

Universitatea “Babeş-Bolyai” Cluj Napoca
Facultatea de Matematică și Informatică

METODA ELEMENTULUI FINIT:
APLICAȚII LA ECUAȚIA MEDIULUI POROS

referat în cadrul programului de pregătire a doctoratului

Conducător științific
Prof. Dr. Gheorghe Coman

Doctorand
Iuliu Sorin Pop

1997

Cuprins

Capitolul 1. Introducere	2
Capitolul 2. Ecuația mediului poros	4
1. Problema model	4
2. Metode numerice	8
Capitolul 3. Scheme numerice bazate pe principiul de maxim	12
1. Ideea de bază	12
2. Metoda Euler implicită	13
3. Metoda semi-implicită	23
4. Analiza discretizării complete	26
5. Exemple numerice	32
Capitolul 4. Concluzii	39
Bibliografie	41

CAPITOLUL 1

Introducere

În lucrarea de față sunt prezentate metode numerice specifice ecuațiilor parabolice degenerate. Problema tratată cu precădere, ecuația mediului poros, reprezintă un exemplu clasic pentru acest gen de probleme, dificultățile pe care le ridică analiza atât din punct de vedere teoretic cât și numeric fiind tipice în categoria menționată (regularitate redusă a soluției, frontiere libere). Numeroase modele matematice conduc la ecuații similare celei considerate aici. Dintre aplicațiile concrete vizate amintim în primul rând modelarea mișcării gazelor ideale izotropice printr-un mediu poros respectiv a infiltrațiilor subterane, teoria gazelor ionizate la temperaturi înalte, modele ale fizicii plasmei precum și ale dinamicii de populație. Toate acestea au făcut ca ecuația mediului poros (alături de alte probleme având caracteristici similare) să fie intens studiată în ultimele decenii, rezolvarea lor aproximativă constituind una din provocările tentante ale analizei numerice.

Prin prezentul referat s-a urmarit, totodată, și introducerea în metoda elementului finit. Acest domeniu a devenit unul din centrele de greutate ale analizei numerice moderne, reprezentând practic tehnica standard de discretizare a ecuațiilor diferențiale cu derivate parțiale.

În prima parte a lucrării sunt prezentate pe scurt suportul fizic al ecuației mediului poros precum și unele aspecte teoretice legate de ecuațiile parabolice degenerate în general și de ecuația mediului poros în special. Totodată, sunt trecute în revistă unele tehnici de abordare numerică a problemei. Partea a doua este dedicată metodelor numerice bazate pe principiul de maxim. După prezentarea ideii de bază sunt analizate semidiscretizările temporale implicită respectiv semi-implicită împreună cu discretizarea completă. Estimările de eroare obținute dovedesc valabilitatea din punct de vedere teoretic a abordării propuse, iar exemplele numerice constituie un argument suplimentar în acest sens. Concluziile finale, precum și bibliografia utilizată completează acest referat.

Lucrarea de față a fost elaborată în cadrul Centrului Interdisciplinar pentru Calcul Științific al Universității din Heidelberg, Germania (IWR). Doresc să-mi exprim pe această cale recunoștința față de prof. dr. Willi Jäger pentru sprijinul oferit precum și pentru îndrumarea sa generoasă. Discuțiile purtate în tot acest timp cu membrii grupului de lucru (în special cu dr. Wen-An Yong) mi-au fost un real sprijin pentru înțelegerea

unor aspecte mai delicate ale problematii. Sunt îndatorat d-lui dr. Nicolas Neuß pentru ajutorul acordat în vederea acomodării cu programul UG (utilizat la obținerea rezultatelor numerice în cazul bidimensional) precum și pentru sugestiile sale. Nu în ultimul rând, calde mulțumiri Catedrei de Calcul Numeric și Statistic din cadrul Facultății de Matematică, în special conducătorului științific, prof. dr. Gheorghe Coman, pentru sprijinul acordat în vederea continuării studiilor în cadrul IWR precum și pentru permanentele încurajări.

Ecuția mediului poros

În acest capitol sunt cuprinse unele aspecte legate de ecuațiile parabolice degenerate în general și de ecuația mediului poros în particular. De asemenea, este inclusă și o descriere succintă a unora dintre algoritmi apăruiți în literatură.

1. Problema model

Această primă secțiune cuprinde o scurtă introducere în problema model studiată, precum și în unele generalizări ale acesteia.

1.1. Enunțarea problemei. Problema considerată pe tot parcursul acestei lucrări drept referință se obține în urma modelării mișcării unui gaz ideal izentropic printr-un mediu poros omogen. Legile care guvernează această mișcare sunt cele de mai jos [8].

Ecuția de stare: $p = p_0 \rho^\alpha$,

unde $p = p(x, t)$ reprezintă presiunea, $\rho = \rho(x, t)$ densitatea iar $p_0 \in \mathbb{R}^+$ și $\alpha \geq 1$ sunt constante ($x \in \mathbb{R}^d, d \geq 1$).

Legea conservării masei: $\kappa \partial_t \rho + \operatorname{div}(\rho v) = 0$,

unde prin $v = v(x, t)$ am desemnat vectorul vitezei iar $\kappa \in \mathbb{R}^+$ este porozitatea mediului (fracțiunea volumului disponibilă gazului).

Legea lui Darcy: $\nu v = -\mu \nabla p$,

$\nu, \mu \in \mathbb{R}^+$ fiind vâscozitatea gazului respectiv permeabilitatea mediului

Eliminând viteza și presiunea din ecuațiile de mai sus rezultă *ecuația mediului poros*

$$(1.1) \quad \partial_t u = \Delta(u^m), \quad m = 1 + \alpha,$$

în care u reprezintă densitatea scalată și prin urmare este o mărime pozitivă.

Această ecuație prezintă câteva din caracteristicile tipice ale ecuațiilor parabolice *degenerate* (ecuații de tip parabolic care pot degenera în alt tip - eliptic ori hiperbolic - datorită anulării difuziei). În regiunile în care u este mărginită inferior deasupra lui 0, problema mai sus menționată este de tip parabolic regulat (ori uniform), dar devine degenerată în vecinătatea punctelor în care u se anulează. Frontiera mulțimii în care u este nulă reprezintă *frontiera liberă*, iar apariția acesteia se datorează dispariției difuziei. Acesta reprezintă cea mai importantă caracteristică a ecuațiilor parabolice degenerate și poate fi

interpretată și sub forma unei viteze finite de propagare a perturbațiilor, în contrast cu ecuațiile regular-parabolice (cum ar fi ecuația căldurii), unde această viteză este infinită. Un alt aspect care distinge ecuațiile parabolice degenerare de cele nedegenerate constă în regularitatea redusă a soluțiilor. În general, deși există suficiente rezultate legate de (Hölder-) continuitatea soluției u , nu același lucru se poate spune despre derivata (slabă) a ei. Pentru cazul nostru particular există exemple de soluții care nu sunt incluse în $H^1(\Omega)$ (spațiul *Sobolev* al funcțiilor care admit derivate generalizate de ordinul 1 la pătrat sumabile pe domeniul Ω), acest lucru putând fi acceptat doar pentru neliniaritatea u^m .

Numeroase lucrări studiază aceste aspecte sub diverse forme. Pentru o prezentare exhaustivă a ecuației mediilor poroase vom face referire la [4]. Încheiem această succintă prezentare cu un exemplu clasic de soluție exactă dat de G. I. Barenblatt, soluție radial simetrică având ambele trăsături descrise mai sus. În [6] se demonstrează că, pornind cu date inițiale potrivite (reprezentând valoarea funcției de mai jos la momentul $t = 0$), ecuația (1.1) este satisfăcută de

$$(1.2) \quad u(t, x) = (t + 1)^{-\frac{d}{(md+2-d)}} \left\{ \left[1 - \frac{m-1}{2m(dm+2-d)} \left(\frac{x}{(t+1)^{\frac{1}{md+2-d}}} \right)^2 \right]_+ \right\}^{\frac{1}{m-1}}.$$

În acest caz d este dimensiunea spațiului, iar prin $[x]_+$ am desemnat valoarea maximă dintre x și 0. Așa cum se poate remarca din (1.2), soluția Barenblatt are suport compact în orice moment finit.

Vom continua acum cu prezentarea pe scurt a unei categorii generale de probleme de tipul celei de mai sus. Pentru aceasta este suficient să ne referim la lucrările lui H. W. Alt și S. Luckhaus ([1]) ori J. Kačur ([20]), unde sunt considerate probleme de valori inițiale cu valori la frontieră pentru ecuația cvasiliniară degenerată de tip parabolic

$$(1.3) \quad \begin{aligned} \partial_t b(u) - \nabla a(b(u), \nabla \beta(u)) &= f(u) && \text{în } Q_T \equiv (0, T) \times \Omega, \\ b(u)(x, 0) &= b(x, u_0) && \text{în } \Omega, \\ u &= u^D && \text{pe } (0, T) \times \Gamma, \\ a(b(u), \nabla u) \cdot \nu &= 0 && \text{pe } (0, T) \times (\partial\Omega \setminus \Gamma). \end{aligned}$$

Funcțiile b și β sunt continue, monoton crescătoare și se anulează în 0. Despre a se presupune că generează un operator continuu și monoton în variabila corespunzătoare gradientului (nu neapărat liniar) iar $f(\cdot)$ este, de asemenea, o funcțională continuă. În plus, acestea satisfac o condiție de creștere limitată, care în cazul nostru se rezumă la o creștere cel mult liniară în $\beta(u)$. Pentru toate detaliile ne vom referi la lucrările citate mai sus.

Să remarcăm că în particular, luând $b(u) = [u]_+^{\frac{1}{m}}$, $\beta(u) = u$ și $a(b, \nabla u) = \nabla u$ obținem ecuația mediului poros (aceeași situație putând fi obținută și direct, pentru $b(u) = u$ și $\beta(u) = u^m$). Contextul este însă mult mai larg, cuprinzând modele din filtrarea nestaționară, probleme Stefan (de topire), putând fi extins și la inegalități variaționale.

1.2. Cadrul funcțional. Pe tot parcursul acestui referat vom considera mulțimea $\Omega \in \mathbb{R}^n$ deschisă, mărginită, conexă având frontieră lipschitziană. Timpul T este considerat finit, iar prin I desemnăm intervalul $(0, T)$. De asemenea, vom folosi mulțimea $Q_T = I \times \Omega$. Ne vom mărgini la cazul $\partial\Omega = \Gamma$ (deci condiții la frontieră de tip Dirichlet).

Vom trece acum la formularea slabă a problemei de mai sus. Pentru aceasta facem uz de următoarele notații ([14]). Funcțiile având puterea a p -a a modulului integrabilă (Lebesgue) pe Ω sunt notate cu $L^p \equiv L^p(\Omega)$, iar cele esențial mărginite cu $L^\infty \equiv L^\infty(\Omega)$. Prin $H^{1,p} \equiv H^{1,p}(\Omega)$ am desemnat spațiul Sobolev al funcțiilor ale căror derivate de ordin întâi sunt în L^p . În cazul $p = 2$ (spațiu Hilbert), indicele al doilea este omis. $L^p(I, V)$ conține funcțiile având puterea a p -a a modulului (Bochner-) integrabilă în raport cu prima variabilă (timpul) pe intervalul I și luând valori în spațiul Banach V . Pe parcursul acestei lucrări V va reprezenta de regulă mulțimea $H_0^1(\Omega)$, adică subspațiul lui $H^1(\Omega)$ având elemente a căror urmă la frontiera domeniului Ω este nulă. Prin (\cdot, \cdot) înțelegem produsul scalar uzual pe $L^2(\Omega)$, ori dualitatea dintre $H_0^1(\Omega)$ și $H^{-1}(\Omega)$ (acesta din urmă fiind notat adeseori cu V^*). Totodată, $\|\cdot\|_1$, $\|\cdot\|_\infty$, $|\cdot|$, $\|\cdot\|_1$, $\|\cdot\|_{-1}$, $\|\cdot\|_{L^p(I, V)}$ reprezintă normele, respectiv seminormele uzuale în spațiile L^2 , L^∞ , $H^1(\Omega)$, $H^{-1}(\Omega)$, $L^p(I, V)$. În cele ce urmează C va desemna orice constantă generică pozitivă.

DEFINIȚIA 1.1. O funcție măsurabilă $u : Q_T \longrightarrow \mathbb{R}$ este o soluție variațională a problemei în (1.3) dacă și numai dacă

- (i) $b(u) \in L^2(Q_T)$, $\partial_t b(x, u) \in L^2(I, V^*)$, $\beta(u) \in L^2(I, V)$;
- (ii) $\int_I (\partial_t b(u), v) = - \int_{Q_T} (b(u) - b(u_0)) \cdot \partial_t v$, pentru orice $v \in L^2(I, V) \cap L^\infty(Q_T)$ cu $\partial_t v \in L^\infty(Q_T)$, $v(T) = 0$;
- (iii) $\int_I (\partial_t b(u), v) + \int_I (a(u, \nabla \beta(u)), \nabla v) = \int_I (f(u), v)$, pentru orice $v \in L^2(I, V)$.

În [1] și [20] este demonstrată existența unei soluții slabe pentru problema de mai sus. Unicitatea poate fi obținută numai presupunând unele proprietăți suplimentare ale ecuației (1.3).

Întrucât în lucrarea de față ne-am propus să tratăm din punct de vedere numeric ecuația mediului poros, revenim la un caz particular al problemei anterioare, care are însă unele proprietăți mai bune ale soluției slabe. Concret, vom considera problema model

$$(1.4) \quad \begin{aligned} \partial_t u - \Delta \beta(u) &= 0, & \text{în } Q_T \equiv (0, T) \times \Omega, \\ u(x, 0) = u_0(x) &\geq 0, & \text{în } \Omega, \\ u &= 0, & \text{pe } \partial\Omega. \end{aligned}$$

Similar celor de mai sus, definiția soluției slabe a problemei (1.4) este

DEFINIȚIA 1.2. O funcție măsurabilă $u : Q_T \rightarrow \mathbb{R}$ este o soluție variațională a problemei în (1.4) dacă și numai dacă

$$u \in L^\infty(0, T; L^2(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)), \quad \beta(u) \in L^2(0, T; H_0^1(\Omega)), \quad u(0) = u_0$$

și pentru orice $\varphi \in H_0^1(\Omega)$ are loc

$$(1.5) \quad (\partial_t u(t), \varphi) + (\nabla \beta(u(t)), \nabla \varphi) = 0$$

pentru aproape orice $t \in (0, T)$.

În condițiile presupuse a fi satisfăcute de către β , problema de mai sus are o unică soluție (conform [1], [18] și a referințelor citate în aceste lucrări).

Problema principală care apare în cazul analizei numerice a ecuațiilor parabolice degenerate nu dipare nici în acest caz. Considerând că $\beta : \mathbb{R} \rightarrow \mathbb{R}$ este o funcție netedă strict crescătoare putem întâlni și aici situația difuziei neliniare degenerate. Degenerarea apare în cazul $\beta'(u) = 0$ pentru o anumită valoare u . Pe tot parcursul lucrării, fără a restrânge generalitatea, presupunem că $\beta(0) = 0$. Pentru a fi mai exacti, vom face următoarele presupuneri

$$(A1) \quad u_0 \geq 0, \quad u_0 \in L^\infty(\Omega) \text{ și } \beta(u_0) \in H_0^1(\Omega).$$

$$(A2) \quad \beta \text{ este continuu diferențiabilă, } \beta(0) = 0, \beta'(u) \geq 0 \text{ și } \beta' \text{ poate să se anuleze cel mult într-un singur punct, pe care îl presupunem a fi } 0.$$

OBSERVAȚIA 1.1. În condițiile în care u_0 satisface (A1) se poate obține mai mult despre soluția u (conform [1], [4], [18], [27] și a referințelor citate în aceste lucrări). Sunt adevărate următoarele proprietăți legate de mărginirea și regularitatea lui u , respectiv a lui $\beta(u)$

$$(1.6) \quad u \in L^\infty(Q_T), \quad \partial_t u \in L^\infty(I, H^{-1}(\Omega)), \quad \beta(u) \in L^\infty(I, H_0^1(\Omega)) \cap H^1(I, L^2(\Omega)).$$

Mai mult, are loc principiul comparației, prin urmare dacă datele inițiale și cele de pe frontieră sunt pozitive, aceeași proprietate o va avea și soluția însăși. Având în vedere modelul fizic de la care am pornit, pozitivitatea datelor inițiale și pe frontieră reprezintă

o proprietate naturală. Ea va juca un rol important și în capitolul următor. Totodată, acest principiu ne arată că u este esențial mărginită.

