



# Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών  
Επιστημών

Τομέας Μαθηματικών

## Αριθμητική Ανάλυση για Σχεδόν γραμμικά Παραβολικά Προβλήματα

Συγγραφέας:  
Παναγιώτης Ηλία

Επιβλέπων Καθηγητής:  
Χρυσάφινος Κωνσταντίνος

19 Ιουλίου 2021



# Περιεχόμενα

Περιεχόμενα	i
<b>1 Η Μέθοδος Πεπερασμένων Στοιχείων Galerkin για την εξίσωση της θερμότητας</b>	<b>1</b>
1.1 Εισαγωγή . . . . .	1
1.2 Χώροι Sobolev και ασθενής παράγωγος . . . . .	2
1.3 Ελλειπτικό πρόβλημα . . . . .	4
1.4 Η εξίσωση της θερμότητας . . . . .	11
<b>2 Γενικότερες Παραβολικές Εξισώσεις</b>	<b>35</b>
2.1 Εισαγωγή . . . . .	35
2.2 Ημιδιακριτό πρόβλημα . . . . .	38
2.3 Πλήρως διακριτό πρόβλημα . . . . .	50
<b>3 Σχεδόν γραμμικές παραβολικές εξισώσεις</b>	<b>56</b>
3.1 Εισαγωγή . . . . .	56
3.2 Κατασκευή ημιδιακριτού προβλήματος . . . . .	58
3.3 Ανάλυση σφάλματος ημιδιακριτού προβλήματος . . . . .	59
3.4 Κατασκευή πλήρους διακριτού προβλήματος . . . . .	65
3.5 Ανάλυση σφάλματος πλήρους διακριτού προβλήματος . . . . .	66
<b>4 Υπολογιστικά παραδείγματα</b>	<b>70</b>
4.1 Εισαγωγή . . . . .	70

---

4.2	Μέθοδος υπολογισμού τάξης σφάλματος ως προς το χώρο και το χρόνο . . . . .	70
4.3	Πρόβλημα αρχικών/συνοριακών τιμών για την εξίσωση της θερμότητας . . . . .	72
4.4	Πρόβλημα αρχικών/συνοριακών τιμών με γνωστή ακριβή λύση .	75
4.5	Σχεδόν γραμμικό παραβολικό πρόβλημα αρχικών/συνοριακών τιμών . . . . .	78
4.5.1	Μέθοδος σταθερού σημείου . . . . .	78
4.5.2	Μέθοδος Newton . . . . .	79
4.5.3	Υπολογιστικό παράδειγμα . . . . .	80
	<b>Κώδικες Python</b>	<b>84</b>
	<b>Βιβλιογραφία</b>	<b>92</b>

# Κεφάλαιο 1

## Η Μέθοδος Πεπερασμένων Στοιχείων Galerkin για την εξίσωση της θερμότητας

### 1.1 Εισαγωγή

Σε αυτό το εισαγωγικό κεφάλαιο θα μελετήσουμε την κλασική Μέθοδο Πεπερασμένων Στοιχείων Galerkin για την εύρεση μιας προσεγγιστικής λύσης του προβλήματος αρχικών/συνοριακών τιμών για την εξίσωση της θερμότητας

$$\begin{aligned}u_t - \Delta u &= f \text{ στο } \Omega, \text{ για } t > 0, \\u &= 0 \text{ στο } \partial\Omega, \text{ για } t > 0, \\u(\cdot, 0) &= v \text{ στο } \Omega,\end{aligned}\tag{1.1}$$

όπου  $\Omega \subset \mathbb{R}^d$  με ομαλό σύνορο  $\partial\Omega$ . Επίσης η  $u = u(x, t)$  καλείται λύση του προβλήματος, το  $u_t$  συμβολίζει την μερική παράγωγο της  $u$  ως προς  $t$   $\partial u/\partial t$  και  $\Delta = \sum_{j=1}^d \partial^2/\partial x_j^2$  τον Λαπλασιανό τελεστή.

Πριν ξεκινήσουμε την μελέτη μας για το πρόβλημα (1.1) θα δούμε σύντομα κάποια βασικά πράγματα για την κλασική Μέθοδο Πεπερασμένων Στοιχείων (ΜΠΣ) Galerkin στο αντίστοιχο στατικό πρόβλημα, το οποίο είναι το πρόβλημα Dirichlet για την εξίσωση Poisson

$$\begin{aligned}-\Delta u &= f \text{ στο } \Omega, \\u &= 0 \text{ στο } \partial\Omega.\end{aligned}\tag{1.2}$$

## 1.2 Χώροι Sobolev και ασθενής παράγωγος

Πριν ασχοληθούμε με την κατασκευή της ΜΠΣ για τα προβλήματα (1.2) και (1.1), χρειάζεται να αναφέρουμε κάποια βασικά στοιχεία για την παραγωγή και τους χώρους Sobolev. Όπως γνωρίζουμε από την βασική ανάλυση, δεν είναι όλες οι συνεχείς συναρτήσεις παραγωγίσιμες. Σε αυτή την παράγραφο θα ορίσουμε μια νέα έννοια "παραγωγίσιμης" η οποία θα μας επιτρέψει να γενικεύσουμε την έννοια της παραγώγου μιας συνάρτησης. Για να το κάνουμε αυτό, πρέπει πρώτα να αναφέρουμε δύο βοηθητικούς ορισμούς.

**Ορισμός 1.1.** Έστω η συνάρτηση  $f : \Omega \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ , όπου  $\Omega$  ανοικτό σύνολο. Καλούμε στήριγμα ή φορέα (*support*) της  $f$  το σύνολο  $\overline{\{x \in \Omega : f(x) \neq 0\}}$ , όπου  $\bar{A}$  ενός  $A \subseteq \mathbb{R}^d$  συμβολίζει την κλειστή θήκη. Συμβολίζουμε με  $\text{supp} f$  τον φορέα της συνάρτησης  $f$ .

**Ορισμός 1.2.** Έστω  $\Omega \subseteq \mathbb{R}^d$  ανοικτό. Συμβολίζουμε με  $C_0^\infty(\Omega)$  την οικογένεια των συναρτήσεων  $f : \Omega \rightarrow \mathbb{R}$ , οι οποίες είναι απείρως διαφορίσιμες και έχουν συμπαγή φορέα στο  $\Omega$ .

Παρατηρούμε ότι αφού το  $\Omega \subseteq \mathbb{R}^d$  είναι ανοικτό, και ο φορέας κάθε συνάρτησης  $f \in C_0^\infty(\Omega)$  είναι συμπαγές υποσύνολο του  $\Omega$ , άρα κλειστό και φραγμένο. Συνεπώς  $f|_{\partial\Omega} = 0$ . Ο χώρος  $C_0^\infty(\Omega)$  είναι ένας άπειρης διάστασης διανυσματικός χώρος, ο οποίος παίζει πολύ σημαντικό ρόλο στην σύγχρονη θεωρία των συναρτήσεων. Συγκεκριμένα, χρησιμοποιείται για την γενίκευση της έννοιας της παραγώγου μιας συνάρτησης.

**Ορισμός 1.3.** Έστω  $\Omega \subset \mathbb{R}^d$  ανοικτό σύνολο. Μια συνάρτηση  $g : \Omega \rightarrow \mathbb{R}$  καλείται *ασθενής μερική παράγωγος* της  $f : \Omega \rightarrow \mathbb{R}$  στην κατεύθυνση  $x_i$  αν

$$\int_{\Omega} g \varphi dx = - \int_{\Omega} f \frac{\partial \varphi}{\partial x_i} dx, \quad \forall \varphi \in C_0^\infty(\Omega). \quad (1.3)$$

**Θεώρημα 1.1.** Έστω  $\Omega \subset \mathbb{R}^d$  ανοικτό, και  $f : \Omega \rightarrow \mathbb{R}$  παραγωγίσιμη συνάρτηση. Τότε η κλασική παράγωγος της  $f$  την οποία συμβολίζουμε με  $f'$  είναι και ασθενής παράγωγος της. Το αντίστροφο δεν ισχύει γενικά.

Από εδώ και στο εξής θα συμβολίζουμε την ασθενή παράγωγο μιας συνάρτησης  $f : \Omega \rightarrow \mathbb{R}$  με  $Df$ .

Αφού γενικεύσαμε την έννοια της παραγώγου και μπορούμε να την χρησιμοποιούμε και για συναρτήσεις που δεν είναι διαφορίσιμες με την κλασική έννοια, είμαστε σε θέση να ορίσουμε τους πιο βασικούς χώρους συναρτήσεων που θα χρησιμοποιήσουμε.

**Ορισμός 1.4.** Έστω  $\Omega \subseteq \mathbb{R}^d$  ανοικτό. Ορίζουμε την οικογένεια συναρτήσεων  $L^2(\Omega) = L^2$  ως εξής

$$L^2(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R} \mid \|f\|_{L^2(\Omega)} := \left( \int_{\Omega} f^2(x) dx \right)^{1/2} < \infty \right\}. \quad (1.4)$$

Δηλαδή η  $L^2(\Omega)$  είναι η οικογένεια όλων των τετραγωνικά ολοκληρώσιμων συναρτήσεων. Η ποσότητα  $\|f\|_{L^2(\Omega)}$  καλείται  $L^2$ -νόρμα της  $f$ , και αποτελεί νόρμα αυτού του χώρου.

**Ορισμός 1.5.** Έστω  $\Omega \subseteq \mathbb{R}^d$  ανοικτό. Ορίζουμε την οικογένεια συναρτήσεων  $H^r = H^r(\Omega) = W_2^r(\Omega)$  ως εξής

$$H^r(\Omega) := \left\{ f \in L^2(\Omega) \mid \|f\|_r = \|f\|_{H^r} := \left( \sum_{|\alpha| \leq r} \|D^\alpha f\|^2 \right)^{1/2} < \infty \right\}, \quad (1.5)$$

όπου  $\alpha = (\alpha_1, \dots, \alpha_d)$ ,  $D^\alpha = (\partial/\partial x_1)^{\alpha_1} \dots (\partial/\partial x_d)^{\alpha_d}$  συμβολίζει μια παράγωγο ως προς  $x$  τάξης  $|\alpha| = \sum_{j=1}^d \alpha_j$  έτσι ώστε το άθροισμα στην σχέση (1.5) να περιέχει όλες τις μερικές παραγώγους τάξης το πολύ  $r$ .

Γνωρίζουμε ότι οι παραπάνω χώροι συναρτήσεων είναι διανυσματικοί χώροι. Ο χώρος  $L^2$  είναι ένα παράδειγμα χώρου Lebesgue, και οι χώροι  $H^r$ , όπου  $r$  θετικός ακέραιος καλούνται χώροι Sobolev. Ο απλούστερος χώρος Sobolev είναι ο

$$H^1 = H^1(\Omega) := \{f \in L^2(\Omega) \mid Df \in L^2(\Omega)\}. \quad (1.6)$$

Επίσης ορίζουμε τον υπόχωρο

$$H_0^1 = H_0^1(\Omega) := \{f \in H^1(\Omega) \mid f|_{\partial\Omega} = 0\}. \quad (1.7)$$

Να σημειώσουμε ότι με  $f|_{\partial\Omega}$  συμβολίζουμε για ευκολία τον τελεστή ίχνους  $tr f$  της συνάρτησης  $f$ . Για κάθε συνάρτηση  $v \in H_0^1(\Omega) = H_0^1$ , γνωρίζουμε ότι οι συναρτήσεις  $v$  και  $\nabla v$  ανήκουν στον  $L^2$  και παίρνουν την τιμή μηδέν στο  $\partial\Omega$ . Επίσης οι νόρμες  $\|\nabla v\|$  και  $\|v\|_1$  είναι ισοδύναμες (βλ. [1] ή [2]). Συνεπώς ισχύει η εξής σχέση

$$c\|v\|_1 \leq \|\nabla v\| \leq \|v\|_1, \quad \forall v \in H_0^1, \text{ με } c > 0, \quad (1.8)$$

όπου η σταθερά  $c$  εξαρτάται από το χωρίο  $\Omega$ .

### 1.3 Ελλειπτικό πρόβλημα

Για να κατασκευάσουμε την ΜΠΣ Galerkin για το πρόβλημα Dirichlet (1.2), πρέπει πρώτα να το γράψουμε σε ασθενή μορφή. Για να το κάνουμε αυτό πολλαπλασιάζουμε την ΜΔΕ με μια συνάρτηση ελέγχου  $\varphi \in H_0^1$  και ολοκληρώνουμε στο  $\Omega$ , παίρνοντας

$$-\int_{\Omega} \Delta u \varphi \, dx = \int_{\Omega} f \varphi \, dx, \quad \forall \varphi \in H_0^1. \quad (1.9)$$

Εφαρμόζοντας το Θεώρημα του Gauss (divergence theorem) στο αριστερό μέλος παίρνουμε

$$\int_{\Omega} \nabla u \cdot \nabla \varphi \, dx - \int_{\partial\Omega} \varphi \nabla u \cdot \vec{n} \, dS = \int_{\Omega} f \varphi \, dx, \quad \forall \varphi \in H_0^1. \quad (1.10)$$

όπου  $\vec{n}$  είναι το μοναδιαίο κάθετο διάνυσμα στο  $\partial\Omega$ . Αφού  $\varphi = 0$  στο  $\partial\Omega$ , έχουμε τελικά

$$(\nabla u, \nabla \varphi) = (f, \varphi), \quad \forall \varphi \in H_0^1, \quad (1.11)$$

όπου χρησιμοποιήσαμε τα  $L^2$  εσωτερικά γινόμενα

$$(v, w) = \int_{\Omega} v w \, dx, \quad (\nabla v, \nabla w) = \int_{\Omega} \nabla v \cdot \nabla w \, dx. \quad (1.12)$$

Η συνάρτηση  $u$  η οποία ικανοποιεί την σχέση (1.11) καλείται ασθενής λύση του προβλήματος (1.2). Τώρα θα παρουσιάσουμε κάποιους ορισμούς που θα μας φανούν χρήσιμοι στην συνέχεια.

**Ορισμός 1.6.** Ορίζουμε το  $\alpha(\cdot, \cdot) : H_0^1 \times H_0^1 \rightarrow \mathbb{R}$  ως

$$\alpha(u, \varphi) := (\nabla u, \nabla \varphi) = \int_{\Omega} \nabla u \cdot \nabla \varphi \, dx, \quad \forall u, \varphi \in H_0^1, \quad (1.13)$$

το οποίο είναι το αριστερό μέλος της ασθενούς μορφής (1.11). Το  $\alpha(u, \varphi)$  καλείται *διγραμμική μορφή*.

**Ορισμός 1.7.** Ορίζουμε το  $l_f : H_0^1 \rightarrow \mathbb{R}$  ως

$$l_f(\varphi) := (f, \varphi) = \int_{\Omega} f \varphi \, dx, \quad \forall \varphi \in H_0^1, \quad (1.14)$$

το οποίο είναι το δεξί μέλος της ασθενούς μορφής (1.11). Το συναρτησιακό  $l_f(\varphi)$  καλείται *γραμμική μορφή*.



**Ορισμός 1.8.** Έστω  $\Omega \subseteq \mathbb{R}^d$  ανοικτό χωρίο, και τη γραμμική μορφή  $l_f : H_0^1 \rightarrow \mathbb{R}$ . Ορίζουμε τον δυϊκό χώρο  $H^{-1}$  του  $H_0^1$  ως

$$H^{-1} := \{f : \Omega \rightarrow \mathbb{R} \mid f \text{ συνεχές και γραμμικό}\}. \quad (1.15)$$

Το επόμενο Θεώρημα (βλ. [3]) μας λέει ότι υπάρχει μοναδική λύση που ικανοποιεί την ασθενή μορφή (1.11) του προβλήματος μας.

**Θεώρημα 1.2. (Lax-Milgram)** Έστω  $\alpha(\cdot, \cdot) : H_0^1 \times H_0^1 \rightarrow \mathbb{R}$  μια συνεχής διγραμμική μορφή και έστω  $l_f : H_0^1 \rightarrow \mathbb{R}$  μια γραμμική πιεστική μορφή τ.ω  $l_f \in H^{-1}$ . Τότε η ασθενής μορφή (1.11) έχει μοναδική λύση.

Για την απόδειξη του παραπάνω Θεωρήματος παραπέμπουμε στο [2]. Επίσης υποθέτοντας ότι έχουμε ομαλό σύνορο  $\partial\Omega$ , η λύση  $u$  είναι δύο τάξεις ομαλότερη στον  $L^2$  από ότι η συνάρτηση στο δεξί μέλος. Αυτό μπορεί να εκφραστεί μαθηματικά ως εξής

$$\|u\|_{m+2} \leq C \|\Delta u\|_m \leq C \|f\|_m, \quad \forall m \geq -1 \quad (1.16)$$

όπου  $C$  θετική σταθερά. Η παραπάνω ανισότητα καλείται *ελλειπτική κανονικότητα*. Επίσης από το Θεώρημα Sobolev, συμπεραίνουμε σε συνδυασμό με την σχέση (1.16) ότι η λύση  $u$  ανήκει στον  $C^\infty$  αν η  $f$  ανήκει στον  $C^\infty$  (βλ. [4]).

Να σημειώσουμε εδώ ότι, για  $m = -1, 0$  ισχύει η σχέση (1.16) μόνο στην περίπτωση κυρτού πολυγωνικού χωρίου  $\Omega$ .

Για να ορίσουμε την λύση πεπερασμένων στοιχείων του ελλειπτικού προβλήματος (1.2), πρέπει να ορίσουμε έναν πεπερασμένης διάστασης υπόχωρο του  $H_0^1$ . Εδώ θα παρουσιάσουμε την πιο απλή περίπτωση, η οποία είναι ένας υπόχωρος του  $H_0^1$  κατά τμήματα γραμμικών συναρτήσεων στο  $\Omega$ .

Έστω τώρα  $d = 2$  και  $\Omega$  ένα κυρτό δυσδιάστατο χωρίο στον χώρο με ομαλό σύνορο  $\partial\Omega$ , και έστω  $\mathcal{T}_h$  μια διαμέριση του  $\Omega$  σε τρίγωνα  $\tau$  τέτοια ώστε καμία κορυφή να μην πέφτει σε πλευρά άλλου τριγώνου. Επίσης πρέπει να ισχύει ότι  $\cup_{\tau \in \mathcal{T}_h} \tau = \Omega_h \subset \Omega$  και οι συνοριακές πλευρές του  $\Omega_h$  να βρίσκονται στο  $\partial\Omega$ . Το  $h$  ορίζεται ως εξής

$$h := \max_{\tau \in \mathcal{T}_h} \{diam(\tau)\}. \quad (1.17)$$

Όπως γίνεται εύκολα αντιληπτό όσο η τριγωνοποίηση  $\mathcal{T}_h$  γίνεται πυκνότερη τόσο η παράμετρος  $h$  μειώνεται. Επίσης υποθέτουμε ότι οι γωνίες των τριγώνων  $\tau \in \mathcal{T}_h$  είναι κάτω φραγμένες από μια θετική σταθερά, η οποία είναι

ανεξάρτητη του  $h$ , και ότι η τριγωνοποίηση είναι *σχεδόν ομοιόμορφη*, δηλαδή ότι όλα τα τρίγωνα της έχουν περίπου το ίδιο μέγεθος. Αυτό μπορεί να εκφραστεί απαιτώντας η περιοχή των  $\tau \in \mathcal{T}_h$  να φράσσεται από το μέγεθος  $ch^2$ , με  $c > 0$ , ανεξάρτητο του  $h$ .

Έστω τώρα  $S_h \subset H_0^1$  το οποίο ορίζεται ως εξής

$$S_h := \{\chi \in H_0^1 \mid \chi \in C(\Omega), \text{ γραμμικό σε κάθε } \tau \in \mathcal{T}_h, \text{ και } \chi|_{\partial\Omega} = 0\}. \quad (1.18)$$

Έστω επίσης  $\{P_j\}_{j=1}^{N_h}$  οι κορυφές οι οποίες βρίσκονται στο εσωτερικό της τριγωνοποίησης  $\mathcal{T}_h$ . Τότε κάθε συνάρτηση  $\chi \in S_h$ , προσδιορίζεται πλήρως από τις τιμές που παίρνει στα σημεία  $P_j$ , δηλαδή εξαρτάται μόνο από  $N_h$  το πλήθος παραμέτρους. Τώρα έστω  $\Phi_j$  οι κατά τμήματα γραμμικές συναρτήσεις τύπου στέγης/πυραμίδας στον  $S_h$ , οι οποίες παίρνουν την τιμή 1 στο σημείο  $P_j$  και την τιμή 0 σε όλα τα υπόλοιπα. Τότε οι συναρτήσεις  $\{\Phi_j\}_{j=1}^{N_h}$  αποτελούν μια βάση για τον υπόχωρο  $S_h$ . Συνεπώς κάθε  $\chi \in S_h$  γράφεται με μοναδικό τρόπο ως εξής

$$\chi(x) = \sum_{j=1}^{N_h} \alpha_j \Phi_j(x), \text{ με } \alpha_j = \chi(P_j). \quad (1.19)$$

Έστω μια ομαλή συνάρτηση  $v$  στο  $\Omega$  τέτοια ώστε  $v|_{\partial\Omega} = 0$ . Μπορούμε να προσεγγίσουμε αυτή την συνάρτηση με την παρεμβάλλουσα της  $I_h v \in S_h$ , η οποία μπορεί να οριστεί ως η συνάρτηση στον  $S_h$ , της οποίας οι τιμές στους εσωτερικούς κόμβους της τριγωνοποίησης  $\mathcal{T}_h$  συμπίπτουν με τις τιμές της  $v$  στα ίδια σημεία, δηλαδή

$$I_h v(x) = \sum_{j=1}^{N_h} v(P_j) \Phi_j(x). \quad (1.20)$$

Για περισσότερες λεπτομέρειες σχετικά με την παρεμβάλλουσα παραπέμπουμε τον αναγνώστη στο [5]. Χρησιμοποιώντας τον παραπάνω συμβολισμό, οι ακόλουθες εκτιμήσεις για την παρεμβάλλουσα είναι γνωστές.

$$\|I_h v - v\| \leq Ch^2 \|v\|_2 \text{ και } \|\nabla(I_h v - v)\| \leq Ch \|v\|_2, \quad \forall v \in H^2 \cap H_0^1. \quad (1.21)$$

Οι παραπάνω εκτιμήσεις λαμβάνονται δείχνοντας τις σε κάθε τριγωνάκι  $\tau \in \mathcal{T}_h$ , και αθροίζοντας τα επιμέρους κομμάτια παίρνουμε το ζητούμενο (βλ. [1], [2] ή [5]).

Στην γενική περίπτωση όπου  $\Omega \subset \mathbb{R}^d$ , και ο  $S_h$  είναι πεπερασμένης διάστασης υπόχωρος του  $H_0^1$ , για κάποιο θετικό ακέραιο  $r \geq 2$  και μικρό  $h$ , ισχύει η

παρακάτω εκτίμηση.

$$\inf_{\chi \in S_h} \{ \|v - \chi\| + h \|\nabla(v - \chi)\| \} \leq Ch^s \|v\|_s, \text{ για } 1 \leq s \leq r, \quad (1.22)$$

όπου  $v \in H^s \cap H_0^1$ . Ο ακέραιος  $r$  καλείται τάξη ακρίβειας του υπόχωρου πεπερασμένης διάστασης  $S_h$ . Το παράδειγμα που αναφέραμε πιο πάνω λαμβάνεται αν θέσουμε στην (1.22)  $d = r = 2$ . Για  $r > 2$ , ο  $S_h$  συνήθως αποτελείται από κατά τμήματα πολυώνυμα τάξης το πολύ  $r - 1$ , σε μια τριγωνοποίηση  $\mathcal{T}_h$ . Για παράδειγμα, για  $r = 3$  έχουμε κατά τμήματα τετραγωνικές συναρτήσεις στον  $S_h$ . Επίσης εκτίμηση όπως την (1.22) ισχύει και για τον τελεστή της παρεμβάλουσας  $I_h : H^r \cap H_0^1 \rightarrow S_h$ , δηλαδή

$$\|I_h v - v\| + h \|\nabla(I_h v - v)\| \leq Ch^s \|v\|_s, \text{ για } 1 \leq s \leq r. \quad (1.23)$$

Έστω τώρα ότι ο υπόχωρος πεπερασμένης διάστασης  $S_h$  βασίζεται σε μια σχεδόν ομοιόμορφη τριγωνοποίηση  $\mathcal{T}_h$ , και ότι αποτελείται από κατά τμήματα πολυώνυμα τάξης το πολύ  $r - 1$ . Τότε ισχύει η *αντίστροφη ανισότητα*

$$\|\nabla \chi\| \leq Ch^{-1} \|\chi\|, \quad \forall \chi \in S_h. \quad (1.24)$$

Η παραπάνω ανισότητα αποδεικνύεται παίρνοντας την αντίστοιχη νόρμα σε κάθε στοιχείο (τριγωνάκι)  $\tau$  της τριγωνοποίησης  $\mathcal{T}_h$  στο τετράγωνο. Στην συνέχεια χρησιμοποιώντας έναν μετασχηματισμό σε ένα τρίγωνο αναφοράς, και χρησιμοποιώντας το γεγονός ότι όλες οι νόρμες σε ένα πεπερασμένης διάστασης χώρο είναι ισοδύναμες, αθροίζουμε σε όλα τα στοιχεία, παίρνουμε την ρίζα και έτσι έχουμε το ζητούμενο (για λεπτομέρειες παραπέμπουμε στο βιβλίο [2]).

Από τις υποθέσεις της (1.22) παρατηρούμε ότι η μέγιστη τάξη σύγκλισης για τις συναρτήσεις και το gradient τους είναι  $O(h^r)$  και  $O(h^{r-1})$ , αντίστοιχα. Τώρα θα κατασκευάσουμε προσεγγίσεις με την χρήση της μεθόδου πεπερασμένων στοιχείων, της λύσης του προβλήματος Dirichlet (1.2) οι οποίες θα έχουν τάξη σύγκλισης ανάλογη με αυτή που μόλις αναφέραμε.

Θέτοντας  $\varphi = \chi \in S_h$  στην ασθενή μορφή (1.11) παίρνουμε

$$(\nabla u, \nabla \chi) = (f, \chi), \quad \forall \chi \in S_h \quad (1.25)$$

Τώρα αντί να ψάχνουμε λύση σε ένα άπειρης διάστασης χώρο όπως ο  $H_0^1$ , επιλέγουμε να ψάξουμε μια προσέγγιση αυτής της λύσης σε έναν πεπερασμένης διάστασης υπόχωρο του τον  $S_h$ . Με αυτό τον τρόπο παίρνουμε την ακόλουθη μέθοδο η οποία καλείται Μέθοδος Πεπερασμένων Στοιχείων Galerkin, η οποία

πήρε το όνομα της από τον σπουδαίο Ρώσο μαθηματικό Boris Grigorievich Galerkin (1871-1945).

Να βρεθεί  $u_h \in S_h$  τέτοιο ώστε

$$(\nabla u_h, \nabla \chi) = (f, \chi), \quad \forall \chi \in S_h. \quad (1.26)$$

Αφαιρώντας την ασθενή μορφή (1.11) από την μέθοδο (1.26) παίρνουμε την σχέση

$$(\nabla(u_h - u), \nabla \chi) = 0, \quad \forall \chi \in S_h, \quad (1.27)$$

η οποία καλείται *ορθογωνιότητα Galerkin*, και μας λέει ότι το σφάλμα που παίρνουμε προσεγγίζοντας την ακριβή λύση  $u$  από την  $u_h$ , είναι κάθετο στον  $S_h$ , ως προς το εσωτερικό γινόμενο  $(\nabla v, \nabla w)$ .

Αφού κάθε  $\chi \in S_h$  γράφεται με μοναδικό τρόπο ως γραμμικός συνδυασμός των συναρτήσεων βάσης, τότε η μέθοδος (1.26) μπορεί να γραφτεί ως εξής

Να βρεθεί  $u_h \in S_h$  τέτοιο ώστε

$$(\nabla u_h, \nabla \Phi_k) = (f, \Phi_k), \quad \text{για } k = 1, \dots, N_h. \quad (1.28)$$

Επίσης αφού και η  $u_h \in S_h$ , τότε ισχύει ότι  $u_h(x) = \sum_{j=1}^{N_h} \alpha_j \Phi_j(x)$ , και έτσι η μέθοδος (1.26), γράφεται ισοδύναμα ως ένα γραμμικό σύστημα  $N_h$  εξισώσεων με  $N_h$  αγνώστους, ως ακολούθως

Να βρεθούν οι συντελεστές  $\alpha_j$  τέτοιοι ώστε

$$\sum_{j=1}^{N_h} \alpha_j (\nabla \Phi_j, \nabla \Phi_k) = (f, \Phi_k), \quad \text{για } k = 1, \dots, N_h. \quad (1.29)$$

$$\mathcal{A}\alpha = \tilde{f}, \quad (1.30)$$

όπου ο  $\mathcal{A} = (a_{jk})$  με στοιχεία  $a_{jk} = (\nabla \Phi_j, \nabla \Phi_k)$  καλείται *πίνακας ακαμψίας*. Επίσης  $\tilde{f} = (f_k)$  το διάνυσμα με στοιχεία  $f_k = (f, \Phi_k)$ , και  $\alpha$  το διάνυσμα με τους αγνώστους συντελεστές  $\alpha_j$ . Οι διαστάσεις του πίνακα είναι  $N_h \times N_h$  και των δύο διανυσμάτων  $N_h \times 1$ , όπου  $N_h$  ο αριθμός των εσωτερικών κόμβων του πλέγματος μας. Να σημειώσουμε επίσης ότι ο πίνακας  $\mathcal{A}$  είναι πίνακας Gram και συνεπώς θετικά ορισμένος και συμμετρικός, άρα και αντιστρέψιμος. Επομένως το διακριτό μας πρόβλημα έχει μοναδική λύση. Για να δείξουμε ότι ο πίνακας ακαμψίας  $\mathcal{A}$  είναι θετικά ορισμένος, παρατηρούμε ότι

$$\sum_{j=1}^d \sum_{k=1}^d a_{jk} \xi_j \xi_k = \|\nabla(\sum_{j=1}^d \xi_j \Phi_j)\|^2 \geq 0. \quad (1.31)$$

Η ισότητα ισχύει αν και μόνο αν  $\nabla(\sum_{j=1}^d \xi_j \Phi_j) \equiv 0$ . Άρα  $\sum_{j=1}^d \xi_j \Phi_j = 0$  και συνεπώς  $\xi_j = 0$ , για  $j = 1, \dots, N_h$ .

