively. (If $(x_1, x_2) \in \Omega$ then (82) is identical with (2), where σ_{ij} is expressed by means of (8), (10), (12).) Let us multiply (37) by -1 and to the both sides let us add the expression

$$\Delta t^2\, a_h\left(\sum_{j=0}^{\nu} \beta_j\, \mathbf{r}_d^{m+j}, \mathbf{v}\right) + c_4(\Delta^2 \mathbf{r}_d^m, \mathbf{v})_h - c_3\, \Delta t^2\left(\sum_{j=0}^{\nu} \beta_j\, \eta_d^{m+j} - T_r, v_{i,i}\right)_h . \quad (83)$$

Let us simplify the left-hand side of the obtained relation by means of (60). On the right-hand side let us express the first term of (83) by means of (52) and (82); in the second term of (83) let us express $\mathbf{r}_d^m$ by means of (73). Further, let us add to the right-hand side zero in the form

$$c_3\,\Delta t^2\left[\left(\sum_{j=0}^{\nu}\beta_j\,\tilde{T}^{m+j}_{,i},v_i\right)_0+\left(\sum_{j=0}^{\nu}\beta_j\,\xi^{m+j},v_{i,i}\right)_0+\right.$$
$$\left.+\left(\sum_{j=0}^{\nu}\beta_j\,\eta_d^{m+j}-T_r,v_{i,i}\right)_0\right]=0\,,\quad \mathbf{v}\in[V_{h0}]^2\,. \quad (84)$$

Multiplying the resulting relation by $c_2\,\Delta t^{-2}$ we obtain

$$c_2\,a_h\left(\sum_{j=0}^{\nu}\beta_j\,\mathbf{e}^{m+j},\mathbf{v}\right)-c_2\,c_3\left(\sum_{j=0}^{\nu}\beta_j\,\varepsilon^{m+j},v_{i,i}\right)_h+$$
$$+\,c_2\,c_4\,\Delta t^{-2}(\Delta^2\,\mathbf{e}^m,\mathbf{v})_h=c_2\,c_4\,\Delta t^{-2}(\mathbf{y}^m,\mathbf{v})_h$$
$$-\,c_2\,c_4\,\Delta t^{-2}(\Delta^2\mathbf{s}^m,\mathbf{v})_h+c_2\,c_3\left(\sum_{j=0}^{\nu}\beta_j\,\xi^{m+j},v_{i,i}\right)_0$$
$$+\,c_2\,c_3\left[\left(\sum_{j=0}^{\nu}\beta_j\,\eta_d^{m+j}-T_r,v_{i,i}\right)_0-\left(\sum_{j=0}^{\nu}\beta_j\,\eta_d^{m+j}-T_r,v_{i,i}\right)_h\right]$$
$$+\,c_2\,c_3\left[\left(\sum_{j=0}^{\nu}\beta_j\,\tilde{T}^{m+j}_{,i},v_i\right)_0-\left(\sum_{j=0}^{\nu}\beta_j\,T^{m+j}_{,i},v_i\right)_h\right]\quad\forall\mathbf{v}\in[V_{h0}]^2 \quad (85)$$

where

$$\mathbf{y}^m=\Delta^2\tilde{\mathbf{u}}^m-\Delta t^2\sum_{j=0}^{\nu}\beta_j\,\ddot{\tilde{\mathbf{u}}}^{m+j}\,. \quad (86)$$

Let us denote for the sake of brevity

$$\hat{w}=\sum_{j=0}^{\nu}\beta_j\,\varepsilon^{m+j}\in W_{h0}\,,\quad \hat{\mathbf{v}}=\sum_{j=0}^{\nu}\beta_j\,\mathbf{e}^{m+j}\in[V_{h0}]^2\,, \quad (87)$$

$$\bar{w}=\sum_{j=0}^{\nu}\alpha_j\,\varepsilon^{m+j}\in W_{h0}\,,\quad \bar{\mathbf{v}}=\sum_{j=0}^{\nu}\alpha_j\,\mathbf{e}^{m+j}\in[V_{h0}]^2\,. \quad (88)$$