Prin (A2) se garantează faptul că β are o inversă β^{-1} , care este continuu diferentiabilă. În această situație, pentru orice valoare (mică) $\varepsilon > 0$, vor exista două constante C_1 și C_2 (depinzând de ε) astfel încât să aibă loc

$$(1.7) \quad 0 < C_1(\varepsilon) \leq (\beta^{-1})'(\theta) \leq C_2(\varepsilon) < \infty, \quad \forall \theta \in [\beta(\varepsilon), \beta(M)].$$

În acest caz M este o constantă care ulterior vom presupune că depășește valoarea datelor inițiale și a celor pe frontieră. Așadar, vom presupune că valoarea inițială u_0 satisface inegalitatea $\|u_0\|_\infty \leq M$.

În general va exista o dependență regulată a constantei C_1 de ε în sensul în care tinderea lui ε către 0 nu va afecta valoarea constantei C_1 . În sprijinul acestei afirmații vom veni cu cazul ecuației (1.1). Prin urmare vom putea utiliza în continuare C_1 în locul lui $C_1(\varepsilon)$.

OBSERVAȚIA 1.2. Așa cum rezultă atât din definiția soluției slabe, cât și din proprietățile cuprinse în relația (1.6), regularitatea lui $\beta(u)$ este mai bună decât cea a soluției u . Această observație justifică abordarea numerică bazată pe obținerea în prealabil a necunoscutei $\theta = \beta(u)$, ulterior deducându-se valoarea corespunzătoare pentru u . Este de așteptat și o estimare mai bună a erorii în raport cu θ .

2. Metode numerice

Continuăm acum cu prezentarea unor aspecte legate de rezolvarea aproximativă a ecuațiilor considerate până acum. Există diferite abordări numerice ale problemelor de tipul celor de mai sus. Așa cum afirmam anterior, problemele apar datorită degenerării ecuației parabolice, fapt care conduce și la apariția frontierelor libere (în general frontiera domeniului în care soluția este strict pozitivă). În vecinătatea acelei mulțimi gradientul soluției devine infinit. Acest fapt se răsfrânge atât asupra analizei numerice cât și a rezolvării numerice efective. Acest ultim aspect obligă la o discretizare extrem de fină a regiunilor menționate, generând creșterea excesivă a dimensiunii problemelor discrete, de unde și un consum mare de resurse (memorie și timp de calcul). Totodată, în cazul metodelor de discretizare temporală explicite sau semi-implicite, pasul de timp este supus restricțiilor legate de discretizarea spațială (datorită condițiilor de stabilitate).

Putem distinge două mari clase de scheme numerice pentru rezolvarea ecuațiilor parabolice degenerate. Prima se bazează pe urmărirea evoluției frontierei libere (cum ar fi, de exemplu, schema concepută de Di Benedetto și Hoff [12]). Metodele din cea de-a

doua clasă încearcă să determine pe parcursul obținerii soluției unde se situează frontiera liberă. În cele ce urmează ne vom concentra asupra celor din urmă. Fără a intra în detalii vom menționa faptul că și în acest ultim caz se poate realiza o urmărire indirectă a frontierei libere prin metodele de discretizare adaptivă (cum ar fi cele dezvoltate în [5], [24], [25], [36] și lucrările citate acolo).

Și în clasa schemelor care detectează singure frontiera liberă putem distinge mai multe categorii. O primă abordare determină direct soluția u (cum sunt cele descrise în [2], [36]). A doua posibilitate se bazează pe proprietatea de regularitate amintită în observația 1.2 și determină inițial necunoscuta θ asociată lui $\beta(u)$, urmând ca apoi u să fie reconstituită printr-o postprocesare (acesta fiind cazul schemelor cuprinse, spre exemplu, în [16], [17], [23], [32] ori [37]). O ultimă categorie de algoritmi determină ambele necunoscute simultan, rezultând un procedeu neliniar ([3], [13], [19], [24], [26]). În unele cazuri aceste abordări conduc la estimări de eroare optimale ([19], [31]), însă schemele sunt mai greu de implementat practic necesitând rezolvarea unor sisteme neliniare, ori chiar și probleme având dimensiunea dublă.

Legat de discretizarea propriu-zisă, există și aici diferite abordări. Mergând pe linia adoptată în monografiile dedicate numericii ecuațiilor parabolice (cum ar fi, de exemplu [35]), inițial s-a propus tratarea în prealabil a spațiului, de aici rezultând un sistem de ecuații ordinare rezolvabil prin metode specifice acestora. Dar prin acest procedeu nu se ține seama de structura problemei inițiale, anume faptul că s-a pornit de la o ecuație cu derivate parțiale. În special în cazul apariției unor factori de transport (derivate de ordin întâi), acest lucru conduce la eforturi suplimentare dacă se urmărește o rezolvare eficientă. Din această cauză, în ultimii ani a câștigat tot mai mult teren abordarea inversă (discretizarea temporală, rezultând un șir de probleme eliptice, urmată de cea spațială).

Una din posibilitățile de investigare a problemelor de tipul celor în (1.4) constă în regularizarea parabolică, care poate fi realizată în diverse moduri. O tehnică larg utilizată constă în modificarea neliniarității β ori a coeficientului difuziei β' astfel încât acesta din urmă să nu se mai anuleze. Dintre schemele numerice obținute prin acest procedeu amintim cele din [16], [17], [23], [25], [26] ori [32].

Pe de altă parte, conform [27], putem afirma că, în cazul în care sunt căutate soluții pozitive, dificultățile datorate degenerării ecuației pot fi evitate printr-o perturbare a datelor inițiale și la frontieră astfel încât soluția obținută este mărginită inferior deasupra lui 0, punctul de degenerare. Această posibilitate a fost descrisă în [29], [37] și este prezentată pe larg în capitolul următor.

Înceiem această secțiune cu prezentarea schemei de aproximare propusă de W. Jäger și J. Kačur - [16], [17]. Facem acest lucru atât datorită eficienței metodei - ilustrată și în [5] - și a generalității ei (fiind aplicabilă problemei definite în (1.3)), cât și datorită similitudinii schemei tratată în capitolul următor cu cea de mai jos.

Fie $n \in \mathbb{N}$ și $\tau = T/n$ pasul de timp, u^i o aproximație a soluției u la momentul $t_k = k \cdot \tau$. Aproximația θ^k a lui $\beta(u^k)$ va fi obținută în momentul t_k rezolvând ecuația eliptică

$$(2.1) \quad \begin{aligned} \lambda_k(\theta^k - \beta(u^{k-1})) - \tau \nabla a(u^{k-1}, \nabla \theta^k) &= \tau f(u^{k-1}), \\ \theta^k|_{\partial\Omega} &= 0, \end{aligned}$$

unde $\lambda_k \in L^\infty(\Omega)$ satisface condiția de convergență

$$(2.2) \quad \left| \lambda_k - \frac{b_n(u^{k-1}) + \mu_k(\theta^k - \beta(u^{k-1}))}{\theta^k - \beta(u^{k-1})} \right| < \tau$$

cu funcția adițională de relaxare μ_k satisfăcând

$$(2.3) \quad \frac{\alpha}{L_\beta} \leq \mu_k \leq \min \left\{ K, \frac{\beta^{-1}(\beta(u^{k-1}) + \alpha(\theta^k - \beta(u^{k-1}))) - u^{k-1}}{\theta^k - \beta(u^{k-1})} \right\}.$$

Aici $K, \alpha \in (0, 1)$ sunt parametri ai metodei, L_β constanta Lipschitz a funcției β , iar

$$b_n(u) = b(u) + \tau^d u, \quad d \in (0, 1/2).$$

Valoarea lui u^k se obține din

$$(2.4) \quad u_k = u^{k-1} + \mu_k(\theta^k - \beta(u^{k-1})).$$

Pentru completitudine vom menționa faptul că u^0 este dat de condițiile inițiale.

În schema de mai sus, pentru $\theta^k = \beta(u^{k-1})$, raporturile din (2.2) și (2.3) se înlocuiesc cu $\partial_u b_n(u^{k-1}) \cdot \min \{K, \alpha/\beta'(u^{k-1})\}$, respectiv $\min \{K, \alpha/\beta'(u^{k-1})\}$. În [17] este propusă și o metodă iterativă de rezolvare a problemei eliptice neliniare din (2.1) bazată pe actualizarea succesivă a necunoscutelor $\theta^k, \lambda_k, \mu_k$.

Legat de convergența schemei mai sus menționate, pentru cazul general poate fi obținută doar existența unui subsir al șirului soluțiilor $\{\theta^n\}_{n \in \mathbb{N}}$ care tinde în $L^s(Q_T)$, $s < 2$ la una din soluțiile problemei definite în (1.3). În cazul ecuației mediului poros, în [15] este demonstrată convergența metodei prin obținerea unei estimări a erorii.

În [32], [33] M. Slodička propune o metodă semi-implicită de discretizare a ecuației definite în (1.4) cu partea dreaptă nenulă (termeni sursă). Regularizarea problemei se face luând drept neliniaritate funcția $\beta_\varepsilon(u) = \beta(u) + \varepsilon u$, necunoscuta fiind $\theta = \beta_\varepsilon(u)$. Metoda obținută este similară cu cea descrisă în [17], cu diferența că în (2.1), în locul

lui λ_k va apărea λ_{k-1} . Totodată, estimarea erorii de aproximare este inferioară celei din cazul precedent.

CAPITOLUL 3

Scheme numerice bazate pe principiul de maxim

Acest capitol este dedicat exclusiv rezolvării numerice a ecuației mediului poros pornind de la principiul de maxim. Vom aborda inițial discretizarea temporală, unde vom studia schema implicită precum și pe cea semi-implicită, iar apoi vom continua cu analiza discretizării complete. Pentru susținerea efectivității schemelor propuse vom apela la estimările de eroare obținute precum și la exemple numerice.

1. Ideea de bază

În cele ce urmează vom considera ecuația definită în (1.4)

$$(1.1) \quad \begin{aligned} \partial_t u - \Delta \beta(u) &= 0, & \text{în } Q_T &\equiv (0, T) \times \Omega, \\ u(x, 0) = u_0(x) &\geq 0, & \text{în } \Omega, \\ u &= 0, & \text{pe } \partial\Omega, \end{aligned}$$

iar când vorbim de o soluție a ei ne gândim la soluția slabă așa cum rezultă din definiția 1.2. De asemenea, considerăm adevărate presupunerile (A1) și (A2) enunțate în capitolul precedent.

Așa cum remarcam în observațiile 1.1 și 1.2, dacă luăm în considerare necunoscuta $\theta = \beta(u)$, beneficiem de avantajele unei regularități mai bune. Acesta este motivul pentru care în numeroase scheme numerice se obține inițial θ , iar ulterior este dedusă valoarea corespunzătoare lui u . În cele ce urmează vom respecta și noi acest principiu.

Totodată am susținut faptul că este convenabilă transformarea ecuației inițiale în una parabolică regulată, acest lucru putându-se face prin modificarea neliniarității β ori a coeficientului difuziei β' , astfel încât valorile acestuia să nu se mai anuleze. În prima din cele două observații menționate mai sus afirmăm că pentru u - deci și pentru $\beta(u)$ - are loc principiul comparației. Prin urmare, în cazul în care se caută soluții pozitive (lucru natural având în vedere modelul de la care s-a pornit), dificultățile cauzate de degenerarea ecuației pot fi evitate printr-o perturbare a datelor inițiale respectiv la frontieră, soluția rezultată rămânând nenulă ([27]). Conform literaturii consultate până acum, această tehnică nu a mai fost aplicată în vederea obținerii unei scheme de rezolvare numerică.

Abordarea propusă constă în rezolvarea problemei definită în (1.1) cu date perturbate local (acolo unde este necesar) în locul celor prescrise. Principiul de maxim ne garantează faptul că soluția problemei perturbate va lua valori mărginite deasupra lui 0 – punctul de degenerare, fapt ce ne permite aplicarea ulterioară a oricărei metode numerice pentru ecuații parabolice regulate. Acesta este ideea de bază a metodelor propuse mai jos.

Având în vedere regularitatea redusă în raport cu variabila t (timpul) a soluției u , respectiv $\beta(u)$ - așa cum rezultă din observațiile 1.1 și 1.2 - ne vom limita la o metodă de discretizare în timp cu un pas. Concret, vom considera metodele Euler implicită și semi-implicită, iar discretizarea spațială va fi făcută cu ajutorul metodei elementului finit, anume o variantă a acesteia care asigură conservarea principiului de maxim și la nivel discret.

2. Metoda Euler implicită

Vom considera acum discretizarea temporală prin metoda Euler implicită a problemei definite în (1.1), în condițiile perturbării locale ale datelor inițiale și la frontieră.

2.1. Prezentarea metodei. Fie n un număr natural și $\tau = T/n$. Aplicând metoda Euler implicită obținem

$$(2.1) \quad \begin{aligned} \beta^{-1}(\theta^{k+1}) - \beta^{-1}(\theta^k) &= \tau \Delta \theta^{k+1}, \\ \theta^{k+1}|_{\partial\Omega} &\equiv \beta(\varepsilon) \end{aligned}$$

unde $k = \overline{0, n-1}$ și $\theta^0 = \beta(u_0) + \beta(\varepsilon)v_0$, ε fiind un număr strict pozitiv (dar mic) iar $v_0 \in H^1(\Omega)$ o funcție artificială. Prin θ^k am aproximat $\theta(t_k) = \beta(u(t_k))$, cu $t_k = k\tau$. Pentru obținerea aproximației necunoscutei inițiale u vom folosi inversa funcției β astfel încât

$$u^{k+1} = \beta^{-1}(\theta^{k+1}),$$

unde $k = \overline{0, n-1}$.

Alegerea funcției v_0 depinde de datele inițiale originale $u_0(x)$ după cum urmează

$$(2.2) \quad v_0|_{\{u_0 \geq \varepsilon\}} \equiv 0, \quad v_0 \leq 1 \text{ și } \beta(u_0) + \beta(\varepsilon)v_0 \geq \beta(\varepsilon).$$

Inegalitățile considerate aici, ca și pe întreg cuprinsul acestui referat atunci când implică mărimile u și θ , trebuie luate în sensul de aproape peste tot (a.p.t.). Un exemplu tipic pentru v_0 este

$$(2.3) \quad v_0 := \left[1 - \frac{\beta(u_0)}{\beta(\varepsilon)} \right]_+.$$

Având în vedere faptul că am considerat inițial valori la frontieră nule, este necesar ca acestea să fie ridicate la $\beta(\varepsilon)$ în (2.1).

Este de interes să menționăm aici faptul că schema propusă mai sus are un grad de similaritate cu metoda Jäger-Kačur pentru o alegere particulară a parametrilor. Astfel, dacă în inegalitatea (2.3) se impune egalitatea lui μ_k cu termenul din partea dreaptă, luând λ_k identic cu μ_k astfel obținut (lucru recomandat de autori în cazul ecuației considerate de noi) și forțând α să ia valoarea 1 (valoare neadmisă teoretic, ea nepermițând obținerea convergenței în abordarea autorilor, însă cu rezultate practice foarte bune) vom ajunge la situația de mai sus abstractie făcând de controlul necesar al lui μ_k . În algoritmul propus în [17], în cazul în care μ_k depășește o valoare prescrisă K , el va fi redus la acea valoare. În cazul nostru, datorită perturbării datelor, controlul unei astfel de valori nu mai este necesar.

OBSERVAȚIA 2.1. Ca și în (2.2), datele inițiale sunt perturbate doar local. Aceasta aduce un plus de eficiență metodei, însă rezultatele ce vor fi enunțate mai jos rămân valabile și în cazul unei perturbări globale (cum ar fi, de exemplu, $v_0 \equiv 1$).

Vom trece acum la analiza schemei propuse în (2.1). Aceasta include și prezentarea unei modalități de rezolvare iterativă a problemelor neliniare rezultate.

2.2. Analiza problemelor eliptice. Această parte este dedicată analizei problemelor definite în (2.1), care, pentru orice k , reprezintă o problemă Dirichlet atașată unei ecuații eliptice neliniare. Pentru început vom scrie aceste probleme în formă variațională luând

$$V = \{\varphi \in H_0^1(\Omega) : 0 \leq \varphi \leq \beta(M) - \beta(\varepsilon), \text{ a.p.t.}\}$$

iar pentru orice $\psi \in V$ și $\varphi \in H_0^1(\Omega)$,

$$\begin{aligned} a(\psi, \varphi) &= (\beta^{-1}(\psi + \beta(\varepsilon)) - \varepsilon, \varphi) + \tau \langle \psi, \varphi \rangle, \\ f_k(\varphi) &= (\beta^{-1}(\theta^k) - \varepsilon, \varphi), \end{aligned}$$

apartenența lui θ^k la $\beta(\varepsilon) + V$ urmând a fi arătată ulterior. Teorema lui Riesz arată că V este închisă, iar convexitatea acestei mulțimi este evidentă. Forma variațională a problemei în (2.1) este dată de

$$(2.4) \quad a(\theta^{k+1} - \beta(\varepsilon), \varphi) = f_k(\varphi), \quad \text{pentru orice } \varphi \in H_0^1(\Omega).$$

(A1) garantează apartenența lui θ^0 la $\beta(\varepsilon) + V$. Presupunând că și θ^k are aceeași proprietate pentru un $k \geq 0$ poate fi utilizat procedeul inducției matematice.