Να σημειώσουμε εδώ ότι τα στοιχεία του πίνακα ακαμψίας  $\mathcal{A}$  αν και στην πράξη υπολογίζονται αριθμητικά, θεωρητικά μπορούν να υπολογιστούν ακριβώς όταν ο χώρος λύσεων  $S_h$  αποτελείται από κατά τμήματα πολυωνυμικές συναρτήσεις. Δυστυχώς όμως δεν συμβαίνει το ίδιο με τα στοιχεία του δεξιού μέλους  $\tilde{f}$ , τα οποία πλην ελαχίστων απλών εξαιρέσεων χρειάζεται χρήση αριθμητικής ολοκλήρωσης για να υπολογιστούν.

Έστω τώρα ότι η ακριβής λύση του προβλήματος Dirichlet είναι αρκούντως ομαλή έτσι ώστε οι νόρμες στο δεξί μέλος της ανισότητας (1.22) να είναι πεπερασμένες. Επίσης παρατηρούμε ότι αφού απαιτούμε το  $\partial\Omega$  να είναι ομαλό, σύμφωνα με την ανισότητα (1.16), η λύση  $u$  του προβλήματος Dirichlet (1.2) έχει την επιθυμητή τάξη ομαλότητας, αν υποθέσουμε ότι η συνάρτηση  $f$  στο δεξί μέλος είναι αρκούντως ομαλή. Συγκεκριμένα,  $u \in H^r \cap H_0^1$  αν  $f \in H^{r-2}$ , και η λύση  $u \in C^\infty$  και  $f \in C^\infty$ . Έχοντας αυτά υπόψη μας, θα αποδείξουμε τις ακόλουθες εκτιμήσεις σφάλματος για την μέθοδο που αναλύσαμε παραπάνω. Το παρακάτω Θεώρημα το μελετήσαμε από το βιβλίο [6] σελ.5-6.

**Θεώρημα 1.3.** Έστω ότι ισχύει η ανισότητα (1.22), και έστω ότι  $u_h$  και  $u$  οι λύσεις των (1.26) και (1.2) αντίστοιχα. Τότε, για  $1 \leq s \leq r$  ισχύει ότι

$$\|u_h - u\| \leq Ch^s \|u\|_s \text{ και} \quad (1.32)$$

$$\|\nabla u_h - \nabla u\| \leq Ch^{s-1} \|u\|_s. \quad (1.33)$$

*Απόδειξη.* Ξεκινάμε την απόδειξη από την εκτίμηση σφάλματος για την κλίση (gradient) (1.33).

$$\begin{aligned} \|\nabla(u_h - u)\|^2 &= \|\nabla(u - u_h)\|^2 = \int_{\Omega} |\nabla(u - u_h)|^2 dx \\ &= \int_{\Omega} \nabla(u - u_h) \cdot \nabla(u - u_h) dx \\ &= \int_{\Omega} \nabla(u - u_h) \cdot \nabla(u - \chi) dx + \int_{\Omega} \nabla(u - u_h) \cdot \nabla(\chi - u_h) dx. \end{aligned}$$

Αφού  $(\chi - u_h) \in S_h$ , τότε από την ορθογωνιότητα Galerkin (1.27) έχουμε ότι το δεύτερο ολοκλήρωμα ισούται με μηδέν και άρα

$$\|\nabla(u_h - u)\|^2 = \int_{\Omega} \nabla(u - u_h) \cdot \nabla(u - \chi) dx \leq \|\nabla(u - u_h)\| \|\nabla(u - \chi)\|.$$

Επομένως έχουμε

$$\|\nabla(u_h - u)\| \leq \|\nabla(\chi - u)\|, \quad \forall \chi \in S_h.$$

Άρα από την (1.22) παίρνουμε

$$\|\nabla(u_h - u)\| \leq \inf_{\chi \in S_h} \|\nabla(\chi - u)\| \leq Ch^{s-1} \|u\|_s, \quad (1.34)$$

το οποίο είναι το ζητούμενο.

Τώρα θα αποδείξουμε την εκτίμηση σφάλματος (1.32), χρησιμοποιώντας το γνωστό δυϊκό πρόβλημα. Έστω  $e = u_h - u$  το σφάλμα, και  $\psi \in H^2 \cap H_0^1$  η λύση του προβλήματος

$$\begin{aligned} -\Delta\psi &= e, \text{ στο } \Omega, \\ \psi &= 0, \text{ στο } \partial\Omega. \end{aligned} \quad (1.35)$$

Από την ανισότητα (1.16) η λύση  $\psi$  είναι ομαλότερη κατά δύο παραγώγους στον  $L^2$  από το δεξί μέλος  $e$ , δηλαδή

$$\|\psi\|_2 \leq C \|\Delta\psi\| = C \|e\|. \quad (1.36)$$

όπου  $C > 0$  σταθερά που εξαρτάται από το χωρίο  $\Omega$ .

Για κάθε  $\psi_h \in S_h$  έχουμε

$$\|e\|^2 = (e, e) = -(e, \Delta\psi) = (\nabla e, \nabla\psi) = (\nabla(u_h - u), \nabla(\psi - \psi_h)),$$

όπου η τελευταία ισότητα ισχύει λόγω της ορθογωνιότητας Galerkin (1.32).

Άρα έχουμε

$$\|e\|^2 = (e, e) = (\nabla(u_h - u), \nabla(\psi - \psi_h)) \leq \|\nabla(u_h - u)\| \|\nabla(\psi - \psi_h)\|, \quad \forall \psi_h \in S_h.$$

Χρησιμοποιώντας την σχέση (1.34) έχουμε

$$\|e\|^2 \leq Ch^{s-1} \|u\|_s \|\nabla(\psi - \psi_h)\|, \quad \forall \psi_h \in S_h. \quad (1.37)$$

Χρησιμοποιώντας και πάλι την σχέση (1.34) με  $s = 2$  παίρνουμε

$$\|e\|^2 \leq Ch^{s-1} \|u\|_s Ch \|\psi\|_2. \quad (1.38)$$

Εφαρμόζοντας τώρα την ανισότητα (1.36) έχουμε

$$\|e\|^2 \leq Ch^{s-1} \|u\|_s Ch \|e\|. \quad (1.39)$$

Τέλος διαιρώντας και τις δύο πλευρές με  $\|e\|$  παίρνουμε το ζητούμενο.  $\square$

## 1.4 Η εξίσωση της θερμότητας

Αφού μελετήσαμε το ελλειπτικό πρόβλημα, έχουμε τα θεμέλια για να προχωρήσουμε στο πρόβλημα αρχικών συνοριακών τιμών (1.1) για την εξίσωση της θερμότητας. Όπως και στην ελλειπτική περίπτωση ξεκινάμε γράφοντας το πρόβλημα μας σε ασθενή μορφή. Δηλαδή πολλαπλασιάζουμε την ΜΔΕ με μια συνάρτηση ελέγχου  $\varphi \in H_0^1$ , ολοκληρώνουμε στο  $\Omega$  και εφαρμόζουμε το divergence theorem, ακριβώς όπως κάναμε και στο ελλειπτικό πρόβλημα. Έτσι παίρνουμε

$$(u_t, \varphi) + (\nabla u, \nabla \varphi) = (f, \varphi), \quad \forall \varphi \in H_0^1, \quad t \in (0, T], \quad (1.40)$$

όπου  $(v, w)$  και  $(\nabla v, \nabla w)$  ορίζονται στην (1.12). Η συνάρτηση  $u = u(x, t)$  καλείται ασθενής λύση του (1.1) στο  $[0, \bar{t}]$  αν  $u \in L^2(H_0^1(\Omega)) \cap H^1(H^{-1}(\Omega))$ , και αν  $u(\cdot, 0) = v$ . Παρόμοια με προηγουμένως, αν το σύνορο  $\partial\Omega$  είναι ομαλό, η ασθενής λύση είναι ομαλή δεδομένου ότι τα δεδομένα είναι ομαλές συναρτήσεις. Επίσης όταν  $t = 0$ , η ασθενής λύση πρέπει να ικανοποιεί ταυτόχρονα και την συνοριακή αλλά και την αρχική συνθήκη, δηλαδή πρέπει  $v|_{\partial\Omega} = 0$ . Ανάλογα με την (1.16) εδώ ισχύει η parabolic inequality, τέτοια ώστε αν  $u^{(j)} = (\partial/\partial t)^j u$  και  $C = C_{m, \bar{t}}$ ,

$$\sum_{j=0}^{m+1} \int_0^{\bar{t}} \|u^{(j)}\|_{2(m-j)+2}^2 dt \leq C \left( \|v\|_{2m+1}^2 + \sum_{j=0}^m \int_0^{\bar{t}} \|f^{(j)}\|_{2(m-j)}^2 dt \right), \quad (1.41)$$

για  $m \geq 0$ . Για  $m = 0$ , απαιτούμε  $u(0) = v = 0$  στο  $\partial\Omega$ . Για  $m = 1$ , αφού  $u_t(t) = 0$  στο  $\partial\Omega$  για  $t > 0$ , για ομαλότητα απαιτούμε  $u_t(0) = g := \Delta v + f(0) = 0$  στο  $\partial\Omega$ . Παρόμοια για  $u^{(m)}(0)$  για  $m \geq 0$ . Για λεπτομέρειες παραπέμπουμε στο βιβλίο [4].

Θα ασχοληθούμε με την ανάλυση της μεθόδου για το πρόβλημα (1.1), σε δύο βήματα. Πρώτα θα προσεγγίσουμε την λύση  $u(x, t)$  με μια συνάρτηση  $u_h(x, t)$ , η οποία για σταθερό  $t$ , ανήκει στον πεπερασμένης διάστασης γραμμικό χώρο  $S_h \subset H_0^1$ . Αυτή η συνάρτηση θα είναι λύση ενός συστήματος ΣΔΕ στον χρόνο, το οποίο θα είναι ανεξάρτητο της παραμέτρου  $h$ . Η προσεγγιστική λύση αυτής της οπτικής καλείται χωρικά διακριτή ή ημιδιακριτή λύση. Όπως και στην ελλειπτική περίπτωση η ημιδιακριτή λύση βασίζεται στην ασθενή μορφή του προβλήματος (1.1).

Στην συνέχεια για να πάρουμε την πλήρως διακριτή λύση, διακριτοποιούμε το σύστημα των ΣΔΕ στον χρόνο, χρησιμοποιώντας πεπερασμένες διαφορές για να προσεγγίσουμε την χρονική παράγωγο.

Πάμε λοιπόν να ασχοληθούμε με το ημιδιακριτό πρόβλημα, το οποίο ορίζεται ως εξής.

Να βρεθεί  $u_h(t) = u_h(\cdot, t) \in S_h$  για  $t \geq 0$ , τέτοιο ώστε

$$(u_{h,t}, \chi) + (\nabla u_h, \nabla \chi) = (f, \chi), \quad \forall \chi \in S_h, \quad t > 0, \quad \text{με } u_h(0) = v_h, \quad (1.42)$$

όπου  $v_h$  είναι μια προσέγγιση της  $v$  στον  $S_h$ . Αφού η μέθοδος (1.42) ισχύει για κάθε  $\chi \in S_h$ , τότε ισοδύναμα έχουμε

Να βρεθεί  $u_h(t) = u_h(\cdot, t) \in S_h$  για  $t \geq 0$ , τέτοιο ώστε

$$(u_{h,t}, \Phi_k) + (\nabla u_h, \nabla \Phi_k) = (f, \Phi_k), \quad t > 0, \quad \text{για } k = 1, \dots, N_h, \quad (1.43)$$

όπου  $\{\Phi_j\}_{j=1}^{N_h}$  οι συναρτήσεις βάσης του  $S_h$ . Επίσης αφού  $u_h(t) \in S_h$ , τότε το παραπάνω πρόβλημα μπορεί να γραφτεί ως ένα σύστημα  $N_h$  εξισώσεων με  $N_h$  αγνώστους ως εξής

Να βρεθούν οι συντελεστές  $\alpha_j(t)$  της  $u_h(x, t) = \sum_{j=1}^{N_h} \alpha_j(t) \Phi_j(x)$  τέτοιοι ώστε να ισχύει

$$\sum_{j=1}^{N_h} \alpha_j'(t) (\Phi_j, \Phi_k) + \sum_{j=1}^{N_h} \alpha_j(t) (\nabla \Phi_j, \nabla \Phi_k) = (f, \Phi_k), \quad \text{για } k = 1, \dots, N_h, \quad (1.44)$$

όπου  $\gamma_j$  οι συντελεστές της δοσμένης προσέγγισης της αρχικής συνθήκης  $v_h$ , έτσι ώστε  $\alpha_j(0) = \gamma_j$  για  $j = 1, \dots, N_h$ . Το σύστημα (1.44) γράφεται σε μορφή πινάκων ως ακολούθως:

$$\mathcal{B}\alpha'(t) + \mathcal{A}\alpha(t) = \tilde{f}(t), \quad \text{για } t > 0, \quad \text{με } \alpha(0) = \gamma, \quad (1.45)$$

όπου ο  $\mathcal{B} = (b_{jk})$  με στοιχεία  $b_{jk} = (\Phi_j, \Phi_k)$  καλείται πίνακας μάζας, και ο  $\mathcal{A} = (a_{jk})$  με στοιχεία  $a_{jk} = (\nabla \Phi_j, \nabla \Phi_k)$  ο πίνακας ακαμψίας ο οποίος είναι ο ίδιος όπως και στην περίπτωση του ελλειπτικού προβλήματος. Επίσης όπου  $\tilde{f} = (f_k)$  το διάνυσμα με στοιχεία  $f_k = (f, \Phi_k)$ , και  $\alpha(t)$  το διάνυσμα με στοιχεία τους άγνωστους συντελεστές  $\alpha_j(t)$  και  $\gamma = (\gamma_k)$ , παρόμοια με την ελλειπτική περίπτωση. Οι διαστάσεις των δύο πινάκων είναι  $N_h \times N_h$  και των διανυσμάτων  $N_h \times 1$ .

Αφού οι πίνακες ακαμψίας  $\mathcal{A}$  και μάζας  $\mathcal{B}$  είναι συμμετρικοί και θετικά ορισμένοι, τότε είναι και αντιστρέψιμοι. Επομένως το παραπάνω σύστημα ΣΔΕ μπορεί να γραφτεί ισοδύναμα

$$\alpha'(t) + \mathcal{B}^{-1}\mathcal{A}\alpha(t) = \mathcal{B}^{-1}\tilde{f}(t), \quad \text{για } t > 0, \quad \text{με } \alpha(0) = \gamma. \quad (1.46)$$



Άρα η παραπάνω διαφορική εξίσωση έχει μοναδική λύση για  $t > 0$ .

Αρχικά θα αποδείξουμε μια εκτίμηση σφάλματος ανάμεσα στην ημιδιακριτή και ακριβή λύση στην  $L^2$  νόρμα. Για να το κάνουμε αυτό χρειάζεται να ορίσουμε πρώτα την ελλειπτική ή αλλιώς *Ritz* προβολή μιας συνάρτησης στον  $S_h$ , την οποία συμβολίζουμε με  $R_h$ .

**Ορισμός 1.9.** Ορίζουμε την *Ritz* προβολή ως την ορθογώνια προβολή στον  $S_h$  ως προς το εσωτερικό γινόμενο  $\nabla v, \nabla w$ , έτσι ώστε

$$(\nabla R_h v, \nabla \chi) = (\nabla v, \nabla \chi), \quad \forall \chi \in S_h, \quad \forall v \in H_0^1. \quad (1.47)$$

Παρατηρούμε ότι η  $R_h v$  είναι η λύση της Μεθόδου Πεπερασμένων Στοιχείων για το αντίστοιχο ελλειπτικό πρόβλημα που έχει ως ακριβή λύση την συνάρτηση  $v$ , αφού η σχέση (1.47) συμπίπτει με την ορθογωνιότητα Galerkin (1.27). Άρα από το Θεώρημα 1.3, συμπεραίνουμε άμεσα το ακόλουθο Πόρισμα.

**Πόρισμα 1.1.** Υποθέτουμε ότι ισχύει η σχέση (1.22). Τότε έχουμε

$$\|R_h v - v\| + h\|\nabla(R_h v - v)\| \leq Ch^s \|v\|_s, \quad \text{για } v \in H^s \cap H_0^1, 1 \leq s \leq r. \quad (1.48)$$

Μια κομβικής σημασίας στρατηγική που θα χρησιμοποιήσουμε πολύ συχνά στις αναλύσεις σφάλματος που θα κάνουμε είναι να γράφουμε το σφάλμα του εκάστοτε παραβολικού προβλήματος σαν άθροισμα δύο όρων.

$$u_h(t) - u(t) = \theta(t) + \rho(t), \quad \text{όπου } \theta = u_h - R_h u, \quad \rho = R_h u - u, \quad (1.49)$$

τους οποίους θα φράσουμε ξεχωριστά. Όπως γίνεται εύκολα αντιληπτό ο όρος  $\rho(t)$  είναι το σφάλμα του ελλειπτικού προβλήματος με λύση την ακριβή λύση του παραβολικού προβλήματος. Συνεπώς βρήκαμε ήδη ένα φράγμα του στην ανάλυση σφάλματος που κάναμε στην ενότητα 1.3. Άρα μας μένει να φράξουμε σε κάθε περίπτωση το  $\theta(t)$ , το οποίο θα είναι το κύριο μας θέμα σε κάθε ανάλυση σφάλματος.

Θέτοντας τώρα  $\chi = R_h v$  στην (1.47), συμπεραίνουμε ότι η *Ritz* προβολή είναι ευσταθής στον  $H_0^1$ , δηλαδή

$$\|\nabla R_h v\| \leq \|\nabla v\|, \quad \forall v \in v \in H_0^1. \quad (1.50)$$

Τώρα είμαστε έτοιμοι να αποδείξουμε την εκτίμηση σφάλματος για το ημιδιακριτό πρόβλημα. Το παρακάτω Θεώρημα το μελετήσαμε από το βιβλίο [6] σελ.8-9.

**Θεώρημα 1.4.** Έστω  $u_h$  και  $u$  λύσεις των προβλημάτων (1.42) και (1.1) αντίστοιχα. Επίσης υποθέτουμε ότι  $v = 0$  στο  $\partial\Omega$ . Τότε ισχύει η εξής εκτίμηση σφάλματος.

$$\|u_h(t) - u(t)\| \leq \|v_h - v\| + Ch^r \left( \|v\|_r + \int_0^t \|u_t\|_r ds \right), \text{ για } t \geq 0. \quad (1.51)$$

*Απόδειξη.* Αρχικά γράφουμε το σφάλμα στη μορφή του αθροίσματος (1.49). Επομένως χρησιμοποιώντας το Πρόρισμα 1, το Θεμελιώδες Θεώρημα του Απειροστικού Λογισμού και την τριγωνική ανισότητα έχουμε

$$\|\rho(t)\| \leq Ch^r \|u(t)\|_r \leq Ch^r \left( \|v\|_r + \int_0^t \|u_t\|_r ds \right). \quad (1.52)$$

Τώρα μας απομένει να φράξουμε το  $\theta(t)$ . Χρησιμοποιώντας τους ορισμούς του  $\theta$  (1.49) και της Ritz προβολής (1.47) όπως και την ασθενή μορφή (1.40) και την ΜΠΣ (1.42) μας έχουμε

$$\begin{aligned} & (\theta_t, \chi) + (\nabla\theta, \nabla\chi) \\ &= (u_{h,t}, \chi) + (\nabla u_h, \nabla\chi) - (R_h u_t, \chi) - (\nabla R_h u, \nabla\chi) \\ &= (f, \chi) - (R_h u_t, \chi) - (\nabla u, \nabla\chi) = (u_t - R_h u_t, \chi). \end{aligned} \quad (1.53)$$

Εξ'ορισμού  $\rho = R_h u - u$ , και αφού ισχύει ότι  $(R_h u)_t = R_h u_t$ , τότε έχουμε

$$(\theta_t, \chi) + (\nabla\theta, \nabla\chi) = -(\rho_t, \chi), \quad \forall \chi \in S_h, \quad t > 0. \quad (1.54)$$

Αφού  $\theta \in S_h$ , μπορούμε να θέσουμε στην (1.54) το  $\chi = \theta$  και έτσι παίρνουμε

$$(\theta_t, \theta) + \|\nabla\theta\|^2 = -(\rho_t, \theta), \quad \text{για } t > 0. \quad (1.55)$$

Ο πρώτος όρος του αριστερού μέλους της (1.55) γράφεται ισοδύναμα

$$\begin{aligned} (\theta_t, \theta) &= \int_{\Omega} \theta_t \theta dx = \frac{1}{2} \int_{\Omega} 2\theta_t \theta dx \\ &= \frac{1}{2} \int_{\Omega} \frac{d}{dt} \theta^2 dx = \frac{1}{2} \frac{d}{dt} \int_{\Omega} \theta^2 dx \\ &= \frac{1}{2} \frac{d}{dt} \|\theta\|^2. \end{aligned} \quad (1.56)$$

Αφού ο δεύτερος όρος του αριστερού μέλους της (1.55)  $\|\nabla\theta\|^2$  είναι μη αρνητικός τότε έχουμε

$$\frac{1}{2} \frac{d}{dt} \|\theta\|^2 = \|\theta\| \frac{d}{dt} \|\theta\| \leq \|\rho_t\| \|\theta\|. \quad (1.57)$$

Επομένως διαγράφοντας και από τις δύο πλευρές το  $\|\theta\|$  και ολοκληρώνοντας στο  $[0, t]$  έχουμε ότι

$$\int_0^t \frac{d}{ds} \|\theta(s)\| ds \leq \int_0^t \|\rho_t\| ds \quad (1.58)$$

$$\|\theta(t)\| \leq \|\theta(0)\| + \int_0^t \|\rho_t\| ds.$$

Να σημειώσουμε ότι στην περίπτωση που ισχύει ότι  $\|\theta\| = 0$  για κάποιο  $t$ , τότε δεν χρειάζεται να εκτιμήσουμε τον όρο  $\theta$  γιατί η προσεγγιστική λύση  $u_h(t)$  ισούται με την  $R_h u(t)$ . Έπειτα χρησιμοποιώντας το Πρόρισμα 1 μπορούμε να φράξουμε περαιτέρω τον όρο  $\|\theta(0)\|$  ως εξής

$$\|\theta(0)\| = \|v_h - R_h v\| \leq \|v_h - v\| + \|R_h v - v\| \leq \|v_h - v\| + Ch^r \|v\|_r. \quad (1.59)$$

Επίσης και πάλι χρησιμοποιώντας το Πρόρισμα 1, μπορούμε να φράξουμε περαιτέρω και τον δεύτερο όρο της (1.58)

$$\|\rho_t\| = \|R_h u_t - u_t\| \leq Ch^r \|u_t\|_r. \quad (1.60)$$

Άρα αποδείξαμε το ζητούμενο φράγμα για το  $\|\theta(t)\|$ . Επομένως έχουμε από την τριγωνική ανισότητα και τα φράγματα που δείξαμε για τα  $\|\theta(t)\|$  και  $\|\rho(t)\|$  έχουμε

$$\begin{aligned} \|u_h(t) - u(t)\| &= \|\theta(t) + \rho(t)\| \leq \|\theta(t)\| + \|\rho(t)\| \\ &\leq \|v_h - v\| + Ch^r \left( \|v\|_r + \int_0^t \|u_t\|_r ds \right), \text{ για } t \geq 0. \end{aligned} \quad (1.61)$$

□

Στην παραπάνω απόδειξη πετάξαμε τον όρο  $\|\nabla\theta\|^2$  της σχέσης (1.55) επειδή είναι μη αρνητικός. Όμως μπορούμε αν τον χρησιμοποιήσουμε διαφορετικά ώστε να δείξουμε ότι η επίδραση που έχουν τα αρχικά δεδομένα στο σφάλμα τείνει στο μηδέν καθώς το  $t$  τείνει στο  $\infty$  (βλ. [6] σελ.10).

**Θεώρημα 1.5.** Έστω  $u_h$  και  $u$  οι λύσεις των προβλημάτων (1.42) και (1.1) αντίστοιχα. Επίσης υποθέτουμε ότι  $v = 0$  στο  $\partial\Omega$ . Τότε ισχύει η εξής εκτίμηση σφάλματος

$$\|u_h(t) - u(t)\| \leq Ch^r \left( e^{-\lambda_1 t} \|v\|_r + \|u(t)\|_r + \int_0^t e^{-\lambda_1(t-s)} \|u_t(s)\|_r ds \right), \quad (1.62)$$

για  $t \geq 0$ .

*Απόδειξη.* Η απόδειξη είναι η ίδια με αυτή του Θεωρήματος 1.4 μέχρι την σχέση (1.56). Τώρα θεωρούμε με  $\lambda_1$  την μικρότερη ιδιοτιμή του  $-\Delta$ , με συνοριακές συνθήκες Dirichlet. Τότε εφαρμόζοντας την διαδικασία κατασκευής της ασθενούς λύσης του ελλειπτικού προβλήματος που εξηγήσαμε στην προηγούμενη ενότητα, καταλαβαίνουμε ότι ισχύει η ανισότητα

$$\|\nabla v\|^2 \geq \lambda_1 \|v\|^2, \quad \forall v \in H_0^1. \quad (1.63)$$

Συνεπώς από τις σχέσεις (1.55) και (1.57) έχουμε

$$\|\theta\| \frac{d}{dt} \|\theta\| + \lambda_1 \|\theta\|^2 \leq \|\rho_t\| \|\theta\|. \quad (1.64)$$

Διαιρώντας και από τις δύο πλευρές με  $\|\theta\|$ , παίρνουμε

$$\frac{d}{dt} \|\theta\| + \lambda_1 \|\theta\| \leq \|\rho_t\|. \quad (1.65)$$

Να σημειώσουμε και πάλι ότι στην περίπτωση που ισχύει ότι  $\|\theta\| = 0$  για κάποιο  $t$ , τότε δεν χρειάζεται να εκτιμήσουμε τον όρο  $\theta$  γιατί η προσεγγιστική λύση  $u_h(t)$  ισούται με την  $R_h u(t)$ . Η παραπάνω σχέση αποτελεί μια ΣΔΕ (αν ήταν ισότητα). Συνεπώς λύνοντας την με την μέθοδο του ολοκληρωτικού παράγοντα παίρνουμε

$$\|\theta(t)\| \leq e^{-\lambda_1 t} \|\theta(0)\| + \int_0^t e^{-\lambda_1(t-s)} \|\rho_t(s)\| ds. \quad (1.66)$$

Χρησιμοποιώντας τις σχέσεις (1.59) και (1.60) φράζουμε τον πρώτο και δεύτερο όρο του δεξιού μέλους αντίστοιχα.

$$\|\theta(t)\| \leq e^{-\lambda_1 t} \|v_h - v\| + Ch^r \left( e^{-\lambda_1 t} \|v\|_r + \int_0^t e^{-\lambda_1(t-s)} \|u_t(s)\|_r ds \right). \quad (1.67)$$

Χρησιμοποιώντας το πρώτο μέρος της (1.52) παίρνουμε

$$\|u_h(t) - u(t)\| \leq Ch^r \left( e^{-\lambda_1 t} \|v\|_r + \|u(t)\|_r + \int_0^t e^{-\lambda_1(t-s)} \|u_t(s)\|_r ds \right).$$

□

Ένα χρήσιμο αποτέλεσμα που θα χρησιμοποιήσουμε στην απόδειξη του επόμενου Θεωρήματος αλλά και γενικά σε όλη την μελέτη μας είναι η ανισότητα Young.

**Λήμμα 1.1 (Ανισότητα Young).** Έστω  $\alpha \geq 0$  και  $\beta \geq 0$  δύο μη αρνητικοί ακέραιοι. Έστω επίσης  $p > 1$  και  $q > 1$  δύο πραγματικοί αριθμοί τέτοιοι ώστε  $\frac{1}{p} + \frac{1}{q} = 1$ . Τότε ισχύει η ανισότητα

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q},$$

η οποία καλείται *ανισότητα Young*. Η ισότητα ισχύει εάν και μόνο εάν  $\alpha^p = \beta^q$ .