Let us set $w=\hat{w}$ in (79) and $\mathbf{v}=\bar{\mathbf{v}}$ in (85) and sum up the obtained relations. (We do it in order to eliminate the term depending on both ε^m and $\mathbf{e}^m$.) After summing the result from $m=0$ to $m=s-\nu\,(s\leqslant M)$ we obtain

$$\sum_{m=0}^{s-\nu} (c_1 c_3 A^m + c_3 \Delta t B^m + c_2 H^m + c_2 c_4 \Delta t^{-2} J^m) = \sum_{m=0}^{s-\nu} (c_1 c_3 D^m + c_2 c_3 E^m + c_3 \Delta t F^m + c_1 c_3 G^m + c_2 c_4 \Delta t^{-2}(K^m - L^m) + c_2 c_3(M^m + N^m + P^m)) \tag{89}$$

where

$$A^m = (\overline{w}, \hat{w})_h^{**}, \quad B^m = D_h(\hat{w}, \hat{w}) \tag{90a, b}$$

$$D^m = (\pi^m - \omega^m, \hat{w})_h, \quad E^m = (\lambda_{i,i}^m - \rho_{i,i}^m, \hat{w})_h \tag{90d, e}$$

$$F^m = (f, \hat{w})_h - (f, \hat{w})_h^*, \quad f = \sum_{j=0}^{\nu} \beta_j Q^{m+j} \tag{90f}$$

$$G^m = (g, \hat{w})_h^{**} - (g, \hat{w})_h, \quad g = \sum_{j=0}^{\nu} \alpha_j \eta_d^{m+j} \tag{90g}$$

$$H^m = a_h(\hat{\mathbf{v}}, \overline{\mathbf{v}}), \quad J^m = (\Delta^2 \mathbf{e}^m, \overline{\mathbf{v}})_h \tag{90h, j}$$

$$K^m = (\mathbf{y}^m, \overline{\mathbf{v}})_h, \quad L^m = (\Delta^2 \mathbf{s}^m, \overline{\mathbf{v}})_h \tag{90k, l}$$

$$M^m = \left(\sum_{j=0}^{\nu} \beta_j \xi^{m+j}, \overline{v}_{i,i}\right)_0 \tag{90m}$$

$$N^m = (f, \overline{v}_{i,i})_0 - (f, \overline{v}_{i,i})_h, \quad f = \sum_{j=0}^{\nu} \beta_j \eta_d^{m+j} - T_r \tag{90n}$$

$$P^m = (g_{,i}, \overline{v}_i)_0 - (g_{,i}, \overline{v}_i)_h, \quad g = \sum_{j=0}^{\nu} \beta_j \tilde{T}^{m+j}. \tag{90p}$$

In the case of Theorem 2 we have $c_1 = c_4 = 0$, $c_3 = 1$, $Q \equiv 0$; thus (89) simplifies to the relation

$$\sum_{m=0}^{s-\nu} (c_2^{-1} \Delta t B^m + H^m) = \sum_{m=0}^{s-\nu} (E^m + M^m + N^m + P^m). \tag{91}$$

We estimate the left-hand sides of (89) and (91) from below and the right-hand sides from abowe. We start with relation (91). In order to estimate $\sum E^m$ let us write

$$E^m = (\lambda_{i,i}^m - \rho_{i,i}^m, \hat{w})_0 + [(\lambda_{i,i}^m, \hat{w})_h - (\lambda_{i,i}^m, \hat{w})_0] + \left[\left(\sum_{j=0}^{\nu} \alpha_j \tilde{u}_{i,i}^{m+j}, \hat{w}\right)_0 - \left(\sum_{j=0}^{\nu} \alpha_j u_{i,i}^{m+j}, \hat{w}\right)_h\right] - \left[\left(\sum_{j=0}^{\nu} \alpha_j r_{di,i}^{m+j}, \hat{w}\right)_0 - \left(\sum_{j=0}^{\nu} \alpha_j r_{di,i}^{m+j}, \hat{w}\right)_h\right]. \tag{92}$$

It holds

$$(\lambda^m_{i,i} - \rho^m_{i,i}, \hat{w})_0 \leqslant C(\| \boldsymbol{\lambda}^m \|_1 + \| \boldsymbol{\rho}^m \|_1) \| \hat{w} \|_1 .$$

Relation (81_1), Calderon's theorem and Taylor's theorem imply

$$\| \boldsymbol{\lambda}^m \|_1 \leqslant C \left\| \sum_{j=0}^{\nu} (\alpha_j \mathbf{u}^{m+j} - \Delta t \beta_j \dot{\mathbf{u}}^{m+j}) \right\|_{n+2,\Omega} \leqslant C \Delta t^{\nu+1} .$$