$\theta^{k+1} \in \beta(\varepsilon) + V$ se obține printr-o iterație bazată pe formularea din (2.4) în locul schemei numerice definite în (2.1). Pentru aceasta fie constanta

$$(2.5) \quad K \geq C_2(\varepsilon)$$

și, pentru orice $\psi, \varphi \in H_0^1(\Omega)$ și $\underline{\psi} \in V$, definim

$$\begin{aligned} a_K(\psi, \varphi) &= K(\psi, \varphi) + \tau \langle \psi, \varphi \rangle, \\ f_K(\underline{\psi}; \varphi) &= K(\underline{\psi}, \varphi) + (\beta^{-1}(\theta^k) - \beta^{-1}(\underline{\psi} + \beta(\varepsilon)), \varphi). \end{aligned}$$

Schema iterativă este introdusă printr-un operator pe V . În acest scop, pentru orice $\psi \in V$, aplicând lema Lax-Milgram se obține câte un unic element $T\psi \in H_0^1(\Omega)$, lement ce satisface

$$(2.6) \quad a_K(T\psi, \varphi) = f_K(\psi; \varphi) \quad \forall \varphi \in H_0^1(\Omega).$$

Astfel s-a definit un operator T pe V cu valori în $H_0^1(\Omega)$. Mai jos vom arăta că, alegând o normă convenabilă, T este o contracție pe mulțimea închisă V . Așadar, iterația de mai jos este bine definită

$$(2.7) \quad \psi^{i+1} = T\psi^i$$

cu $i \geq 0$ și $\psi^0 \in V$. O alegere convenabilă pentru ψ^0 este $\psi^0 = \theta^k - \beta(\varepsilon)$, iar pentru θ^{k+1}

$$\theta^{k+1} = \beta(\varepsilon) + \lim_{i \rightarrow \infty} \psi^i.$$

Existența acestei limite și apartenența lui θ^{k+1} la $\beta(\varepsilon) + V$ va rezulta în urma aplicării teoremei de punct fix lui T .

Un prim pas în demonstrarea acestor proprietăți este făcut prin următorul principiu de maxim

LEMA 2.1. *Dacă (A2) are loc și $\theta^k \in \beta(\varepsilon) + V$, atunci $TV \subset V$.*

Demonstrație. Fie $\psi \in V$. Arătăm, pentru început, că $f_K(\psi; \varphi) \geq 0$ pentru orice $\varphi \geq 0$. De fapt, întrucât $\theta^k \geq \beta(\varepsilon)$, conform (A2) are loc $\beta^{-1}(\theta^k) \geq \beta^{-1}(\beta(\varepsilon))$. Atunci,

$$\begin{aligned} f_K(\psi; \varphi) &= (K\psi - \beta^{-1}(\psi + \beta(\varepsilon)) + \beta^{-1}(\beta(\varepsilon)), \varphi) + (\beta^{-1}(\theta^k) - \beta^{-1}(\beta(\varepsilon)), \varphi) \\ &= (K\psi - (\beta^{-1})'(\theta)\psi, \varphi) + (\beta^{-1}(\theta^k) - \beta^{-1}(\beta(\varepsilon)), \varphi) \\ &\geq ((K - (\beta^{-1})'(\theta))\psi, \varphi) \geq 0. \end{aligned}$$

S-a ținut cont de faptul că $\theta \in [\beta(\varepsilon), \beta(\varepsilon) + \psi]$ și de inegalitatea (2.5), deci

$$(\beta^{-1})'(\theta) \leq C_2(\varepsilon) \leq K.$$

Luând $\varphi = [-T\psi]_+ \in H_0^1(\Omega)$ suntem conduși la

$$\begin{aligned} 0 &\leq f_K(\psi; \varphi) = a_K(T\psi, \varphi) \\ &= K(T\psi, [-T\psi]_+) + \tau \langle T\psi, [-T\psi]_+ \rangle \\ &= K(-[-T\psi]_+, [-T\psi]_+) + \tau \langle -[-T\psi]_+, [-T\psi]_+ \rangle \leq 0. \end{aligned}$$

Aceasta arată că

$$([-T\psi]_+, [-T\psi]_+) = 0.$$

Prin urmare, $[-T\psi]_+ \equiv 0$ a.p.t., deci, $T\psi \geq 0$ a.p.t..

Raționând similar cu $\beta(M) - \beta(\varepsilon) - T\psi$ și $\beta(M) - \beta(\varepsilon) - \psi$ în locul lui $T\psi$ și ψ rezultă inegalitatea $T\psi \leq \beta(M) - \beta(\varepsilon)$ și cu aceasta demonstrația este completă. \diamond

Acum poate fi arătată contractivitatea operatorului T pe mulțimea închisă V într-o normă aleasă convenabil.

LEMA 2.2. *În condițiile lemei 2.1 există o normă pe $H_0^1(\Omega)$ echivalentă celei uzuale, astfel încât T este o contracție pe mulțimea închisă V .*

Să remarcăm faptul că nu apare nici o restricție legată de ordinul de mărime a pasului de timp τ .

Demonstrație. Este ușor de observat faptul că $\|\varphi\|_K \equiv \sqrt{a_K(\varphi, \varphi)}$ reprezintă o normă pe $H_0^1(\Omega)$, care este echivalentă celei uzuale. În cele ce urmează vom arăta că, în raport cu această normă, T este contractivă pe V . Fie $\psi_1, \psi_2 \in V$. Din (2.6) se deduce că

$$\begin{aligned} a_K(T\psi_1 - T\psi_2, \varphi) &= K(\psi_1 - \psi_2, \varphi) - (\beta^{-1}(\psi_1 + \beta(\varepsilon)) - \beta^{-1}(\psi_2 + \beta(\varepsilon)), \varphi) \\ &= ((K - (\beta^{-1})'(\theta))(\psi_1 - \psi_2), \varphi) \\ &\leq (K - C_1) \|\psi_1 - \psi_2\| \|\varphi\| \\ &\leq (K - C_1) \frac{\|\psi_1 - \psi_2\|_K}{\sqrt{K}} \cdot \frac{\|\varphi\|_K}{\sqrt{K}} \\ &\leq \left(1 - \frac{C_1}{K}\right) \|\psi_1 - \psi_2\|_K \|\varphi\|_K. \end{aligned}$$

Luând $\varphi = T\psi_1 - T\psi_2$ rezultă relația

$$\|T\psi_1 - T\psi_2\|_K \leq \left(1 - \frac{C_1}{K}\right) \|\psi_1 - \psi_2\|_K,$$

ceea ce arată că T este o contracție. \diamond

OBSERVAȚIA 2.2. Demonstrația de mai sus arată și ordinul de convergență al iterației propuse în (2.7), anume $1 - O(1/C_2(\varepsilon))$. Pentru ε mic este aproape de 1. De aceea ne vom referi la această variantă mai mult sub aspectul ei teoretic. Împreună cu discretizarea spațială am considerat o alternativă la iterația de mai sus.

OBSERVAȚIA 2.3. Rezultatul anterior implică un principiu de maxim și pentru soluția problemei definite în (2.4). Dar acest fapt putea fi obținut direct, într-o manieră similară demonstrației lemei 2.1.

Prin teorema anterioară am arătat că problema pusă în (2.4) are o soluție unică în mulțimea convexă și închisă V . De fapt, așa cum rezultă de mai jos, unicitatea se extinde asupra întregului $H_0^1(\Omega)$, existența putând fi arătată cu ajutorul lemei Lax-Milgram neliniară.

LEMA 2.3. *Problema definită în (2.4) are cel mult o soluție $\theta^{k+1} \in \beta(\varepsilon) + H_0^1(\Omega)$.*

Demonstrație. Fie $\psi, \varphi \in H_0^1(\Omega)$. Datorită monotoniei lui β^{-1} are loc

$$\begin{aligned} a(\psi, \psi - \varphi) - a(\varphi, \psi - \varphi) &= (\beta^{-1}(\psi + \beta(\varepsilon)) - \beta^{-1}(\varphi + \beta(\varepsilon)), \psi - \varphi) + \tau |\psi - \varphi|_1^2 \\ &\geq \tau |\psi - \varphi|_1^2. \end{aligned}$$

Din aceasta rezultă monotonia strictă a formei a , ceea ce implică unicitatea eventualei soluții. \diamond

OBSERVAȚIA 2.4. Ecuația originală (1.1) poate degenera numai în $u = 0$, pentru care $\beta'(0) = 0$. Datorită principiului de maxim demonstrat mai sus avem garanția faptului că, dacă datele inițiale și valorile la frontieră sunt mărginite deasupra lui 0, același lucru va fi valabil și pentru soluție. Aceasta este ideea de la care am pornit în abordarea degenerării.

OBSERVAȚIA 2.5. Iterația se bazează pe operatorul T , care a fost utilizat și în [34], pp. 96. Dar în cele de mai sus am arătat că, în cadrul definit de lema 2.2, T este o contracție, ceea ce simplifică mult demonstrația. Utilizând monotonia lui T putem realiza o demonstrație similară celei din [34].

2.3. Estimarea erorii. Vom trece acum la demonstrarea convergenței metodei propuse în (2.1) prin estimarea erorii dintre soluția numerică θ și mărimea $\beta(u)$ – densitatea scalată, unde u este soluția exactă a problemei (1.1). Pentru aceasta vom demonstra următorul rezultat de stabilitate

TEOREMA 2.4. *Presupunând (A1) și (A2), pentru orice $p < n$ are loc*

$$(2.8) \quad \tau \left(|\theta^{p+1}|_1^2 + \sum_{k=0}^p |\theta^{k+1} - \theta^k|_1^2 \right) + 2C_1 \sum_{k=0}^p \|\theta^{k+1} - \theta^k\|^2 \leq \tau |\theta^0|_1^2.$$

Demonstrație. Considerând $\phi = \theta^{k+1} - \theta^k \in H_0^1(\Omega)$ în (2.4), se obține

$$(\beta^{-1}(\theta^{k+1}) - \beta^{-1}(\theta^k), \theta^{k+1} - \theta^k) + \tau \langle \theta^{k+1}, \theta^{k+1} - \theta^k \rangle = 0.$$

Are loc

$$2 \langle \theta^{k+1}, \theta^{k+1} - \theta^k \rangle = |\theta^{k+1}|_1^2 - |\theta^k|_1^2 + |\theta^{k+1} - \theta^k|_1^2,$$

în timp ce monotonia lui β^{-1} și (1.7) conduc la

$$(2.9) \quad (\beta^{-1}(\theta^{k+1}) - \beta^{-1}(\theta^k), \theta^{k+1} - \theta^k) \geq C_1 \|\theta^{k+1} - \theta^k\|^2 \geq 0.$$

Din precedentele două relații se deduce că

$$(2.10) \quad \frac{\tau}{2} \left(|\theta^{k+1}|_1^2 - |\theta^k|_1^2 + |\theta^{k+1} - \theta^k|_1^2 \right) + C_1 \|\theta^{k+1} - \theta^k\|^2 \leq 0.$$

Însumând (2.10) pentru $k = \overline{0, p}$ se ajunge la inegalitatea din (2.8). \diamond

Totodată, în cele ce urmează se va face uz de următoarea lemă.

LEMA 2.5. *Pentru orice $u \geq 0$, fie*

$$u^\varepsilon = \beta^{-1}(\beta(u) + \beta(\varepsilon)).$$

Dacă β satisface (A2), atunci are loc

$$(2.11) \quad 0 < u^\varepsilon - u \leq \varepsilon + C_2(\varepsilon)\beta(\varepsilon),$$

unde $C_2(\varepsilon)$ a fost definit în (1.7).

Demonstrație. Datorită faptului că β^{-1} și β sunt crescătoare iar $\varepsilon > 0$, are loc

$$u = \beta^{-1}(\beta(u)) < \beta^{-1}(\beta(u) + \beta(\varepsilon)) \equiv u^\varepsilon.$$

Pe de altă parte, din teorema lui Taylor se deduce că

$$u^\varepsilon \equiv \beta^{-1}(\beta(u) + \beta(\varepsilon)) = \beta^{-1}(\beta(\max(u, \varepsilon))) + \beta(\min(u, \varepsilon))(\beta^{-1})'(\theta),$$

unde θ reprezintă o combinație convexă între $\beta(u) + \beta(\varepsilon)$ și $\beta(\max(u, \varepsilon))$, ambele mărimi fiind mai mari decât $\beta(\varepsilon)$ întrucât $u \geq 0$. Așadar,

$$\beta^{-1}(\beta(\max(u, \varepsilon))) \leq u + \varepsilon, \quad \beta(\min(u, \varepsilon)) \leq \beta(\varepsilon) \quad \text{and} \quad (\beta^{-1})'(\theta) \leq C_2(\varepsilon),$$

prin urmare

$$u^\varepsilon \leq u + \varepsilon + C_2(\varepsilon)\beta(\varepsilon).$$

\diamond

OBSERVAȚIA 2.6. Dacă pentru β are loc $\varepsilon \cdot \inf \{\beta'(x) : x \in [\varepsilon, M]\} \geq C\beta(\varepsilon)$, proprietate adevărată în cazul $\beta(u) = u^m$ pentru $m \geq 1$, atunci în lema de mai sus se poate renunța la termenul $C_2(\varepsilon)\beta(\varepsilon)$.

Acum poate fi demonstrat rezultatul principal al secțiunii.

TEOREMA 2.6. *Presupunând adevărate (A1) și (A2), dacă u este soluția slabă a problemei (1.1), atunci*

$$\begin{aligned} \|\beta(u) - \theta_\Delta\|_{L^2(Q_T)}^2 + \sum_{k=1}^n \left| \int_{t_{k-1}}^{t_k} [\beta(u(t)) - \theta^k] dt \right|_1^2 + \left| \int_0^T [\beta(u(t)) - \theta_\Delta(t)] dt \right|_1^2 \\ \leq C \{ \tau + \varepsilon^2 + (C_2(\varepsilon)\beta(\varepsilon))^2 + \beta^2(\varepsilon) \}, \end{aligned}$$

unde $\theta_\Delta(t) = \theta^k$ pentru $t \in (t_{k-1}, t_k]$, $k = \overline{1, n}$ și, pentru orice k , θ^k rezolvă problema definită în (2.4).

Demonstrație. Demonstrația se bazează pe procedeul descris în [23], utilizat și în [32], [22] and [15]. Integrând ecuația din (1.5) pentru $t \in (t_{j-1}, t_j)$ se ajunge la

$$(2.12) \quad (u(t_j) - u(t_{j-1}), \varphi) + \left\langle \int_{t_{j-1}}^{t_j} \beta(u(t)) dt, \varphi \right\rangle = 0.$$

Scăzând din (2.12) relația (2.4) și însumând rezultatul obținut pentru $j = \overline{1, k}$ se deduce egalitatea

$$(2.13) \quad (u(t_k) - \beta^{-1}(\theta^k), \varphi) - (u(t_0) - \beta^{-1}(\theta^0), \varphi) + \sum_{j=1}^k \left\langle \int_{t_{j-1}}^{t_j} \beta(u(t)) - \theta^j dt, \varphi \right\rangle = 0.$$

Din $\theta^k \in \beta(\varepsilon) + H_0^1(\Omega)$ și $\beta(u) \in L^2(0, T; H_0^1(\Omega))$ rezultă că $\int_{t_{k-1}}^{t_k} \beta(u) dt + \tau(\beta(\varepsilon) - \theta^k)$ este un element al lui $H_0^1(\Omega)$. Prin urmare, luând $\varphi = \int_{t_{k-1}}^{t_k} (\beta(u^\varepsilon) - \theta^k) dt = \int_{t_{k-1}}^{t_k} \beta(u) dt + \tau(\beta(\varepsilon) - \theta^k)$ și adunând relația de mai sus pentru $k = \overline{1, n}$ se obține

$$(2.14) \quad \begin{aligned} \sum_{k=1}^n \left((u(t_k) - \beta^{-1}(\theta^k), \int_{t_{k-1}}^{t_k} [\beta(u^\varepsilon) - \theta^k] dt) - \sum_{k=1}^n \left((u(t_0) - \beta^{-1}(\theta^0), \int_{t_{k-1}}^{t_k} [\beta(u^\varepsilon) - \theta^k] dt) \right) \right. \\ \left. + \sum_{k=1}^n \sum_{j=1}^k \left\langle \int_{t_{j-1}}^{t_j} [\beta(u) - \theta^j] dt, \int_{t_{k-1}}^{t_k} [\beta(u) - \theta^k + \beta(\varepsilon)] dt \right\rangle \right) = 0. \end{aligned}$$