*Απόδειξη.* Η ζητούμενη ανισότητα προφανώς ισχύει για  $\alpha = 0$  και  $\beta = 0$ . Έστω λοιπόν  $\alpha > 0$  και  $\beta > 0$ . Θέτουμε  $t = 1/p$  και  $(1-t) = 1/q$ . Επειδή η λογαριθμική συνάρτηση είναι κοίλη, γνωρίζουμε ότι ισχύει το εξής

$$\ln(t\alpha^p + (1-t)\beta^q) \geq t \ln(\alpha^p) + (1-t) \ln(\beta^q) \geq \ln \alpha + \ln \beta = \ln(\alpha\beta).$$

Συνεπώς ισχύει ότι

$$e^{\ln(t\alpha^p + (1-t)\beta^q)} \geq e^{\ln(\alpha\beta)}.$$

Επομένως έχουμε το ζητούμενο. □

Τώρα θα αποδείξουμε μια εκτίμηση σφάλματος για την κλίση (gradient) του σφάλματος του ημιδιακριτού προβλήματος (1.42). Το παρακάτω Θεώρημα το μελετήσαμε από το βιβλίο [6] σελ.11-12.

**Θεώρημα 1.6.** Έστω  $u_h$  και  $u$  οι λύσεις των προβλημάτων (1.42) και (1.1) αντίστοιχα. Επίσης υποθέτουμε ότι  $v = 0$  στο  $\partial\Omega$ . Τότε ισχύει η εξής εκτίμηση σφάλματος.

$$\begin{aligned} \|\nabla u_h(t) - \nabla u(t)\| &\leq C \|\nabla v_h - \nabla v\| \\ &+ Ch^{r-1} \left( \|v\|_r + \|u(t)\|_r + \left( \int_0^t \|u_s\|_{r-1}^2 ds \right)^{1/2} \right), \text{ για } t \geq 0. \end{aligned} \quad (1.68)$$

*Απόδειξη.* Αρχικά γράφουμε το σφάλμα στην μορφή (1.49). Άρα έχουμε ότι  $\nabla e(t) = \nabla \theta(t) + \nabla \rho(t)$ . Λόγω της τριγωνικής ανισότητας έχουμε

$$\|\nabla e(t)\| \leq \|\nabla \theta(t)\| + \|\nabla \rho(t)\|. \quad (1.69)$$

Ο δεύτερος όρος φράσσεται άμεσα χρησιμοποιώντας το Πόρισμα 1 ως εξής

$$\|\nabla \rho(t)\| = \|\nabla(R_h u(t) - u(t))\| \leq Ch^{r-1} \|u(t)\|_r. \quad (1.70)$$

Τώρα για να φράξουμε τον πρώτο όρο, χρησιμοποιούμε την σχέση (1.54) θέτοντας  $\chi = \theta_t$  και παίρνουμε

$$\|\theta_t\|^2 + \frac{1}{2} \frac{d}{dt} \|\nabla\theta\|^2 = -(\rho_t, \theta_t) \leq \frac{1}{2} \|\rho(t)\|^2 + \frac{1}{2} \|\theta_t\|^2. \quad (1.71)$$

Άρα έχουμε

$$\frac{1}{2} \|\theta_t\|^2 + \frac{1}{2} \frac{d}{dt} \|\nabla\theta\|^2 \leq \frac{1}{2} \|\rho(t)\|^2. \quad (1.72)$$

Αφού ο όρος  $\frac{1}{2} \|\theta_t\|^2$  είναι μη αρνητικός τον πετάμε και έχουμε

$$\frac{d}{dt} \|\nabla\theta\|^2 \leq \|\rho(t)\|^2. \quad (1.73)$$

Ολοκληρώνοντας την παραπάνω σχέση στο  $[0, t]$  παίρνουμε

$$\|\nabla\theta(t)\|^2 \leq \|\nabla\theta(0)\|^2 + \int_0^t \|\rho(s)\|^2 ds. \quad (1.74)$$

Χρησιμοποιώντας το πρώτο μέρος της σχέσης (1.59) παίρνουμε

$$\|\nabla\theta(t)\|^2 \leq (\|\nabla(v_h - v)\| + \|\nabla(R_h v - v)\|)^2 + \int_0^t \|\rho(s)\|^2 ds. \quad (1.75)$$

Από το Πόρισμα 1 έχουμε

$$\|\nabla\theta(t)\|^2 \leq 2\|\nabla(v_h - v)\|^2 + Ch^{2r-2} \left( \|v\|_r^2 + \int_0^t \|u_t\|_{r-1}^2 ds \right). \quad (1.76)$$

Συνεπώς χρησιμοποιώντας την ανισότητα Young και παίρνοντας την τετραγωνική ρίζα της παραπάνω εξίσωσης έχουμε

$$\|\nabla\theta(t)\| \leq C\|\nabla(v_h - v)\| + Ch^{r-1} \left( \|v\|_r + \left( \int_0^t \|u_t\|_{r-1}^2 ds \right)^{1/2} \right). \quad (1.77)$$

Τέλος συνδυάζουμε τις σχέσεις (1.69), (1.70) και (1.77) παίρνουμε το ζητούμενο.  $\square$

Παρατηρούμε ότι αν  $v_h = I_h v$  ή  $v_h = R_h v$ , τότε από (1.23) ή από Πόρισμα 1 αντίστοιχα, ο πρώτος όρος της (1.68) φράσσεται από τον δεύτερο. Στην ειδική περίπτωση των σχεδόν ομοιόμορφων τριγωνοποιήσεων  $\mathcal{T}_h$  του χωρίου  $\Omega$ , ισχύει το παρακάτω Θεώρημα (βλ. [6] σελ.12).

**Θεώρημα 1.7.** Έστω μια σχεδόν ομοιόμορφη οικογένεια τριγωνοποιήσεων  $\mathcal{T}_h$  του χωρίου  $\Omega$ , και έστω  $u_h$  και  $u$  λύσεις των προβλημάτων (1.42) και (1.1) αντίστοιχα. Επίσης υποθέτουμε ότι  $v = 0$  στο  $\partial\Omega$ . Τότε ισχύει η εξής εκτίμηση σφάλματος.

$$\|\nabla u_h(t) - \nabla u(t)\| \leq Ch^{r-1} \left( \|v\|_r + \int_0^t \|u_t(s)\|_r ds \right). \quad (1.78)$$

*Απόδειξη.* Αφού έχουμε σχεδόν ομοιόμορφη τριγωνοποίηση  $\mathcal{T}_h$  του χωρίου  $\Omega$ , τότε ισοδύναμα ισχύει η αντίστροφη ανισότητα (1.24). Με την χρήση της (1.24) έχουμε  $\forall \chi \in S_h$  ότι

$$\begin{aligned} \|\nabla u_h(t) - \nabla u(t)\| &\leq \|\nabla(u_h(t) - \chi)\| + \|\nabla\chi - \nabla u(t)\| \\ &\leq Ch^{-1}\|u_h(t) - \chi\| + \|\nabla\chi - \nabla u(t)\| \\ &\leq Ch^{-1}\|u_h(t) - u(t)\| + Ch^{-1}(\|\chi - u(t)\| + h\|\nabla\chi - \nabla u(t)\|). \end{aligned} \quad (1.79)$$

Τώρα με την χρήση της (1.22) και του Θεωρήματος 1.4 έχουμε για κατάλληλο  $\chi \in S_h$

$$\|\chi - u(t)\| + h\|\nabla\chi - \nabla u(t)\| \leq Ch^r\|v\|_r. \quad (1.80)$$

Συνδυάζοντας της σχέσεις (1.79) και (1.80) παίρνουμε το ζητούμενο.  $\square$

Παρατηρούμε ότι αν υποθέσουμε ότι  $v_h = R_h v$  τότε  $\theta(0) = 0$ . Άρα από την σχέση (1.74) και το Πρόσχημα 1, έχουμε για την κλίση του  $\theta = u_h - R_h u$  το ακόλουθο φράγμα.

$$\|\nabla\theta(t)\| \leq C \left( \int_0^t \|\rho_t\|^2 ds \right)^{1/2} \leq Ch^r \left( \int_0^t \|u_t\|^2 ds \right)^{1/2}. \quad (1.81)$$

Επομένως αν κάνουμε αυτή την επιλογή για το  $v_h$  έχουμε ότι η τάξη σύγκλισης του  $\nabla\theta(t)$  είναι  $O(h^r)$ . Δηλαδή το  $\nabla u_h$  προσεγγίζει καλύτερα το  $R_h u$  από ότι είναι δυνατόν να προσεγγίσει το  $\nabla u$ . Αυτό το φαινόμενο καλείται *superconvergence*.

Η φύση της μεθόδου Galerkin μας παραπέμπει στην χρήση της  $L^2$ -νόρμας στις αναλύσεις σφάλματος, αφού γράφεται σε  $L^2$  εσωτερικά γινόμενα. Παρόλα αυτά, στην βιβλιογραφία των ΜΠΣ υπάρχουν αναλύσεις σφάλματος σε διάφορες νόρμες, όπως για παράδειγμα η  $L_\infty$ -νόρμα και η  $W_\infty^r$ -νόρμα, οι οποίες ορίζονται ως εξής

$$\|v\|_{L_\infty} := \sup_{x \in \Omega} |u(x)|, \quad \|v\|_{W_\infty^r}^r := \max_{|\alpha| \leq r} \|D^\alpha v\|_{L_\infty}, \quad (1.82)$$

όπου  $L_\infty = L_\infty(\Omega)$  και  $W_\infty^r = W_\infty^r(\Omega)$ . Εδώ θα περιοριστούμε στην ανάλυση σφάλματος με την  $L^2$ -νόρμα.

Τώρα θα μελετήσουμε κάποιες απλές πλήρως διακριτές μεθόδους, χρησιμοποιώντας πεπερασμένες διαφορές για να προσεγγίσουμε την χρονική παράγωγο του προβλήματος (1.1). Αρχικά ορίζουμε το χρονικό βήμα  $k$  τέτοιο ώστε  $t = t_n = nk$ , όπου  $n$  είναι μη αρνητικός ακέραιος. Επίσης θα συμβολίζουμε με  $U^n = U_h^n \in S_h$  την προσέγγιση της  $u(t_n)$  την οποία θέλουμε να υπολογίσουμε.

Θα ξεκινήσουμε με την έμμεση *Euler Galerkin* μέθοδο, η οποία ορίζεται αντικαθιστώντας την χρονική παράγωγο στην (1.42) με την οπίσθόδρομη διαφορά την οποία ορίζουμε ως  $\bar{\partial}U^n = (U^n - U^{n-1})/k$ .

$$(\bar{\partial}U^n, \chi) + (\nabla U^n, \nabla \chi) = (f(t_n), \chi), \quad \forall \chi \in S_h, \quad n \geq 1, \quad (1.83)$$

όπου  $U^0 = v_h$ . Άρα όταν το  $U^{n-1}$  είναι γνωστό, το  $U^n$  υπολογίζεται έμμεσα από την σχέση

$$(U^n, \chi) + k(\nabla U^n, \nabla \chi) = (U^{n-1} + kf(t_n), \chi), \quad \forall \chi \in S_h. \quad (1.84)$$

Χρησιμοποιώντας τους ίδιους συμβολισμούς με αυτούς που χρησιμοποιήσαμε στο ημιδιακριτό πρόβλημα, η έμμεση *Euler Galerkin* μέθοδος, γράφεται στην μορφή πινάκων ως ακολούθως

$$(\mathcal{B} + k\mathcal{A})\alpha^n = \mathcal{B}\alpha^{n-1} + k\tilde{f}(t_n), \quad (1.85)$$

όπου ο πίνακας  $\mathcal{B} + k\mathcal{A}$  είναι συμμετρικός και θετικά ορισμένος και συνεπώς αντιστρέψιμος. Το Θεώρημα που ακολουθεί (βλ. [6] σελ.15) μας δίνει μια εκτίμηση σφάλματος αυτής της μεθόδου.

**Θεώρημα 1.8.** Έστω  $U^n$  και  $u$  οι λύσεις των (1.83) και (1.1) αντίστοιχα. Εάν  $\|v_h - v\| \leq Ch^r \|v\|_r$  και αν  $v = 0$  στο  $\partial\Omega$ , τότε ισχύουν οι εξής εκτιμήσεις.

$$\|U^n - u(t_n)\| \leq Ch^r \left( \|v\|_r + \int_0^{t_n} \|u_t\|_r ds \right) + k \int_0^{t_n} \|u_{tt}\| ds, \quad \text{για } n \geq 0. \quad (1.86)$$

και

$$\|\nabla(U^n - u(t_n))\| \leq Ch^{r-1} \left( \|v\|_r + \int_0^{t_n} \|u_t\|_r ds \right) + k \int_0^{t_n} \|u_{tt}\| ds, \quad \text{για } n \geq 0. \quad (1.87)$$



Απόδειξη. Ανάλογα με προηγούμενως γράφουμε το σφάλμα ως το άθροισμα

$$U^n - u(t_n) = (U^n - R_h u(t_n)) + (R_h u(t_n) - u(t_n)) = \theta^n + \rho^n, \quad (1.88)$$

όπου  $\theta^n = \theta(t_n)$  και  $\rho^n = \rho(t_n)$ . Ο όρος  $\rho^n$  φράσσεται όπως και προηγούμενως από την σχέση (1.52). Επομένως για να ολοκληρώσουμε την απόδειξη αρκεί να φράξουμε το  $\theta^n$ . Χρησιμοποιώντας τους ορισμούς του  $\theta^n$  (1.88) και της Ritz προβολής (1.47) όπως και την ασθενή μορφή (1.40) και την ΜΠΣ (1.83) έχουμε

$$\begin{aligned} (\bar{\partial}\theta^n, \chi) + (\nabla\theta^n, \nabla\chi) &= (\bar{\partial}U^n, \chi) + (\nabla U^n, \nabla\chi) \\ &\quad - (\bar{\partial}R_h u(t_n), \chi) - (\nabla R_h u(t_n), \nabla\chi) \\ &= (f(t_n), \chi) - (\bar{\partial}R_h u(t_n), \chi) - (\nabla u(t_n), \nabla\chi) \\ &= (u_t(t_n), \chi) - (\bar{\partial}R_h u(t_n), \chi) = (u_t(t_n) - \bar{\partial}R_h u(t_n), \chi) \end{aligned} \quad (1.89)$$

Συνοψίζοντας έχουμε

$$(\bar{\partial}\theta^n, \chi) + (\nabla\theta^n, \nabla\chi) = -(\omega^n, \chi), \quad \forall \chi \in S_h, \quad n \geq 1, \quad (1.90)$$

όπου

$$\omega^n := R_h \bar{\partial}u(t_n) - u_t(t_n) = (R_h - I)\bar{\partial}u(t_n) + (\bar{\partial}u(t_n) - u_t(t_n)) = \omega_1^n + \omega_2^n. \quad (1.91)$$

Θέτοντας  $\chi = \theta^n$  στην (1.90) παίρνουμε

$$(\bar{\partial}\theta^n, \theta^n) + \|\nabla\theta^n\|^2 = -(\omega^n, \theta^n). \quad (1.92)$$

Να σημειώσουμε ότι στην περίπτωση που ισχύει  $\|\theta^n\| = 0$  τότε δεν χρειάζεται να εκτιμήσουμε το  $\theta$  για τον ίδιο λόγο που εξηγήσαμε και προηγούμενως.

Αφού  $\|\nabla\theta^n\|^2$  είναι μη αρνητικό τότε έχουμε

$$(\bar{\partial}\theta^n, \theta^n) \leq \|\omega^n\| \|\theta^n\|. \quad (1.93)$$

Ισοδύναμα

$$\|\theta^n\|^2 \leq (\theta^{n-1}, \theta^n) + k\|\omega^n\| \|\theta^n\| \leq \|\theta^{n-1}\| \|\theta^n\| + k\|\omega^n\| \|\theta^n\|. \quad (1.94)$$

Διαιρώντας και τις δύο πλευρές με  $\|\theta^n\|$  παίρνουμε

$$\|\theta^n\| \leq \|\theta^{n-1}\| + k\|\omega^n\|. \quad (1.95)$$

Εφαρμόζοντας την παραπάνω σχέση επαναληπτικά παίρνουμε

$$\|\theta^n\| \leq \|\theta^0\| + k \sum_{j=1}^n \|\omega^j\| \leq \|\theta^0\| + k \sum_{j=1}^n \|\omega_1^j\| + k \sum_{j=1}^n \|\omega_2^j\|. \quad (1.96)$$

Όπως δείξαμε προηγουμένως έχουμε για το  $\theta^0 = \theta(0)$  ότι

$$\|\theta(0)\| \leq \|v_h - v\| + Ch^r \|v\|_r \leq Ch^r \|v\|_r. \quad (1.97)$$

Άρα μας μένει να φράξουμε τους δύο τελευταίους όρους. Αρχικά γράφουμε

$$\omega_1^j = (R_h - I)k^{-1} \int_{t_{j-1}}^{t_j} u_t ds = k^{-1} \int_{t_{j-1}}^{t_j} (R_h - I)u_t ds, \quad (1.98)$$

και έτσι έχουμε από το Πόρισμα 1

$$k \sum_{j=1}^n \|\omega_1^j\| \leq \sum_{j=1}^n \int_{t_{j-1}}^{t_j} Ch^r \|u_t\|_r ds = Ch^r \int_0^{t_n} \|u_t\|_r ds. \quad (1.99)$$

Για τον τελευταίο όρο έχουμε

$$k\omega_2^j = u(t_j) - u(t_{j-1}) - ku_t(t_j) = - \int_{t_{j-1}}^{t_j} (s - t_{j-1})u_{tt}(s) ds. \quad (1.100)$$

Για να δείξουμε την παραπάνω σχέση ξεκινάμε ανάποδα.

$$\begin{aligned} - \int_{t_{j-1}}^{t_j} (s - t_{j-1})u_{tt}(s) ds &= -(s - t_{j-1})u_t(s) \Big|_{t_{j-1}}^{t_j} + \int_{t_{j-1}}^{t_j} u_t(s) ds \\ &= -(t_j - t_{j-1})u_t(t_j) + 0 + u(t_j) - u(t_{j-1}) = u(t_j) - u(t_{j-1}) - ku_t(t_j). \end{aligned}$$

Χρησιμοποιώντας λοιπόν την (1.100) έχουμε

$$k \sum_{j=1}^n \|\omega_2^j\| \leq \sum_{j=1}^n \left\| \int_{t_{j-1}}^{t_j} (s - t_{j-1})u_{tt}(s) ds \right\| \leq k \int_0^{t_n} \|u_{tt}\| ds. \quad (1.101)$$

Συνοψίζοντας τα φράγματα (1.97), (1.99) και (1.101) έχουμε

$$\|\theta^n\| \leq Ch^r \left( \|v\|_r + \int_0^{t_n} \|u_t\|_r ds \right) + k \int_0^{t_n} \|u_{tt}\| ds, \text{ για } n \geq 0. \quad (1.102)$$

Άρα συνδυάζοντας την παραπάνω ανισότητα με το φράγμα που βρήκαμε για τον όρο  $\rho^n$  παίρνουμε το ζητούμενο. Υπενθυμίζουμε ότι  $\partial\theta^n = (\theta^n - \theta^{n-1})/k$ .

Για να δείξουμε μια εκτίμηση για το  $\|\nabla\theta^n\|$ , θέτουμε  $\chi = \bar{\partial}\theta^n$  στην (1.90) και χρησιμοποιώντας την ανισότητα Young παίρνουμε

$$\|\bar{\partial}\theta^n\|^2 + (\nabla\theta^n, \nabla\bar{\partial}\theta^n) = -(\omega^n, \bar{\partial}\theta^n) \leq \|\omega^n\|\|\theta^n\| \leq \frac{1}{2}\|\omega^n\|^2 + \frac{1}{2}\|\bar{\partial}\theta^n\|^2.$$

Συνεπώς έχουμε ότι

$$\frac{1}{2}\|\bar{\partial}\theta^n\|^2 + (\nabla\theta^n, \nabla\bar{\partial}\theta^n) \leq \frac{1}{2}\|\omega^n\|^2.$$

Αφού  $\|\bar{\partial}\theta^n\|^2 \geq 0$  μπορούμε απλά να "πετάξουμε" αυτόν τον όρο και έτσι έχουμε ότι

$$(\nabla\theta^n, \nabla\bar{\partial}\theta^n) \leq \frac{1}{2}\|\omega^n\|^2.$$

Εφαρμόζοντας και πάλι την ανισότητα Young έχουμε

$$\begin{aligned} (\nabla\theta^n, \nabla\bar{\partial}\theta^n) &= \frac{1}{k}(\nabla\theta^n, \nabla(\theta^n - \theta^{n-1})) = \frac{1}{k}(\|\nabla\theta^n\|^2 - (\nabla\theta^n, \nabla\theta^{n-1})) \\ &\geq \frac{1}{k}(\|\nabla\theta^n\|^2 - \|\nabla\theta^n\|\|\nabla\theta^{n-1}\|) \\ &\geq \frac{1}{k}(\|\nabla\theta^n\|^2 - \frac{1}{2}\|\nabla\theta^n\|^2 - \frac{1}{2}\|\nabla\theta^{n-1}\|^2) \\ &= \frac{1}{2}(\frac{1}{k}\|\nabla\theta^n\|^2 - \frac{1}{k}\|\nabla\theta^{n-1}\|^2) = \frac{1}{2}\bar{\partial}\|\theta^n\|^2. \end{aligned}$$

Άρα παίρνουμε τελικά την σχέση

$$\bar{\partial}\|\nabla\theta^n\|^2 \leq \|\omega^n\|^2. \quad (1.103)$$

Αν ισχύει ότι  $\nabla\theta^0 = 0$ , τότε έχουμε

$$\|\nabla\theta^n\|^2 \leq k \sum_{j=1}^n \|\omega^j\|^2 \leq Ch^{2s} \int_0^{t_n} \|u_t\|_s^2 dt + Ck^2 \int_0^{t_n} \|u_{tt}\|^2 dt, \quad (1.104)$$

για  $1 \leq s \leq r$ . Μαζί με την εκτίμηση (1.70) παίρνουμε για την κλίση του σφάλματος την εξής εκτίμηση (1.87) για  $s = r - 1$ . Συνεπώς

$$\|\nabla(U^n - u(t_n))\| \leq C(u)(h^{r-1} + k). \quad (1.105)$$

□

Να σημειώσουμε ότι υποθέσαμε ότι η ακριβής λύση  $u$  ικανοποιεί της κατάλληλες συνθήκες ομαλότητας έτσι ώστε τα ολοκληρώματα στο δεξί μέλος

της εκτίμησης (1.86) να είναι πεπερασμένα. Παρατηρούμε ότι λόγω της μη συμμετρικής επιλογής στην διακριτοποίηση του χρόνου η έμμεση Euler Galerkin μέθοδος είναι μόλις πρώτης τάξης ως προς  $k$ . Αυτό μπορούμε να το λύσουμε χρησιμοποιώντας την *Crank-Nicolson Galerkin μέθοδο*, στην οποία η διακριτοποίηση γίνεται με συμμετρικό τρόπο γύρω από το σημείο  $t_{n-\frac{1}{2}} = (n - \frac{1}{2})k$ . Με αυτό τον τρόπο έχουμε δεύτερης τάξης σύγκλιση ως προς  $k$ . Πιο συγκεκριμένα, θέτουμε  $\hat{U}^n = \frac{1}{2}(U^n + U^{n-1})$  και ορίζουμε  $U^n \in S_h$  και έτσι έχουμε την μέθοδο

$$(\bar{\partial}U^n, \chi) + (\nabla\hat{U}^n, \nabla\chi) = (f(t_{n-\frac{1}{2}}), \chi), \quad \forall \chi \in S_h, \text{ για } n \geq 0, \quad (1.106)$$

με  $U^0 = v_h$ . Κάνοντας κάποιες απλές αλγεβρικές πράξεις βλέπουμε ότι Crank-Nicolson Galerkin μέθοδος γράφεται σε μορφή πινάκων ως εξής,

$$(\mathcal{B} + \frac{1}{2}k\mathcal{A})\alpha^n = (\mathcal{B} - \frac{1}{2}k\mathcal{A})\alpha^{n-1} + k\tilde{f}(t_{n-\frac{1}{2}}), \quad (1.107)$$

όπου οι πίνακες  $\mathcal{A}$  και  $\mathcal{B}$  καθώς και τα διανύσματα  $\alpha^n$  και  $\tilde{f}(t_n)$  ορίζονται ακριβώς όπως και προηγουμένως. Παρατηρούμε επίσης ότι ο πίνακας  $\mathcal{B} + \frac{1}{2}k\mathcal{A}$  είναι συμμετρικός και θετικά ορισμένος. Συνεπώς το παραπάνω σύστημα έχει μοναδική λύση.

Μία εκτίμηση σφάλματος της Crank-Nicolson Galerkin μεθόδου είναι η ακόλουθη. Το παρακάτω Θεώρημα το μελετήσαμε από το βιβλίο [6] σελ.16-17.