As $\sum \alpha_j \mathbf{r}_d^{m+j}$ is the discrete Ritz approximation of $\sum \alpha_j \tilde{\mathbf{u}}^{m+j}$ relations (73_1), (81_2), Lemma 2 and Taylor's theorem imply

$$\| \boldsymbol{\rho}^m \|_1 \leqslant C h^n \left\| \sum_{j=0}^{\nu} \alpha_j \mathbf{u}^{m+j} \right\|_{n+2,\Omega} \leqslant C \Delta t h^n .$$

In estimating the remaining terms on the right-hand side of (92) we use the estimates

$$| (f_{,i}, w)_h - (\tilde{f}_{,i}, w)_0 | \leqslant C h^n \| f \|_{n+1,\Omega} \| w \|_1 ,$$

$$| (r_{di,j}, w)_h - (r_{di,j}, w)_0 | \leqslant C h^{n+1} \| \mathbf{u} \|_{n+2,\Omega} \| w \|_1$$

which follow from [5, Theorems 4.1.5 and 4.4.5], [10, Lemma 7] and from Lemma 2. Then we use Taylor's theorem and find that the sum of the last three terms in (92) is bounded by $C h^n \Delta t \| \hat{w} \|_1$. Thus

$$\sum_{m=0}^{s-\nu} E^m \leqslant C \Delta t (h^n + \Delta t^\nu) \sum_{m=0}^{s-\nu} \left\| \sum_{j=0}^{\nu} \beta_j \varepsilon^{m+j} \right\|_1 . \tag{93}$$

Estimate (93) holds also in the case of Theorem 1 with $\nu = 2$.

In order to estimate $\sum M^m$ in a proper form we use so called summation by parts. Let us write

$$\bar{\mathbf{v}} = \Delta \mathbf{f}^m \tag{94}$$

where, according to (88_2) and (41)-(45),

$$\mathbf{f}^m = \mathbf{e}^m (\nu = 1) , \quad \mathbf{f}^m = \theta \Delta \mathbf{e}^m + \mathbf{e}^m (\nu = 2) . \tag{95}$$

Then, according to $(90m)$,

$$\sum_{m=0}^{s-\nu} M^m = - \sum_{m=1}^{s-\nu} \left(\sum_{j=0}^{\nu} \beta_j \Delta \xi^{m-1+j}, f^m_{i,i} \right)_0 + \\ + \left(\sum_{j=0}^{\nu} \beta_j \xi^{s-\nu+j}, f^{s-\nu+1}_{i,i} \right)_0 - \left(\sum_{j=0}^{\nu} \beta_j \xi^j, f^0_{i,i} \right)_0 . \tag{96}$$

Relations (73_2), (57), Lemma 1 and Taylor's theorem imply

$$\| \xi^m \|_0 \leqslant Ch^{p+1} , \quad \| \Delta\xi^m \|_0 \leqslant Ch^{p+1} \Delta t .$$

Thus it follows from (94)-(96)

$$\sum_{m=0}^{s-\nu} M^m \leqslant Ch^n S_\nu \tag{97}$$

where we denote for the sake of brevity

$$S_\nu = \Delta t \sum_{m=0}^{s-1} \| \mathbf{e}^m \|_1 + \sum_{j=0}^{\nu-1} (\| \mathbf{e}^{s-j} \|_1 + \| \mathbf{e}^j \|_1) . \tag{98}$$

The estimate

$$\sum_{m=0}^{s-\nu} (N^m + P^m) \leqslant Ch^n S_\nu \tag{99}$$

can be obtained by means of (90*n*), (90*p*), summation by parts, standard devices of the analysis of numerical integration, Calderon's and Taylor's theorems.

According to the first inequality (66), (90*b*) and (87_1),

$$\Delta t \sum_{m=0}^{s-\nu} B^m \geqslant C \, \Delta t \sum_{m=0}^{s-\nu} \left\| \sum_{j=0}^{\nu} \beta_j \, \varepsilon^{m+j} \right\|_1^2 . \tag{100}$$

Because of expression (98) we shall need the estimate of $\sum H^m$ from below in the form

$$\sum_{m=0}^{s-\nu} H^m \geqslant C \sum_{j=0}^{\nu-1} (\| \mathbf{e}^{s-j} \|_1^2 - \| \mathbf{e}^j \|_1^2) . \tag{101}$$