Acum trebuie estimați fiecare din cei trei termeni din (2.14), notați cu T_1 , T_2 , respectiv T_3 . Pentru T_3 , întrucât $\beta(\varepsilon)$ este constantă, identitatea

$$2 \sum_{k=1}^n a_k \left(\sum_{j=1}^k a_j \right) = \left(\sum_{k=1}^n a_k \right)^2 + \sum_{k=1}^n a_k^2$$

conduce la

$$(2.15) \quad \begin{aligned} T_3 &= \sum_{k=1}^n \sum_{j=1}^k \left\langle \int_{t_{j-1}}^{t_j} [\beta(u) - \theta^j] dt, \int_{t_{k-1}}^{t_k} [\beta(u) - \theta^k] dt \right\rangle \\ &= \frac{1}{2} \left| \sum_{k=1}^n \int_{t_{k-1}}^{t_k} [\beta(u(t)) - \theta^k] dt \right|_1^2 + \frac{1}{2} \sum_{k=1}^n \left| \int_{t_{k-1}}^{t_k} [\beta(u(t)) - \theta^k] dt \right|_1^2. \end{aligned}$$

T_1 poate fi descompus în modul următor.

$$\begin{aligned}
T_1 &= \sum_{k=1}^n \int_{t_{k-1}}^{t_k} (u(t_k) - \beta^{-1}(\theta^k), \beta(u^\varepsilon(t)) - \theta^k) dt \\
&= \sum_{k=1}^n \int_{t_{k-1}}^{t_k} (u(t_k) - u(t), \beta(u^\varepsilon(t)) - \theta^k) dt \\
&\quad + \sum_{k=1}^n \int_{t_{k-1}}^{t_k} (u(t) - u^\varepsilon(t), \beta(u^\varepsilon(t)) - \theta^k) dt \\
&\quad + \sum_{k=1}^n \int_{t_{k-1}}^{t_k} (u^\varepsilon(t) - \beta^{-1}(\theta^k), \beta(u^\varepsilon(t)) - \theta^k) dt \\
&\equiv T_{11} + T_{12} + T_{13}.
\end{aligned}$$

Din teorema 2.4 rezultă că $\theta_\Delta - \beta(\varepsilon) \in L^2(0, T; H_0^1(\Omega))$. Deoarece $\partial_t u \in L^2(0, T; H^{-1}(\Omega))$ și $\beta(u) \in L^2(0, T; H_0^1(\Omega))$, are loc

$$\begin{aligned}
(2.16) \quad |T_{11}| &= \left| \sum_{k=1}^n \int_{t_{k-1}}^{t_k} (u(t_k) - u(t), \beta(u^\varepsilon(t)) - \theta^k) dt \right| \\
&= \left| \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \left(\int_t^{t_k} \partial_t u ds, \beta(u^\varepsilon(t)) - \theta^k \right) dt \right| \\
&\leq \tau \|\partial_t u\|_{L^2(0, T; H^{-1}(\Omega))} \|\beta(u) - \theta_\Delta + \beta(\varepsilon)\|_{L^2(0, T; H_0^1(\Omega))} \leq C\tau.
\end{aligned}$$

Pentru T_{12} , lema 2.5 conduce la

$$\begin{aligned}
(2.17) \quad |T_{12}| &\leq \sum_{k=1}^n \int_{t_{k-1}}^{t_k} |(u(t) - u^\varepsilon(t), \beta(u^\varepsilon(t)) - \theta^k)| dt \\
&\leq \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|u^\varepsilon - u\|_\infty \|\beta(u^\varepsilon) - \theta^k\|_{L^1} \\
&\leq (\varepsilon + C_2(\varepsilon)\beta(\varepsilon)) \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\beta(u^\varepsilon) - \theta^k\|_{L^1} \\
&\leq \frac{\varepsilon + C_2(\varepsilon)\beta(\varepsilon)}{\delta} \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\beta(u^\varepsilon) - \theta^k\|^2 + \delta C(\varepsilon + C_2(\varepsilon)\beta(\varepsilon)).
\end{aligned}$$

În cele de mai sus s-a utilizat mărginirea lui Ω precum și inegalitatea $ab \leq \frac{a^2}{\delta} + \frac{\delta b^2}{4}$, δ fiind un număr pozitiv ce urmează a fi determinat ulterior. Estimarea lui T_{13} este mai simplă, așa cum rezultă din

$$\begin{aligned}
(2.18) \quad T_{13} &= \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \int_{\Omega} (u^\varepsilon - \beta^{-1}(\theta^k))(\beta(u^\varepsilon) - \theta^k) dx dt \\
&= \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \int_{\Omega} (\beta^{-1})'(\theta)(\beta(u^\varepsilon) - \theta^k)^2 \\
&\geq C_1 \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\beta(u^\varepsilon(t)) - \theta^k\|^2 dt,
\end{aligned}$$

cu θ între $\beta(u^\varepsilon)$ și θ^k .

Pentru a majora T_2 se procedează într-o manieră similară celei aplicate în cazul lui T_{12} . Conform (2.2) are loc

$$\beta(u_0) \leq \theta^0 \equiv \beta(u_0) + \beta(\varepsilon)v_0 \leq \beta(u_0) + \beta(\varepsilon) \equiv \beta(u_0^\varepsilon),$$

care, împreună cu lema 2.5 conduce la

$$(2.19) \quad 0 \leq \beta^{-1}(\theta^0) - u_0 \leq u_0^\varepsilon - u_0 \leq \varepsilon + C_2(\varepsilon)\beta(\varepsilon).$$

Prin urmare,

$$(2.20) \quad \begin{aligned} |T_2| &\leq \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\beta^{-1}(\theta^0) - u_0\|_\infty \|\beta(u^\varepsilon) - \theta^k\|_{L^1} \\ &\leq \frac{\varepsilon + C_2(\varepsilon)\beta(\varepsilon)}{\delta} \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\beta(u^\varepsilon) - \theta^k\|^2 + \delta C (\varepsilon + C_2(\varepsilon)\beta(\varepsilon)). \end{aligned}$$

Luând în considerare toate inegalitățile din (2.15) - (2.20), relația (2.14) arată că

$$\begin{aligned} 0 &= T_{11} + T_{12} + T_{13} + T_2 + T_3 \\ &\geq -C\tau - 2 \frac{\varepsilon + C_2(\varepsilon)\beta(\varepsilon)}{\delta} \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\beta(u^\varepsilon) - \theta^k\|^2 \\ &\quad - 2\delta C (\varepsilon + C_2(\varepsilon)\beta(\varepsilon)) + C_1 \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\beta(u^\varepsilon(t)) - \theta^k\|^2 dt \\ &\quad + \frac{1}{2} \left| \sum_{k=1}^n \int_{t_{k-1}}^{t_k} [\beta(u(t)) - \theta^k] dt \right|_1^2 + \frac{1}{2} \sum_{k=1}^n \left| \int_{t_{k-1}}^{t_k} [\beta(u(t)) - \theta^k] dt \right|_1^2. \end{aligned}$$

Cu δ ales corespunzător se obține

$$(2.21) \quad \begin{aligned} \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\beta(u^\varepsilon) - \theta^k\|^2 + \left| \int_0^T [\beta(u(t)) - \theta_\Delta(t)] dt \right|_1^2 + \sum_{k=1}^n \left| \int_{t_{k-1}}^{t_k} [\beta(u(t) - \theta^k)] dt \right|_1^2 \\ \leq C \{ \tau + \varepsilon^2 + (C_2(\varepsilon)\beta(\varepsilon))^2 \}. \end{aligned}$$

În final mai remarcăm faptul că

$$\begin{aligned} \|\beta(u^\varepsilon) - \theta^k\| &= \|\beta(u) - \theta^k + \beta(\varepsilon)\| \\ &\geq \left| \|\beta(u) - \theta^k\| - \|\beta(\varepsilon)\| \right| = \left| \|\beta(u) - \theta^k\| - C\beta(\varepsilon) \right|, \end{aligned}$$

așadar

$$2\|\beta(u^\varepsilon) - \theta^k\|^2 \geq \|\beta(u) - \theta^k\|^2 - C\beta^2(\varepsilon).$$

Relația de mai sus, împreună cu (2.21), demonstrează rezultatul dorit. \diamond

OBSERVAȚIA 2.7. Dacă β are proprietatea $\beta(u) + \beta(v) \leq \beta(u+v)$ - care este adevărată în cazul $\beta(u) = u^m$ cu $m \geq 1$, inegalitatea din lema 2.5 devine

$$0 < u^\varepsilon - u \leq \varepsilon.$$

Același lucru este valabil și în situația descrisă în observația 2.6. Totodată, este de așteptat ca $\beta(\varepsilon)$ să aibă un ordin de mărime mai mic decât ε . Prin urmare, eroarea de aproximare devine

$$\begin{aligned} \|\beta(u) - \theta_\Delta\|_{L^2(Q_T)}^2 + \sum_{k=1}^n \left| \int_{t_{k-1}}^{t_k} [\beta(u(t)) - \theta^k] dt \right|_1^2 + \left| \int_0^T [\beta(u(t)) - \theta_\Delta(t)] dt \right|_1^2 \\ \leq C(\tau + \varepsilon^2). \end{aligned}$$

OBSERVAȚIA 2.8. Din rezultatul precedent se poate deduce faptul că schema descrisă în (2.1) are, din punct de vedere teoretic, un comportament cel puțin la fel de bun ca și cele propuse în [23], [16] and [32]. Schema tratată în [23] are ordinul de convergență $\tau^{1/2}$, iar cea din [32] este de ordin $\tau^{1/6}$. Pentru metoda Jäger-Kačur ([17]) s-a demonstrat în [15] o estimare proporțională cu $\tau^{1/4}$. Recent, pentru schema propusă în [19], aplicându-se tehnica descrisă în [31], s-a putut obține ordinul optimal ($O(\tau)$), însă, așa cum spuneam și în capitolul precedent, schema este dificil de implementat.

O consecință imediată a rezultatului precedent o constituie estimarea erorii pentru necunoscuta u . Are loc

TEOREMA 2.7. În condițiile teoremei 2.6, dacă u este soluția slabă a problemei (1.1), atunci pentru orice $k = 1, \dots, n$ are loc

$$\|u(t_k) - \beta^{-1}(\theta_k)\|_{-1} \leq C \{ \tau + \varepsilon^2 + (C_2(\varepsilon)\beta(\varepsilon))^2 + \beta^2(\varepsilon) \}.$$

Demonstrație. Demonstrația pornește de la relația dată în (2.13). Pentru orice $k \leq n$ și $\varphi \in H_0^1(\Omega)$ are loc

$$|(u(t_k) - \beta^{-1}(\theta^k), \varphi)| \leq |(u(t_0) - \beta^{-1}(\theta^0), \varphi)| + \left| \sum_{j=1}^k \left\langle \int_{t_{j-1}}^{t_j} \beta(u(t)) - \theta^j dt, \varphi \right\rangle \right|,$$

prin urmare

$$|(u(t_k) - \beta^{-1}(\theta^k), \varphi)| \leq \|\varphi\|_1 \left[\|u(t_0) - \beta^{-1}(\theta^0)\| + \left| \sum_{j=1}^k \int_{t_{j-1}}^{t_j} \beta(u(t)) - \theta^j \right|_1 \right].$$

Ținând seama de inegalitățile cuprinse în (2.19) și (2.21) precum și de definiția normei $\|\cdot\|_{-1}$ se obține rezultatul dorit. \diamond

Și în acest caz, în condițiile enunțate în observația 2.7, se poate renunța la ultimii doi termeni ai inegalității de mai sus.

3. Metoda semi-implicită

Vom considera acum discretizarea temporală a problemei definite în (1.1) printr-o metodă de tip Euler semi-implicită. Rămânem în aceleași condițiile de perturbare locală a datelor inițiale și la frontieră.

3.1. Prezentarea metodei. Fie n un număr natural și $\tau = T/n$. Pornind de la schema implicită definită în (2.1) putem construi următoarea variantă linearizată, care apare în mod natural

$$(3.1) \quad \begin{aligned} \sigma_k(\theta^{k+1} - \theta^k) &= \tau \Delta \theta^{k+1}, \\ \theta^{k+1}|_{\partial\Omega} &\equiv \beta(\varepsilon), \\ \sigma_{k+1} &= (\beta^{-1})'(\theta^{k+1}) \end{aligned}$$

pentru orice $k = \overline{0, n-1}$, unde $\theta^0 = \beta(u_0) + \beta(\varepsilon)v_0$ și $\sigma_0 = (\beta^{-1})'(\theta^0)$. Funcția v_0 este construită ca și în secțiunea precedentă (un exemplu a fost dat în relația (2.2)), iar valorile la frontieră sunt modificate cu $\beta(\varepsilon)$.

OBSERVAȚIA 3.1. Pentru analiza acestui algoritm vom face o presupunere suplimentară celor două din secțiunea precedentă ((A1) și (A2)). Vom considera și β' ca fiind Lipschitz continuă, iar L_β va reprezenta o constantă superioară celor lipschitziene atât pentru β cât și pentru β' .

OBSERVAȚIA 3.2. Ca și în (2.2), datele inițiale sunt perturbate doar local. Aceasta aduce un plus de eficiență metodei, însă rezultatele ce vor fi enunțate mai jos rămân valabile și în cazul unei perturbări globale - cum ar fi, de exemplu, $v_0 \equiv 1$.

OBSERVAȚIA 3.3. După cum rezultă din schema de mai sus, la fiecare pas de timp se rezolvă o ecuație liniară de tip eliptic, cu toate avantajele care decurg din aceasta față de schema implicită neliniară. Prețul plătit pentru simplitate constă într-o estimare a erorii inferioară celei precedente. Însă acest lucru este mai puțin observabil din punct de vedere practic.

Așa cum remarcăm și în cazul schemei anterioare, există un grad de similaritate cu schema propusă de această dată de M. Slodiča în [32], unde regularizarea problemei este făcută printr-o perturbare globală a funcției β . Dar pentru schema de mai sus nu mai este necesară o astfel de modificare, fapt datorat principiului de maxim. Principiul este valabil și în acest caz și dă sens metodei propuse prin eliminarea riscului ca valoarea lui σ_k să devină infinită.

3.2. Estimarea erorii. În cele ce urmează vom demonstra convergența schemei semi-implicite dând din nou o estimare a erorii. Întrucât în demonstrațiile de mai jos sunt folosite aceleași idei ca și în cazul implicit, vom insista doar asupra aspectelor diferite care apar în noua situație. Vom porni și aici de la formularea variațională a problemelor ce apar în (3.1), care pentru fiecare k sunt eliptice și liniare cu condiții la frontieră de tip Dirichlet

$$(3.2) \quad (\sigma_k(\theta^{k+1} - \theta_k), \varphi) + \tau \langle \theta^{k+1}, \varphi \rangle = 0$$

pentru orice $\varphi \in H_0^1(\Omega)$ și $k \geq 1$, cu $\sigma_k = (\beta^{-1})'(\theta^k)$, $\theta^k \in \beta(\varepsilon) + V$ și V definit în secțiunea precedentă.

Pentru orice $k \leq n$, dacă σ_k satisface a.p.t. în Ω inegalitățile

$$(3.3) \quad 0 < C_1 \leq \sigma_k \leq C_2(\varepsilon) < \infty,$$

existența și unicitatea soluției $\theta^{k+1} \in H^1(\Omega)$ este garantată de lema Lax-Milgram. Demonstrând apartenența lui θ^{k+1} la submulțimea $\beta(\varepsilon) + V$, inegalitățile care apar în (1.7) garantează faptul că și σ_{k+1} satisface (3.3). Ținând seama de modul în care au fost alese datele inițiale, conform (1.7), (3.3) are loc pentru $k = 0$. Aplicând acum principiul inducției matematice se obține existența și unicitatea soluției problemei (3.2) împreună cu apartenența ei la $\beta(\varepsilon) + V$ - de unde și proprietatea (3.3) - pentru orice $k \leq n$. Totodată, prin aceasta s-a arătat că schema semi-implicită are sens.

Afirmația anterioară este susținută de următoarea leamnă analogă lemei 2.1

LEMA 3.1. *Dacă (A2) are loc în forma extinsă din observația 3.1 și $\theta^k \in \beta(\varepsilon) + V$, atunci $\theta^k \in \beta(\varepsilon) + V$.*

Demonstrație. Demonstrația este similară celei date pentru lema 2.1. Problema definită în (3.2) poate fi scrisă în forma echivalentă

$$(3.4) \quad (\sigma_k(\theta^{k+1} - \beta(\varepsilon)), \varphi) + \tau \langle \theta^{k+1}, \varphi \rangle = (\sigma_k(\theta^k - \beta(\varepsilon)), \varphi)$$

pentru orice $\varphi \in H_0^1(\Omega)$.

Luând φ din mulțimea V , are loc $(\sigma_k(\theta^k - \beta(\varepsilon)), \varphi) \geq 0$. Scriind egalitatea (3.4) cu $\varphi = [\beta(\varepsilon) - \theta^{k+1}]_+$ se obține ca și în lema 2.1 inegalitatea $\theta^{k+1} - \beta(\varepsilon) \geq 0$ a.p.t.. A doua inegalitate rezultă analog, ceea ce încheie demonstrația. \diamond

Stabilitatea aproximației este cuprinsă în următoarea teoremă.