**Θεώρημα 1.9.** Έστω  $U^n$  και  $u$  οι λύσεις των (1.106) και (1.1) αντίστοιχα. Έστω επίσης ότι  $\|v_h - v\| \leq Ch^r\|v\|_r$  και  $v = 0$  στο  $\partial\Omega$ . Τότε έχουμε για  $n \geq 0$  ότι

$$\|U^n - u(t_n)\| \leq Ch^r \left( \|v\|_r + \int_0^{t_n} \|u_t\|_r \right) + Ck^2 \int_0^{t_n} (\|u_{ttt}\| + \|\Delta u_{tt}\|) ds. \quad (1.108)$$

*Απόδειξη.* Ο όρος  $\rho^n$  του σφάλματος  $U^n - u(t_n) = \theta^n + \rho^n$  φράσσεται ακριβώς όπως και προηγουμένως. Άρα μας απομένει να φράξουμε το  $\theta^n$ . Χρησιμοποιώντας τους ίδιους ορισμούς που χρησιμοποιήσαμε και στην

προηγούμενη απόδειξη έχουμε  $\forall \chi \in S_h$  ότι

$$\begin{aligned}
 (\bar{\partial}\theta^n, \chi) + (\nabla\hat{\theta}^n, \nabla\chi) &= (\bar{\partial}U^n, \chi) + (\nabla\hat{U}^n, \nabla\chi) - (\bar{\partial}R_h u(t_n), \chi) \\
 &- (\nabla R_h \hat{u}(t_n), \nabla\chi) = (f(t_{n-\frac{1}{2}}), \chi) - (\bar{\partial}R_h u(t_n), \chi) - (\nabla\hat{u}(t_n), \nabla\chi) \\
 &= (u_t(t_{n-\frac{1}{2}}) - \Delta u(t_{n-\frac{1}{2}}), \chi) - (\bar{\partial}R_h u(t_n), \chi) - (\nabla\hat{u}(t_n), \nabla\chi) \\
 &= (u_t(t_{n-\frac{1}{2}}) - \bar{\partial}R_h u(t_n), \chi) - (\Delta u(t_{n-\frac{1}{2}}), \chi) + (\Delta\hat{u}(t_n), \chi) \\
 &= -(\bar{\partial}R_h u(t_n) - u_t(t_{n-\frac{1}{2}}), \chi) - (\Delta(u(t_{n-\frac{1}{2}}) - \frac{1}{2}(u(t_n) + u(t_{n-1}))), \chi).
 \end{aligned} \tag{1.109}$$

Συνοψίζοντας έχουμε

$$(\bar{\partial}\theta^n, \chi) + (\nabla\hat{\theta}^n, \nabla\chi) = -(\omega^n, \chi), \text{ για } \chi \in S_h, \ n \geq 1, \tag{1.110}$$

όπου

$$\begin{aligned}
 \omega^n &= (R_h - I)\bar{\partial}u(t_n) + (\bar{\partial}u(t_n) - u_t(t_{n-\frac{1}{2}})) \\
 &+ \Delta(u(t_{n-\frac{1}{2}}) - \frac{1}{2}(u(t_n) + u(t_{n-1}))) = \omega_1^n + \omega_2^n + \omega_3^n.
 \end{aligned} \tag{1.111}$$

Θέτοντας  $\chi = \hat{\theta}^n$  στην (1.110) παίρνουμε,

$$(\bar{\partial}\theta^n, \hat{\theta}^n) + \|\nabla\hat{\theta}^n\|^2 = -(\omega^n, \hat{\theta}^n) = -\frac{1}{2}((\omega^n, \theta^n) + (\omega^n, \theta^{n-1})). \tag{1.112}$$

Αφού ο όρος  $\|\nabla\hat{\theta}^n\|^2$  είναι μη αρνητικός, τον πετάμε και έτσι έχουμε ότι

$$(\bar{\partial}\theta^n, \hat{\theta}^n) \leq \frac{1}{2}\|\omega^n\|(\|\theta^n\| + \|\theta^{n-1}\|). \tag{1.113}$$

Ισοδύναμα έχουμε

$$\|\theta^n\|^2 - \|\theta^{n-1}\|^2 \leq k\|\omega^n\|(\|\theta^n\| + \|\theta^{n-1}\|), \tag{1.114}$$

αφού

$$\begin{aligned}
 (\bar{\partial}\theta^n, \hat{\theta}^n) &= \frac{1}{2k}(\theta^n - \theta^{n-1}, \theta^n + \theta^{n-1}) \\
 &= \frac{1}{2k}(\|\theta^n\|^2 + (\theta^n, \theta^{n-1}) - (\theta^{n-1}, \theta^n) + \|\theta^{n-1}\|^2) \\
 &= \frac{1}{2k}(\|\theta^n\|^2 + \|\theta^{n-1}\|^2).
 \end{aligned}$$

Διαιρώντας με  $(\|\theta^n\| + \|\theta^{n-1}\|)$  και στις δύο πλευρές παίρνουμε

$$\|\theta^n\| \leq \|\theta^{n-1}\| + k\|\omega^n\| \leq \|\theta^{n-1}\| + k(\|\omega_1^n\| + \|\omega_2^n\| + \|\omega_3^n\|), \text{ για } n \geq 1. \quad (1.115)$$

Εφαρμόζοντας την παραπάνω ανισότητα επαναληπτικά παίρνουμε

$$\|\theta^n\| \leq \|\theta^0\| + k \sum_{j=1}^n (\|\omega_1^j\| + \|\omega_2^j\| + \|\omega_3^j\|), \text{ για } n \geq 1. \quad (1.116)$$

Τους όρους με τα  $\theta^0$  και  $\omega_1^j$  τους έχουμε ήδη φράξει προηγουμένως. Άρα μας μένει να φράξουμε τους όρους με τα  $\omega_2^j$  και  $\omega_3^j$ . Έχουμε λοιπόν επειδή  $|s - t_{j-1}| \leq k$  για  $t \in (t_{j-1}, t_{j-\frac{1}{2}})$  και  $|s - t_j| \leq k$  για  $t \in (t_{j-\frac{1}{2}}, t_j)$  ότι

$$\begin{aligned} k\|\omega_2^j\| &= \|u(t_j) - u(t_{j-1}) - ku(t_{j-\frac{1}{2}})\| \\ &= \frac{1}{2} \left\| \int_{t_{j-1}}^{t_{j-\frac{1}{2}}} (s - t_{j-1})^2 u_{ttt}(s) ds + \int_{t_{j-\frac{1}{2}}}^{t_j} (s - t_j)^2 u_{ttt}(s) ds \right\| \\ &\leq Ck^2 \int_{t_{j-1}}^{t_j} \|u_{ttt}\| ds, \quad \forall s \in (t_{j-1}, t_j). \end{aligned} \quad (1.117)$$

αφού

$$k\|\omega_2^j\| = k\|\bar{\partial}u(t_j) - u_t(t_{j-\frac{1}{2}})\| = \|u(t_j) - u(t_{j-1}) - ku(t_{j-\frac{1}{2}})\|$$

και

$$\begin{aligned} &\frac{1}{2} \int_{t_{j-1}}^{t_{j-\frac{1}{2}}} (s - t_{j-1})^2 u_{ttt}(s) ds + \frac{1}{2} \int_{t_{j-\frac{1}{2}}}^{t_j} (s - t_j)^2 u_{ttt}(s) ds \\ &= \frac{1}{2} (s - t_{j-1})^2 u_{tt}(s) \Big|_{t_{j-1}}^{t_{j-\frac{1}{2}}} - \int_{t_{j-1}}^{t_{j-\frac{1}{2}}} (s - t_{j-1}) u_{tt}(s) ds \\ &+ \frac{1}{2} (s - t_j)^2 u_{tt}(s) \Big|_{t_{j-\frac{1}{2}}}^{t_j} - \int_{t_{j-\frac{1}{2}}}^{t_j} (s - t_j) u_{tt}(s) ds \\ &= \frac{1}{2} (t_{j-\frac{1}{2}} - t_{j-1})^2 u_{tt}(t_{j-\frac{1}{2}}) - (s - t_{j-1}) u_t(s) \Big|_{t_{j-1}}^{t_{j-\frac{1}{2}}} + \int_{t_{j-1}}^{t_{j-\frac{1}{2}}} u_t(s) ds \\ &- \frac{1}{2} (t_{j-\frac{1}{2}} - t_j)^2 u_{tt}(t_{j-\frac{1}{2}}) - (s - t_j) u_t(s) \Big|_{t_{j-\frac{1}{2}}}^{t_j} + \int_{t_{j-\frac{1}{2}}}^{t_j} u_t(s) ds \\ &= \frac{k^2}{8} u_{tt}(t_{j-\frac{1}{2}}) - (t_{j-\frac{1}{2}} - t_{j-1}) u_t(t_{j-\frac{1}{2}}) + u(t_{j-\frac{1}{2}}) - u(t_{j-1}) \\ &- \frac{k^2}{8} u_{tt}(t_{j-\frac{1}{2}}) + (t_{j-\frac{1}{2}} - t_j) u_t(t_{j-\frac{1}{2}}) + u(t_j) - u(t_{j-\frac{1}{2}}) \\ &= -(t_{j-\frac{1}{2}} - t_{j-1} - t_{j-\frac{1}{2}} + t_j) u_{tt}(t_{j-\frac{1}{2}}) + u(t_j) - u(t_{j-1}) \end{aligned}$$

$$= u(t_j) - u(t_{j-1}) - (t_j - t_{j-1})u_t(t_{j-\frac{1}{2}}) = u(t_j) - u(t_{j-1}) - ku_t(t_{j-\frac{1}{2}}).$$

Παρόμοια έχουμε ότι

$$k\|\omega_3^j\| = k\|\Delta(u(t_{j-\frac{1}{2}}) - \frac{1}{2}(u(t_j) + u(t_{j-1})))\| \leq Ck^2 \int_{t_{j-1}}^{t_j} \|\Delta u_{tt}\| ds. \quad (1.118)$$

Συνδυάζοντας τις (1.117) και (1.118) παίρνουμε

$$k \sum_{j=1}^n (\|\omega_2^j\| + \|\omega_3^j\|) \leq \int_0^{t_n} (\|u_{ttt}\| + \|\Delta u_{tt}\|) ds. \quad (1.119)$$

Συνοψίζοντας όλα τα παραπάνω αποδείξαμε το ζητούμενο.  $\square$

Να σημειώσουμε ότι η μέθοδος Crank-Nicolson απαιτεί η λύση να είναι ομαλότερη από την μέθοδο που θα αναλύσουμε στην συνέχεια. Ένας άλλος τρόπος να πάρουμε δεύτερης τάξης σύγκλιση ως προς τον χρόνο είναι να προσεγγίσουμε την χρονική παράγωγο με μια δεύτερης τάξης οπίσθόδρομη διαιρεμένη διαφορά. Για παράδειγμα

$$\bar{D}U^n = \bar{\partial}U^n + \frac{1}{2}k\bar{\partial}^2U^n = (\frac{3}{2}U^n - 2U^{n-1} + \frac{1}{2}U^{n-2})/k. \quad (1.120)$$

Χρησιμοποιώντας το Θεώρημα Taylor, για μια ομαλή συνάρτηση  $u$  παίρνουμε ότι  $\bar{D}u(t_n) = u_t(t_n) + O(k^2)$  καθώς το  $k \rightarrow 0$ . Με αυτό τον τρόπο παίρνουμε την πάρακατω ΜΠΣ Galerkin.

$$(\bar{D}U^n, \chi) + (\nabla U^n, \nabla \chi) = (f(t_n), \chi), \quad \forall \chi \in S_h, \quad n \geq 2. \quad (1.121)$$

Παρατηρούμε ότι αυτή η μέθοδος για να λειτουργήσει χρειάζεται σε κάθε βήμα τις τιμές της  $U^n$  σε τρία χρονικά επίπεδα, σε αντίθεση με τις προηγούμενες μεθόδους που μελετήσαμε που χρειάζονται μόνο δύο. Ξεκινάμε λοιπόν από το  $n = 2$ , επειδή δεν μπορούμε να έχουμε αρνητικό χρονικό επίπεδο. Για να υπολογίσουμε την τιμή  $U^1$  χρησιμοποιούμε την έμμεση Euler μέθοδο στο πρώτο βήμα. Δηλαδή για  $n = 1$  θέτουμε

$$(\bar{\partial}U^1, \chi) + (\nabla U^1, \nabla \chi) = (f(t_1), \chi), \quad \forall \chi \in S_h. \quad (1.122)$$

Να σημειώσουμε επίσης ότι η μεθοδος (1.121) γράφεται σε μορφή πινάκων ως εξής

$$(\frac{3}{2}\mathcal{B} + k\mathcal{A})\alpha^n = 2\mathcal{B}\alpha^{n-1} - \frac{1}{2}\mathcal{B}\alpha^{n-2} + k\tilde{f}(t_n), \quad \text{για } n \geq 2, \quad (1.123)$$

όπου οι παραπάνω πίνακες και διανύσματα είναι οι ίδιοι όπως στις υπόλοιπες μεθόδους. Επίσης αφού ο πίνακας  $(\frac{3}{2}\mathcal{B} + k\mathcal{A})$  είναι συμμετρικός και θετικά ορισμένος, τότε είναι αντιστρέψιμος και έτσι η μέθοδος έχει μοναδική λύση.

Το παρακάτω Θεώρημα (βλ. [6] σελ.18-20) μας λέει ότι η παραπάνω μέθοδος έχει σφάλμα της τάξης  $O(h^r + k^2)$ .

**Θεώρημα 1.10.** Έστω  $U^n$  και  $u$  οι λύσεις των (1.121) και (1.1) αντίστοιχα. Επίσης  $U^0 = v_h$  και η  $U^1$  ορίζεται από την (1.122). Τότε αν  $\|v_h - v\| \leq Ch^r \|v\|_r$  και  $v = 0$  στο  $\partial\Omega$  έχουμε ότι

$$\begin{aligned} \|U^n - u(t_n)\| &\leq Ch^r \left( \|v\|_r + \int_0^{t_n} \|u_t\|_r ds \right) \\ &+ Ck \int_0^k \|u_{tt}\| ds + Ck^2 \int_0^{t_n} \|u_{ttt}\| ds, \text{ για } n \geq 0. \end{aligned} \quad (1.124)$$

*Απόδειξη.* Γράφοντας και πάλι το σφάλμα μας στην μορφή του αθροίσματος  $U^n - u(t_n) = \theta^n + \rho^n$ , έχουμε να φράξουμε μόνο τον όρο  $\theta^n$ , αφού για τον δεύτερο όρο έχουμε ήδη δείξει την εκτίμηση που θέλουμε. Έχουμε λοιπόν ότι

$$(\bar{D}\theta^n, \chi) + (\nabla\theta^n, \nabla\chi) = -(\omega^n, \chi), \text{ για } n \geq 2, \quad (1.125)$$

όπου

$$\omega^n = \bar{D}R_h u^n - u_t^n = (R_h - I)\bar{D}u^n + (\bar{D}u^n - u_t^n) = \omega_1^n + \omega_2^n, \text{ για } n \geq 2, \quad (1.126)$$

αφού

$$\begin{aligned} (\bar{D}\theta^n, \chi) + (\nabla\theta^n, \nabla\chi) &= (\bar{D}U^n, \chi) + (\nabla U^n, \nabla\chi) \\ &- (\bar{D}R_h u^n, \chi) - (\nabla R_h u^n, \nabla\chi) \\ &= (f(t_n), \chi) - (\bar{D}R_h u^n, \chi) - (\nabla u^n, \nabla\chi) \\ &= (u_t^n, \chi) - (\bar{D}R_h u^n, \chi) = -(\bar{D}R_h u^n - u_t^n, \chi) \\ &= -(\omega^n, \chi), \text{ για } n \geq 2, \end{aligned}$$

Επίσης

$$(\bar{\partial}\theta^1, \chi) + (\nabla\theta^1, \nabla\chi) = -(\omega^1, \chi). \quad \forall \chi \in S_h, \quad (1.127)$$

όπου

$$\omega^1 = \bar{\partial}R_h u^1 - u_t^1 = (R_h - I)\bar{\partial}u^1 + (\bar{\partial}u^1 - u_t^1) = \omega_1^1 + \omega_2^1. \quad (1.128)$$

από την (1.91). Έστω τώρα ο διαφορικός τελεστής  $\delta_l \theta^n = \theta^n - \theta^{n-l}$  για  $l = 1, 2$ . Μπορούμε χρησιμοποιώντας άλγεβρα να γράψουμε

$$k\bar{D}\theta^n = 2\delta_1\theta^n - \frac{1}{2}\delta_2\theta^n, \quad (1.129)$$



αφού

$$\begin{aligned} k\bar{D}\theta^n &= \frac{3}{2}\theta^n - 2\theta^{n-1} + \frac{1}{2}\theta^{n-2} = 2\theta^n - \frac{1}{2}\theta^n - 2\theta^{n-1} + \frac{1}{2}\theta^{n-2} \\ &= 2(\theta^n - \theta^{n-1}) - \frac{1}{2}(\theta^n - \theta^{n-1}) = 2\delta_1\theta^n - \frac{1}{2}\delta_2\theta^n. \end{aligned}$$

Επίσης έχουμε ότι

$$\begin{aligned} 2(\delta_l\theta^n, \theta^n) &= (\theta^n - \theta^{n-l}, \theta^n) + (\theta^n - \theta^{n-l}, \theta^n) \\ &= (\theta^n - \theta^{n-l}, \theta^n) + (\theta^n - \theta^{n-l}, \theta^n - \theta^{n-l}) + (\theta^n - \theta^{n-l}, -\theta^{n-l}) \\ &= \|\delta_l\theta^n\|^2 + \|\theta^n\|^2 + (\theta^n, \theta^{n-l}) - (\theta^n, \theta^{n-l}) - \|\theta^{n-l}\|^2 \\ &= (\|\theta^n\|^2 - \|\theta^{n-l}\|^2) + \|\delta_l\theta^n\|^2. \end{aligned}$$

Άρα ισχύει ότι

$$2(\delta_l\theta^n, \theta^n) = \delta_l\|\theta^n\|^2 + \|\delta_l\theta^n\|^2. \quad (1.130)$$

Χρησιμοποιώντας τις σχέσεις (1.129) και (1.130) παίρνουμε

$$(k\bar{D}\theta^n, \theta^n) = \delta_1\|\theta^n\|^2 - \frac{1}{4}\delta_2\|\theta^n\|^2 + \|\delta_1\theta^n\|^2 - \frac{1}{4}\|\delta_2\theta^n\|^2, \quad \text{για } n \geq 2. \quad (1.131)$$

Αντικαθιστώντας στην παραπάνω εξίσωση το  $n$  με  $j$  και αθροίζοντας από το 2 μέχρι το  $n$ , παίρνουμε το τηλεσκοπικό άθροισμα

$$\sum_{j=2}^n (\delta_1\|\theta^j\|^2 - \frac{1}{4}\delta_2\|\theta^j\|^2) = \frac{3}{4}\|\theta^n\|^2 - \frac{1}{4}\|\theta^{n-1}\|^2 - \frac{3}{4}\|\theta^1\|^2 + \frac{1}{4}\|\theta^0\|^2.$$

Επίσης αφού ισχύει ότι  $\delta_2\theta^n = \delta_1\theta^n + \delta_1\theta^{n-1}$ , παίρνουμε για τους υπόλοιπους όρους ότι

$$\begin{aligned} \sum_{j=2}^n (\|\delta_1\theta^j\|^2 - \frac{1}{4}\|\delta_2\theta^j\|^2) &\geq \sum_{j=2}^n (\|\delta_1\theta^j\|^2 - \frac{1}{4}(\|\delta_1\theta^j\| + \|\delta_1\theta^{j-1}\|)^2) \\ &\geq \frac{1}{2} \sum_{j=2}^n (\|\delta_1\theta^j\|^2 - \|\delta_1\theta^{j-1}\|^2) = \frac{1}{2}(\|\delta_1\theta^n\|^2 - \|\delta_1\theta^1\|^2). \end{aligned}$$

Επομένως έχουμε ότι

$$\begin{aligned} k(\bar{\partial}\theta^1, \theta^1) + k \sum_{j=2}^n (\bar{D}\theta^j, \theta^j) \\ &\geq \frac{1}{2}(\|\theta^1\|^2 - \|\theta^0\|^2 + \|\delta_1\theta^1\|^2) + \frac{1}{2}(\|\delta_1\theta^n\|^2 - \|\delta_1\theta^1\|^2) \\ &+ (\frac{3}{4}\|\theta^n\|^2 - \frac{1}{4}\|\theta^{n-1}\|^2 - \frac{3}{4}\|\theta^1\|^2 - \frac{1}{4}\|\theta^0\|^2) \\ &\geq \frac{3}{4}\|\theta^n\|^2 - \frac{1}{4}\|\theta^{n-1}\|^2 - \frac{1}{4}\|\theta^1\|^2 - \frac{1}{4}\|\theta^0\|^2. \end{aligned} \quad (1.132)$$

Επίσης συνδυάζοντας τις σχέσεις (1.125) και (1.127) έχουμε

$$k(\bar{\partial}\theta^1, \theta^1) + k \sum_{j=2}^n (\bar{D}\theta^j, \theta^j) + k \sum_{j=1}^n (\nabla\theta^j, \nabla\theta^j) = -k \sum_{j=1}^n (\omega^j, \theta^j). \quad (1.133)$$

Συνδυάζοντας τις δύο παραπάνω σχέσεις παίρνουμε

$$\|\theta^n\|^2 \leq \frac{1}{3}(\|\theta^{n-1}\|^2 + \|\theta^1\|^2 + \|\theta^0\|^2) + \frac{4}{3}k \sum_{j=1}^n \|\omega^j\| \|\theta^j\|. \quad (1.134)$$

Έστω τώρα  $m$  τέτοιο ώστε  $\|\theta^m\| = \max_{0 \leq j \leq n} \|\theta^j\|$ . Τότε

$$\|\theta^m\|^2 \leq \frac{1}{3}(\|\theta^m\| + \|\theta^1\| + \|\theta^0\| + 4k \sum_{j=1}^n \|\omega^j\|) \|\theta^m\|. \quad (1.135)$$

Συνεπώς

$$\|\theta^n\| \leq \|\theta^m\| \leq \frac{1}{2}(\|\theta^1\| + \|\theta^0\|) + 2k \sum_{j=1}^n \|\omega^j\|. \quad (1.136)$$

Από προηγούμενως γνωρίζουμε ότι  $\|\theta^1\| \leq \|\theta^0\| + k\|\omega^1\|$ . Επομένως παίρνουμε τελικά την ανισότητα

$$\|\theta^n\| \leq \|\theta^0\| + 2k \sum_{j=2}^n \|\omega^j\| + \frac{5}{2}k\|\omega^1\|, \quad \text{για } n \geq 1. \quad (1.137)$$

Τώρα μας απομένει να φράξουμε τους δύο τελευταίους όρους μιας και τον πρώτο τον έχουμε ήδη φράξει σε προηγούμενη απόδειξη. Χρησιμοποιώντας λοιπόν το Θεώρημα Taylor και γράφοντας τους υπολοίπωντες όρους σε μορφή ολοκληρώματος και υποθέτοντας ότι η λύση είναι κατάλληλα ομαλή έχουμε ότι

$$\begin{aligned} k\bar{D}w^j &= \frac{3}{2}w^j - 2w^{j-1} + \frac{1}{2}w^{j-2} \\ &= \frac{3}{2}w^j - 2 \left( w^j - kw_t^j + \int_{t_{j-1}}^{t_j} u_{tt}(s)(s - t_{j-1}) ds \right) \\ &\quad + \frac{1}{2} \left( w^j - 2kw_t^j + \int_{t_{j-2}}^{t_j} u_{tt}(s)(s - t_{j-2}) ds \right) \\ &= kw_t^j - 2 \int_{t_{j-1}}^{t_j} u_{tt}(s)(s - t_{j-1}) ds + \frac{1}{2} \int_{t_{j-2}}^{t_j} u_{tt}(s)(s - t_{j-2}) ds \end{aligned}$$

$$\begin{aligned}
 &= ku_t^j - 2(s - t_{j-1})u_t(s) \Big|_{t_{j-1}}^{t_j} - 2 \int_{t_{j-1}}^{t_j} u_t(s) ds \\
 &+ \frac{1}{2}(s - t_{j-2})u_t(s) \Big|_{t_{j-2}}^{t_j} + \frac{1}{2} \int_{t_{j-2}}^{t_j} u_t(s) ds \\
 &= -2 \int_{t_{j-1}}^{t_j} u_t(s) ds + \frac{1}{2} \int_{t_{j-2}}^{t_j} u_t(s) ds.
 \end{aligned}$$

και

$$\begin{aligned}
 k\omega_2^j &= \bar{D}u^j - u_t^j = \frac{3}{2}u^j - 2u^{j-1} + \frac{1}{2}u^{j-2} - u_t^j \\
 &= \frac{3}{2}u^j - 2 \left( u^j - ku_t^j + \frac{k^2}{2}u_{tt}^j - \frac{k^3}{6}u_{ttt}^j + \frac{1}{6} \int_{t_{j-1}}^{t_j} u_{tttt}(s)(s - t_{j-1})^3 ds \right) \\
 &+ \frac{1}{2} \left( u^j - 2ku_t^j + 2k^2u_{tt}^j - \frac{8}{6}k^3u_{ttt}^j + \frac{1}{6} \int_{t_{j-2}}^{t_j} u_{tttt}(s)(s - t_{j-2})^3 ds \right) - ku_t^j \\
 &= -\frac{1}{3}k^3u_{ttt}^j - \frac{1}{3} \int_{t_{j-1}}^{t_j} u_{tttt}(s)(s - t_{j-1})^3 ds + \frac{1}{12} \int_{t_{j-2}}^{t_j} u_{tttt}(s)(s - t_{j-2})^3 ds \\
 &= -\frac{1}{3}k^3u_{ttt}^j - \frac{1}{3}(s - t_{j-1})^3 u_{ttt}(s) \Big|_{t_{j-1}}^{t_j} + \int_{t_{j-1}}^{t_j} u_{ttt}(s)(s - t_{j-1})^2 ds \\
 &+ \frac{1}{12}(s - t_{j-2})^3 u_{ttt}(s) \Big|_{t_{j-2}}^{t_j} - \frac{1}{4} \int_{t_{j-2}}^{t_j} u_{tttt}(s)(s - t_{j-2})^2 ds \\
 &= \int_{t_{j-1}}^{t_j} u_{ttt}(s)(s - t_{j-1})^2 ds - \frac{1}{4} \int_{t_{j-2}}^{t_j} u_{ttt}(s)(s - t_{j-2})^2 ds, \quad \forall s \in (t_{j-2}, t_j).
 \end{aligned}$$

Συνεπώς για  $j \geq 2$  έχουμε ότι

$$k\|\omega_1^j\| \leq Ch^r k \|\bar{D}u^j\|_r \leq Ch^r \int_{t_{j-2}}^{t_j} \|u_t\|_r ds, \quad (1.138)$$

$$k\|\omega_2^j\| \leq Ck^2 \int_{t_{j-2}}^{t_j} \|u_{ttt}\| ds, \quad (1.139)$$

χρησιμοποιώντας ότι  $|s - t_{j-1}| \leq k$  για  $s \in (t_{j-1}, t_j)$  και ότι  $|s - t_{j-2}| \leq 2k$  για  $s \in (t_{j-2}, t_j)$ .

Παρόμοια

$$k\|\omega_1^1\| + k\|\omega_2^1\| \leq Ch^r \int_0^k \|u_t\|_r ds + k \int_0^k \|u_{tt}\| ds. \quad (1.140)$$

Επομένως έχουμε ότι

$$k \sum_{j=1}^n \|\omega^j\| \leq Ch^r \int_0^{t_n} \|u_t\|_r + k \int_0^k \|u_{tt}\| ds + Ck^2 \int_0^{t_n} \|u_{ttt}\|. \quad (1.141)$$

Χρησιμοποιώντας την τελευταία σχέση στην (1.137) μαζί με το ήδη γνωστό φράγμα για τον όρο  $\|\theta^0\|$  έχουμε το ζητούμενο.  $\square$

Στις παραπάνω πλήρως διακριτές μεθόδους χρησιμοποιήσαμε σταθερό χρονικό βήμα  $k$ . Τώρα θα δείξουμε μια εκτίμηση σφάλματος για μια πλήρως διακριτή μέθοδο με μεταβλητό χρονικό βήμα. Αυτή η μέθοδος είναι μια παραλλαγή της έμμεσης Euler μεθόδου.

Έστω  $0 = t_0 < t_1 < \dots, t_n < \dots$  μια διαμέριση του θετικού ημιάξονα στον χρόνο, και έστω  $k_n = t_n - t_{n-1}$ . Η ΜΠΣ που μόλις περιγράψαμε ορίζεται ως εξής

Να βρεθεί  $U^n \in S_h$  τέτοιο ώστε

$$(\bar{\partial}_n U^n, \chi) + (\nabla U^n, \nabla \chi) = (f(t_n), \chi), \quad \forall \chi \in S_h, \text{ για } n \geq 1, \quad (1.142)$$

με  $U^0 = v_h$ , όπου  $\bar{\partial}_n U^n = (U^n - U^{n-1})/k_n$ . Αν  $k_n$  σταθερό τότε το παρακάτω Θεώρημα συμπίπτει με το Θεώρημα 1.8. Το Θεώρημα που ακολουθεί το μελετήσαμε από το βιβλίο [6] σελ.20-21.

**Θεώρημα 1.11.** Έστω  $U^n$  και  $u$  λύσεις των (1.142) και (1.1) αντίστοιχα, με  $U^0 = v_h$  τέτοιο ώστε  $\|v_h - v\| \leq Ch^r \|v\|_r$  και  $v = 0$  στο  $\partial\Omega$ . Τότε έχουμε για  $n \geq 0$  ότι

$$\|U^n - u(t_n)\| \leq Ch^r \left( \|v\|_r + \int_0^{t_n} \|u_t\|_r ds \right) + \sum_{j=1}^n k_j \int_{t_{j-1}}^{t_j} \|u_{tt}\| ds. \quad (1.143)$$

*Απόδειξη.* Παρόμοια με την απόδειξη του Θεωρήματος 1.8 έχουμε

$$(\bar{\partial}_n \theta^n, \chi) + (\nabla \theta^n, \nabla \chi) = -(\omega^n, \chi), \quad \forall \chi \in S_h, \text{ για } n \geq 1, \quad (1.144)$$

όπου

$$\omega^n = (R_h - I)\bar{\partial}_n u^n + (\bar{\partial}_n u^n - u_t^n) = \omega_1^n + \omega_2^n. \quad (1.145)$$

Εντελώς ανάλογα με την (1.95) στο Θεώρημα 1.8 έχουμε ότι

$$\|\theta^n\| \leq \|\theta^{n-1}\| + k_n \|\omega^n\|.$$

Συνεπώς ανάλογα με την (1.95) έχουμε ότι

$$\|\theta^n\| \leq \|\theta^0\| + \sum_{j=1}^n k_j (\|\omega_1^j\| + \|\omega_2^j\|). \quad (1.146)$$

Τώρα έχουμε αντικαθιστώντας στις σχέσεις (1.99) και (1.101) το  $k$  με  $k_j$  ότι

$$\sum_{j=1}^n k_j \|\omega_1^j\| \leq \sum_{j=1}^n \int_{t_{j-1}}^{t_j} Ch^r \|u_t\|_r ds = Ch^r \int_0^{t_n} \|u_t\|_r ds, \text{ και} \quad (1.147)$$

$$\sum_{j=1}^n k_j \|\omega_2^j\| \leq \sum_{j=1}^n \left\| \int_{t_{j-1}}^{t_j} (s - t_{j-1}) u_{tt}(s) ds \right\| \leq \int_0^{t_n} k_j \|u_{tt}\| ds, \quad (1.148)$$

αφού  $|s - t_{j-1}| \leq k_j$  για  $s \in (t_{j-1}, t_j)$  Συνδυάζοντας τα παραπάνω με τις γνωστές εκτιμήσεις για τα  $\rho^n$  και  $\theta^0$  έχουμε το ζητούμενο.  $\square$

Παρατηρούμε ότι η παραπάνω εκτίμηση σφάλματος μας λέει ότι για να έχουμε καλή εκτίμηση σε σημεία όπου έχουμε μεγάλη τιμή για το  $\|u_{tt}\|$  πρέπει να χρησιμοποιούμε μικρότερα χρονικά βήματα.