In the case $\nu = 1$ we have $\theta \leqslant 1/2$; thus using (41), (87), (88), (90*h*) and (67) we obtain easily

$$H^m \geqslant C(\| \mathbf{e}^{m+1} \|_1^2 - \| \mathbf{e}^m \|_1^2)$$

from where (101) follows. In the case $\nu = 2$, $\delta > 0$ using (42), (90*h*) and (67) we can derive in a way similar to [10, p. 430] :

$$\sum_{m=0}^{s-2} H^m \geqslant R - C \sum_{j=0}^{1} \| \mathbf{e}^j \|_1^2 \tag{102}$$

where

$$R = \frac{1}{2}(\theta^2 + \delta) \| \mathbf{e}^s \|^2 + \frac{1}{2}[(\theta - 1)^2 + \delta] \| \mathbf{e}^{s-1} \|^2 - \| \mathbf{e}^s \| . \| \gamma \mathbf{e}^{s-1} \|$$

with $\| \mathbf{w} \|^2 = a_h(\mathbf{w}, \mathbf{w})$ and $\gamma = \theta(\theta - 1) + \delta$. In the case of Theorem 2

$$R = \frac{5}{4} \| \mathbf{e}^s \|^2 + \frac{1}{4} \| \mathbf{e}^{s-1} \|^2 - \| \mathbf{e}^s \| . \| \mathbf{e}^{s-1} \| .$$

Using the inequality

$$| ab | \leqslant \frac{\tau}{2} a^2 + \frac{1}{2\tau} b^2 \tag{103}$$

with $\tau = 2$ we obtain $R \geqslant \| \mathbf{e}^s \|^2/4$. Using (103) with $\tau = 2/5$ we obtain $R \geqslant \| \mathbf{e}^{s-1} \|^2/20$. Thus $R \geqslant (\| \mathbf{e}^s \|^2 + \| \mathbf{e}^{s-1} \|^2)/40$ and, according to (67), $R \geqslant C(\| \mathbf{e}^s \|_1^2 + \| \mathbf{e}^{s-1} \|_1^2)$. Inserting this result into (102) we obtain (101). (In the case of Theorem 1 we can derive (101) similarly.)

Now we sketch the proof of Theorem 1. The terms depending on E^m, M^m, N^m and P^m are estimated in (93), (97) and (99). In the case of K^m and L^m we do first summation by parts. Then in the case of D^m, F^m, G^m, K^m and L^m we add to each term and subtract from each term the corresponding continuous form $(.,.)_0$; e.g.

$$G^m = (g, \hat{w})_h^{**} - (g, \hat{w})_0 + (g, \hat{w})_0 - (g, \hat{w})_h .$$

Then we use the standard results of the analysis of numerical integration. At the end we estimate the continuous forms

$$(\pi^m - \omega^m, w)_0 , \quad (\Delta^{k-2} \mathbf{y}^m, \mathbf{v})_0 , \quad (\Delta^k \mathbf{s}^m, \mathbf{v})_0$$

where $k = 2, 3$. We do it by means of the estimates

$$\| \pi^m \|_0 \leqslant C \, \Delta t^{q+1} , \quad \| \omega^m \|_0 \leqslant C \, \Delta t h^n ,$$

$$\| \Delta^{k-2} \mathbf{y}^m \|_0 \leqslant C \, \Delta t^{q+k} , \quad \| \Delta^k \mathbf{s}^m \|_1 \leqslant C \, \Delta t^k \, h^n$$

which can be obtained by means of Lemmas 1, 2 and Calderon's and Taylor's theorems. The result is that the right-hand side of (89) (briefly R.H.S.) satisfies the estimate

$$\text{R.H.S.} \leqslant C(\Delta t^q + h^n) \left[S_2 + \Delta t \sum_{m=0}^{s-2} \left\| \sum_{j=0}^{2} \beta_j \, \varepsilon^{m+j} \right\|_1 \right] \tag{104}$$

where S_2 is given by (98). The estimate

$$\sum_{m=0}^{s-2} A^m \geqslant C \left(\| \varepsilon^s \|_0^2 - \sum_{j=0}^{1} \| \varepsilon^j \|_0^2 \right) \tag{105}$$

follows from [10, pp. 429-431] and from (68).