TEOREMA 3.2. *Presupunând (A1) și (A2) în forma extinsă, dacă θ^k rezolvă problema (3.2), pentru orice $p < n$ are loc*

$$(3.5) \quad \tau \left(|\theta^{p+1}|_1^2 + \sum_{k=0}^p |\theta^{k+1} - \theta^k|_1^2 \right) + 2 \sum_{k=0}^p \|(\theta^{k+1} - \theta^k)\sqrt{\sigma_k}\|^2 \leq \tau |\theta^0|_1^2.$$

Demonstrație. Luând în (2.4) $\varphi = \theta^{k+1} - \theta^k \in H_0^1(\Omega)$ și însumând egalitățile rezultate pentru $k = \overline{0, p}$ se obține ca și în cazul teoremei 2.4 rezultatul dorit. \diamond

Acum se poate trece la estimarea erorii de aproximare pentru schema propusă în (3.1).

TEOREMA 3.3. *În condițiile teoremei 3.2, dacă u este soluția slabă a problemei (1.1), atunci*

$$\sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\beta(u(t)) - \theta^k\|^2 dt \leq C \{ \tau + \varepsilon^2 + (\tau C_2(\varepsilon))^2 + (C_2(\varepsilon)\beta(\varepsilon))^2 + \beta^2(\varepsilon) \}.$$

unde $\theta_\Delta(t) = \theta^k$ pentru $t \in (t_{k-1}, t_k]$, $k = \overline{1, n}$ și θ^k rezolvă problema definită în (3.2).

Demonstrație. Se pot urma ideile demonstrației date teoremei 2.6. Parcurgând aceiași pași pentru ecuațiile scrise în (1.5) și (3.2) se ajunge la

$$(3.6) \quad \sum_{k=1}^n \sum_{i=1}^k \left\langle \left(u(t_i) - u(t_{i-1}) - \sigma_{i-1}(\theta^i - \theta^{i-1}), \int_{t_{k-1}}^{t_k} [\beta(u^\varepsilon) - \theta^k] dt \right) \right. \\ \left. + \sum_{k=1}^n \sum_{j=1}^k \left\langle \int_{t_{j-1}}^{t_j} [\beta(u) - \theta^j] dt, \int_{t_{k-1}}^{t_k} [\beta(u) - \theta^k + \beta(\varepsilon)] dt \right\rangle \right\rangle = 0.$$

Pentru fiecare $\theta, \psi \in \beta(\varepsilon) + V$ poate fi definită funcția

$$(3.7) \quad \sigma(\theta, \psi) = \int_0^1 (\beta^{-1})'(s\bar{\theta} + (1-s)\psi) ds,$$

pentru care se observă că, oricare ar fi $\theta, \psi \in \beta(\varepsilon) + V$, are loc

$$\sigma(\theta, \psi)(\theta - \psi) = \beta^{-1}(\theta) - \beta^{-1}(\psi).$$

Adunând și scăzând $\beta^{-1}(\theta^i) - \beta^{-1}(\theta^{i-1}) = \sigma(\theta^i, \theta^{i-1})$ în primul termen al egalității (3.6) se ajunge la o relație asemănătoare celei din (2.14), cu menționarea apariției termenului suplimentar

$$(3.8) \quad T \equiv \sum_{k=1}^n \sum_{i=1}^k \left((\sigma(\theta^i, \theta^{i-1}) - \sigma_{i-1})(\theta^i - \theta^{i-1}), \int_{t_{k-1}}^{t_k} [\beta(u^\varepsilon) - \theta^k] dt \right).$$

Lipschitz continuitatea lui β' împreună cu relațiile (1.7), (3.3) implică inegalitatea

$$(3.9) \quad |\sigma(\theta^i, \theta^{i-1}) - \sigma_{i-1}| \leq L_\beta C_2(\varepsilon)^2 \sigma_{i-1} |\theta^i - \theta^{i-1}|.$$

De aceea, având în vedere mărginirile a.p.t. ale lui $\beta(u)$ și θ^i , $i = \overline{1, n}$, datorită stabilității demonstrate în (2.5), T poate fi estimat după cum urmează

$$\begin{aligned}
 (3.10) \quad |T| &\leq L_\beta C_2(\varepsilon)^2 \sum_{k=1}^n \sum_{i=1}^k \left(\sigma_{i-1} |\theta^i - \theta^{i-1}|^2, \left| \int_{t_{k-1}}^{t_k} [\beta(u^\varepsilon) - \theta^k] dt \right| \right) \\
 &\leq C\tau L_\beta C_2(\varepsilon)^2 \sum_{k=1}^n \sum_{i=1}^k \|\sqrt{\sigma_{i-1}}(\theta^i - \theta^{i-1})\|^2 \\
 &\leq C\tau L_\beta C_2(\varepsilon)^2.
 \end{aligned}$$

Restul demonstrației este identic cu cea pentru cazul implicit. \diamond

OBSERVAȚIA 3.4. Schema semi-implicită are un ordin de convergență mai mic, depinzând de proprietățile lui β . În cazul ecuației mediului poros obținem o eroare de ordinul $\tau^{1/2m}$. Rezultatele numerice arată însă că această estimare este prea pesimistă.

Ca și în cazul anterior, teorema de mai jos se referă la eroarea în raport cu u .

TEOREMA 3.4. *În condițiile teoremei 3.3, dacă u este soluția slabă a problemei (1.1), atunci pentru orice $k = 1, \dots, n$ are loc*

$$\|u(t_k) - \beta^{-1}(\theta_k)\|_{-1} \leq C \{ \tau + \varepsilon^2 + \tau C_2(\varepsilon)^2 + (C_2(\varepsilon)\beta(\varepsilon))^2 + \beta^2(\varepsilon) \}.$$

4. Analiza discretizării complete

Această secțiune este dedicată discretizării complete a problemei descrise în (1.1). În cele ce urmează ne vom referi doar la metoda implicită, cazul semi-explicit putând fi tratat analog.

4.1. Discretizarea spațială. Trecem acum la discretizarea spațială a problemelor semidiscrete prin metoda elementului finit. Pentru diferitele aspecte teoretice legate de aceasta vom face referire la [9] ori [28]. În dauna metodei standard s-a adoptat varianta condensării masei - *lumped mass method*, [35] - pentru că permite conservarea principiului de maxim și la nivelul discret. Există numeroase lucrări care fac uz de această abordare în tratarea unor probleme de acest gen - [2], [5], [13], [24], [25], [13], [26], [33], [36].

Considerăm familia $\{S_h\}_h$ de descompuneri *regulare* ([9]) ale închiderii domeniului $\overline{\Omega} \subset \mathbb{R}^d$ în simplexe d -dimensionale (închise); h se referă la ordinul de mărime al elementelor. Pentru simplitate se presupune că $\overline{\Omega} = \cup_{T \in S_h} T$. Prin urmare este neglijată eroarea datorată nedescompunerii exacte a domeniului în elemente de tipul celui precizat. O analiză completă în acest sens poate fi găsită, de exemplu, în [26]. Prin

$$V_h = \{ \varphi \in C(\overline{\Omega}) : \varphi|_T \text{ este liniară pentru orice } T \in S_h, \varphi|_{\partial\Omega} = 0 \}$$

am definit submulțimea elementelor finite d-simplicial liniare având valori nule la frontieră. Acestui spațiu liniar îi atașăm următorul produs scalar

$$(4.1) \quad (\chi, \varphi)_h = \sum_{T \in S_h} \int_T \Pi_h(\chi\varphi) dx = \sum_{T \in S_h} \frac{\text{meas}(T)}{d+1} \sum_{k=0}^d \chi(A_k^T) \varphi(A_k^T),$$

pentru oricare două elemente d-simplicial liniare χ, φ . Prin Π_h am desemnat operatorul local de interpolare liniară, iar $A_k^T, k = \overline{0, d}$ sunt vârfurile elementului T . $(\cdot, \cdot)_h$ generează pe V_h o normă - notată cu $\|\cdot\|_h$ - și este echivalent cu produsul scalar uzual (\cdot, \cdot) .

Principiul de maxim joacă un rol esențial în abordările propuse în acest capitol. De aceea vom presupune că fiecare descompunere S_h este *acută*. Prin aceasta se conservă proprietatea dorită de noi ([10]).

Relațiile de mai jos reflectă legătura dintre cele două produse scalare menționate anterior. Pentru orice $\chi, \varphi \in V_h$ are loc

$$(4.2) \quad \begin{aligned} C_1 \|\chi\| &\leq \|\chi\|_h \leq C_2 \|\chi\|, \\ |(\chi, \varphi)_h| &\leq C \|\chi\| \|\varphi\|, \\ |(\chi, \varphi) - (\chi, \varphi)_h| &\leq C \|\chi\| \|\varphi\|, \\ |(\chi, \varphi) - (\chi, \varphi)_h| &\leq Ch \|\chi\|_1 \|\varphi\|, \\ |(\chi, \varphi) - (\chi, \varphi)_h| &\leq Ch^2 \|\chi\|_1 \|\varphi\|_1, \end{aligned}$$

unde C, C_1 și C_2 sunt constante generice. Pentru demonstrație ne vom referi la [11], [30].

Având definit cadrul de mai sus se poate da formularea variațională a problemei discretizate complet. Ea reprezintă rezultatul discretizării spațiale a problemelor neliniare de tip eliptic definite în (2.4).

$$(4.3) \quad \begin{aligned} \theta_h^k &\in \beta(\varepsilon) + V_h, \\ (\Pi_h \beta^{-1}(\theta_h^k) - \Pi_h \beta^{-1}(\theta_h^{k-1}), \varphi)_h + \tau \langle \theta_h^k, \varphi \rangle &= 0, \end{aligned}$$

pentru orice $\varphi \in V_h$ și $i \geq 1$, cu θ_h^0 definit mai jos.

Pornind de la datele inițiale perturbate conform descrierii făcute în (2.1), (2.3) putem construi următoarea funcție din $\beta(\varepsilon) + V_h$

$$(4.4) \quad u_h^0 = I_h \beta^{-1}(\theta^0), \quad \theta_h^0 = \Pi_h \beta(u_h^0)$$

unde I_h este proiectorul pe V_h definit pentru orice $\theta \in L^2(\Omega)$ prin

$$(4.5) \quad (I_h \theta, \varphi)_h = (\theta, \varphi) \text{ oricare ar fi } \varphi \in V_h.$$

Este ușor de arătat că u_h^0 este cuprins între ε și M . Are loc $\theta_h^0 \in \beta(\varepsilon) + V_h$

Menționăm faptul că această alegere este dictată de o analiză mai ușoară a problemei, după cum va rezulta în subsecțiunea corespunzătoare. Din punct de vedere practic poate fi luată orice altă aproximare a datelor inițiale perturbate, însă rezultatele obținute cu varianta propusă mai sus sunt mai bune.

4.2. Metoda iterativă. În această parte propunem o metodă iterativă de rezolvare a problemelor neliniare discrete definite în (4.3). Această abordare reprezintă o alternativă la iterația propusă în (2.7) care, așa cum afirmam în observația 2.2, are un ordin de convergență redus.

Cu $\sigma(\theta, \psi)$ definit în (3.7), întrucât $V_h \subset V$, pentru fiecare $k = \overline{1, n}$ putem considera șirul problemelor discrete

$$(4.6) \quad \begin{aligned} \bar{\theta}_h^i &\in \beta(\varepsilon) + V_h, \\ (\sigma(\bar{\theta}_h^{i-1}, \theta_h^k)(\bar{\theta}_h^i - \theta_h^k), \varphi)_h + \tau < \bar{\theta}_h^i, \varphi > &= 0 \end{aligned}$$

pentru orice $\varphi \in V_h$, unde $\bar{\theta}_h^0 = \theta_h^k$, $\sigma(\bar{\theta}_h^0, \theta_h^k) = (\beta^{-1})'(\theta_h^k)$ și $\theta_h^k \in \beta(\varepsilon) + V_h$. Datele inițiale descrise în (4.4) completează această definiție.

Problemele de mai sus reprezintă discretizarea următoarei secvențe de probleme eliptice liniare

$$(4.7) \quad \begin{aligned} \bar{\theta}^i &\in \beta(\varepsilon) + H_0^1(\Omega), \\ (\sigma(\bar{\theta}^{i-1}, \theta^k)(\bar{\theta}^i - \theta^k), \varphi) + \tau < \bar{\theta}^i, \varphi > &= 0, \end{aligned}$$

pentru orice $\varphi \in H_0^1(\Omega)$ și $i \geq 1$, unde $\bar{\theta}^0 = \theta^k$, $\sigma(\bar{\theta}^0, \theta^k) = (\beta^{-1})'(\theta^k)$.

Aceste probleme sunt bine definite doar în cazul în care $\theta_h^k \in \beta(\varepsilon) + V_h$, respectiv $\theta^k \in \beta(\varepsilon) + V$. Întrucât datele inițiale au aceste proprietăți, demonstrația pentru $k > 0$ se poate face prin inducție matematică. Pentru aceasta este necesar să arătăm convergența șirurilor $\{\bar{\theta}_h^i\}_{i \in \mathbb{N}}$, respectiv $\{\bar{\theta}^i\}_{i \in \mathbb{N}}$, în mulțimile închise $\beta(\varepsilon) + V_h$, respectiv $\beta(\varepsilon) + V$. Limitele corespunzătoare acestor șiruri vor fi chiar soluțiile la momentul următor.

Trebuie să demonstrăm că, pornind cu θ_h^k , respectiv θ^k având proprietățile dorite de mărginire, șirul soluțiilor are proprietăți similare și converge. Cazul semidiscret este mai dificil de tratat. O idee în acest sens este furnizată în [21], dar nu vom insista asupra ei aici întrucât suntem interesați de discretizarea completă. Ne vom limita la problemele definite în (4.6).

Fie $\{\psi_j\}_{j=\overline{1, N_h}}$ baza canonică în spațiul V_h atașată descompunerii S_h a domeniului; N_h este numărul nodurilor interioare ale triangularizării. Prin urmare, pentru orice nod A_j și orice funcție ψ_i , $i, j = \overline{1, N_h}$, are loc

$$(4.8) \quad \psi_i(A_j) = \delta_{ij},$$

unde δ_{ij} este simbolul lui Kronecker. Orice element $\chi \in V_h$ poate fi reprezentat în această bază sub forma $\chi = \sum_{i=1}^{N_h} \chi(P_i)\psi_i$.

Luând în ecuația (4.6) drept funcții test tocmai elementele bazei canonice, ținând seama de comprimarea masei, pentru fiecare $i > 0$ se ajunge la următorul sistem liniar asociat problemei definite în (4.6)

$$(4.9) \quad (M_h^{i-1} + \tau A_h)\bar{\Theta}^i = M_h^{i-1}\Theta^k,$$

unde $\bar{\Theta}^i, \Theta^k$ sunt vectori conținând valorile nodale ale lui $\bar{\theta}_h^i, \theta_h^k$. A_h reprezintă matricea de discretizare a operatorului - în cazul nostru Laplace - în baza considerată (*stiffness matrix*). Elementele ei sunt $(A_h)_{ij} = (\nabla\psi_i, \nabla\psi_j)$ cu $i, j = \overline{1, N_h}$. M_h^i este matricea diagonală obținută prin condensarea masei înmulțită cu valoarea nodală a lui $\sigma(\bar{\theta}_h^i, \theta_h^k)$, respectiv

$$(M_h^i)_{lj} = \delta_{lj}\sigma(\bar{\theta}_h^i(A_l), \theta_h^k(A_l)) \cdot \frac{meas(\Delta_l)}{d+1},$$

unde $l, j = \overline{1, N_h}$ iar prin Δ_l am desemnat reuniunea d -simplexelor cu proprietatea că A_l este nod al acestora, iar $meas(\Delta_l)$ reprezintă volumul acestei mulțimi.

Datorită presupunerilor făcute asupra triangularizării S_h , $M_h^i + \tau A_h$ este o M -matrice ([10]). Are loc așadar varianta discretă a principiului de maxim, deci dacă $\theta_h^k \in \beta(\varepsilon) + V_h$, același lucru se poate spune pentru întregul șir $\{\bar{\theta}_h^i, i \geq 0\}$. Prin urmare ecuațiile din (4.6) au sens.