Θα τελειώσουμε αυτό το κεφάλαιο δείχνοντας την κατασκευή της ημιδιακριτής ΜΠΣ για το πρόβλημα αρχικών συνοριακών τιμών

$$\begin{aligned} u_t - \Delta u + u &= f \text{ στο } \Omega, \text{ για } t > 0, \\ \frac{\partial u}{\partial n} &= 0 \text{ στο } \partial\Omega, \text{ για } t > 0, \\ u(\cdot, 0) &= v \text{ στο } \Omega, \end{aligned} \quad (1.149)$$

όπου  $\partial u / \partial n$  η παράγωγος στην κατεύθυνση του μοναδιαίου κάθετου διανύσματος στο  $\partial\Omega$ . Το αντίστοιχο στατικό πρόβλημα είναι το εξής

$$\begin{aligned} -\Delta u + u &= f \text{ στο } \Omega, \\ \frac{\partial u}{\partial n} &= 0 \text{ στο } \partial\Omega. \end{aligned} \quad (1.150)$$

Να σημειώσουμε ότι ο όρος  $u$  έχει ουσιαστική σημασία στην ανάλυση του προβλήματος για αυτό και αναφέρουμε το συγκεκριμένο πρόβλημα ξεχωριστά. Για να γράψουμε το παραπάνω πρόβλημα σε ασθενή μορφή πολλαπλασιάζουμε την ΜΔΕ με μια συνάρτηση ελέγχου  $\varphi \in H^1$  (δηλαδή χωρίς να απαιτούμε η  $\varphi = 0$  στο  $\partial\Omega$ ), ολοκληρώνουμε στο  $\Omega$  και παίρνουμε

$$(-\Delta u + u, \varphi) + \int_{\partial\Omega} \frac{\partial u}{\partial n} \varphi ds = (f, \varphi), \quad \forall \varphi \in H^1. \quad (1.151)$$

Αντιστρόφως αν η συνάρτηση  $u$  είναι ομαλή τότε η παραπάνω σχέση συνεπάγεται το πρόβλημα (1.150). Χρησιμοποιώντας την φόρμουλα του Green έχουμε την ασθενή μορφή

$$(\nabla u, \nabla \varphi) + (u, \varphi) = (f, \varphi), \quad \forall \varphi \in H^1. \quad (1.152)$$

Παρατηρούμε σε αυτή την περίπτωση ότι σε αντίθεση με προηγουμένως, η συνοριακή συνθήκη είναι μια συνέπεια της ασθενούς μορφής, και για αυτό τον λόγο η σ.σ Neumann καλείται και φυσική συνοριακή συνθήκη. Αντίθετα η σ.σ Dirichlet καλείται ουσιώδης συνοριακή συνθήκη.

Τώρα υποθέτουμε ότι ο χώρος των υποψήφιων προσεγγιστικών λύσεων  $S_h$  είναι υπόχωρος του  $H^1$  και ικανοποιεί την σχέση (1.22) όταν  $v \in H^s$ . Άρα το αντίστοιχο διακριτό στατικό πρόβλημα είναι το εξής

$$(\nabla u_h, \nabla \chi) + (u_h, \chi) = (f, \chi), \quad \forall \chi \in S_h. \quad (1.153)$$

Για το παραπάνω πρόβλημα γίνεται μια ανάλογη ανάλυση σφάλματος όπως αυτή του Θεωρήματος 1.3. Τέλος έχουμε ότι το ημιδιακριτό πρόβλημα που αντιστοιχεί στο (1.150) είναι

$$(u_{h,t}, \chi) + (\nabla u_h, \nabla \chi) + (u_h, \chi) = (f, \chi), \quad \forall \chi \in S_h, \quad t > 0, \quad u_h(0) = v_h, \quad (1.154)$$

Για την ανάλυση σφάλματος της παραπάνω μεθόδου καθώς και για της πλήρως διακριτής εκδοχής του ακολουθούμε την ίδια λογική με αυτή που χρησιμοποιήσαμε και όταν είχαμε σ.σ Dirichlet.

## Κεφάλαιο 2

# Γενικότερες Παραβολικές Εξισώσεις

### 2.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα μελετήσουμε μια γενικότερη ομάδα παραβολικών εξισώσεων, στις οποίες ο ελλειπτικός τελεστής θα εξαρτάται αυτή την φορά όχι μόνο από τον χώρο αλλά και από τον χρόνο, θα περιέχει όρους χαμηλότερης τάξης και δεν θα είναι απαραίτητα συμμετρικός και θετικά ορισμένος. Για να μην έχουμε πιθανή εκθετική αύξηση στην εκτίμηση του σφάλματος και για υπολογιστικούς λόγους περιορίζομαστε σε ένα πεπερασμένο χρονικό διάστημα.

Έστω το πρόβλημα αρχικών συνοριακών τιμών

$$\begin{aligned} u_t + A(t)u &= f \text{ στο } \Omega, \text{ για } t \in J, \\ u &= 0 \text{ στο } \partial\Omega, \text{ για } t \in J, \quad u(0) = v \text{ στο } \Omega, \end{aligned} \quad (2.1)$$

όπου  $\Omega \subset \mathbb{R}^d$ , με κατάλληλα ομαλό σύνορο  $\partial\Omega$ , και  $d$  θετικός ακέραιος. Επίσης  $J = (0, \bar{t}]$  και  $A(t)$  συμβολίζει τον ελλειπτικό τελεστή που ορίζεται ως εξής

$$A(t)u := - \sum_{j=1}^d \sum_{k=1}^d \frac{\partial}{\partial x_j} \left( a_{jk} \frac{\partial u}{\partial x_k} \right) + \sum_{j=1}^d a_j \frac{\partial u}{\partial x_j} + a_0 u, \quad (2.2)$$

όπου  $a_{jk}$ ,  $a_j$  και  $a_0$  είναι  $C^\infty$  συναρτήσεις στο  $\bar{\Omega} \times \bar{J}$ . Επίσης  $a_{jk} = a_{kj}$  και

$$\sum_{j=1}^d \sum_{k=1}^d a_{jk}(x, t) \xi_j \xi_k \geq c_0 |\xi|^2, \text{ με } c_0 > 0, \text{ για } (x, t) \in \bar{\Omega} \times \bar{J}.$$

Σχετίζοντας τον τελεστή  $A(t)$  με την διγραμμική μορφή

$$A(t; v, w) := \int_{\Omega} \left( \sum_{j=1}^d \sum_{k=1}^d a_{jk} \frac{\partial v}{\partial x_k} \frac{\partial w}{\partial x_j} + \sum_{j=1}^d a_j \frac{\partial v}{\partial x_j} w + a_0 v w \right) dx, \quad (2.3)$$

μπορούμε να γράψουμε το πρόβλημα (2.1) στην ασθενή μορφή

$$\begin{aligned} (u_t, \varphi) + A(t; u, \varphi) &= (f, \varphi), \quad \forall \varphi \in H_0^1, \quad t \in J, \\ u(0) &= v, \end{aligned} \quad (2.4)$$

ακολουθώντας την γνωστή διαδικασία που περιγράψαμε στο προηγούμενο κεφάλαιο. Αυτή η διγραμμική μορφή δεν είναι απαραίτητα θετικά ορισμένη, όμως ισχύει η ακόλουθη ανισότητα (βλ. [6] σελ.56).

**Πρόταση 2.1. (Ανισότητα Gårding's)** Έστω  $A(t; v, w)$  που ορίζεται από την (2.3). Τότε ισχύει η ανισότητα

$$A(t; v, v) \geq c \|v\|_1^2 - \kappa \|v\|^2, \quad \forall v \in H_0^1, \quad \text{με } c > 0, \quad \kappa \in \mathbb{R}. \quad (2.5)$$

*Απόδειξη.* Για κάθε  $v \in H_0^1$  ισχύει ότι

$$\begin{aligned} &A(t; v, v) + \kappa \|v\|^2 \\ &= \int_{\Omega} \left( \sum_{j=1}^d \sum_{k=1}^d a_{jk} \frac{\partial v}{\partial x_k} \frac{\partial v}{\partial x_j} + \sum_{j=1}^d a_j v \frac{\partial v}{\partial x_j} + (a_0 + \kappa) v^2 \right) dx \\ &= \int_{\Omega} \left( \sum_{j=1}^d \sum_{k=1}^d a_{jk} \frac{\partial v}{\partial x_k} \frac{\partial v}{\partial x_j} + \left( a_0 + \kappa - \frac{1}{2} \sum_{j=1}^d \frac{\partial a_j}{\partial x_j} \right) v^2 \right) dx \\ &\geq c \|v\|_1^2, \quad \text{με } c > 0, \quad \text{αν } \kappa > \sup_{\Omega \times J} \left( \sum_{j=1}^d \frac{\partial a_j}{\partial x_j} - a_0 \right). \end{aligned}$$

Άρα έχουμε το ζητούμενο. Να σημειώσουμε ότι χρησιμοποιήσαμε ολοκλήρωση κατά μέρη στην δεύτερη ισότητα.  $\square$

Το ακόλουθο πολύ γνωστό Λήμμα θα μας φανεί ιδιαίτερα χρήσιμο στην συνέχεια της μελέτης μας.

**Λήμμα 2.1. (Gronwall)** Έστω  $\varphi(t)$  μια συνεχής συνάρτηση στο διάστημα  $[0, T]$ . Υποθέτουμε ότι υπάρχει  $\kappa = \kappa(\tau) > 0$  και μια άλλη συνεχής συνάρτηση  $f(t)$  στο  $[0, T]$  τέτοια ώστε να ισχύει

$$\varphi(t) \leq f(t) + \int_0^t \kappa(\tau) \varphi(\tau) d\tau, \quad \forall t \in [0, T].$$



Τότε έχουμε ότι

$$\varphi(t) \leq f(t) + \int_0^t \kappa(\tau) f(\tau) \exp\left(\int_\tau^t \kappa(s) ds\right) d\tau.$$

Επιπλέον αν η  $f'(t) \geq 0$ , τότε ισχύει

$$\varphi(t) \leq f(t) \exp\left(\int_0^t \kappa(\tau) d\tau\right).$$

Απόδειξη. Έστω

$$R(t) = \int_0^t \kappa(\tau) \varphi(\tau) d\tau.$$

Τότε έχουμε ότι  $R'(t) = \kappa(t)\varphi(t) \leq \kappa(t)f(t) + \kappa(t)R(t)$  σχεδόν παντού στο  $(0, +\infty)$ . Έτσι έχουμε ότι  $R'(t) - \kappa(t)R(t) \leq \kappa(t)f(t)$  σχεδόν παντού στο  $(0, +\infty)$ . Επομένως

$$\frac{d}{ds} \exp\left(-\int_0^s \kappa(\tau) d\tau\right) R(s) \leq \exp\left(-\int_0^s \kappa(\tau) d\tau\right) \kappa(s)f(s),$$

σχεδόν παντού στο  $(0, +\infty)$ . Ολοκληρώνοντας ως προς  $t$  παίρνουμε

$$\exp\left(-\int_0^t \kappa(\tau) d\tau\right) R(t) \leq \int_0^t \exp\left(-\int_0^s \kappa(\tau) d\tau\right) \kappa(s)f(s) ds.$$

Επομένως ισχύει ότι

$$R(t) \leq \int_0^t \exp\left(\int_s^t \kappa(\tau) d\tau\right) \kappa(s)f(s) ds,$$

το οποίο αποδεικνύει το ζητούμενο.  $\square$

Γνωρίζουμε ότι το πρόβλημα (2.1) έχει μοναδική λύση η οποία ανήκει σε έναν χώρο  $H^s$ , όπου  $s$  θετικός ακέραιος, όπως και οι χρονικές της παραγώγους για  $t \in J$ . Επίσης η λύση ικανοποιεί την (1.41), δεδομένου ότι η  $f$  και  $v$  είναι αρκούντως ομαλές, και ικανοποιούν τις συνθήκες συμβατότητας στο  $\partial\Omega$  όταν  $t = 0$ . Για λεπτομέρειες παραπέμπουμε στο βιβλίο [4]. Εμείς θα αρκεστούμε στην απόδειξη της ακόλουθης εκτίμησης ευστάθειας στον  $L^2$ .

**Πρόταση 2.2.** Έστω  $u$  λύση του προβλήματος (2.1). Επίσης υποθέτουμε ότι ισχύει η (1.41). Τότε ισχύει

$$\|u(t)\| \leq C \left( \|v\| + \int_0^t \|f\| ds \right), \text{ για } t \in \bar{J}, \quad (2.6)$$

όπου η σταθερά  $C > 0$  εξαρτάται από το χωρίο  $\Omega$ , από το άνω άκρο  $\bar{t}$  του διαστήματος  $J$  και από το χρονικό βήμα  $k$ .

Απόδειξη. Όπως είπαμε προηγουμένως, το πρόβλημα (2.1) γράφεται στην ασθενή μορφή (2.4). Θέτοντας λοιπόν στην ασθενή μορφή  $\varphi = u$  παίρνουμε

$$(u_t, u) + A(t; u, u) = (f, u) \quad (2.7)$$

Χρησιμοποιώντας την ανισότητα Gårding's παίρνουμε

$$\frac{1}{2} \frac{d}{dt} \|u\|^2 + c \|u\|_1^2 \leq \|f\| \|u\| + \kappa \|u\|^2. \quad (2.8)$$

Αφού ο όρος  $c \|u\|_1^2$  είναι μη αρνητικός μπορούμε να τον πετάξουμε. Στην συνέχεια διαιρώντας με τον όρο  $\|u\|$  και τις δύο πλευρές και ολοκληρώνοντας στο  $[0, t]$  παίρνουμε

$$\|u(t)\| \leq \|v\| + \int_0^t \|f\| ds + \kappa \int_0^t \|u\| ds. \quad (2.9)$$

Από το Λήμμα του Gronwall προκύπτει η ζητούμενη ανισότητα.  $\square$

## 2.2 Ημιδιακριτό πρόβλημα

Το χρονικά εξαρτούμενο πρόβλημα Dirichlet που αντιστοιχεί στο πρόβλημα (2.1) ορίζεται ως ακολούθως

$$A_\kappa(t)u := A(t)u + \kappa u = f \text{ στο } \Omega, \text{ με } u = 0 \text{ στο } \partial\Omega, \text{ για } \bar{t} \in J, \quad (2.10)$$

το οποίο ακολουθώντας την γνωστή μας πλέον διαδικασία γράφεται σε ασθενή μορφή ως εξής

$$A_\kappa(t; u, \varphi) := A(t; u, \varphi)u + \kappa(u, \varphi) = (f, \varphi), \quad \forall \varphi \in H_0^1, \text{ για } t \in J. \quad (2.11)$$

**Ορισμός 2.1.** Ορίζουμε με  $T(t) : L^2 \rightarrow H^2 \cap H_0^1$  τον τελεστή λύσης του προβλήματος (2.10)-(2.11). Με άλλα λόγια ο  $T(t)$  ορίζεται από την σχέση

$$A_\kappa(t; T(t)f, \varphi) = (f, \varphi), \quad \forall \varphi \in H_0^1, \text{ για } t \in \bar{J}. \quad (2.12)$$

Επισημαίνουμε ότι η διγραμμική μορφή  $A_\kappa$  έχει χωροχρονικά εξαρτούμενους συντελεστές.

Ακολουθώντας τον παραπάνω συμβολισμό η ελλειπτική κανονικότητα γράφεται και στην μορφή

$$\|T(t)f\|_s \leq C \|f\|_{s-2}, \text{ για } s \geq 2, \text{ για } t \in \bar{J}. \quad (2.13)$$

Θέτοντας τώρα  $\check{u} = e^{-t\kappa}u$  ως νέα εξαρτημένη μεταβλητή στο πρόβλημα (2.1) έχουμε

$$\check{u}_t + A_\kappa(t)\check{u} = \check{f}, \quad (2.14)$$

όπου  $\check{f} = e^{-t\kappa}f$ . Ισοδύναμα σε ασθενή μορφή έχουμε

$$(\check{u}_t, \varphi) + A_\kappa(t; \check{u}, \varphi) = (\check{f}, \varphi), \quad \forall \varphi \in H_0^1, \text{ για } t \in J. \quad (2.15)$$

Μπορούμε επίσης να γράψουμε το πρόβλημα μας και στην μορφή

$$T(t)\check{u}_t + \check{u} = T(t)\check{f}, \quad \text{για } t \in J, \text{ με } \check{u}(0) = v. \quad (2.16)$$

Για να κατασκευάσουμε προσεγγιστικές λύσεις για το πρόβλημα (2.1), υποθέτουμε ότι ο  $S_h$  είναι ένας πεπερασμένης διάστασης υπόχωρος του  $L^2$  και ορίζουμε  $T_h(t) : L^2 \rightarrow S_h$  ως την προσέγγιση του  $T(t)$ , η οποία ικανοποιεί κάποιες ιδιότητες τις οποίες θα δούμε στην συνέχεια, και ορίζεται από την παρακάτω σχέση

$$T_h = R_h T, \quad (2.17)$$

όπου  $R_h$  η ελλειπτική προβολή που ορίσαμε στο προηγούμενο κεφάλαιο. Το ημιδιακριτό μας πρόβλημα λοιπόν μπορεί να γραφτεί ως εξής

Να βρεθεί μια συνάρτηση  $u_h : J \rightarrow S_h$  τέτοια ώστε

$$u_h(t) = e^{\kappa t} \check{u}_h(t), \quad \text{όπου } T_h \check{u}_{h,t} + \check{u}_h = T_h \check{f}, \quad \text{για } t \in J, \quad \check{u}_h(0) = v_h. \quad (2.18)$$

Παρατηρούμε ότι για  $t > 0$ , τα  $\check{u}$  και  $\check{u}_h$  στις παραπάνω εξισώσεις αυξάνονται εκθετικά όσο αυξάνονται τα  $u$  και  $u_h$ , όταν το  $\kappa > 0$ . Το παραπάνω ημιδιακριτό πρόβλημα γράφεται και στην μορφή

$$(\check{u}_{h,t}, \chi) + A_\kappa(\check{u}_h, \chi) = (\check{f}, \chi), \quad \forall \chi \in S_h, \quad t \in J,$$

το οποίο με την σειρά του, αφού  $u_h(t) = e^{\kappa t} \check{u}_h(t)$  είναι ισοδύναμο με την κλασσική ασθενή μορφή του προβλήματος μας η οποία είναι η εξής

$$(u_{h,t}, \chi) + A(u_h, \chi) = (f, \chi), \quad \forall \chi \in S_h, \quad t \in J.$$

Για λόγους συντομίας από εδώ και στο εξής θα παραλείψουμε την μεταβλητή  $t$  από τους συμβολισμούς μας. Για παράδειγμα θα συμβολίζουμε το  $A(t)$  με  $A$ , το  $A(t; v, w)$  με  $A(v, w)$ , το  $T_h(t)$  με  $T_h$  κ.ο.κ.

Τώρα θα αναφέρουμε τις συνθήκες που απαιτούνται για τον τελεστή  $T_h$  έτσι ώστε η προσέγγιση  $u_h$  που ορίσαμε στην (2.18) να είναι μια καλή προσέγγιση της ακριβούς λύσης  $u$  του (2.1). Το παρακάτω Θεώρημα το μελετήσαμε από το βιβλίο [6] σελ.57-59.

**Θεώρημα 2.1.** Έστω ο τελεστής λύσης  $T$  του προβλήματος (2.10)-(2.11) και  $T_h$  η προσέγγιση του  $T$ . Έστω επίσης ότι  $t \in J$ , και  $C$  σταθερά ανεξάρτητη του  $t$ . Επιπρόσθετα συμβολίζουμε με  $'$  την παράγωγο ως προς  $t$ . Τότε ισχύουν οι εξής ιδιότητες για τον προσεγγιστικό τελεστή λύσης  $T_h$ .

- i.  $(f, T_h f) \geq 0$  για  $f \in L^2$  και  $(\chi, T_h \chi) > 0$  για  $0 \neq \chi \in S_h$ .
- ii. Για  $r \geq 2$  ακέραιο και  $2 \leq s \leq r$  ισχύει ότι

$$\|(T_h - T)f\| + \|(T_h' - T')f\| \leq Ch^s \|f\|_{s-2}, \quad \text{για } f \in H^{s-2}.$$

- iii.  $|(T_h f, g) - (f, T_h g)| \leq C(f, T_h f)^{1/2} \|T_h g\|$ , για  $f, g \in L^2$ .

*Απόδειξη.* Εξ'ορισμού ο τελεστής  $T_h$  ικανοποιεί την ΜΠΣ η οποία είναι το διακριτό ανάλογο της σχέσης (2.12), δηλαδή

$$A_\kappa(t; T_h(t)f, \chi) = (f, \chi), \quad \forall \chi \in S_h, \quad \text{για } t \in J. \quad (2.19)$$

Από την Πρόταση 2.1 έχουμε ότι για κάθε  $f \in L^2$  ισχύει ότι

$$(f, T_h f) = A_\kappa(T_h f, T_h f) \geq c \|T_h f\|_1^2 \geq 0. \quad (2.20)$$

Άρα αφού  $S_h \subset L^2$  ισχύει επίσης ότι

$$(\chi, T_h \chi) = A_\kappa(T_h \chi, T_h \chi) \geq c \|T_h \chi\|_1^2 \geq 0. \quad (2.21)$$

Η ισότητα ισχύει αν και μόνο αν  $\chi = 0$ . Συνεπώς δείξαμε το (i). Πάμε τώρα να αποδείξουμε την (ii). Από τις σχέσεις (1.22) και (2.13) γνωρίζουμε ότι ισχύει το εξής

$$\|(T_h - T)f\| + h \|(T_h - T)f\|_1 \leq Ch^s \|f\|_{s-2}, \quad \text{για } 2 \leq s \leq r, \quad (2.22)$$

και απομένει να δείξουμε το αντίστοιχο αποτέλεσμα για την χρονική παράγωγο. Για να το κάνουμε αυτό θέτουμε  $w = Tf$ ,  $w_h = T_h f$  και  $e = w_h - w$ . Συνεπώς έχουμε ότι  $(T_h' - T')f = e_t$ . Θέτοντας  $\varphi = \chi \in S_h$  στην (2.12) και αφαιρώντας την από την (2.19) παίρνουμε

$$A_\kappa(e, \chi) = 0, \quad \forall \chi \in S_h, \quad \text{για } t \in J. \quad (2.23)$$

Παραγωγίζοντας την παραπάνω εξίσωση ως προς  $t$  παίρνουμε

$$A_\kappa(e_t, \chi) + A'(e, \chi) = 0, \quad \forall \chi \in S_h, \quad \text{για } t \in J, \quad (2.24)$$

όπου  $A'(\cdot, \cdot)$  η διγραμμική μορφή που προκύπτει από την  $A(\cdot, \cdot)$  όταν την παραγωγίσουμε ως προς  $t$ . Υπενθυμίζουμε ότι οι συντελεστές  $a_{jk}$ ,  $a_j$  και  $a_0$  είναι ομαλές συναρτήσεις ως προς  $t$ . Παρατηρούμε επίσης ότι  $A'_\kappa(\cdot, \cdot) = A'(\cdot, \cdot)$ . Επομένως, χρησιμοποιώντας την ανισότητα Gårding's έχουμε για κάθε  $\chi \in S_h$  ότι

$$c\|e_t\|_1^2 \leq A_\kappa(e_t, e_t) = A_\kappa(e_t, e_t + \chi) + A'(e, e_t + \chi) - A'(e, e_t). \quad (2.25)$$

Συνεπώς έχουμε ότι

$$\|e_t\|_1^2 \leq C(\|e_t\|_1 + \|e\|_1) \inf_{\chi \in S_h} \|w_t - \chi\|_1 + C\|e\|_1\|e_t\|_1. \quad (2.26)$$

Διαιρώντας και τις δύο πλευρές με  $\|e_t\|_1$  και χρησιμοποιώντας τώρα την σχέση (2.22) καθώς επίσης και την (1.22) παίρνουμε

$$\|e_t\|_1 \leq C(\|e\|_1 + \inf_{\chi \in S_h} \|w_t - \chi\|_1) \leq Ch^{s-1}(\|f\|_{s-2} + \|w_t\|_s), \quad 2 \leq s \leq r. \quad (2.27)$$

Αφού  $w_t \in H_0^1$  είναι η λύση του προβλήματος Dirichlet

$$A_\kappa(w_t, \varphi) = -A'(w, \varphi), \quad \forall \varphi \in H_0^1, \quad (2.28)$$

τότε ισχύει ότι  $\|w_t\|_s \leq C\|w\|_s$ . Επίσης από την (2.13) ισχύει ότι  $\|w\|_s \leq C\|f\|_{s-2}$ . Άρα έχουμε ότι

$$\|e_t\|_1 \leq h^{s-1}\|f\|_{s-2}, \quad \text{για } 2 \leq s \leq r. \quad (2.29)$$

Για να δείξουμε το φράγμα που θέλουμε για την  $L^2$ -νόρμα του  $e_t$ , υποθέτουμε ότι  $A_\kappa^*$  ο προσαρτημένος πίνακας του  $A_\kappa$  και  $\psi$  η λύση του προβλήματος

$$\begin{aligned} A_\kappa^* \psi &= \varphi \text{ στο } \Omega, \\ \psi &= 0 \text{ στο } \partial\Omega. \end{aligned} \quad (2.30)$$

Τότε έχουμε χρησιμοποιώντας την (2.24) ότι

$$(e_t, \varphi) = A_\kappa(e_t, \psi) = A_\kappa(e_t, \psi - \chi) + A'(e, \psi - \chi) - A'(e, \psi), \quad \forall \chi \in S_h.$$

Εφαρμόζοντας τώρα την φόρμουλα του Green για την ελλειπτική τελεστή του τελευταίου όρου και γνωρίζοντας ότι το ολοκλήρωμα στο σύνορο  $\partial\Omega$  μηδενίζεται αφού  $e|_{\partial\Omega} = 0$  παίρνουμε

$$\begin{aligned} |(e_t, \varphi)| &\leq C(\|e_t\|_1 + \|e\|_1) \inf_{\chi \in S_h} \|\psi - \chi\|_1 + C\|e\|\|\psi\|_2 \\ &\leq C(h(\|e_t\|_1 + \|e\|_1) + \|e\|)\|\psi\|_2. \end{aligned}$$

Από την ελλειπτική κανονικότητα  $\|\psi\|_2 \leq C\|\varphi\|$ , σε συνδυασμό με τα φράγματα (2.22) και (2.29) παίρνουμε

$$\|e_t\| \leq C(h\|e_t\|_1 + \|e\|_1) + \|e\| \leq Ch^s\|f\|_{s-2}, \text{ για } 2 \leq s \leq r. \quad (2.31)$$

Επομένως έχουμε αποδείξει το (ii). Μας απομένει τώρα να αποδείξουμε το (iii). Αρχικά θέτουμε  $v_h = T_h f$  και  $w_h = T_h g$ . Έχουμε λοιπόν ότι

$$\begin{aligned} (T_h f, g) - (f, T_h g) &= A_\kappa(T_h g, T_h f) - A_\kappa(T_h f, T_h g) \\ &= \int_\Omega \sum_{j=1}^d a_j \left( \frac{\partial w_h}{\partial x_j} v_h - \frac{\partial v_h}{\partial x_j} w_h \right) dx \\ &= \int_\Omega \sum_{j=1}^d a_j \left( \frac{\partial w_h}{\partial x_j} v_h - \frac{\partial v_h}{\partial x_j} w_h + 2 \frac{\partial v_h}{\partial x_j} w_h - 2 \frac{\partial v_h}{\partial x_j} w_h \right) dx, \\ &= - \int_\Omega \sum_{j=1}^d 2a_j \frac{\partial v_h}{\partial x_j} w_h dx \\ &\quad + \int_\Omega \sum_{j=1}^d a_j \left( \frac{\partial w_h}{\partial x_j} v_h + \frac{\partial v_h}{\partial x_j} w_h \right) dx \\ &= - \int_\Omega \sum_{j=1}^d \left( a_j \frac{\partial v_h}{\partial x_j} w_h + \frac{\partial a_j}{\partial x_j} v_h w_h \right) dx, \end{aligned} \quad (2.32)$$

αφού  $v_h \in S_h$ . Επομένως έχουμε ότι

$$\begin{aligned} |(T_h f, g) - (f, T_h g)| &\leq C\|v_h\|_1\|w_h\| = C\|T_h f\|_1\|T_h g\| \\ &\leq C(f, T_h f)^{1/2}\|T_h g\|, \end{aligned} \quad (2.33)$$

το οποίο είναι το ζητούμενο.  $\square$

Θα συνεχίσουμε με την ανάλυση σφάλματος στην  $L^2$ -νόρμα για το πρόβλημα αρχικών συνοριακών τιμών (2.1). Να σημειώσουμε ότι σε αυτή την περίπτωση δεν είναι απαραίτητο να ισχύει η συνθήκη (iii) του Θεωρήματος 2.1. Το παρακάτω Θεώρημα το μελετήσαμε από το βιβλίο [6] σελ.60.