Using (42) and the fact that $\theta \geqslant 1/2$ we obtain from (90j)

$$J^m = (\Delta \mathbf{e}^{m+1} - \Delta \mathbf{e}^m, \theta\, \Delta \mathbf{e}^{m+1} + (1-\theta)\, \Delta \mathbf{e}^m)_h \geqslant$$
$$\geqslant \frac{1}{2}(\Delta \mathbf{e}^{m+1}, \Delta \mathbf{e}^{m+1})_h - \frac{1}{2}(\Delta \mathbf{e}^m, \Delta \mathbf{e}^m)_h .$$

Thus, according to (69),

$$\sum_{m=0}^{s-2} J^m \geqslant - C \| \Delta \mathbf{e}^0 \|_1^2 . \tag{106}$$

Combining (89), (100), (101), (104)-(106), using several times inequality (103) and then the discrete form of Gronwall's lemma we obtain

$$\| \varepsilon^s \|_0 + \| \mathbf{e}^s \|_1 \leqslant C \Big\{ \Delta t^q + h^n + \Delta t^{-1} \| \Delta \mathbf{e}^0 \|_1 +$$
$$+ \sum_{j=0}^{1} (\| \varepsilon^j \|_0 + \| \mathbf{e}^j \|_1) \Big\} . \tag{107}$$

Inequalities (74), (76) and (107) imply (59).

In the case $\delta = 0$ it is impossible to prove (101). However, in the case $\delta = 0$, $\theta = 1/2$ it holds, according to (61), (87), (88),

$$\hat{\mathbf{v}} = \frac{1}{2}(\mathbf{e}^{m+3/2} + \mathbf{e}^{m+1/2}) , \quad \bar{\mathbf{v}} = \Delta \mathbf{e}^{m+1/2} .$$

Thus

$$\sum_{m=0}^{s-2} H^m \geqslant C \{ \| \mathbf{e}^{s-1/2} \|_1^2 - \| \mathbf{e}^{1/2} \|_1^2 \} . \tag{108}$$

According to (61) and (95), $\mathbf{f}^m = \mathbf{e}^{m+1/2}$ if $\theta = 1/2$. Thus we can estimate the right-hand side of (89) (briefly R.H.S.) in the form

$$\text{R.H.S.} \leqslant C(\Delta t^q + h^n) \left[\bar{S}_2 + \Delta t \sum_{m=0}^{s-2} \left\| \sum_{j=0}^{2} \beta_j\, \varepsilon^{m+j} \right\|_1 \right] \tag{109}$$

where

$$\bar{S}_2 = \| \mathbf{e}^{s-1/2} \|_1 + \| \mathbf{e}^{1/2} \|_1 + \Delta t \sum_{m=0}^{s-2} \| \mathbf{e}^{m+1/2} \|_1 . \tag{110}$$

Combining (61), (89), (100), (105), (106), (108) and (109), using several times inequality (103) and then the discrete form of Gronwall's lemma we find that

$\| \varepsilon^s \|_0 + \| \mathbf{e}^{s-1/2} \|_1$ is bounded by the right-hand side of (107). From here we easily find that (62) is bounded by the right-hand side of (59). Theorem 1 is proved.

Now we sketch the proof of Theorem 2 in the case $\nu = 1$: In (91) we sum only from $m = 1$ to $m = s - 1$. In the case $m = 0$, as $u_{i,i}^0 = 0$ we can replace in (79) and (81) all expressions of the type Δf^0 by f^1. Setting $w = \varepsilon^1$ in such a changed relation (79) and $\mathbf{v} = \mathbf{e}^1$ in (85) and summing up the obtained relations we get a relation which will be added to (91). Using the fact that $u_{i,i}^1 = \Delta u_{i,i}^0$ and modifying a little the preceding devices we obtain (64). In the case $\nu = 2$ the proof is similar.

It should be noted that the term $\| \mathbf{u}^0 \|_{n+2,\Omega} h^n \Delta t^{-1/2}$ is a consequence of estimating in the first step.

Remark 7 : If we modify the proof of Theorem 2 and use the Ritz approximation $\mathbf{r}$ instead of the discrete Ritz approximation $\mathbf{r}_d$ we can weaken assumption (63) and assume $u_i \in C^{\nu+1}(H^{n+1}(\Omega))$ only. Similar results can be obtained in case of $c_1 > 0$, $c_4 = 0$.