Vom schița acum demonstrația convergenței șirului $\{\bar{\theta}_h^i, i \geq 0\}$ β' . Pentru aceasta avem nevoie ca β' să aibă proprietatea cerută în observația 3.1. Fie $i > 1$. Înmulțind relația scrisă în (4.9) cu $(M_h^{i-1})^{-1}$ și scăzând rezultatul obținut pentru $i - 1$ din cel corespunzător lui i , se ajunge la egalitatea

$$(4.10) \quad (I_{N_h} + \tau D_h^{i-1} A_h) \cdot (\bar{\Theta}^i - \bar{\Theta}^{i-1}) = \tau(D_h^{i-1} - D_h^{i-2}) \cdot A_h \cdot \bar{\Theta}^{i-1},$$

unde prin D_h^i am desemnat inversa lui M_h^i iar I_{N_h} este matricea unitate de ordin N_h . Cu notația

$$\bar{e}^i = \bar{\Theta}^i - \bar{\Theta}^{i-1},$$

în situația în care β și β' îndeplinesc condițiile cerute, poate fi arătat că, pentru orice $i = \overline{1, N_h}$, elementele matricii diferență $(D_h^{i-1} - D_h^{i-2})$ satisfac inegalitatea

$$(4.11) \quad |(D_h^{i-1} - D_h^{i-2})_{jj}| \leq L_\beta C_2(\varepsilon) \frac{d+1}{meas(\Delta_j)} |(\bar{e}^{i-1})_j|.$$

Înmulțind în (4.10) cu inversa matricii $(I_{N_h} + \tau D_h^{i-1} A_h)$ și trecând la norma matricială generată de cea vectorială l_2 - euclidiană - are loc

$$(4.12) \quad \|\bar{e}^i\| \leq \tau \|(I_{N_h} + \tau D_h^{i-1} A_h)^{-1}\| \|D_h^{i-1} - D_h^{i-2}\| \|A_h\| \|\bar{\Theta}^{i-1}\|.$$

Întrucât pe spațiile finit dimensionale oricare două norme sunt echivalente, inegalitatea de mai sus se păstrează și în cazul normei vectoriale l_∞ până la o constantă C . Așadar,

$$(4.13) \quad \|\bar{e}^i\|_\infty \leq C\tau \|(I_{N_h} + \tau D_h^{i-1} A_h)^{-1}\|_\infty \|D_h^{i-1} - D_h^{i-2}\|_\infty \|A_h\|_\infty \|\bar{\Theta}^{i-1}\|_\infty.$$

Inegalitatea din (4.11) conduce la

$$(4.14) \quad \|D_h^{i-1} - D_h^{i-2}\|_\infty \leq Ch^{-d} L_\beta C_2(\varepsilon) \|\bar{e}^{i-1}\|_\infty,$$

iar pentru A_h , întrucât descompunerea S_h este regulată, are loc

$$(4.15) \quad \|A_h\|_\infty \leq Ch^{d-2}.$$

Estimarea normei $\|(I_{N_h} + \tau D_h^{i-1} A_h)^{-1}\|_\infty$ poate fi făcută pornind de la definiția normei generate de norma l^∞ . Prin reducere la absurd se obține

$$(4.16) \quad \|(I_{N_h} + \tau D_h^{i-1} A_h)^{-1}\|_\infty \leq 1.$$

Incluzând aceste inegalități în (4.13) și ținând cont de principiul de maxim dedus anterior pentru a mărgini superior $\|\bar{\Theta}^{i-1}\|_\infty$, rezultă relația

$$(4.17) \quad \|\bar{e}^i\|_\infty \leq C\tau h^{-2} C_2(\varepsilon) \|\bar{e}^{i-1}\|_\infty.$$

Prin urmare s-a demonstrat convergența șirului soluțiilor. Are loc

LEMA 4.1. *Pentru τ suficient de mic și $\theta_h^k \in \beta(\varepsilon) + V_h$, șirul soluțiilor problemelor definite în (4.6) este convergent în $\beta(\varepsilon) + V_h$.*

OBSERVAȚIA 4.1. Rezultatul de mai sus justifică existența limitei $\theta_h^{k+1} = \lim_{i \rightarrow \infty} \bar{\theta}_h^i$. Este ușor de verificat faptul că θ_h^{k+1} rezolvă problema neliniară definită în (4.3).

OBSERVAȚIA 4.2. În exemplele numerice a fost adoptat în special schema iterativă descrisă acum. Pentru fiecare $k = \overline{1, n}$, prima iterație este corespondentul discret al metodei semi-implicite. De fapt, în majoritatea exemplurilor, nu a fost nevoie de iterații suplimentare celei inițiale.

OBSERVAȚIA 4.3. Restricția care rezultă relativ la pasul de timp τ este extrem de severă și nu este susținută practic. Cu toate acestea, în cazul în care iterația de mai sus nu converge, poate fi adoptată varianta discretă a schemei propuse în (2.7). Într-o manieră similară celei din cazul semidiscret se poate obține convergența iterațiilor fără a impune o restricție asupra parametrilor de discretizare.

4.3. Estimarea erorii. Trecem acum la analiza erorii de discretizare pentru metoda propusă în (4.3). Pentru aceasta vom considera următoarea problemă ajutătoare

$$(4.18) \quad \begin{aligned} u_h^k &\in \varepsilon + V_h, \\ (u_h^k - u_h^{k-1}, \varphi)_h + \tau < \Pi_h \beta(u_h^k), \varphi > &= 0, \end{aligned}$$

pentru orice $\varphi \in V_h$ și $i \geq 1$, cu u_h^0 definit în (4.4).

Vom arăta mai jos că această problemă nu reprezintă decât o formă echivalentă a celei definite în (4.3). Prin urmare, rezultatele obținute pentru problema auxiliară vor fi valabile și în cazul celei inițiale.

OBSERVAȚIA 4.4. Principiul de maxim este satisfăcut în cazul ambelor probleme, iar demonstrația acestei afirmații poate fi făcută fie analog demonstrațiilor din secțiunile precedente, fie ca rezultat al procedurilor iterative. La fel se poate justifica existența și unicitatea soluțiilor celor două seturi de probleme, în cazul nefolosirii iterațiilor putând fi aplicată lema Lax-Milgram neliniară.

Următoarea leamnă susține aserțiunea referitoare la echivalența problemelor.

LEMA 4.2. *Pentru orice $k = \overline{1, n}$, dacă $\theta_h^k \in \beta(\varepsilon) + V_h$ și $u_h^k \in \varepsilon + V_h$ sunt soluțiile problemelor definite în (4.3), respectiv (4.18), alegând datele inițiale conform (4.4), are loc*

$$(4.19) \quad \theta_h^k = \Pi_h \beta(u_h^k).$$

Demonstrație. Scăzând din (4.18) relația (4.3) și însumând rezultatul obținut pentru $j = \overline{1, k}$ se ajunge la

$$(4.20) \quad (u_h^k - \beta^{-1}(\theta_h^k), \varphi)_h - (u_h^0 - \beta^{-1}(\theta_h^0), \varphi)_h + \tau \sum_{j=1}^k \langle \Pi_h \beta(u_h^j) - \theta_h^j, \varphi \rangle = 0.$$

Dar, conform alegerii datelor inițiale, θ_h^0 și u_h^0 fiind din V_h , are loc $\theta_h^0 = \Pi_h \beta(u_h^0)$, deci $u_h^0 = \Pi_h \beta^{-1}(\theta_h^0)$. Rezultă de aici că al doilea termen în (4.20) este nul.

Întrucât $\Pi_h \beta(u_h^k) - \theta_h^k$ este, pentru orice k , un element din V_h , se poate lua $\varphi = \Pi_h \beta(u_h^k) - \theta_h^k$ în (4.20). Însumând rezultatul obținut pentru $k = \overline{1, n}$ și ținând seama de observația anterioară rezultă egalitatea

$$(4.21) \quad \sum_{k=1}^n (u_h^k - \beta^{-1}(\theta_h^k), \Pi_h \beta(u_h^k) - \theta_h^k)_h + \tau \sum_{k=1}^n \sum_{j=1}^k \langle \Pi_h \beta(u_h^j) - \theta_h^j, \Pi_h \beta(u_h^k) - \theta_h^k \rangle = 0.$$

Este ușor de remarcat pozitivitatea celor doi termeni din partea stângă, ceea ce implică egalitatea fiecăruia cu 0. Având în vedere definiția produsului scalar $(\cdot, \cdot)_h$, rezultă că,

pentru orice k , are loc

$$u_h^k(A) = \beta^{-1}(\theta_h^k)(A)$$

pentru fiecare nod al triangularizării S_h . Cu aceasta demonstrația este încheiată. \diamond

Evaluarea erorii de aproximare rezultă acum urmând ideile cuprinse în [13], lucrând însă cu problema auxiliară din (4.18). O demonstrație alternativă poate fi dată și prin inducție, această abordare necesitând o atenție sporită la constantele ce se însumează. Nu vom mai da detalii aici, mulțumindu-ne să prezentăm rezultatul final.

TEOREMA 4.3. *Presupunând adevărate (A1) și (A2), dacă u este soluția slabă a problemei în (1.1) și, pentru orice $k > 0$, θ_h^k rezolvă ecuația (4.3), luând $h = O(\tau)$ are loc*

$$\sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\beta(u(t)) - \theta_h^k\|^2 dt \leq C \{ \tau + \varepsilon^2 + (C_2(\varepsilon)\beta(\varepsilon))^2 + \beta^2(\varepsilon) \}.$$

OBSERVAȚIA 4.5. Restricția impusă parametrului h este de ordin pur teoretic. Ea contravine condiției impuse lui τ în secțiunea precedentă și nu este verificată practic. Abordări diferite ale estimărilor de eroare - cum ar fi cea din [33] - conduc la eliminarea oricărei restricții de acest gen, dar rezultatul va depinde de $C_2(\varepsilon)$. Acest fapt alterează ordinul de convergență al metodei semidiscrete, ceea ce nu este cazul în rezultatul de mai sus.

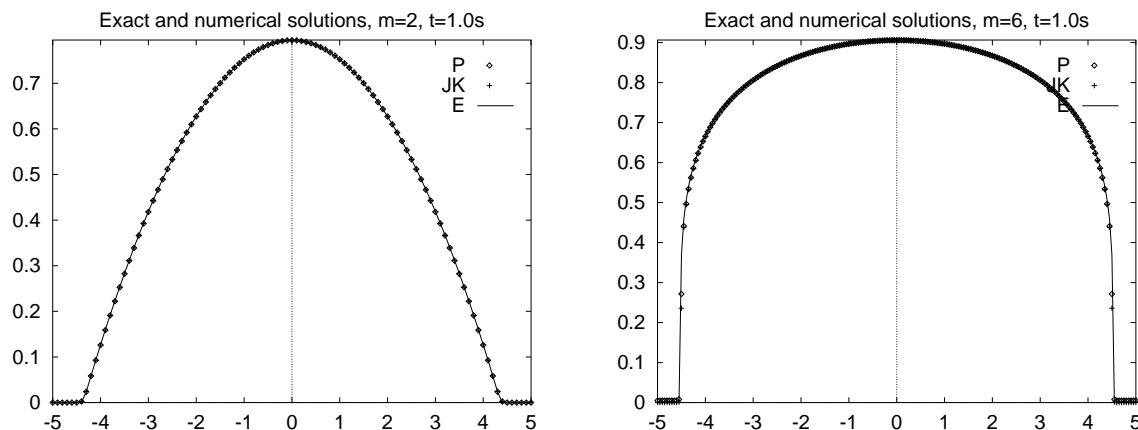
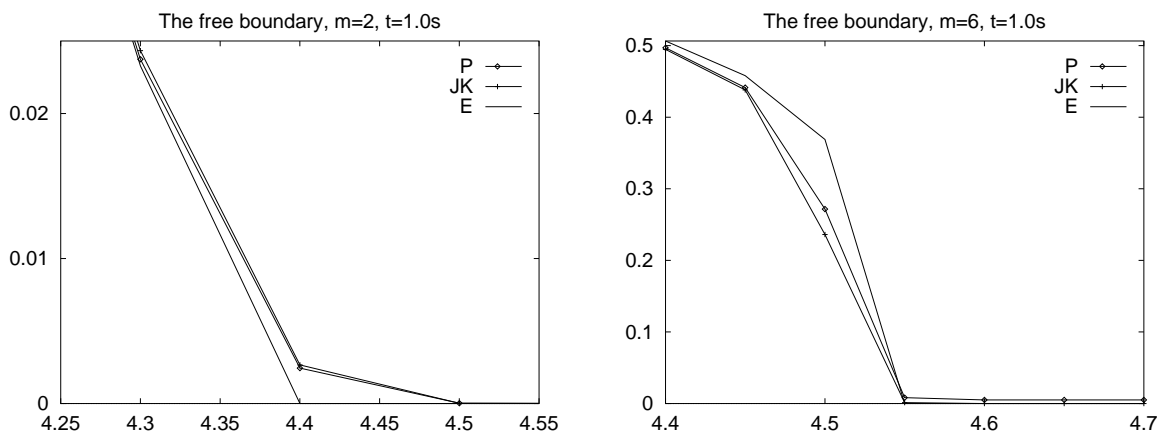
OBSERVAȚIA 4.6. Toate observațiile făcute în cazul semidiscret rămân valabile și în acest context, incluzînd și estimarea erorii referitoare la u .

5. Exemple numerice

Această secțiune este dedicată susținerii efectivității metodelor propuse prin exemple numerice. Problema test considerată a fost definită în (1.1), iar o soluție exactă este dată în (1.2). Exemplele se referă la aceasta situație.

Se poate remarca faptul că metoda însăși sugerează utilizarea extrapolării în raport cu ε . Deși nu avem o justificare riguroasă pentru aceasta, rezultatele numerice arată că aproximația obținută în acest mod este mai bună, dar implică un efort suplimentar de calcul.

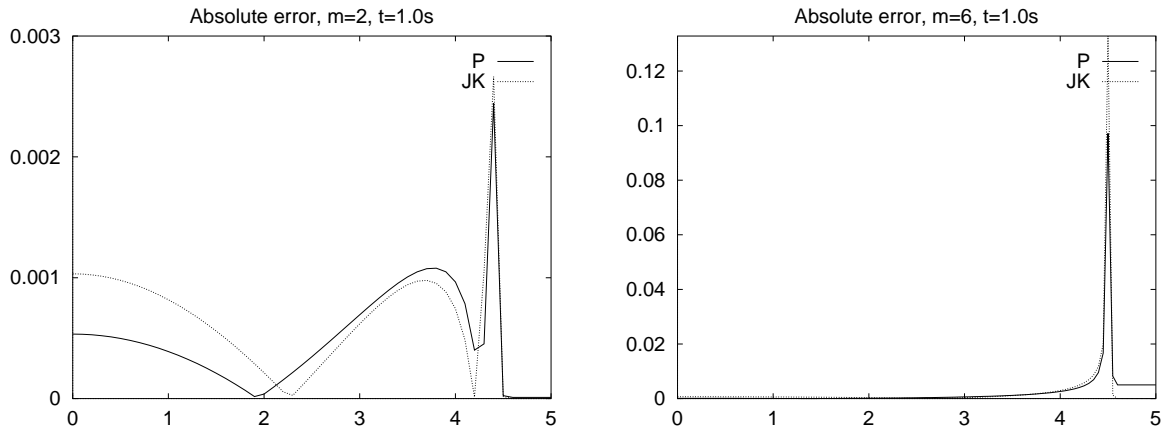
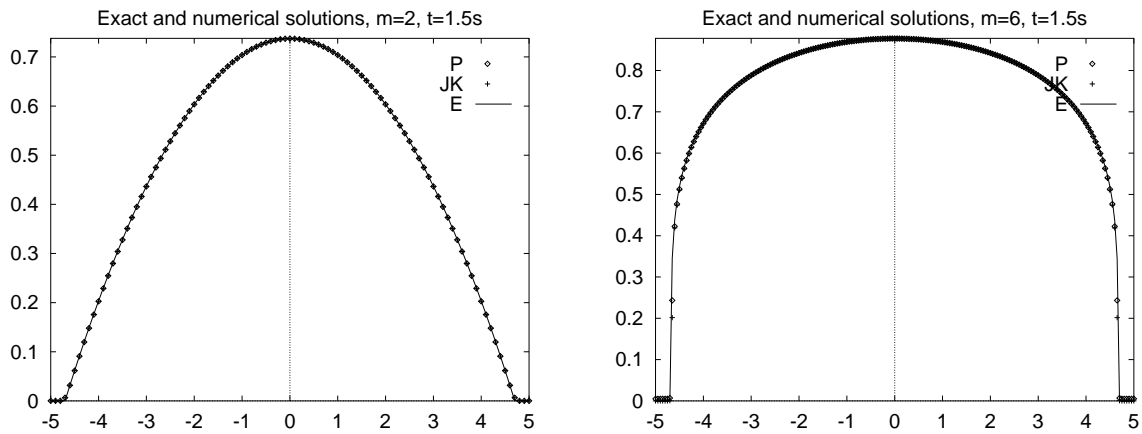
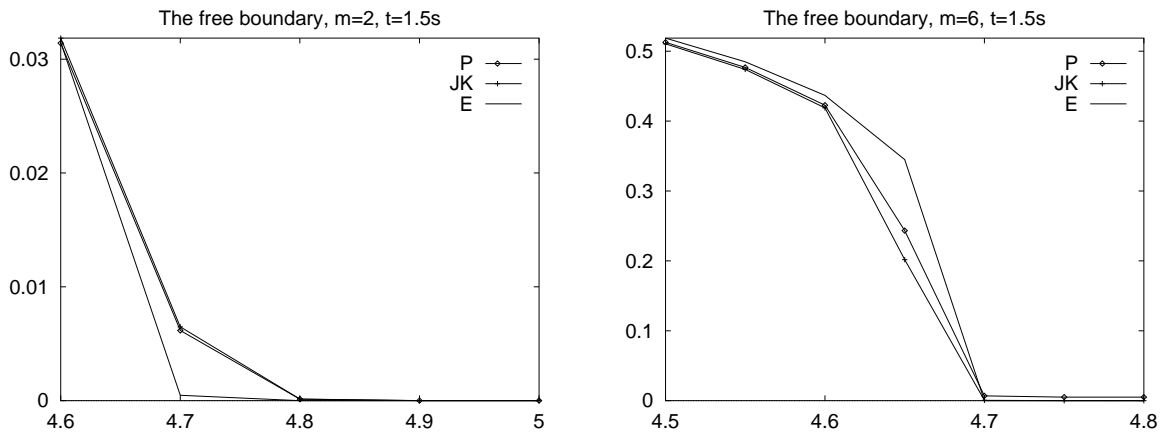
Figurile prezintă eroarea absolută - corespunzătoare normei L^∞ , dar analiza a fost făcută pentru norma L^2 . Acest lucru este vizibil și prin faptul că o rafinare a discretizării spațiale nu are un efect proporțional asupra erorii absolute, aceasta restrângându-și, totuși, suportul. Fenomenul este explicabil prin faptul că erorile cele mai pronunțate apar în apropierea frontierei libere, deci ponderea măsurii acestei regiuni este mai mică în raport cu cea a întregului domeniu.