**Θεώρημα 2.2.** Έστω ότι οι συνθήκες (i) και (ii) του Θεωρήματος 2.1 ικανοποιούνται. Τότε έχουμε για το σφάλμα του ημιδιακριτού προβλήματος (2.18) την ακόλουθη εκτίμηση.

$$\|u_h(t) - u(t)\| \leq C\|v_h - v\| + Ch^r \left( \|v\|_r + \int_0^t \|u_t\|_r ds \right), \text{ για } t \in \bar{J}. \quad (2.34)$$

Απόδειξη. Αρχικά θέτουμε  $e = \check{u}_h - \check{u} = e^{-\kappa t}(u_h - u)$  και έτσι έχουμε

$$\begin{aligned} T_h e_t + e &= (T_h \check{u}_{h,t} + \check{u}_h) - (T_h \check{u}_t + \check{u}) = T_h \check{f} - (T \check{u}_t + \check{u}) + (T - T_h) \check{u}_t \\ &= T_h \check{f} - T \check{f} + (T - T_h) \check{u}_t = (T - T_h)(\check{u}_t - \check{f}) = (T - T_h) A_\kappa \check{u}. \end{aligned}$$

Άρα έχουμε ότι

$$T_h e_t + e = \rho := (T_h - T) A_\kappa \check{u}. \quad (2.35)$$

Τώρα πολλαπλασιάζουμε την παραπάνω εξίσωση με  $2e_t$ , ολοκληρώνουμε στο  $\Omega$  και έτσι έχουμε

$$2(T_h e_t, e_t) + \frac{d}{dt} \|e\|^2 = 2(\rho, e_t) = 2 \frac{d}{dt} (\rho, e) - 2(\rho_t, e).$$

Στην συνέχεια ολοκληρώνουμε ως προς  $t$ , και πετάμε τον πρώτο όρο του αριστερού μέλους αφού ισχύει η (i), και έτσι παίρνουμε την ανισότητα

$$\begin{aligned} \|e(t)\|^2 &\leq \|e(0)\|^2 + 2\|\rho(t)\| \|e(t)\| + \|\rho(0)\| \|e(0)\| + 2 \int_0^t \|\rho_t\| \|e\| ds \\ &\leq \sup_{s \leq t} \|e(s)\| \left( \|e(0)\| + 4 \sup_{s \leq t} \|\rho(s)\| + 2 \int_0^t \|\rho_t\| ds \right). \end{aligned}$$

Θέτουμε  $\tau$  τέτοιο ώστε  $\|e(\tau)\| = \sup_{s \leq t} \|e(s)\|$ . Συνεπώς

$$\begin{aligned} \|e(t)\| &\leq \|e(\tau)\| \leq \|e(0)\| + 4 \sup_{s \leq t} \|\rho(s)\| + 2 \int_0^t \|\rho_t\| ds \\ &\leq \|e(0)\| + \left( \|\rho(0)\| + \int_0^t \|\rho_t\| ds \right). \end{aligned} \quad (2.36)$$

Χρησιμοποιώντας την συνθήκη (ii) έχουμε

$$\|\rho(0)\| \leq Ch^r \|A_\kappa v\|_{r-2} \leq Ch^r \|v\|_r. \quad (2.37)$$

Επίσης

$$\begin{aligned} \|\rho_t\| &\leq \|(T'_h - T') A_\kappa \check{u}\| + \|(T_h - T)(A' \check{u} + A_\kappa \check{u}_t)\| \\ &\leq Ch^r (\|\check{u}\|_r + \|\check{u}_t\|_r) \leq Ch^r (\|u\|_r + \|u_t\|_r). \end{aligned}$$

Αφού το  $J$  είναι φραγμένο, έχουμε ότι

$$\int_0^t \|\rho_t\| ds \leq Ch^r \left( \|v\|_r + \int_0^t \|u_t\|_r ds \right). \quad (2.38)$$

Συνδυάζοντας τις σχέσεις (2.37) και (2.38) με την (2.36) παίρνουμε το ζητούμενο.  $\square$

Τώρα θα ασχοληθούμε με το ομογενές πρόβλημα αρχικών συνοριακών του οποίου τα αρχικά δεδομένα είναι ασυνεχή ή αλλιώς η  $v$  θα ικανοποιεί περιορισμένες υποθέσεις ομαλότητας.

$$\begin{aligned} u_t + A(t)u &= 0 \text{ στο } \Omega, \text{ για } t \in J, \\ u &= 0 \text{ στο } \partial\Omega, \text{ για } t \in J, \\ u(0) &= v \text{ στο } \Omega. \end{aligned} \quad (2.39)$$

Το ημιδιακριτό ανάλογο του παραπάνω προβλήματος είναι το εξής

Να βρεθεί  $u_h(t) : J \rightarrow S_h$ , τέτοιο ώστε  $\check{u}_h(t) = e^{-\kappa t}u_h(t)$ , το οποίο ικανοποιεί την εξίσωση

$$T_h\check{u}_{h,t} + \check{u}_h = 0, \text{ για } t \in J, \text{ με } \check{u}_h(0) = v_h. \quad (2.40)$$

Για την απόδειξη του κυρίως μας Θεωρήματος για την εκτίμηση σφάλματος θα χρειαστεί να αποδείξουμε πρώτα κάποια βοηθητικά Λήμματα. Ξεκινάμε με την απόδειξη ενός Λήμματος το οποίο είναι παρόμοιο με το Θεώρημα 2.1.

**Λήμμα 2.2.** Υποθέτουμε ότι η συνθήκη (ii) του Θεωρήματος 2.1 ισχύει για  $r = 2$ . Τότε για  $f \in L^2$  ισχύουν οι ανισότητες

$$|(T'_h f, f)| \leq C((T_h f, f) + h^2\|f\|^2), \quad (2.41)$$

και

$$\|T'_h f\| \leq C(\|T_h f\| + h^2\|f\|). \quad (2.42)$$

*Απόδειξη.* Θα αποδείξουμε τα συνεχή ανάλογα των ανισοτήτων (2.41) και (2.42). Δηλαδή τις ανισότητες

$$|(T' f, f)| \leq C(T f, f), \quad (2.43)$$

και

$$\|T' f\| \leq C\|T f\|. \quad (2.44)$$

Οι ζητούμενες ανισότητες προκύπτουν από την συνθήκη (ii) του Θεωρήματος 2.1 ως εξής

$$\begin{aligned} |(T' f, f)| &= |(T' f, f) + ((T'_h - T')f, f)| \leq C((T f, f) + h^2\|f\|^2) \\ &\leq C((T_h f, f) + h^2\|f\|^2), \end{aligned}$$

και

$$\begin{aligned} \|T'_h f\| &= \|T' f + T'_h f - T' f\| \leq \|T' f\| + \|(T'_h - T')f\| \\ &\leq C(\|T f\| + h^2\|f\|) \leq C(\|T_h f\| + h^2\|f\|). \end{aligned}$$



Για να αποδείξουμε τώρα τις συνεχής ανισότητες (2.43) και (2.44) υπενθυμίζουμε τον ορισμό (2.12) και σημειώνουμε ότι ο συζυγής του τελεστή  $T$ , τον οποίο συμβολίζουμε με  $T^* : L^2 \rightarrow H^2 \cap H_0^1$  ορίζεται από την σχέση

$$A_\kappa(\varphi, T^*g) = (\varphi, g), \text{ για } \varphi \in H_0^1. \quad (2.45)$$

Δηλαδή ο  $T^*$  είναι η λύση του προβλήματος Dirichlet με τον ελλειπτικό τελεστή  $A_\kappa^*$ . Συνεπώς έχουμε

$$(Tf, g) = A_\kappa(Tf, T^*g) = (f, T^*g), \quad \forall f, g \in L^2. \quad (2.46)$$

Παραγωγίζοντας την (2.12) παίρνουμε ότι  $A_\kappa(T'f, \varphi) + A'(Tf, \varphi) = 0$ . Επομένως έχουμε ότι

$$\begin{aligned} |(T'f, f)| &= |A_\kappa(T'f, T^*f)| = |A'(Tf, T^*f)| \leq C\|Tf\|_1\|T^*f\|_1 \\ &\leq C(f, Tf)^{1/2}(T^*, f)^{1/2} = C(f, Tf). \end{aligned}$$

Άρα δείξαμε την (2.43). Μας απομένει λοιπόν να αποδείξουμε την (2.44). Γνωρίζουμε ότι για  $\varphi \in L^2$  ισχύει η σχέση

$$(T'f, \varphi) = A_\kappa(T'f, T^*\varphi) = -A'(Tf, T^*\varphi). \quad (2.47)$$

Χρησιμοποιώντας την φόρμουλα του Green παίρνουμε

$$|(T'f, \varphi)| \leq C\|Tf\|\|T^*\varphi\|_2 \leq C\|Tf\|\|\varphi\|,$$

το οποίο είναι η ζητούμενη ανισότητα.  $\square$

**Λήμμα 2.3.** Υποθέτουμε ότι ισχύουν οι συνθήκες (i), (ii) με  $r = 2$ , και (iii) του Θεωρήματος 2.1. Επίσης έχουμε ότι

$$T_h e_t + e = \rho, \text{ για } t \in J, \text{ με } T_h e(0) = 0. \quad (2.48)$$

Τότε για κάθε  $\varepsilon > 0$  υπάρχει  $C_\varepsilon$  τέτοιο ώστε

$$\|e(t)\| \leq \varepsilon \sup_{s \leq t} (s\|\rho_t(s)\|) + C_\varepsilon \sup_{s \leq t} \|\rho(s)\|, \text{ για } t \in J. \quad (2.49)$$

*Απόδειξη.* Θα αποδείξουμε την ανισότητα

$$\|e(t)\|^2 \leq \frac{\varepsilon^2}{t} \int_0^t s^2 \|\rho_t(s)\|^2 ds + C_\varepsilon \left( \|\rho(t)\|^2 + \frac{1}{t} \int_0^t \|\rho\|^2 ds \right), \quad (2.50)$$

η οποία συνεπάγεται το ζητούμενο, αφού παίρνοντας το *supremum* του δεξιού μέλους και έπειτα την τετραγωνική ρίζα και στις δύο πλευρές παίρνουμε την

σχέση (2.49). Για να δείξουμε λοιπόν την παραπάνω σχέση πολλαπλασιάζουμε την (2.48) με  $2te_t$ , ολοκληρώνουμε στο  $\Omega$  και έτσι κάνοντας κάποιες απλές αλγεβρικές πράξεις παίρνουμε

$$2t(T_h e_t, e_t) + \frac{d}{dt}(t\|e\|^2) = 2\frac{d}{dt}(t(\rho, e)) - 2(\rho, e) - 2t(\rho_t, e) + \|e\|^2.$$

Χρησιμοποιώντας την συνθήκη (i) πετάμε τον πρώτο όρο του αριστερού μέλους. Ολοκληρώνοντας ως προς  $t$  και παίρνουμε

$$t\|e(t)\|^2 \leq \int_0^t s^2 \|\rho_t\|^2 ds + C_\varepsilon \left( t\|\rho(t)\|^2 + \int_0^t (\|\rho\|^2 + \|e\|^2) ds \right).$$

Βλέποντας την παραπάνω εξίσωση, παρατηρούμε ότι για να ολοκληρώσουμε την απόδειξη της (2.50), χρειάζεται να δείξουμε ότι

$$\int_0^t \|e\|^2 ds \leq C \int_0^t \|\rho\|^2 ds. \quad (2.51)$$

Για να αποδείξουμε λοιπόν την παραπάνω σχέση πολλαπλασιάζουμε την (2.48) με  $2e$ , ολοκληρώνουμε στο  $\Omega$  και κάνοντας κάποιες απλές αλγεβρικές πράξεις παίρνουμε

$$\frac{d}{dt}(T_h e, e) + \|e\|^2 = 2(\rho, e) + (T'_h e, e) + ((T_h e, e_t) - (T_h e_t, e)).$$

Μπορούμε να φράξουμε τον τελευταίο όρο του δεξιού μέλους χρησιμοποιώντας την συνθήκη (iii) του Θεωρήματος 2.1 καθώς επίσης και την (2.48).

$$\begin{aligned} |(T_h e, e_t) - (T_h e_t, e)| &\leq C(T_h e, e)^{1/2} \|T_h e_t\| \leq C(T_h e, e)^{1/2} (\|\rho\| + \|e\|) \\ &\leq C(T_h e, e) + C\|\rho\|^2 + \frac{1}{2}\|e\|^2. \end{aligned}$$

Επίσης από το Λήμμα 2.2, για μικρό  $h$ , έχουμε ότι

$$|(T'_h e, e)| \leq C(T_h e, e) + \frac{1}{4}\|e\|^2.$$

Επομένως έχουμε ότι

$$\frac{d}{dt}(T_h e, e) + \|e\|^2 \leq (\|\rho\|^2 + (T_h e, e)). \quad (2.52)$$

Στην συνέχεια χρησιμοποιούμε το Λήμμα του Gronwall, τις υποθέσεις ότι  $T_h e(0) = 0$ , το  $J$  είναι φραγμένο και ότι  $(T_h e, e) \geq 0$  από την συνθήκη (i) και έτσι παίρνουμε την σχέση (2.51). Άρα δείξαμε ότι ισχύει η σχέση (2.50) και συνεπώς το ζητούμενο.  $\square$

Το τελευταίο Λήμμα που θα χρειαστούμε είναι το εξής.

**Λήμμα 2.4.** Για κάθε  $j \geq 0$  έχουμε για τη λύση του προβλήματος (2.39) με  $u(0 = v)$  ότι

$$\|D_t^j u(t)\| \leq C t^{-j} \|v\|, \text{ για } t \in J. \quad (2.53)$$

*Απόδειξη.* Θα δείξουμε το ζητούμενο μόνο για την περίπτωση όπου το  $j = 1$ . Για  $j = 0$  η ζητούμενη ανισότητα προκύπτει από την ευστάθεια της  $u$  (2.6). Για τιμές του  $j$  μεγαλύτερες τις μονάδας παραπέμπουμε στο [7]. Υποθέτουμε τώρα ότι  $\kappa = 0$  έτσι ώστε να μην χρειαστεί να κάνουμε και πάλι τον μετασχηματισμό  $\tilde{u}(t) = e^{-\kappa t} u(t)$ . Θέτοντας  $f = 0$  και  $\varphi = 2t^2 u_t$  στην (2.4) και παραγωγίζοντας ως προς τη μεταβλητή  $t$  παίρνουμε

$$\frac{d}{dt}(t^2 \|u_t\|^2) + 2t^2 (A(u_t, u_t) + A'(u, u_t)) = 2t \|u_t\|^2.$$

Ολοκληρώνοντας την παραπάνω εξίσωση και χρησιμοποιώντας την ανισότητα Gårding's με  $\kappa = 0$  μαζί με άλλες γνωστές εκτιμήσεις έχουμε ότι

$$t^2 \|u_t(t)\|^2 + \int_0^t s^2 \|u_t\|_1^2 ds \leq C \int_0^t \|u\|_1^2 ds + C \int_0^t s \|u_t\|^2 ds.$$

Ολοκληρώνοντας τώρα την σχέση (2.8) θέτοντας πρώτα την  $f = 0$  παρατηρούμε ότι ο πρώτος όρος του δεξιού μέλους της παραπάνω ανισότητας φράσσεται από την ποσότητα  $C \|v\|^2$ . Για να φράξουμε τον δεύτερο όρο στο δεξί μέλος παρατηρούμε από την σχέση (2.32) ότι

$$|A(v, w) - A(w, v)| \leq C \|v\|_1 \|w\|.$$

Επομένως

$$\frac{d}{dt} A(u, u) = A(u, u_t) + A(u_t, u) + A'(u, u) \leq 2A(u, u_t) + \|u_t\|^2 + C \|u\|_1^2.$$

Θέτοντας τώρα  $f = 0$  και  $\varphi = 2u_t$  στην (2.4) παίρνουμε

$$\|u_t\|^2 + \frac{d}{dt} A(u, u) \leq C \|u\|_1^2.$$

Πολλαπλασιάζοντας την παραπάνω σχέση με  $t$  και ολοκληρώνοντας παίρνουμε

$$\int_0^t s \|u_t\|^2 ds + t \|u(t)\|_1^2 \leq C \int_0^t \|u\|_1^2 ds \leq C \|v\|^2.$$

Συνοψίζοντας δείξαμε ότι  $t^2 \|u_t(t)\| \leq C \|v\|$ , το οποίο είναι το ζητούμενο για  $j = 1$ .  $\square$

Στο ακόλουθο Θεώρημα θα αποδείξουμε μια εκτίμηση σφάλματος για μη ομαλά αρχικά δεδομένα, για συνεχή κατά τμήματα γραμμικές συναρτήσεις, στην οποία έχουμε δεύτερη τάξη σύγκλισης ως προς τη χωρική παράμετρο  $h$ . Το παρακάτω Θεώρημα το μελετήσαμε από το βιβλίο [6] σελ.61-65.

**Θεώρημα 2.3.** Υποθέτουμε ότι ισχύουν οι συνθήκες (i), (ii) με  $r = 2$ , και (iii) του Θεωρήματος 2.1. Επίσης υποθέτουμε ότι  $v_h = P_h v$ , όπου  $P_h v$  η ορθογώνια προβολή του  $v$  στον  $S_h$  ως προς το εσωτερικό γινόμενο στον  $L^2$ . Τότε έχουμε για το σφάλμα της ημιδιακριτής μεθόδου (2.40) ότι

$$\|u_h(t) - u(t)\| \leq Ch^2 t^{-1} \|v\|, \text{ για } t \in J. \quad (2.54)$$

*Απόδειξη.* Χρησιμοποιώντας τον συμβολισμό  $e = \tilde{u}_h - \tilde{u} = e^{-\kappa t}(u_h - u)$  για το σφάλμα, έχουμε το εξής

$$T_h e_t + e = (T_h \tilde{u}_{h,t} + \tilde{u}_h) - (T_h \tilde{u}_t + \tilde{u}) = -(T \tilde{u}_t + \tilde{u}) + (T - T_h) \tilde{u}_t$$

Αφού ισχύει η ΜΔΕ  $T \tilde{u}_t + \tilde{u} = 0$ , τότε έχουμε ότι

$$T_h e_t + e = \rho = -(T_h - T) \tilde{u}_t, \text{ για } t \in J. \quad (2.55)$$

Παρατηρούμε επίσης ότι  $v_h = P_h v$  και  $e(0) = P_h v - v$  είναι ορθογώνιο στον  $S_h$ , αφού χρησιμοποιώντας την συνθήκη (iii) του Θεωρήματος 2.1, παίρνουμε ότι

$$\begin{aligned} |(T_h e(0), \chi)| &= |(T_h e(0), \chi) - (e(0), T_h \chi)| \\ &\leq C(T_h e(0), e(0))^{1/2} \|T_h \chi\| = 0, \quad \forall \chi \in S_h. \end{aligned}$$

Τώρα θα αποδείξουμε την σχέση

$$\|e(t)\| \leq Ct^{-1} \sup_{s \leq t} (s^2 \|\rho_t\| + s \|\rho\| + \|\tilde{\rho}\| + h^2 \|e\|), \text{ για } t \in J, \quad (2.56)$$

όπου  $\tilde{\rho}(t) = \int_0^t \rho ds$ . Για να αποδείξουμε λοιπόν την παραπάνω σχέση, θέτουμε  $w = te$ . Το  $w$  ικανοποιεί την σχέση

$$T_h w_t + w = T_h(e + te_t) = t(T_h e_t + e) + T_h e$$

Χρησιμοποιώντας την (2.55) παίρνουμε

$$T_h w_t + w = \omega = t\rho + T_h e. \quad (2.57)$$

Επομένως έχουμε ότι

$$\|w_t\| \leq \|t\rho_t + \rho + T_h e_t + T_h' e\| \leq C(t\|\rho_t\| + \|\rho\| + \|e\|).$$

Από το Λήμμα 2.3, αφού  $w(0) = 0$  έχουμε για κατάλληλο  $\varepsilon$  και για  $t \in J$  ότι

$$\begin{aligned} \|w(t)\| &\leq \varepsilon \sup_{s \leq t} (s \|\omega_t(s)\|) + C_\varepsilon \sup_{s \leq t} \|\omega(s)\| \\ &\leq \frac{1}{2} \sup_{s \leq t} \|w(s)\| + C \sup_{s \leq t} (s^2 \|\rho_t\| + s \|\rho\| + \|T_h e\|). \end{aligned}$$

Συνεπώς

$$\|w(t)\| \leq C \sup_{s \leq t} (s^2 \|\rho_t\| + s \|\rho\| + \|T_h e\|). \quad (2.58)$$

Τώρα ολοκληρώνουμε την (2.48) και αφού  $T_h e(0) = 0$  έχουμε ότι

$$\begin{aligned} \int_0^t (T_h e_t + e) ds &= \int_0^t \rho ds \\ \int_0^t T_h e_t ds + \int_0^t e ds &= \tilde{\rho}(t) \\ T_h e \Big|_0^t - \int_0^t T_h' e ds + \tilde{e} &= \tilde{\rho}(t). \end{aligned}$$

Επομένως

$$T_h e + \tilde{e} \equiv T_h \tilde{e}_t + \tilde{e} = \tilde{\rho}(t) + \int_0^t T_h' e ds. \quad (2.59)$$

Αφού  $\tilde{e}(0) = 0$ , εφαρμόζοντας και πάλι το Λήμμα 2.3 παίρνουμε ότι

$$\|\tilde{e}(t)\| \leq \varepsilon \sup_{s \leq t} (s \|\tilde{\rho}_t\| + s \|T_h' e\|) + C_\varepsilon \sup_{s \leq t} \left( \|\tilde{\rho}\| + \left\| \int_0^s T_h' e dy \right\| \right).$$

Έπειτα για να εκτιμήσουμε το  $T_h' e$  χρησιμοποιούμε το Λήμμα 2.2.

$$\begin{aligned} \|\tilde{e}(t)\| &\leq \varepsilon \sup_{s \leq t} (s \|\rho\| + Cs \|T_h e\| + Csh^2 \|e\|) + C_\varepsilon \sup_{s \leq t} \|\tilde{\rho}\| \\ &\quad + C_\varepsilon \int_0^t (\|T_h e\| + h^2 \|e\|) dy. \end{aligned}$$

Συνδυάζοντας την παραπάνω ανισότητα με την σχέση (2.59) έχουμε ότι

$$\begin{aligned} \|T_h e(t)\| &\leq \|\tilde{e}\| + \|\tilde{\rho}\| + \left\| \int_0^t T_h' e ds \right\| \\ &\leq \varepsilon C \bar{t} \sup_{s \leq t} \|T_h e(s)\| + C_\varepsilon \sup_{s \leq t} (s \|\rho\| + \|\tilde{\rho}\| + h^2 \|e\|) + C_\varepsilon \int_0^t \|T_h e\| ds. \end{aligned}$$

Διαλέγοντας  $\varepsilon$  τέτοιο ώστε  $\varepsilon C \bar{t} < 1$ , τότε για  $t \in \bar{J}$  έχουμε ότι

$$\|T_h e(t)\| \leq C \sup_{s \leq t} (s \|\rho\| + \|\tilde{\rho}\| + h^2 \|e\|) + C \int_0^t \|T_h e\| ds.$$

Εφαρμόζοντας το Λήμμα του Gronwall παίρνουμε

$$\|T_h e(t)\| \leq C \sup_{s \leq t} (s \|\rho\| + \|\tilde{\rho}\| + h^2 \|e\|). \quad (2.60)$$

Για να εκτιμήσουμε τον τελευταίο όρο στο δεξί μέλος της (2.58) χρησιμοποιούμε την (2.60). Επίσης, αφού  $w = te$  τότε έχουμε την σχέση (2.56). Τώρα χρησιμοποιώντας το Λήμμα 2.4 και την συνθήκη (ii) του Θεωρήματος 2.1 έχουμε ότι

$$s \|\rho(s)\| = s \|(T_h - T)\ddot{u}_t(s)\| \leq Ch^2 s \|\ddot{u}_t(s)\| \leq Ch^2 \|v\|, \quad (2.61)$$

και

$$\begin{aligned} s^2 \|\rho_t(s)\| &\leq s^2 \|(T'_h - T')\ddot{u}_t(s)\| + s^2 \|(T_h - T)\ddot{u}_{tt}\| \\ &\leq Ch^2 s^2 (\|\ddot{u}_t(s)\| + \|\ddot{u}_{tt}(s)\|) \leq Ch^2 \|v\|. \end{aligned} \quad (2.62)$$

Επίσης αφού

$$\tilde{\rho}(t) = - \int_0^t (T_h - T)\ddot{u}_t ds = -(T_h - T)\ddot{u}(s) \Big|_0^t + \int_0^t (T'_h - T')\ddot{u} ds,$$

τότε έχουμε ότι

$$\|\tilde{\rho}(s)\| = Ch^2 \sup_{y \leq s} \|\ddot{u}(y)\| \leq Ch^2 \|v\|. \quad (2.63)$$

Γνωρίζοντας ότι οι τελεστές λύσης είναι ευσταθής παίρνουμε

$$\|e(s)\| \leq \|\ddot{u}_h(s)\| + \|\ddot{u}(s)\| \leq 2\|v\|. \quad (2.64)$$

Εφαρμόζοντας τις σχέσεις (2.61), (2.62), (2.63) και (2.64) στην (2.56) παίρνουμε ότι

$$\|e(t)\| \leq Ch^2 \|v\|.$$

Συνεπώς δείξαμε το ζητούμενο.  $\square$

## 2.3 Πλήρως διακριτό πρόβλημα

Όπως και στο πρώτο κεφάλαιο, θα μελετήσουμε την πλήρως διακριτή εκδοχή του γενικού παραβολικού προβλήματος (2.1). Έστω λοιπόν  $S_h$  ένας πεπερασμένης διάστασης υπόχωρος του  $H_0^1$ . Αρχικά γράφουμε το πρόβλημα μας στην κλασσική ασθενή μορφή.

Να βρεθεί  $u_h : J \rightarrow S_h$  τέτοιο ώστε

$$(u_{h,t}, \chi) + A(t; u_h, \chi) = (f, \chi), \quad \forall \chi \in S_h, \quad t \in J, \quad (2.65)$$

με  $u_h(0) = v_h$ , όπου  $v_h$  μια προσέγγιση της  $v$  στον  $S_h$ . Έστω τώρα το χρονικό βήμα  $k$  έτσι ώστε  $t = t_n = nk$ , όπου  $n$  μη αρνητικός ακέραιος. Όπως και στο προηγούμενο κεφάλαιο, θα συμβολίζουμε με  $U^n = U_h^n \in S_h$  την προσέγγιση της ακριβούς λύσης  $u(t_n)$  την οποία θέλουμε να υπολογίσουμε.

Εδώ θα ασχοληθούμε με την έμμεση *Euler Galerkin* μέθοδο, η οποία ορίζεται αντικαθιστώντας την χρονική παράγωγο στην (2.65) με την οπισθόδρομη διαφορά την οποία ορίζουμε ως  $\partial U^n = (U^n - U^{n-1})/k$ . Επίσης θέτουμε αυθαίρετα ότι ο ελλειπτικός τελεστής θα εξαρτάται σε κάθε χρονικό βήμα από την χρονική στιγμή  $t_n$

$$(\bar{\partial} U^n, \chi) + A(t_n; U^n, \chi) = (f(t_n), \chi), \quad \forall \chi \in S_h, \quad n \geq 1, \quad (2.66)$$

όπου  $U^0 = v_h$ . Άρα όταν το  $U^{n-1}$  είναι γνωστό, το  $U^n$  υπολογίζεται έμμεσα από την σχέση

$$(U^n, \chi) + kA(t_n; U^n, \chi) = (U^{n-1} + kf(t_n), \chi), \quad \forall \chi \in S_h. \quad (2.67)$$

Για λόγους συντομίας από εδώ και στο εξής θα παραλείπουμε την μεταβλητή  $t$  από τον ελλειπτικό τελεστή. Δηλαδή αντί για  $A(t; v, w)$  θα γράφουμε  $A(v, w)$ . Αφού η παραπάνω μέθοδος ισχύει για κάθε  $\chi \in S_h$  και αφού  $U^n \in S_h$ , τότε παρόμοια με προηγουμένως η παραπάνω μέθοδος γράφεται ισοδύναμα ως εξής

Για κάθε  $n = 1, \dots, N_h$  να βρεθούν οι συντελεστές  $\alpha_j^n \in \mathbb{R}$  της  $U^n = \sum_{j=1}^{N_h} \alpha_j^n \Phi_j$  έτσι ώστε

$$\sum_{j=1}^{N_h} \alpha_j^n (\Phi_j, \Phi_i) + k \sum_{j=1}^{N_h} \alpha_j^n A(\Phi_j, \Phi_i) = \sum_{j=1}^{N_h} \alpha_j^{n-1} (\Phi_j, \Phi_i) + k(f(t_n), \Phi_i), \quad (2.68)$$

για  $i = 1, \dots, N_h$ . Η παραπάνω μέθοδος γράφεται σε μορφή πινάκων ως ακολούθως

$$(\mathcal{B} + k\mathcal{A})\alpha^n = \mathcal{B}\alpha^{n-1} + k\tilde{f}(t_n), \quad \text{για } t_n \in J, \quad (2.69)$$

όπου ο  $\mathcal{B} = (b_{ji})$  με στοιχεία  $b_{ji} = (\Phi_j, \Phi_i)$  καλείται *πίνακας μάζας*, ο  $\mathcal{A} = (a_{ji})$  με στοιχεία  $a_{ji} = A(\Phi_j, \Phi_i)$  καλείται *πίνακας ακαμψίας*, και με  $\tilde{f}(t_n) = (f_i)$  συμβολίζουμε το διάνυσμα με στοιχεία  $(f(t_n), \Phi_i)$ . Επίσης συμβολίζουμε με  $\alpha^n$  το διάνυσμα με στοιχεία τους άγνωστους συντελεστές  $\alpha_j^n$ . Όπως είπαμε παραπάνω μπορεί ο πίνακας  $\mathcal{A}$  να μην είναι απαραίτητα θετικά ορισμένος, αλλά ισχύει η ανισότητα Gårding's.