Remark 8 : The starting values defined in Remark 3 do not spoil the rate of convergence. E.g., according to the definition of $\mathbf{u}_h^0$, we have $\mathbf{e}^0 \equiv \mathbf{0}$; thus

$$\| \Delta \mathbf{e}^0 \|_1 = \| \mathbf{e}^1 \|_1 = \| (\mathbf{r}_d^1 - \mathbf{z}^{\text{apr}}) - (\mathbf{u}^1 - \mathbf{z}) + (\mathbf{u}^1 - \mathbf{z}) \|_1 \leqslant$$
$$\leqslant \| \mathbf{u}^1 - \mathbf{z} \|_1 + Ch^n \| \mathbf{u}^1 - \mathbf{z} \|_{n+2,\Omega} \leqslant C \Delta t^{q+1}$$

because $\mathbf{r}_d^1 - \mathbf{z}^{\text{apr}}$ is the discrete Ritz approximation of $\mathbf{u}^1 - \mathbf{z}$ and $\mathbf{z}$ is defined by (47) or by (48). The estimates of $\| \varepsilon^0 \|_0$ and $\| \varepsilon^1 \|_0$ can be obtained similarly.

Remark 9 : Assumptions (57), (58), (63) can be weakened without loosing the maximum rate of convergence. E.g., in the case of thermoelasticity we can assume

$$T \in C^q(H^{n+2}(\Omega)), \quad \partial^{q+1} T / \partial t^{q+1} \in L^2(H^{n+2}(\Omega))$$
$$u_i \in C^{q+2}(H^{n+2}(\Omega)), \quad \partial^{q+3} u_i / \partial t^{q+3} \in L^2(H^{n+2}(\Omega))$$

instead of (57), (58). The only change in the proof of Theorem 1 is that we use Taylor's theorem with the integral remainder. In the case of Theorem 2 the situation is the same.

REFERENCES

[1] D. AUBRY, J. C. HUJEUX, *Special algorithms for elastoplastic consolidation with finite elements,* Third International Conference on Numerical Methods in Geomechanics (Aachen), 2-6 April 1979.

[2] B. A. BOLEY, J. H. WEINER, *Theory of Thermal Stresses,* John Wiley and Sons, New York-London-Sydney, 1960.

[3] J. R. BOOKER, *A numerical method for the solution of Biot's consolidation theory,* Quart. J. Mech. Appl. Math. 26, 1973, pp. 457-470.

[4] S.-I. CHOU, C.-C. WANG, *Estimates of error in finite element approximate solutions to problems in linear thermoelasticity, Part 1, Computationally coupled numerical schemes.* Arch. Rational Mech. Anal. 76, 1981, pp. 263-299.

[5] P. G. CIARLET, *The Finite Element Method for Elliptic Problems,* North-Holland, Amsterdam, 1978.

[6] T. DUPONT, *L^2-estimates for Galerkin methods for second order hyperbolic equations,* SIAM J. Numer. Anal. 10, 1973, pp. 880-889.

[7] G. FICHERA, *Uniqueness, existence and estimate of the solution in the dynamical problem of thermodiffusion in an elastic solid,* Arch. Mech. (Arch. Mech. Stos.) 26, 1974, pp. 903-920.

[8] C. JOHNSON, *A finite element method for consolidation of clay,* Research Report 77.05 R, Chalmers University of Technology, Göteborg, 1977.

[9] J. KRATOCHVÍL, F. LEITNER, *Solution of problems of stress and strain of fully saturated porous media* (In Czech.) Staveb, Čas. 27, 1979, pp. 23-27.

[10] J. NEDOMA, *The finite element solution of parabolic equations,* Apl. Mat. 23, 1978, pp. 408-438.

[11] J. NEDOMA, *The finite element solution of elliptic and parabolic equations using simplicial isoparametric elements,* RAIRO Anal. Numér. 13, 1979, pp. 257-289.

[12] M. ZLÁMAL, *The finite element method in domains with curved boundaries,* Internat. J. Numer. Methods Engrg. 5, 1973, pp. 367-373.

[13] M. ZLÁMAL, *Curved elements in the finite element method,* II. SIAM J. Numer. Anal. 11, 1974, pp. 347-362.

[14] M. ZLÁMAL, *Finite element methods for nonlinear parabolic equations,* RAIRO Anal. Numér. 11, 1977, No. 1, pp. 93-107.

[15] A. ŽENÍŠEK, *Curved triangular finite C^m-elements,* Apl. Mat. 23, 1978, pp. 346-377.

[16] A. ŽENÍŠEK, *The existence and uniqueness theorem in Biot's consolidation theory.* (To appear in Apl. Mat. 29, 1984.)