FIGURA 1. Soluțiile la momentul $t=1.0s$ FIGURA 2. Frontierele libere la momentul $t=1.0s$

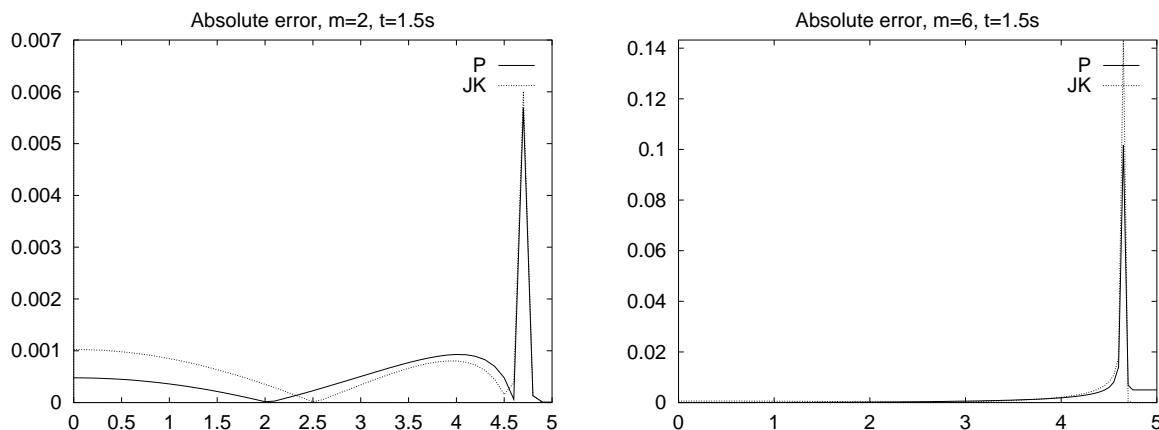
Primele figuri prezintă rezultate obținute în cazul unidimensional. S-a considerat intervalul $[-5.0, 5.0]$, iar pentru neliniaritatea $\beta(u) = u^m$ au fost tratate două situații, $m = 2$ și $m = 6$. v_0 respectă descrierea din (2.3) iar parametrul perturbației are valoarea $\varepsilon = 0.0001$ pentru primul caz și $\varepsilon = 0.005$ pentru cel de-al doilea. Pasul de timp este constant, anume $\tau = 0.01$.

Ecuatiile eliptice rezultate au fost discretizate pe o grilă uniformă cu $h = 0.1$, respectiv $h = 0.05$. Soluția exactă este comparată cu rezultatele numerice furnizate de metoda bazată pe principiul de maxim, respectiv de metoda descrisă în [17]. În cazul ambilor algoritmi s-a efectuat același număr de iterații. Pentru $m = 2$ o execuție a fost suficientă, în timp ce pentru $m = 6$ patru iterații - așa cum au fost descrise în (4.7) - au produs o îmbunătățire a rezultatului.

În toate figurile soluția exactă este notată cu E, cea numerică obținută prin metoda Jäger-Kačur cu JK, iar pentru aproximația bazată pe perturbarea datelor s-a folosit notația P. În figurile 1 și 4 sunt prezentate în paralel graficele soluțiilor pentru $m = 2$ și

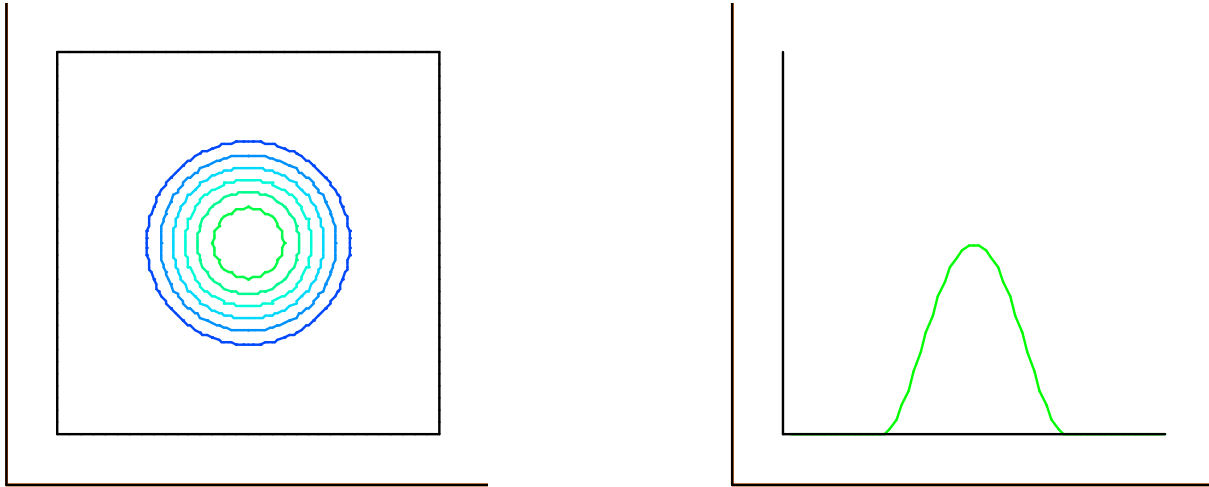
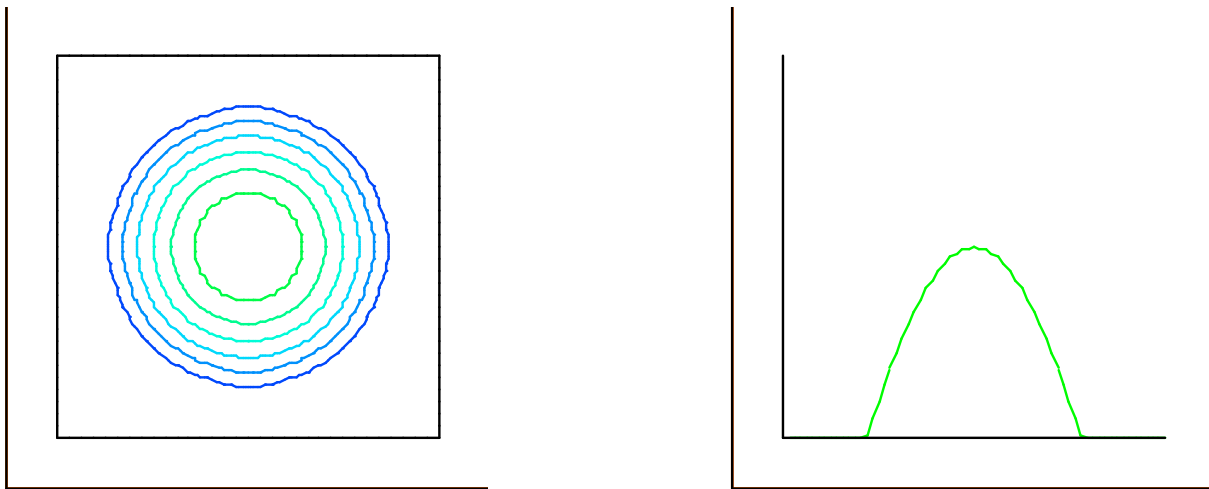
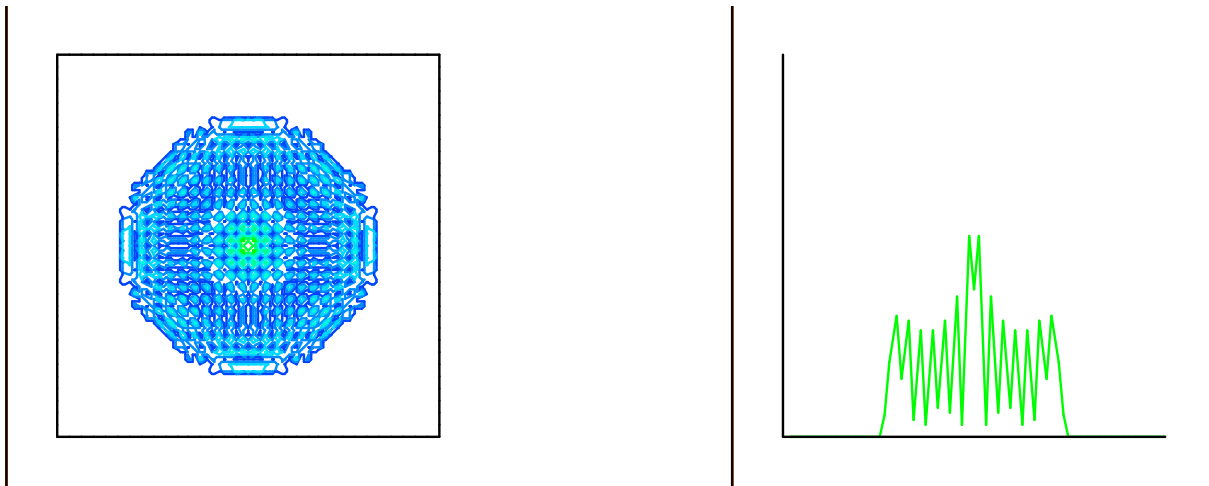
FIGURA 3. Erorile absolute la momentul $t=1.0s$ FIGURA 4. Soluțiile la momentul $t=1.5s$ FIGURA 5. Frontierele libere la momentul $t=1.5s$

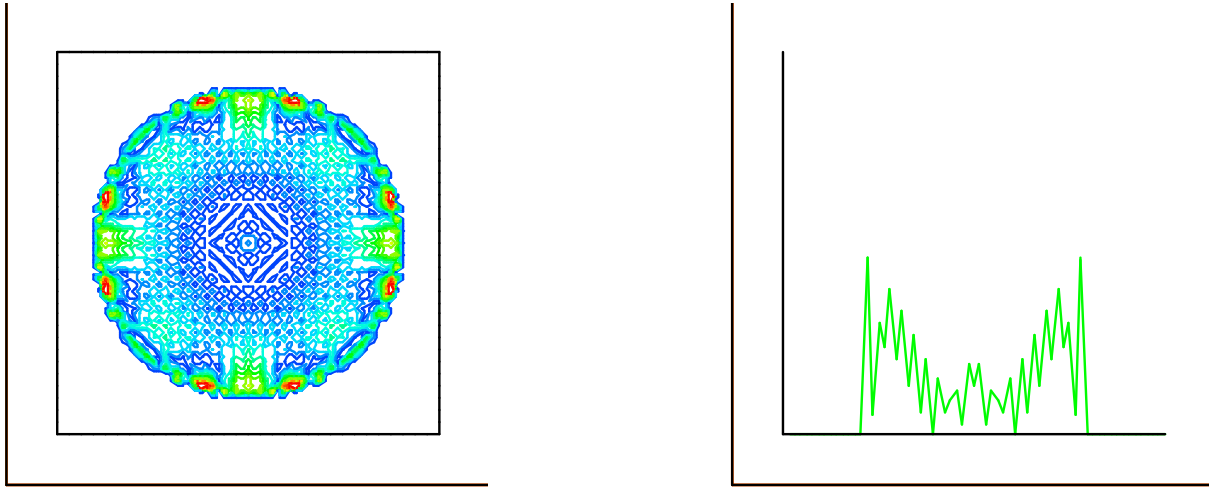
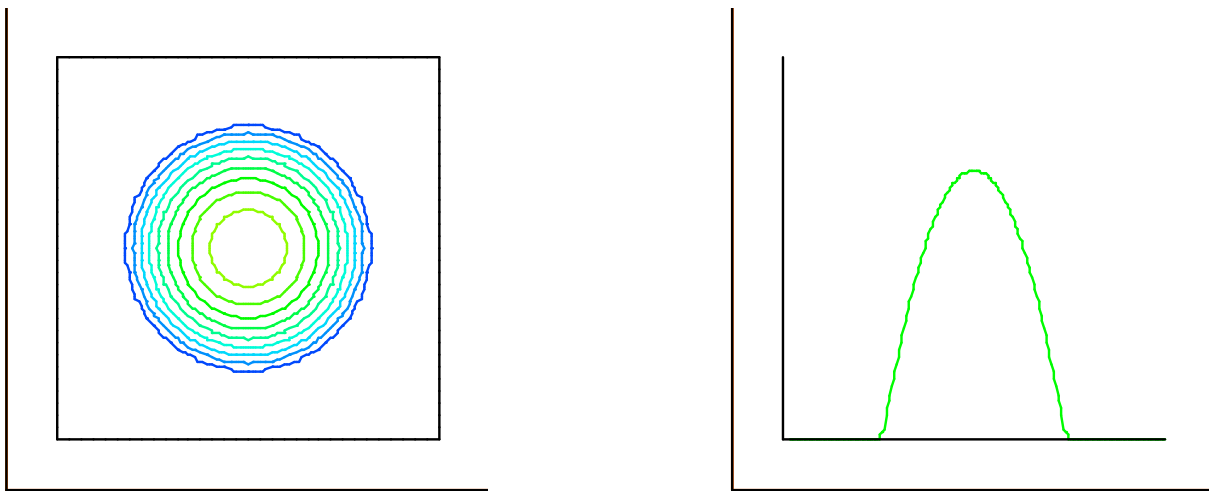
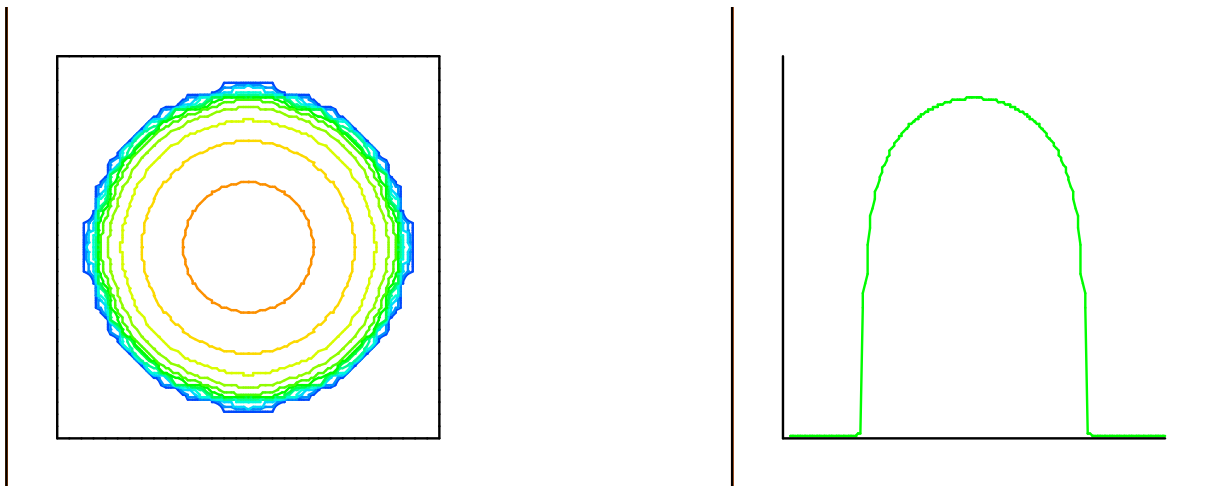
$m = 6$ în două secțiuni de timp, $t = 1s$ și $t = 1.5s$. La scara de reprezentare aleasă nu se poate face nici o distincție între soluția exactă și cele numerice. Prin urmare, figurile 2 și 5 prezintă detaliile din zona frontierei libere, zona critică. Chiar și aici este greu de făcut