Όπως αναφέραμε και στην προηγούμενη ενότητα το χρονικά εξαρτούμενο πρόβλημα Dirichlet που αντιστοιχεί στο πρόβλημα (2.1) ορίζεται ως εξής

$$A_\kappa(t)u := A(t)u + \kappa u = f \text{ στο } \Omega, \text{ με } u = 0 \text{ στο } \partial\Omega, \text{ για } \bar{t} \in J, \quad (2.70)$$

το οποίο ακολουθώντας την γνωστή μας πλέον διαδικασία γράφεται σε ασθενή μορφή ως εξής

$$A_\kappa(t; u, \varphi) := A(t; u, \varphi)u + \kappa(u, \varphi) = (f, \varphi), \quad \forall \varphi \in H_0^1, \text{ για } t \in J. \quad (2.71)$$

Θέτοντας στην παραπάνω ασθενή μορφή  $\varphi = \chi$  το οποίο να ανήκει στον πεπερασμένης διάστασης υπόχωρο  $S_h$  του  $H_0^1$  και ψάχνοντας προσεγγιστική λύση σε αυτό τον υπόχωρο παίρνουμε την εξής ΜΠΣ

$$A_\kappa(t; u, \chi) = (f, \chi), \quad \forall \chi \in S_h, \text{ για } t \in J. \quad (2.72)$$

Τώρα ορίζοντας  $\check{u} = e^{-t\kappa}u$  ως την νέα εξαρτημένη μεταβλητή της (2.1) παίρνουμε

$$\begin{aligned} \check{u}_t + A_\kappa(t)\check{u} &= \check{f}, \text{ στο } \Omega, \text{ για } t \in J, \\ \check{u} &= 0 \text{ στο } \partial\Omega, \text{ για } t \in J, \quad \check{u}(0) = v \text{ στο } \Omega, \end{aligned} \quad (2.73)$$

όπου  $\check{f} = e^{-t\kappa}f$ . Ισοδύναμα σε ασθενή μορφή έχουμε

$$(\check{u}_t, \varphi) + A_\kappa(t; \check{u}, \varphi) = (\check{f}, \varphi), \quad \forall \varphi \in H_0^1, \text{ για } t \in J. \quad (2.74)$$

Με ακριβώς τον ίδιο τρόπο που χρησιμοποιήσαμε για την κατασκευή της (2.66), κατασκευάζουμε και την έμμεση Euler Galerkin μέθοδο για το πρόβλημα (2.73).

$$(\partial \check{U}^n, \chi) + A_\kappa(t_n; \check{U}^n, \chi) = (\check{f}(t_n), \chi), \quad \forall \chi \in S_h, \quad n \geq 1, \quad (2.75)$$

όπου  $\check{U}^n = e^{-t\kappa}u$  και συνεπώς  $\check{U}^0 = e^{-t\kappa}v_h$ . Επισημαίνουμε ότι ο ελλειπτικός τελεστής υπολογίζεται στην χρονική στιγμή  $t = t_n$  όταν θέλουμε να προσεγγίσουμε την  $u(t_n)$ . Ο παρακάτω ορισμός θα μας βοηθήσει στην απόδειξη της εκτίμησης σφάλματος της έμμεσης Euler Galerkin μεθόδου που μόλις αναφέραμε.

**Ορισμός 2.2.** Καλούμε *προβολή Ritz* την απεικόνιση  $R_h : H_0^1 \rightarrow S_h \times J$ , η οποία ορίζεται από την σχέση

$$A_\kappa(R_h u, \chi) = A_\kappa(u, \chi), \quad \forall \chi \in S_h, \quad (2.76)$$

όπου  $A_\kappa$  είναι ο χωροχρονικά εξαρτούμενος ελλειπτικός τελεστής που ορίζεται από την σχέση (2.11).



Το Θεώρημα που ακολουθεί μας δίνει μια εκτίμηση σφάλματος της παραπάνω μεθόδου.

**Θεώρημα 2.4.** Έστω  $U^n$  και  $u$  οι λύσεις των (2.66) και (2.1) αντίστοιχα. Έστω επίσης  $\check{U}^n$  και  $\check{u}$  οι λύσεις των (2.75) και (2.73) αντίστοιχα, και έστω ότι ισχύει η συνθήκη (ii) του Θεωρήματος 2.1. Έαν ισχύει ότι  $\|v_h - v\| \leq Ch^r \|v\|_r$  και αν  $v = 0$  στο  $\partial\Omega$ , τότε

$$\|U^n - u(t_n)\| \leq Ch^r \left( \|v\|_r + \int_0^{t_n} \|u_t\|_r ds \right) + k \int_0^{t_n} \|u_{tt}\| ds, \quad \text{για } n \geq 0. \quad (2.77)$$

*Απόδειξη.* Αρχικά γράφουμε το σφάλμα  $\check{e} = \check{U}^n - \check{u}(t_n)$  ως το άθροισμα

$$e = (\check{U}^n - R_h \check{u}(t_n)) + (R_h \check{u}(t_n) - \check{u}(t_n)) = \theta^n + \rho^n. \quad (2.78)$$

Ο όρος  $\rho^n = \rho(t_n)$  φράσσεται χρησιμοποιώντας την συνθήκη (ii) του Θεωρήματος 2.1 ως εξής

$$\begin{aligned} \|\rho^n\| &\leq Ch^r \|\check{u}(t_n)\|_r \leq Ch^r \left( \|v\|_r + \int_0^t \|\check{u}_t\|_r ds \right) \\ &\leq Ch^r \left( \|v\|_r + \int_0^t \|u_t\|_r ds \right). \end{aligned} \quad (2.79)$$

Άρα μας απομένει να φράξουμε τον όρο  $\theta^n$ . Για να το κάνουμε αυτό παρόμοια με το Θεώρημα 1.8 γράφουμε

$$\begin{aligned} (\bar{\partial}\theta^n, \chi) + A_\kappa(\theta^n, \chi) &= (\bar{\partial}\check{U}^n, \chi) + A_\kappa(\check{U}^n, \chi) - (\bar{\partial}R_h \check{u}(t_n), \chi) - A_\kappa(R_h \check{u}(t_n), \chi) \\ &= (\check{f}(t_n), \chi) - (\bar{\partial}R_h \check{u}(t_n), \chi) - A_\kappa(\check{u}(t_n), \chi) \\ &= (\check{u}_t(t_n), \chi) - (\bar{\partial}R_h \check{u}(t_n), \chi) = (\check{u}_t(t_n) - \bar{\partial}R_h \check{u}(t_n), \chi). \end{aligned}$$

Συνοψίζοντας έχουμε ότι

$$(\bar{\partial}\theta^n, \chi) + A_\kappa(\theta^n, \chi) = -(\omega^n, \chi), \quad \forall \chi \in S_h, \quad n \geq 1, \quad (2.80)$$

όπου

$$\omega^n := R_h \bar{\partial}\check{u}(t_n) - \check{u}_t(t_n) = (R_h - I)\bar{\partial}\check{u}(t_n) + (\bar{\partial}\check{u}(t_n) - \check{u}_t(t_n)) = \omega_1^n + \omega_2^n. \quad (2.81)$$

Θέτοντας  $\chi = \theta^n$  στην (2.80) παίρνουμε

$$(\bar{\partial}\theta^n, \theta^n) + A_\kappa(\theta^n, \theta^n) = -(\omega^n, \theta^n). \quad (2.82)$$

Χρησιμοποιώντας την ανισότητα Gårding's και επιλέγοντας κατάλληλο  $\kappa$ , μπορούμε να θεωρήσουμε ότι ο όρος  $A_\kappa(\theta^n, \theta^n)$  είναι μη αρνητικός. Συνεπώς μπορούμε απλά να το πετάξουμε και έτσι να έχουμε

$$(\bar{\partial}\theta^n, \theta^n) \leq \|\omega^n\| \|\theta^n\|. \quad (2.83)$$

Ισοδύναμα

$$\|\theta^n\|^2 \leq (\theta^{n-1}, \theta^n) + k\|\omega^n\| \|\theta^n\| \leq \|\theta^{n-1}\| \|\theta^n\| + k\|\omega^n\| \|\theta^n\|. \quad (2.84)$$

Διαιρώντας και τις δύο πλευρές με  $\|\theta^n\|$ , παίρνουμε

$$\|\theta^n\| \leq \|\theta^{n-1}\| + k\|\omega^n\|. \quad (2.85)$$

Στην περίπτωση όπου  $\|\theta^n\| = 0$  δεν χρειάζεται να εκτιμήσουμε το  $\theta^n$  για τους ίδιους λόγους που αναφέραμε και στο προηγούμενο κεφάλαιο. Εφαρμόζοντας την παραπάνω σχέση επαναληπτικά παίρνουμε

$$\|\theta^n\| \leq \|\theta^0\| + k \sum_{j=1}^n \|\omega^j\| \leq \|\theta^0\| + k \sum_{j=1}^n \|\omega_1^j\| + k \sum_{j=1}^n \|\omega_2^j\|. \quad (2.86)$$

Για να φράξουμε τον όρο  $\theta(0)$  χρησιμοποιούμε την συνθήκη (ii) του θεωρήματος 2.1 και έτσι παίρνουμε

$$\|\theta(0)\| = \|v_h - R_h\| \leq \|v_h - v\| + \|R_h v - v\| \leq \|v_h - v\| + Ch^r \|v\|_r. \quad (2.87)$$

Συνεπώς

$$\|\theta(0)\| \leq Ch^r \|v\|_r. \quad (2.88)$$

Άρα μας μένει να φράξουμε τους δύο τελευταίους όρους. Αρχικά γράφουμε

$$\omega_1^j = (R_h - I)k^{-1} \int_{t_{j-1}}^{t_j} \ddot{u}_t ds = k^{-1} \int_{t_{j-1}}^{t_j} (R_h - I)\ddot{u}_t ds. \quad (2.89)$$

Χρησιμοποιώντας και πάλι την συνθήκη (ii) του Θεωρήματος 2.1 παίρνουμε

$$k \sum_{j=1}^n \|\omega_1^j\| \leq \sum_{j=1}^n \int_{t_{j-1}}^{t_j} Ch^r \|\ddot{u}_t\|_r ds = Ch^r \int_0^{t_n} \|\ddot{u}_t\|_r ds \leq Ch^r \int_0^{t_n} \|u_t\|_r ds. \quad (2.90)$$

Για τον τελευταίο όρο έχουμε

$$k\omega_2^j = \ddot{u}(t_j) - \ddot{u}(t_{j-1}) - k\ddot{u}_t(t_j) = - \int_{t_{j-1}}^{t_j} (s - t_{j-1})\ddot{u}_{tt}(s) ds. \quad (2.91)$$

Την παραπάνω σχέση την δείξαμε στην απόδειξη του Θεωρήματος 1.8. Έπειτα έχουμε ότι

$$k \sum_{j=1}^n \|\omega_2^j\| \leq \sum_{j=1}^n \left\| \int_{t_{j-1}}^{t_j} (s - t_{j-1}) \check{u}_{tt}(s) ds \right\| \leq k \int_0^{t_n} \|\check{u}_{tt}\| ds \leq k \int_0^{t_n} \|u_{tt}\| ds. \quad (2.92)$$

Συνοψίζοντας τα φράγματα (2.88), (2.90) και (2.92) έχουμε

$$\|\theta^n\| \leq Ch^r \left( \|v\|_r + \int_0^{t_n} \|u_t\|_r ds \right) + k \int_0^{t_n} \|u_{tt}\| ds, \text{ για } n \geq 0. \quad (2.93)$$

Συνδυάζοντας τις σχέσεις (2.79) και (2.93) και ότι  $\|U^n - u(t_n)\| \leq \|e\|$  παίρνουμε το ζητούμενο.  $\square$

## Κεφάλαιο 3

# Σχεδόν γραμμικές παραβολικές εξισώσεις

### 3.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα μελετήσουμε την ανάλυση σφάλματος για ένα γενικό σχεδόν γραμμικό παραβολικό πρόβλημα. Θα ασχοληθούμε κυρίως με ΜΠΣ στις οποίες ο χώρος των υποψήφιων λύσεων είναι ο πεπερασμένης διάστασης υπόχωρος  $S_h$  του  $H_0^1 = H_0^1(\Omega)$  σε κυρτό χωρίο  $\Omega \subset \mathbb{R}^d$ , για  $d \geq 1$  ένα θετικό ακέραιο. Να σημειώσουμε επίσης ότι στο πρόβλημα αρχικών συνοριακών τιμών που θα μελετήσουμε η χρονική μεταβλητή ορίζεται σε ένα φραγμένο διάστημα  $J = (0, t]$ . Αρχικά θα ξεκαθαρίσουμε τι εννοούμε με τον όρο σχεδόν γραμμική ΜΔΕ.

**Ορισμός 3.1.** Έστω  $X \subset \mathbb{R}^d$ , ένα ανοικτό υποσύνολο του  $\mathbb{R}^d$ , για  $d > 1$  ένα θετικό ακέραιο. Έστω επίσης  $x = (x_1, x_2, \dots, x_d) \in X$  οι ανεξάρτητες μεταβλητές, και  $u : X \rightarrow \mathbb{R}$  η εξαρτημένη μεταβλητή μιας ΜΔΕ. Αυτή η ΜΔΕ καλείται *γραμμική* αν μπορεί να γραφτεί στη μορφή

$$\begin{aligned} a(x)u + b_1(x)u_{x_1} + b_2(x)u_{x_2} + \dots + b_d(x)u_{x_d} \\ + c_{11}(x)u_{x_1x_1} + c_{12}(x)u_{x_1x_2} + \dots + c_{dd}(x)u_{x_dx_d} + \dots = f(x). \end{aligned} \quad (3.1)$$

Δηλαδή, αν οι συντελεστές της άγνωστης συνάρτησης  $u$  και όλων των μερικών της παραγώγων εξαρτώνται μόνο από τις ανεξάρτητες μεταβλητές  $x = (x_1, x_2, \dots, x_d)$ , τότε η ΜΔΕ καλείται *γραμμική*. Αν δεν μπορεί η ΜΔΕ να γραφτεί στη μορφή (3.1) τότε καλείται *μη γραμμική*.

**Ορισμός 3.2.** Έστω μια μη γραμμική ΜΔΕ  $k$  τάξης με άγνωστη λύση  $u$ .

- Αν οι συντελεστές των μερικών παραγώγων  $k$  τάξης της  $u$  είναι συναρτήσεις μόνο των ανεξάρτητων μεταβλητών  $x = (x_1, x_2, \dots, x_d)$ , τότε η ΜΔΕ καλείται *ημι-γραμμική*.
- Αν οι συντελεστές των μερικών παραγώγων  $k$  τάξης της  $u$  είναι συναρτήσεις των ανεξάρτητων μεταβλητών  $x = (x_1, x_2, \dots, x_d)$  και/ή μερικών παραγώγων της  $u$  μέχρι  $k - 1$  τάξης (συμπεριλαμβανομένης και της  $u$ ), τότε η ΜΔΕ καλείται *σχεδόν γραμμική*.
- Αν μια μη γραμμική ΜΔΕ δεν είναι σχεδόν γραμμική, τότε καλείται *πλήρως μη γραμμική*.

Να σημειώσουμε ότι μια ημι-γραμμική ΜΔΕ είναι προφανώς και σχεδόν γραμμική. Επίσης θα χρειαστούμε για την ανάλυση μας τον παρακάτω ορισμο.

**Ορισμός 3.3.** Μια συνάρτηση  $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$  καλείται *Lipschitz συνεχής* αν υπάρχει σταθερά  $L \geq 0$ ,  $L \in \mathbb{R}$ , τέτοια ώστε

$$\|f(x) - f(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathbb{R}^d. \quad (3.2)$$

Η μικρότερη σταθερά  $L \geq 0$  που ικανοποιεί την σχέση (3.2) καλείται *σταθερά Lipschitz*.

Έστω λοιπόν  $\Omega \subset \mathbb{R}^d$ , με  $d \geq 1$  ένα θετικό ακέραιο, ένα κυρτό χωρίο με ομαλό σύνορο. Το παραβολικό πρόβλημα αρχικών συνοριακών τιμών που θα μας απασχολήσει σε αυτό το κεφάλαιο είναι το εξής

$$\begin{aligned} u_t - \nabla \cdot (a(u)\nabla u) &= f(u) \text{ στο } \Omega, \quad t \in J, \\ u &= 0 \text{ στο } \partial\Omega, \quad t \in J, \\ u(\cdot, 0) &= v \text{ στο } \Omega, \end{aligned} \quad (3.3)$$

όπου  $a$  και  $f$  είναι συναρτήσεις Lipschitz ως προς  $u$  και ομαλές στο  $\mathbb{R}$  τέτοιες ώστε

$$0 < \mu \leq a(u) \leq M, \quad |a'(u)| + |f'(u)| \leq B, \quad \forall u \in \mathbb{R}. \quad (3.4)$$

Υποθέτουμε ότι το παραπάνω πρόβλημα έχει μοναδική λύση η οποία είναι αρχούντως ομαλή.

### 3.2 Κατασκευή ημιδιακριτού προβλήματος

Έστω  $\mathcal{T}_h$  μια σχεδόν ομοιόμορφη οικογένεια τριγωνοποιήσεων του  $\Omega$  με  $\max_{\tau \in \mathcal{T}_h} \text{diam } \tau \leq h$  και έστω  $S_h$  ο πεπερασμένης διάστασης υπόχωρος του  $H_0^1$ , ο οποίος ορίζεται ως εξής

$$S_h := \{\chi \in H_0^1 \mid \chi \in C(\Omega), \text{ γραμμικό σε κάθε } \tau \in \mathcal{T}_h, \text{ και } \chi|_{\partial\Omega} = 0\}. \quad (3.5)$$

Πολλαπλασιάζοντας την ΜΔΕ με μια συνάρτηση ελέγχου  $\varphi \in S_h$ , ολοκληρώνοντας στο  $\Omega$  και χρησιμοποιώντας την φόρμουλα του Green παίρνουμε την ασθενή μορφή του προβλήματος η οποία είναι η εξής

$$(u_t, \varphi) + (a(u)\nabla u, \nabla \varphi) = (f(u), \varphi), \quad \forall \varphi \in H_0^1, t \in J, \quad (3.6)$$

με  $u(\cdot, 0) = v$ . Θέτοντας  $\varphi = \chi \in S_h$  και περιορίζοντας τον χώρο των υποψήφιων λύσεων, ψάχνοντας την προσεγγιστική λύση στον  $S_h$  παίρνουμε το εξής ημιδιακριτό πρόβλημα

Να βρεθεί  $u_h : J \rightarrow S_h$  τέτοιο ώστε

$$(u_{h,t}, \chi) + (a(u_h)\nabla u_h, \nabla \chi) = (f(u_h), \chi), \quad \forall \chi \in S_h, t \in J, \quad (3.7)$$

με  $u_h(\cdot, 0) = v_h$ , όπου  $v_h$  είναι μια προσέγγιση της  $v$  στον  $S_h$ . Αφού η σχέση (3.7) ισχύει για κάθε  $\chi \in S_h$  τότε ισχύει και για κάθε συνάρτηση πυραμίδας  $\Phi_j$  στο σύνολο  $\{\Phi_j\}_{j=1}^{N_h}$ , το οποίο είναι βάση του  $S_h$ . Επίσης αφού  $u_h \in S_h$  τότε μπορεί να γραφτεί σαν γραμμικός συνδυασμός των συναρτήσεων βάσης  $\{\Phi_j\}_{j=1}^{N_h}$  ως  $u_h(x, t) = \sum_{j=1}^{N_h} \alpha_j(t)\Phi_j(x)$ . Έτσι το πρόβλημα (3.7) γράφεται ισοδύναμα

$$\begin{aligned} \sum_{j=1}^{N_h} \alpha_j'(t)(\Phi_j, \Phi_k) + \sum_{j=1}^{N_h} \alpha_j(t) \left( a \left( \sum_{l=1}^{N_h} \alpha_l(t)\Phi_l \right) \nabla \Phi_j, \nabla \Phi_k \right) \\ = \left( f \left( \sum_{l=1}^{N_h} \alpha_l(t)\Phi_l \right), \Phi_k \right), \quad \text{για } k = 1, \dots, N_h, \end{aligned} \quad (3.8)$$

όπου  $\alpha = \alpha(t) = (\alpha_1(t), \dots, \alpha_{N_h}(t))^T$  το διάνυσμα των αγνώστων συντελεστών,  $\mathcal{B} = (b_{jk})$  και  $\mathcal{A}(\alpha) = (a_{jk}(\alpha))$  οι πίνακες με στοιχεία

$$b_{jk} = (\Phi_j, \Phi_k) \text{ και } a_{jk}(\alpha) = \left( a \left( \sum_{l=1}^{N_h} \alpha_l(t)\Phi_l \right) \nabla \Phi_j, \nabla \Phi_k \right),$$

αντίστοιχα, και το διάνυσμα  $\tilde{f}(\alpha) = (\tilde{f}_1(\alpha), \dots, \tilde{f}_{N_h}(\alpha))^T$  με στοιχεία  $\tilde{f}_k(\alpha) = (f(\sum_{l=1}^{N_h} \alpha_l \Phi_l), \Phi_k)$ . Συνεπώς το σύστημα (3.8) γράφεται σε μορφή πινάκων ισοδύναμα ως

$$\mathcal{B}\alpha' + \mathcal{A}(\alpha)\alpha = \tilde{f}(\alpha), \text{ για } t \in J, \text{ με } \alpha(0) = \gamma, \quad (3.9)$$

όπου  $\gamma$  είναι το διάνυσμα με τις τιμές που παίρνει στους εσωτερικούς κόμβους του πλέγματος μας η αρχική συνθήκη  $v_h$ . Οι πίνακες  $\mathcal{B}$  και  $\mathcal{A}(\alpha)$  είναι θετικά ορισμένοι. Επίσης ισοδύναμα με τις υποθέσεις για τον μη γραμμικό όρο  $a(u)$  και το μη γραμμικό δεξί μέλος  $f(u)$  (3.4), τα  $\mathcal{A}(\alpha)$  και  $\tilde{f}(\alpha)$  είναι Lipschitz συνεχής στον  $\mathbb{R}^{N_h}$ . Μπορεί να αποδειχθεί λοιπόν ότι το σύστημα (3.9) έχει μοναδική λύση για  $t \in J$ , η οποία είναι φραγμένη. Αυτή την προσεγγιστική λύση την βρίσκουμε προσδιορίζοντας τα  $\alpha_n = \alpha_n(t)$  για  $n = 0, 1, \dots$ , χρησιμοποιώντας την επαναληπτική μέθοδο

$$\begin{aligned} \mathcal{B}\alpha'_{n+1} + \mathcal{A}(\alpha_n)\alpha_{n+1} &= \tilde{f}(\alpha_n), \text{ για } t \in J, \text{ για } n \geq 0, \\ \alpha_{n+1}(0) &= \gamma, \text{ για } n \geq 0, \\ \alpha_0(t) &\equiv \gamma \text{ στο } \bar{J}. \end{aligned} \quad (3.10)$$

### 3.3 Ανάλυση σφάλματος ημιδιακριτού προβλήματος

Στόχος μας είναι να εκτιμήσουμε το σφάλμα του ημιδιακριτού προβλήματος (3.7). Θα χρησιμοποιήσουμε την ίδια τεχνική που χρησιμοποιήσαμε στα προηγούμενα δύο κεφάλαια. Δηλαδή θα γράψουμε το σφάλμα ως το άθροισμα δύο όρων ως ακολούθως

$$u_h - u = (u_h - w_h) + (w_h - u) = \theta + \rho, \quad (3.11)$$

όπου  $w_h$  είναι μια ελλειπτική προβολή της ακριβούς λύσης  $u$  στον πεπερασμένης διάστασης χώρο  $S_h$ . Έδω θα χρησιμοποιήσουμε την προβολή  $w_h = w_h(t)$  που ορίζεται από την σχέση

$$(a(u(t))\nabla(w_h(t) - u(t)), \nabla\chi) = 0, \quad \forall \chi \in S_h, \quad t \geq 0. \quad (3.12)$$

Παρατηρούμε ότι το εσωτερικό γινόμενο που ορίζει την  $w_h$  εξαρτάται από την ακριβή λύση  $u$ . Θα χρειαστούμε για την ανάλυση σφάλματος της ημιδιακριτής μας μεθόδου μια εκτίμηση σφάλματος για την προβολή  $w_h$ .

**Λήμμα 3.1.** Έστω  $b = b(x)$  μια ομαλή συνάρτηση στο  $\Omega$  με  $0 < \mu \leq b(x) \leq M$  για  $x \in \Omega$ . Υποθέτουμε ότι  $u \in H^2 \cap H_0^1$  και έστω  $w_h$  που ορίζεται από την σχέση

$$(b\nabla(w_h - u), \nabla\chi) = 0, \quad \forall \chi \in S_h.$$

Τότε

$$\|\nabla(w_h - u)\| \leq C_1 h \|u\|_2 \quad (3.13)$$

και

$$\|w_h - u\| \leq C_0 h^2 \|u\|_2, \quad (3.14)$$

όπου το  $C_1$  εξαρτάται από την τριγωνοποίηση  $\mathcal{T}_h$ , και από τα  $\mu$  και  $M$ . Το  $C_0$  εξαρτάται από ένα άνω φράγμα του  $\nabla b$ .

*Απόδειξη.* Χρησιμοποιώντας τον ορισμό του  $w_h$  έχουμε για  $\chi \in S_h$  ότι

$$\begin{aligned} \mu \|\nabla(w_h - u)\|^2 &\leq (b\nabla(w_h - u), \nabla(w_h - u)) \\ &= (b\nabla(w_h - u), \nabla(\chi - u)) \\ &\leq M \|\nabla(w_h - u)\| \|\nabla(\chi - u)\|, \end{aligned}$$

Θέτοντας  $\chi = I_h u$ , όπου  $I_h u$  η παρεμβάλλουσα της  $u$  στον  $S_h$  και χρησιμοποιώντας την σχέση (1.23) παίρνουμε

$$\|\nabla(w_h - u)\| \leq \frac{M}{\mu} \|\nabla(I_h u - u)\| \leq C_1 h \|u\|_2,$$

η οποία είναι η (3.13). Για να δείξουμε την (3.14) χρησιμοποιούμε το δυϊκό πρόβλημα

$$\begin{aligned} -\nabla \cdot (b\nabla\psi) &\equiv -b\Delta\psi - \nabla b \cdot \nabla\psi = \varphi, \quad \text{στο } \Omega, \\ \psi &= 0, \quad \text{στο } \partial\Omega. \end{aligned} \quad (3.15)$$

Αφού  $\|\psi\| \leq C \|\nabla\psi\| \quad \forall \psi \in H_0^1$ , συμπεραίνουμε ότι

$$\mu \|\nabla\psi\|^2 \leq (b\nabla\psi, \nabla\psi) = (\varphi, \psi) \leq \|\varphi\| \|\psi\| \leq C \|\varphi\| \|\nabla\psi\|.$$

Συνεπώς

$$\|\nabla\psi\| \leq C \|\varphi\|. \quad (3.16)$$

Επομένως για  $\nabla b$  φραγμένο έχουμε ότι

$$\|\psi\|_2 \leq C \|\Delta\psi\| \leq C \|b\Delta\psi\| = C \|\varphi + \nabla b \cdot \nabla\psi\| \leq C \|\varphi\|. \quad (3.17)$$



Επίσης έχουμε χρησιμοποιώντας και πάλι τον ορισμό του  $w_h$  ότι

$$\begin{aligned} (w_h - u, \varphi) &= (b\nabla(w_h - u), \nabla\psi) \\ &= (b\nabla(w_h - u), \nabla(\psi - \chi)) \\ &\leq M\|\nabla(w_h - u)\|\|\nabla(\psi - \chi)\|. \end{aligned}$$

Θέτοντας  $\chi = I_h\psi$  και χρησιμοποιώντας τις σχέσεις (1.23), (3.13) και (3.17) παίρνουμε

$$(w_h - u, \varphi) \leq (Ch\|u\|_2)(Ch\|\psi\|_2) \leq C_0h^2\|u\|_2\|\varphi\|.$$

Τέλος θέτουμε  $\varphi = w_h - u$  και διαιρώντας και τις δύο πλευρές με  $\|w_h - u\|$  παίρνουμε την ζητούμενη ανισότητα (3.14).  $\square$

Τώρα είμαστε σε θέση να αποδείξουμε εύκολα το παρακάτω αποτέλεσμα το οποίο είναι ένα φράγμα για το σφάλμα της ελλειπτικής προβολής  $w_h$  που ορίζεται από την σχέση (3.12), υπό τις κατάλληλες υποθέσεις ομαλότητας για την  $u$ . Από εδώ και στο εξής, για λόγους συντομίας δεν θα αναφέρουμε την εξάρτηση των σταθερών στα φράγματα από την ομαλότητα της ακριβούς λύσης  $u$ .

**Λήμμα 3.2.** Έστω  $w_h$  η ελλειπτική προβολή που ορίζεται από την σχέση (3.12) και  $\rho = w_h - u$ . Τότε έχουμε υπό τις κατάλληλες υποθέσεις ομαλότητας για την ακριβή λύση  $u$  ότι

$$\|\rho(t)\| + h\|\nabla\rho(t)\| \leq C(u)h^2, \text{ για } t \in J, \quad (3.18)$$

και

$$\|\rho_t(t)\| + h\|\nabla\rho_t(t)\| \leq C(u)h^2, \text{ για } t \in J, \quad (3.19)$$

όπου  $C(u)$  ανεξάρτητο του  $t$ .