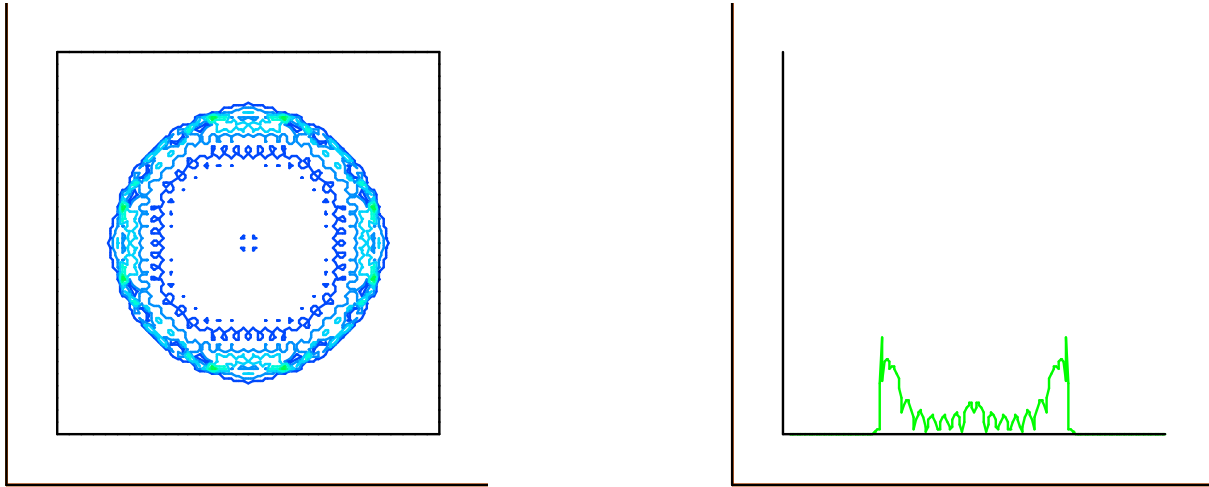
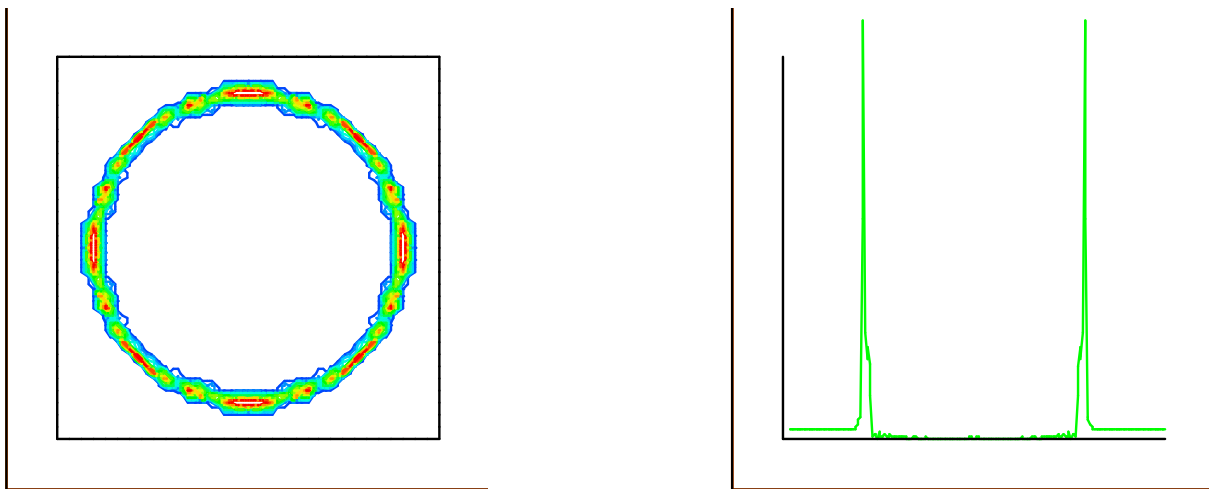
FIGURA 6. Erorile absolute la momentul $t=1.5s$

o distincție între cele două rezultate numerice. Comportamentul similar este susținut și de graficul erorilor absolute, așa cum rezultă din figurile 3 și 6. Ultimele două grafice aparțin erorilor absolute obținute cu și fără extrapolare. Ele sunt notate cu P, respectiv Ex.

Aceleași situații au fost considerate și în cazul bidimensional, cu excepția faptului că acum nu au fost efectuate iterații nici în cazul $m = 6$, iar domeniul bidimensional considerat este $[-7.0, 7.0] \times [-7.0, 7.0]$. Pentru implementarea exemplilor numerice prezentate în acest caz s-a utilizat pachetul de programe *UG* ([7]), disponibil la adresa <http://dom.ica3.uni-stuttgart.de/>. Toate datele prezentate sunt calculate la momentul $t = 1.0s$. Reprezentările grafice au în partea stângă proiecția plană a liniilor de contur, iar în dreapta secțiunea prin origine, paralelă cu axa Oz , a datelor bidimensionale. În figurile 7 și 8 sunt reprezentate aproximațiile mărimii $\theta = \beta(u)$ pentru $m = 2$, respectiv $m = 6$, iar figurile 9 și 10 conțin erorile absolute în aceste cazuri. Următoarele figuri prezintă situațiile similare pentru necunoscuta u . Rezolvarea numerică a problemelor rezultate s-a făcut prin metoda multigrad în patru rafinări succesive ale domeniului, pornind de la o triangularizare uniformă cu un nod interior și opt pe frontieră. Ordinul de mărime al datelor reprezentate este cuprins între 0 și 1 pentru soluții, iar în cazul erorilor după cum urmează. Eroarea în θ este reprezentată pe scala 0 – 0.01, iar pentru cea în u s-a ales 0 – 0.04 în cazul $m = 2$, respectiv 0 – 0.2 pentru $m = 6$. Această ultimă eroare este puternic afectată de interpolarea datelor în zona gradientilor mari. Se poate observa însă faptul că erorile se concentrează în jurul frontierei libere.

FIGURA 7. Soluția numerică θ , $m=2$, $t=1.0s$ FIGURA 8. Soluția numerică θ , $m=6$, $t=1.0s$ FIGURA 9. Eroarea absolută $\beta(u) - \theta$, $m=2$, $t=1.0s$

FIGURA 10. Eroarea absolută $\beta(u) - \theta$, $m=6$, $t=1.0s$ FIGURA 11. Soluția numerică u , $m=2$, $t=1.0s$ FIGURA 12. Soluția numerică u , $m=6$, $t=1.0s$

FIGURA 13. Eroarea absolută $u - \beta^{-1}(\theta)$, $m=2$, $t=1.0s$ FIGURA 14. Eroarea absolută $u - \beta^{-1}(\theta)$, $m=6$, $t=1.0s$

CAPITOLUL 4

Concluzii

Prin lucrarea de față s-a urmărit introducerea autorului în unele aspecte legate de analiza numerică a ecuațiilor diferențiale cu derivate parțiale. S-a abordat un domeniu concret, legat de ecuațiile parabolice degenerare - care apar sub diverse forme ca rezultat al analizei aplicate, în ideea familiarizării prin exercițiu efectiv cu una dintre cele mai răspândite metode de discretizare spațială - metoda elementului finit - precum și cu schemele numerice pentru probleme nestaționare.

Au fost expuse cu precădere rezultatele obținute în urma aplicării acestor metode în cazul problemelor menționate, aceasta în detrimentul unei expuneri exhaustive. Acesta este și motivul pentru care s-a renunțat la definirea tuturor noțiunilor care intervin în referat ori la includerea rezultatelor implicate, rezumându-ne la prezentarea în esență a aspectelor cu caracter de noutate.

Centrul de greutate al lucrării constă în prezentarea algoritmilor bazați pe principiul de maxim. Deși această idee a apărut de foarte mult timp în literatură, conform cunoștințelor autorului ea nu a mai fost folosită până acum în vederea rezolvării numerice a ecuațiilor de tipul celui tratat. Pornind de aici s-au propus două scheme numerice. Cea neliniară se bazează pe o discretizare implicită în timp, pentru ea propunându-se și metode iterative de rezolvare, în timp ce varianta liniară are la bază o discretizare semi-implicită.

Totodată, vom remarca faptul că schemele propuse sunt simple - așadar ușor de implementat - și eficiente. În sprijinul acestei afirmații venim atât cu exemplele numerice, cât și cu analiza teoretică. De menționat este ordinul erorii de aproximare obținut în cazul schemei neliniare, care până anul trecut a fost considerat optimal, fiind cel puțin comparabil cu cel al schemelor pentru care există certitudinea aplicabilității (prin aceasta am făcut referire la algoritmi pentru care există și exemple numerice).

Dintre posibilele extinderi ale metodei vom menționa unele cu caracter imediat. Pot fi considerate condiții la frontieră de tip Neumann ori mixt, însă acestea trebuie să permită un control asupra degenerării de genul celui propus. De asemenea sunt permisi termenii de ordin nul - termenii sursă, fiind eventual necesară o transformare a ecuației în prealabil. Perspective pentru această abordare se deschid și în cazul degenerării cauzate de funcții

având alte caracteristici (cum ar fi, de exemplu, probele Stefan), ori pentru ecuații dublu degenerate.

Nu se poate însă trece peste unele neajunsuri ale acestor scheme. În primul rând, abordarea propusă este posibilă exclusiv în cazul în care este valabil principiul de maxim. De aceea gradul de generalitate al schemei este mai restrâns, ea neputând fi extinsă decât la unele sisteme de ecuații. De asemenea, nu sunt tratați termenii de ordinul întâi, aceștia necesitând o discretizare specifică. Ambele aspecte apar frecvent în matematica aplicată, cum ar fi curgerea multifazică, respectiv transportul și convecția, și necesită un studiu suplimentar. Acestea ar fi direcții generale de urmat pentru viitor.

Un alt aspect extrem de important rămas netratat - care la rândul lui constite unul dintre subiectele demne de a fi luate în considerare în cele ce urmează - îl constituie metodele de discretizare adaptivă, în combinație cu tehnicile de rezolvare eficientă. Toate aceste lucruri reprezintă un vast câmp de lucru, intens studiat la ora actuală.

Bibliografie

- [1] H. W. Alt & S. Luckhaus, Quasilinear Elliptic-Parabolic Differential Equations, *Math. Z.*, 183(1983), 311 - 341.
- [2] G. Amiez & P. A. Gremaud, On a Numerical Approach to Stefan-like Problems, *Numer. Math.*, 59(1991), 71 - 89.
- [3] T. Arbogast & M. F. Wheeler & N.-Y. Zhang, A Nonlinear Mixed Finite Element Method for a Degenerate Parabolic Equation Arising in Flow in Porous Media, *SIAM J. Numer. Anal.*, 33(1996), 1669 - 1687.
- [4] D. G. Aronson, The Porous Medium Equation, in Nonlinear Diffusion Problems, A. Fasano & M. Primicerio (eds.), *Lecture Notes in Mathematics 1224*, Springer Verlag, Berlin, 1985, 1 - 46.
- [5] E. Bänsch, Numerical Experiments with Adaptivity for the Porous Medium Equation, va apărea în *Acta Math. Univ. Comenianae*.
- [6] G. I. Barenblatt, On Some Unsteady Motion of a Liquid or a Gas in a Porous Medium, *Prikl. Math. Meh.*, 16(1952), 67 - 78.
- [7] P. Bastian & K. Birken & K. Johannsen & S. Lang & N. Neuss & H. Rentz-Reichert & C. Wieners, UG - A Flexible Software Toolbox for Solving Partial Differential Equations, to appear in *Computing and Visualization in Science*.
- [8] J. Bear, Dynamics of Fluids in Porous Media, *American Elsevier, New York*, 1972.
- [9] P. G. Ciarlet, The Finite Element Method for Elliptic Problems, *North-Holland, Amsterdam*, 1978.
- [10] P. G. Ciarlet & P. A. Raviart, Maximum Principle and Uniform Convergence for the Finite Element Method, *Comput. Methods Appl. Mech. Engrg.*, 2(1973), 17 - 31.
- [11] J. F. Ciavaldini, Analyse numérique d'un problème de Stefan á deux phases par une méthode d'éléments finis, *SIAM J. Numer. Anal.*, 12(1975), 464 - 487.
- [12] E. Di Benedetto & D. Hoff, An Interface Tracking Algorithm for the Porous Medium Equation, *Trans. Amer. Math. Soc.*, 284(1984), 463 - 500.
- [13] C. M. Elliott, Error Analysis of the Enthalpy Method for the Stefan Problem, *IMA J. Numer. Anal.*, 7(1987), 61 - 71.
- [14] A. Kufner, O. John, S Fučić, Function Spaces, *Academia CSAV, Prague*, 1967.
- [15] A. Handlovičová, Error Estimates of a Linear Approximation Scheme for Nonlinear Diffusion Problems, *Acta Math. Univ. Comenianae*, 61(1992), 27 - 39.
- [16] W. Jäger & J. Kačur, Solution of Porous Medium Type Systems by Linear Approximation Schemes, *Numer. Math.*, 60(1991), 407 - 427.
- [17] W. Jäger & J. Kačur, Solution of Doubly Nonlinear and Degenerate Parabolic Problems by Relaxation Schemes, *M²AN (Math. Model. Numer. Anal.)*, 29(1995), 605 - 627.

- [18] J. W. Jerome, Approximation of Nonlinear Evolution Systems, *Academic Press, New York*, 1983.
- [19] X. Jiang & R. H. Nochetto, Optimal Error Estimates for Semidiscrete Phase Relaxation Models, *M²AN (Math. Model. Numer. Anal.)*, 31(1997), 91 - 120.
- [20] J. Kačur, On a Solution of Degenerate Elliptic-Parabolic Systems in Orlicz-Sobolev Spaces I; II, *Math. Z.*, 203(1990), 153 - 171; 569 - 579.
- [21] J. Kačur, Solution to Strongly Nonlinear Parabolic Problems by a Linear Approximation Scheme, *Preprint M1-94 (1994)*, Comenius University Bratislava, Faculty of Mathematics and Physics.
- [22] J. Kačur & A. Handlovičová & M. Kačurová, Solution of Nonlinear Diffusion Problems by Linear Approximation Schemes, *SIAM J. Numer. Anal.*, 30(1993), 1703 - 1722.
- [23] E. Magenes & R. H. Nochetto & C. Verdi, Energy Error Estimates for a Linear Scheme to Approximate Nonlinear Parabolic Problems, *M²AN (Math. Model. Numer. Anal.)*, 21(1987), 655 - 678.
- [24] R. H. Nochetto & M. Paolini & C. Verdi, An Adaptive Finite Element Method for Two-Phase Stefan Problems in Two Space Dimensions. Part I: Stability and Error Estimates, *Math. Comput.*, 57(1991), 73 - 108.
- [25] R. H. Nochetto & M. Paolini & C. Verdi, A Fully Discrete Adaptive Nonlinear Chernoff Formula, *SIAM J. Numer. Anal.*, 30(1993), 991 - 1914.
- [26] R. H. Nochetto & C. Verdi, Approximation of Degenerate Parabolic Problems Using numerical Integration, *SIAM J. Numer. Anal.*, 25(1988), 784 - 814.
- [27] O. A. Oleinik & A. S. Kalashnikov & Y.-L. Zhou, The Cauchy Problem and Boundary Value Problems for Equations of the Type of Unsteady Filtration, *Izv. Akad. Nauk SSSR Ser. Mat.*, 22(1958), 667 - 704.
- [28] T. Petrilă & C. I. Gherorghiu, Metode element finit și aplicații, *Ed. Acad. R.S.R., București*, 1987.
- [29] I. S. Pop & W. A. Yong, A Maximum Principle Based Numerical Approach to Porous Medium Equations, va apărea în volumul conferinței ALGORTIMY'97.
- [30] P. A. Raviart, The Use of Numerical Integration in finite Element Methods for Solving Parabolic Equations, in Topics in Numerical Analysis, J. J. H. Miller (ed.), *Academic Press, London*, 1973, 233 - 264.
- [31] J. Rulla, Error Analysis for Implicit Approximations to Solutions to Cauchy Problems, *SIAM J. Numer. Anal.*, 33(1996), 68 - 87.
- [32] M. Slodička, Solution of Nonlinear Parabolic Problems by Linearization, *Preprint M3-92 (1992)*, Comenius University Bratislava, Faculty of Mathematics and Physics.
- [33] M. Slodička, On a Numerical Approach to Nonlinear Degenerate Parabolic Problems, *Preprint M6-92 (1992)*, Comenius University Bratislava, Faculty of Mathematics and Physics.
- [34] J. Smoller, Shock Waves and Reaction-Diffusion Equations, *Springer-Verlag, New York*, 1983.
- [35] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, Lecture Notes in Math. 1054, *Springer-Verlag, Berlin*, 1984.
- [36] C. Wagner, Numerical Methods for Diffusion-Reaction-Transport Processes in Unsaturated Porous Media, în preparare.
- [37] W. A. Yong & I. S. Pop, A Numerical Approach to Porous Medium Equations, *Preprint 96-50 (SFB 359) (1996)*, IWR, University of Heidelberg.