*Απόδειξη.* Αφού  $\nabla a(u) = a'(u)\nabla u$  τότε η ανισότητα (3.18) προκύπτει με απευθείας εφαρμογή του Λήμματος 3.1 θέτοντας  $b(x) = a(u(x, t))$ . Για να αποδείξουμε την σχέση (3.19) παραγωγίζουμε την (3.12) και παίρνουμε

$$(a(u)\nabla\rho_t, \nabla\chi) + (a(u)_t\nabla\rho, \nabla\chi) = 0, \quad \forall \chi \in S_h. \quad (3.20)$$

Χρησιμοποιώντας ότι  $a(u)$  και  $a(u)_t$  είναι ομοιόμορφα φραγμένες, δηλαδή ότι για κάποια σταθερά  $C > 0$  ισχύουν τα φράγματα  $|a(u)| \leq C$  και  $|a(u)_t| \leq C$  έχουμε ότι

$$\begin{aligned} \mu\|\nabla\rho_t\|^2 &\leq (a(u)\nabla\rho_t, \nabla\rho_t) \\ &= (a(u)\nabla\rho_t, \nabla(\chi - u_t)) + (a(u)\nabla\rho_t, \nabla(w_{h,t} - \chi)) \\ &= (a(u)\nabla\rho_t, \nabla(\chi - u_t)) + (a(u)_t\nabla\rho, \nabla(\chi - w_{h,t})) \\ &\leq C(\|\nabla\rho_t\|\|\nabla(\chi - u_t)\| + \|\nabla\rho\|\|\nabla(\chi - w_{h,t})\|). \end{aligned}$$

Θέτοντας  $\chi = I_h u_t$  και χρησιμοποιώντας την σχέση (1.21) παίρνουμε

$$\begin{aligned} \mu \|\nabla \rho_t\|^2 &\leq Ch^2 \|u_t\|_2 \|\nabla \rho_t\| + \|\nabla \rho\| (\|\nabla(I_h u_t - u_t)\| + \|\nabla(u_t - w_{h,t})\|) \\ &\leq Ch^2 \|u_t\|_2 \|\nabla \rho_t\| + \|\nabla \rho\| (Ch \|u_t\|_2 + \|\nabla \rho_t\|). \end{aligned} \quad (3.21)$$

Για  $\alpha, \beta \geq 0$  έχουμε ότι  $\alpha^2 - 2\alpha\beta + \beta^2 = (\alpha - \beta)^2 \geq 0$  και συνεπώς

$$2\alpha\beta \leq \alpha^2 + \beta^2.$$

Για να φράξουμε περαιτέρω την σχέση (3.21) χρησιμοποιούμε την παραπάνω ανισότητα και έτσι έχουμε ότι

$$h \|u_t\|_2 \|\nabla \rho_t\| = 2 \left( \frac{h}{\sqrt{\mu}} \|u_t\|_2 \right) \left( \frac{\sqrt{\mu}}{2} \|\nabla \rho_t\| \right) \leq \frac{h^2}{\mu} \|u_t\|_2^2 + \frac{\mu}{4} \|\nabla \rho_t\|^2, \quad (3.22)$$

$$\|\nabla \rho\| h \|u_t\|_2 = 2 \left( \frac{1}{2} \|\nabla \rho\| \right) (h \|u_t\|_2) \leq \frac{1}{4} \|\nabla \rho\|^2 + h^2 \|u_t\|_2^2, \quad (3.23)$$

και

$$\|\nabla \rho\| \|\nabla \rho_t\| = 2 \left( \frac{1}{\sqrt{\mu}} \|\nabla \rho\| \right) \left( \frac{\sqrt{\mu}}{2} \|\nabla \rho_t\| \right) \leq \frac{1}{\mu} \|\nabla \rho\|^2 + \frac{\mu}{4} \|\nabla \rho_t\|^2. \quad (3.24)$$

Συνδυάζοντας τις σχέσεις (3.22)-(3.24) με την (3.21) παίρνουμε

$$\mu \|\nabla \rho_t\|^2 \leq \frac{\mu}{2} \|\nabla \rho_t\|^2 + C(\|\nabla \rho\|^2 + h^2 \|u_t\|_2^2).$$

Συνεπώς έχουμε ότι

$$\frac{\mu}{2} \|\nabla \rho_t\|^2 \leq C(\|\nabla \rho\|^2 + h^2 \|u_t\|_2^2). \quad (3.25)$$

Αφού ισχύει η σχέση (3.18) τότε έχουμε τελικά

$$\|\nabla \rho_t\| \leq C(u)h. \quad (3.26)$$

Για να αποδείξουμε την σχέση (3.19) μας απομένει λοιπόν να εκτιμήσουμε τον όρο  $\|\rho_t(t)\|$ . Θα το κάνουμε αυτό χρησιμοποιώντας το δυϊκό πρόβλημα (3.15) που χρησιμοποιήσαμε και στην απόδειξη του Λήμματος 3.1. Έστω λοιπόν η συνάρτηση  $\psi$  η οποία ικανοποιεί το πρόβλημα (3.15) με  $b = a(u)$ . Συνεπώς χρησιμοποιώντας τον ορισμό της προβολής  $w_h$  και την σχέση (3.20) παίρνουμε

$$\begin{aligned} (\rho_t, \varphi) &= (a(u)\nabla \rho_t, \nabla \psi) = (a(u)\nabla \rho_t, \nabla(\psi - \chi)) \\ &\quad + (a(u)_t \nabla \rho, \nabla(\psi - \chi)) - (\nabla \rho, a(u)_t \nabla \psi). \end{aligned}$$

Επομένως θέτοντας  $\chi = I_h\psi$  και χρησιμοποιώντας ολοκλήρωση κατά μέρη στον τελευταίο όρο παίρνουμε

$$|(\rho_t, \varphi)| \leq C(\|\nabla\rho_t\|_h\|\psi\|_2 + \|\nabla\rho\|_h\|\psi\|_2 + \|\rho\|\|\psi\|_2).$$

Χρησιμοποιώντας τα φράγματα που έχουμε ήδη αποδείξει για τους όρους  $\|\rho\|$ ,  $\|\nabla\rho\|$  και  $\nabla\rho_t$ , και την σχέση (3.17) έχουμε τελικά

$$|(\rho_t, \varphi)| \leq C(u)h^2\|\psi\|_2 \leq C(u)h^2\|\varphi\|.$$

Επομένως

$$\|\rho_t\| \leq C(u)h^2. \quad (3.27)$$

Συνδυάζοντας τις σχέσεις (3.26) και (3.27) παίρνουμε την ζητούμενη ανισότητα.  $\square$

Έστω επίσης ο χώρος  $L^\infty = L^\infty(\Omega)$  ο οποίος ορίζεται από την νόρμα

$$\|v\|_{L^\infty} = \sup_{x \in \Omega} |v(x)|. \quad (3.28)$$

Θα χρειαστούμε μια εκτίμηση για το  $\nabla w_h$ , την οποία θα αποδείξουμε στο παρακάτω Λήμμα.

**Λήμμα 3.3.** Έστω  $w_h$  το οποίο ορίζεται από την σχέση (3.12). Τότε ισχύει ότι

$$\|\nabla w_t(t)\|_{L^\infty} \leq C(u), \text{ για } t \in J. \quad (3.29)$$

*Απόδειξη.* Χρησιμοποιώντας την αντίστροφη ανισότητα (η οποία σε αυτή την περίπτωση ισχύει προφανώς αφού  $\nabla\chi$  είναι σταθερό σε κάθε στοιχείο/τριγωνάκι)

$$\|\nabla\chi\|_{L^\infty} \leq Ch^{-1}\|\nabla\chi\|, \quad \forall \chi \in S_h, \quad (3.30)$$

μαζί με το Λήμμα 3.2 και την γνωστή εκτίμηση (1.23) για την παραμβάλλουσα  $I_h u$  παίρνουμε

$$\begin{aligned} \|\nabla(w_h - I_h u)\|_{L^\infty} &\leq Ch^{-1}\|\nabla(w_h - I_h u)\| \\ &\leq Ch^{-1}(\|\nabla\rho\| + \|\nabla(I_h u - u)\|) \leq C(u). \end{aligned} \quad (3.31)$$

Γνωρίζουμε ότι ισχύει η ανισότητα  $\|\nabla I_h u\|_{L^\infty} \leq C\|\nabla u\|_{L^\infty}$ . Συνδυάζοντας την λοιπόν με την σχέση (3.31) παίρνουμε το ζητούμενο.  $\square$

Τώρα είμαστε έτοιμοι να δείξουμε την εκτίμηση του σφάλματος του ημιδιακριτού προβλήματος στην  $L^2$  νόρμα. Το παρακάτω Θεώρημα το μελετήσαμε από το βιβλίο [6] σελ.235.

**Θεώρημα 3.1.** Έστω  $u_h$  και  $u$  οι λύσεις των προβλημάτων (3.7) και (3.3) αντίστοιχα. Τότε έχουμε υπό τις κατάλληλες υποθέσεις ομαλότητας για την ακριβή λύση  $u$  ότι

$$\|u_h(t) - u(t)\| \leq C\|v_h - v\| + C(u)h^2, \text{ για } t \in \bar{J}. \quad (3.32)$$

*Απόδειξη.* Αρχικά γράφουμε το σφάλμα στην μορφή του αθροίσματος (3.11). Την εκτίμηση του όρου  $\rho$  την παίρνουμε απευθείας από το Λήμμα 3.2. Άρα μας απομένει να φράξουμε τον όρο  $\theta = u_h - w_h$ . Έχουμε λοιπόν χρησιμοποιώντας την σχέση (3.12) ότι  $\forall \chi \in S_h$

$$\begin{aligned} & (\theta_t, \chi) + (a(u_h)\nabla\theta, \nabla\chi) \\ &= (u_{h,t}, \chi) + (a(u_h)\nabla u_h, \nabla\chi) - (w_{h,t}, \chi) - (a(u_h)\nabla w_h, \nabla\chi) \\ &= (f(u_h), \chi) - (w_{h,t}, \chi) - (a(u)\nabla w_h, \nabla\chi) + ((a(u) - a(u_h))\nabla w_h, \nabla\chi) \\ &= (f(u_h), \chi) - (\rho_t, \chi) - (u_t, \chi) - (a(u)\nabla u, \nabla\chi) + ((a(u) - a(u_h))\nabla w_h, \nabla\chi) \\ &= (f(u_h) - f(u), \chi) + ((a(u) - a(u_h))\nabla w_h, \nabla\chi) - (\rho_t, \chi). \end{aligned}$$

Θέτοντας  $\chi = \theta$ , χρησιμοποιώντας το Λήμμα 3.3, και τις σχέσεις (1.8) και (3.4) παίρνουμε

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\theta\|^2 + \mu \|\nabla\theta\|^2 &\leq C(\|u_h - u\|(\|\theta\| + \|\nabla\theta\|) + \|\rho_t\|\|\theta\|) \\ &\leq C(\|\theta\|^2 + \|\rho\|\|\theta\| + \|\theta\|\|\nabla\theta\| + \|\rho\|\|\nabla\theta\|) + \|\rho_t\|\|\theta\|. \end{aligned} \quad (3.33)$$

Χρησιμοποιώντας και πάλι την ανισότητα  $2\alpha\beta \leq \alpha^2 + \beta^2$  παίρνουμε τις εξής ανισότητες

$$\|\rho\|\|\theta\| = 2 \left( \frac{1}{2} \|\rho\| \right) \|\theta\| \leq \frac{1}{4} \|\rho\|^2 + \|\theta\|^2, \quad (3.34)$$

$$\|\theta\|\|\nabla\theta\| = 2 \left( \frac{\sqrt{2}}{2\sqrt{\mu}} \|\theta\| \right) \left( \frac{\sqrt{\mu}}{\sqrt{2}} \|\nabla\theta\| \right) \leq \frac{1}{2\mu} \|\theta\|^2 + \frac{\mu}{2} \|\nabla\theta\|^2, \quad (3.35)$$

$$\|\rho\|\|\nabla\theta\| = 2 \left( \frac{\sqrt{2}}{2\sqrt{\mu}} \|\rho\| \right) \left( \frac{\sqrt{\mu}}{\sqrt{2}} \|\nabla\theta\| \right) \leq \frac{1}{2\mu} \|\rho\|^2 + \frac{\mu}{2} \|\nabla\theta\|^2, \quad (3.36)$$

και

$$\|\rho_t\|\|\theta\| = 2 \left( \frac{1}{2} \|\rho_t\| \right) \|\theta\| \leq \frac{1}{4} \|\rho_t\|^2 + \|\theta\|^2. \quad (3.37)$$

Αντικαθιστώντας τις σχέσεις (3.34)-(3.37) στην (3.33) παίρνουμε

$$\frac{1}{2} \frac{d}{dt} \|\theta\|^2 \leq C(\|\theta\|^2 + \|\rho\|^2 + \|\rho_t\|^2).$$

Ολοκληρώνουμε ως προς  $t$

$$\|\theta(t)\|^2 \leq \|\theta(0)\|^2 + C \int_0^t (\|\theta\|^2 + \|\rho\|^2 + \|\rho_t\|^2) ds,$$

και εφαρμόζοντας το Λήμμα του Gronwall παίρνουμε

$$\|\theta(t)\|^2 \leq C\|\theta(0)\|^2 + C \int_0^t (\|\rho\|^2 + \|\rho_t\|^2) ds, \quad (3.38)$$

όπου τώρα το  $C$  εξαρτάται από το  $t$ . Τώρα χρησιμοποιώντας το Λήμμα 3.1 φράζουμε τον όρο  $\|\theta(0)\|$  ως εξής

$$\|\theta(0)\| \leq \|v_h - v\| + \|w_h(0) - v\| \leq \|v_h - v\| + Ch^2\|v\|_2. \quad (3.39)$$

Συνδυάζοντας τις δυο τελευταίες σχέσεις παίρνουμε

$$\|\theta(t)\| \leq \|v_h - v\| + C(u)h^2. \quad (3.40)$$

Τέλος συνδυάζοντας τα φράγματα που αποδείξαμε για τους όρους  $\theta$  και  $\rho$  έχουμε το ζητούμενο.  $\square$

### 3.4 Κατασκευή πλήρους διακριτού προβλήματος

Θα ολοκληρώσουμε το παρόν κεφάλαιο κάνοντας την ανάλυση σφάλματος ενός πλήρους διακριτού προβλήματος που αντιστοιχεί στο πρόβλημα αρχικών συνοριακών τιμών (3.3). Συγκεκριμένα θα μελετήσουμε την έμμεση Euler Galerkin μέθοδο, η οποία κατασκευάζεται με την διαδικασία που περιγράφεται παρακάτω.

Έστω λοιπόν  $k$  το χρονικό βήμα τέτοιο ώστε  $t_n = nk$  και έστω επίσης  $U^n$  η προσέγγιση της ακριβούς λύσης  $u(t_n)$  στον  $S_h$  την οποία ψάχνουμε να βρούμε λύνοντας την ΜΠΣ. Για να κατασκευάσουμε την έμμεση Euler Galerkin μέθοδο προσεγγίζουμε την χρονική παράγωγο στο ημιδιακριτό πρόβλημα (3.7) με την

οπίσθόδρομη διαφορά  $\bar{\partial}U^n = (U^n - U^{n-1})/k$ . Έτσι η έμμεση Euler Galerkin μέθοδος του προβλήματος αρχικών/συνοριακών τιμών (3.3) είναι η εξής

$$(\bar{\partial}U^n, \chi) + (a(U^n)\nabla U^n, \nabla \chi) = (f(U^n), \chi), \quad \forall \chi \in S_h, \quad t_n \in J, \quad (3.41)$$

με  $U^0 = v_h$ . Αφου η παραπάνω μέθοδος ισχύει για κάθε  $\chi \in S_h$ , τότε ισχύει και για κάθε συνάρτηση βάσης  $\Phi_j$ . Επίσης αφού  $U^n \in S_h$  τότε μπορεί να γραφτεί σαν γραμμικός συνδυασμός των συναρτήσεων βάσης  $\Phi_j$ . Έχουμε λοιπόν ότι η παραπάνω μέθοδος γράφεται ισοδύναμα ως εξής

Για κάθε  $n = 1, \dots, N_h$ , να βρεθούν οι συντελεστές  $\alpha_j^n \in \mathbb{R}$  της  $U^n = \sum_{j=1}^{N_h} \alpha_j^n \Phi_j$  έτσι ώστε

$$\begin{aligned} \sum_{j=1}^{N_h} \alpha_j^n (\Phi_j, \Phi_i) + k \sum_{j=1}^{N_h} \alpha_j^n \left( a \left( \sum_{l=1}^{N_h} \alpha_l^n \Phi_l \right) \nabla \Phi_j, \nabla \Phi_i \right) \\ = \sum_{j=1}^{N_h} \alpha_j^n (\Phi_j, \Phi_i) + k \left( f \left( \sum_{l=1}^{N_h} \alpha_l^n \Phi_l \right), \Phi_i \right), \quad \text{για } i = 1, \dots, N_h. \end{aligned} \quad (3.42)$$

Η παραπάνω μέθοδος γράφεται σε μορφή πινάκων ως ακολούθως

$$(\mathcal{B} + k\mathcal{A}(\alpha^n))\alpha^n = \mathcal{B}\alpha^{n-1} + k\tilde{f}(\alpha^n), \quad \text{για } t_n \in J, \quad (3.43)$$

όπου οι πίνακες μάζας  $\mathcal{B} = (b_{ji})$  και ακαμψίας  $\mathcal{A} = (a_{ji})$  ορίζονται ακριβώς όπως και στην περίπτωση του ημιδιακριτού προβλήματος. Το ίδιο συμβαίνει και με το διάνυσμα  $\tilde{f}(\alpha^n) = (\tilde{f}_i)$ . Να σημειώσουμε επίσης ότι  $\alpha^0 = \gamma$  που είναι το διάνυσμα με τις τιμές που παίρνει στους εσωτερικούς κόμβους του πλέγματος μας η αρχική συνθήκη  $v_h$ .

Γνωρίζουμε ότι η παραπάνω μέθοδος έχει μοναδική λύση. Για περισσότερες λεπτομέρειες σχετικά με την ύπαρξη και μοναδικότητα της μεθόδου παραπέμπουμε στο [6].

### 3.5 Ανάλυση σφάλματος πλήρους διακριτού προβλήματος

Στο Θεώρημα που ακολουθεί θα εκτιμήσουμε το σφάλμα της έμμεσης Euler Galerkin μεθόδου που μόλις κατασκευάσαμε. Το παρακάτω Θεώρημα το μελετήσαμε από το βιβλίο [6] σελ.237-238.

**Θεώρημα 3.2.** Έστω  $U^n$  και  $u$  οι λύσεις των προβλημάτων (3.41) και (3.3) αντίστοιχα. Τότε έχουμε υπό τις κατάλληλες υποθέσεις ομαλότητας για την ακριβή λύση  $u$  ότι

$$\|U^n - u(t_n)\| \leq \|v - v_h\| + C(u)(h^2 + k), \text{ για } t_n \in \bar{J}. \quad (3.44)$$

Απόδειξη. Αρχικά γράφουμε το σφάλμα ως το άθροισμα

$$U^n - u^n = (U^n - w_h^n) + (w_h^n - u^n) = \theta^n + \rho^n, \quad (3.45)$$

όπου συμβολίζουμε με  $u^n = u(t_n)$  και  $w_h^n$  η ελλειπτική προβολή του  $u^n$  που ορίζεται από την σχέση (3.12). Ο όρος  $\rho^n$  φράσσεται απευθείας από το Λήμμα 3.2. Άρα μας απομένει μόνο να φράξουμε τον όρο  $\theta^n$ . Έχουμε λοιπόν για κάθε  $\chi \in S_h$  ότι

$$\begin{aligned} & (\bar{\partial}\theta^n, \chi) + (a(U^n)\nabla\theta^n, \nabla\chi) \\ &= (\bar{\partial}U^n, \chi) + (a(U^n)\nabla U^n, \nabla\chi) - (\bar{\partial}w_h^n, \chi) - (a(U^n)\nabla w_h^n, \nabla\chi) \\ &= (f(U^n), \chi) - (u_t^n, \chi) - (\bar{\partial}w_h^n - u_t^n, \chi) \\ &\quad - (a(u^n)\nabla w_h^n, \nabla\chi) - ((a(U^n) - a(u^n))\nabla w_h^n, \nabla\chi). \end{aligned}$$

Έπειτα χρησιμοποιώντας την σχέση (3.12) και τον ορισμό του  $\rho^n$  παίρνουμε

$$\begin{aligned} & (\bar{\partial}\theta^n, \chi) + (a(U^n)\nabla\theta^n, \nabla\chi) = (f(U^n), \chi) - (u_t^n, \chi) - (\bar{\partial}\rho^n, \chi) \\ &\quad - (\bar{\partial}u^n - u_t^n, \chi) - (a(u^n)\nabla u^n, \nabla\chi) - ((a(U^n) - a(u^n))\nabla w_h^n, \nabla\chi) \\ &= ((f(U^n) - f(u^n)), \chi) - (\bar{\partial}\rho^n, \chi) - (\bar{\partial}u^n - u_t^n, \chi) \\ &\quad - ((a(U^n) - a(u^n))\nabla w_h^n, \nabla\chi) \end{aligned}$$

Αφού οι συναρτήσεις  $a$  και  $f$  είναι Lipschitz, ισχύουν οι ανισότητες (3.4). Χρησιμοποιώντας λοιπόν αυτές τις ανισότητες καθώς και το Λήμμα 3.3, και θέτοντας  $\chi = \theta^n$  έχουμε ότι

$$\begin{aligned} & \frac{1}{2}\bar{\partial}\|\theta^n\|^2 + \mu\|\nabla\theta^n\|^2 \\ &\leq C\|U^n - u^n\|(\|\theta^n\| + \|\nabla\theta^n\|) + (\|\bar{\partial}\rho^n\| + \|\bar{\partial}u^n - u_t^n\|)\|\theta^n\| \\ &\leq C(\|\theta^n\|^2 + \|\theta^n\|\|\nabla\theta^n\| + \|\rho^n\|\|\theta^n\| + \|\rho^n\|\|\nabla\theta^n\|) \\ &\quad + \|\bar{\partial}\rho^n\|\|\theta^n\| + \|\bar{\partial}u^n - u_t^n\|\|\theta^n\|. \end{aligned} \quad (3.46)$$

Χρησιμοποιώντας την ανισότητα  $2\alpha\beta \leq \alpha^2 + \beta^2$  παίρνουμε τις εξής ανισότητες.

$$\|\theta^n\|\|\nabla\theta^n\| = 2 \left( \frac{\sqrt{2}}{2\sqrt{\mu}}\|\theta^n\| \right) \left( \frac{\sqrt{\mu}}{\sqrt{2}}\|\nabla\theta^n\| \right) \leq \frac{1}{2\mu}\|\theta^n\|^2 + \frac{\mu}{2}\|\nabla\theta^n\|^2, \quad (3.47)$$

$$\|\rho^n\| \|\theta^n\| = 2 \left( \frac{1}{2} \|\rho^n\| \right) \|\theta^n\| \leq \frac{1}{4} \|\rho^n\|^2 + \|\theta^n\|^2, \quad (3.48)$$

$$\|\rho^n\| \|\nabla \theta^n\| = 2 \left( \frac{\sqrt{2}}{2\sqrt{\mu}} \|\rho^n\| \right) \left( \frac{\sqrt{\mu}}{\sqrt{2}} \|\nabla \theta^n\| \right) \leq \frac{1}{2\mu} \|\rho^n\|^2 + \frac{\mu}{2} \|\nabla \theta^n\|^2, \quad (3.49)$$

$$\|\bar{\partial} \rho^n\| \|\theta^n\| = 2 \left( \frac{1}{2} \|\bar{\partial} \rho^n\| \right) \|\theta^n\| \leq \frac{1}{4} \|\bar{\partial} \rho^n\|^2 + \|\theta^n\|^2, \quad (3.50)$$

και

$$\|\bar{\partial} u^n - u_t^n\| \|\theta^n\| = 2 \left( \frac{1}{2} \|\bar{\partial} u^n - u_t^n\| \right) \|\theta^n\| \leq \frac{1}{4} \|\bar{\partial} u^n - u_t^n\|^2 + \|\theta^n\|^2. \quad (3.51)$$

Αντικαθιστώντας τις σχέσεις (3.47)-(3.51) στην (3.46) παίρνουμε

$$\bar{\partial} \|\theta^n\|^2 \leq C(\|\theta^n\|^n + R_n), \quad (3.52)$$

όπου

$$R_n = \|\rho^n\|^2 + \|\bar{\partial} \rho^n\|^2 + \|\bar{\partial} u^n - u_t^n\|^2.$$

Αφήνοντας στην μια πλευρά της ανισότητας την νόρμα  $\|\theta^n\|$  έχουμε

$$(1 - Ck) \|\theta^n\|^2 \leq \|\theta^{n-1}\|^2 + CkR_n. \quad (3.53)$$

Για  $k$  αρκετά μικρό γράφουμε ισοδύναμα

$$\|\theta^n\|^2 \leq (1 + Ck) \|\theta^{n-1}\|^2 + CkR_n.$$

Εφαρμόζοντας επαναληπτικά την παραπάνω ανισότητα παίρνουμε

$$\begin{aligned} \|\theta^n\|^2 &\leq (1 + Ck)^n \|\theta^{n-1}\|^2 + Ck \sum_{j=1}^n (1 + Ck)^{n-j} R_j \\ &\leq C \|\theta^0\|^2 + Ck \sum_{j=1}^n R_j, \text{ για } t_n \in J. \end{aligned} \quad (3.54)$$

Αρχικά φράζουμε τον όρο  $\|\theta^0\|$ , χρησιμοποιώντας το Λήμμα 3.1.

$$\|\theta^0\| \leq \|v_h - v\| + \|w_h^0 - v\| \leq \|v_h - v\| + Ch^2 \|v\|_2. \quad (3.55)$$

Τέλος πρέπει να φράξουμε τον όρο  $\sum_{j=1}^n R_j$ . Άρα πρέπει ουσιαστικά να φράξουμε τις ποσότητες  $\|\rho^j\|$ ,  $\|\bar{\partial} \rho^j\|$  και  $\|\bar{\partial} u^j - u_t^j\|$ . Από το Λήμμα 3.2 έχουμε απευθείας ότι

$$\|\rho^j\| \leq C(u)h^2. \quad (3.56)$$



Και πάλι χρησιμοποιώντας το Λήμμα 3.2 έχουμε ότι

$$\|\bar{\partial}\rho^j\| = \|k^{-1} \int_{t_{j-1}}^{t_j} \rho_t ds\| \leq \|k^{-1}(t_j - t_{j-1})C(u)h^2\| \leq C(u)h^2. \quad (3.57)$$

Επίσης

$$\begin{aligned} -k^{-1} \int_{t_{j-1}}^{t_j} (s - t_{j-1})u_{tt}(s) ds &= -k^{-1}(s - t_{j-1})u_t(s)|_{t_{j-1}}^{t_j} + k^{-1} \int_{t_{j-1}}^{t_j} u_t(s) ds \\ &= -u_t(t_j) + k^{-1}(u(t_j) - u(t_{j-1})) \\ &= \frac{u^j - u^{j-1}}{k} - u_t^j = \bar{\partial}u^j - u_t^j. \end{aligned}$$

Συνεπώς έχουμε ότι

$$\|\bar{\partial}u^j - u_t^j\| = \|k^{-1} \int_{t_{j-1}}^{t_j} (s - t_{j-1})u_{tt}(s) ds\| \leq C(u)k. \quad (3.58)$$

Αφού ισχύουν οι σχέσεις (3.56), (3.57) και (3.58), τότε έχουμε ότι

$$R_j \leq C(u)(h^2 + k)^2. \quad (3.59)$$

Αντικαθιστώντας τις σχέσεις (3.55) και (3.59) στην (3.54) παίρνουμε

$$\|\theta^n\| \leq C\|v_h - v\| + C(u)(h^2 + k). \quad (3.60)$$

Συνδυάζοντας τις εκτιμήσεις που βρήκαμε για τους όρους  $\rho^n$  και  $\theta^n$  παίρνουμε το ζητούμενο.  $\square$

Να σημειώσουμε εδώ ότι υπάρχουν πολλές μεθόδους που έχουν καλύτερη ακρίβεια από την έμμεση Euler Galerkin μέθοδο, όμως εμείς επιλέξαμε να μελετήσουμε αυτή θεωρώντας την, την πιο απλή έτσι ώστε να γίνουν πλήρως κατανοητές οι τεχνικές που χρησιμοποιήσαμε στην ανάλυση μας.

## Κεφάλαιο 4

# Υπολογιστικά παραδείγματα

### 4.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα μελετήσουμε κάποια συγκεκριμένα αριθμητικά παραδείγματα προβλημάτων αρχικών/συνοριακών τιμών έτσι ώστε να επιβεβαιώσουμε την τάξη ακρίβειας του σφάλματος που αποδείξαμε στα προηγούμενα κεφάλαια. Θα ξεκινήσουμε με ένα πρόβλημα αρχικών/συνοριακών τιμών για την εξίσωση της θερμότητας. Έπειτα θα ασχοληθούμε με ένα γενικό παραβολικό πρόβλημα του οποίου γνωρίζουμε την ακριβή λύση, η οποία είναι ομαλή όπως και η χρονικά εξαρτούμενη συνάρτηση συντελεστών του ελλειπτικού όρου. Το πρόβλημα αυτό έχει μηδενικές συνοριακές συνθήκες καθώς και συνεχή αρχική συνθήκη. Τέλος θα μελετήσουμε ένα σχεδόν γραμμικό πρόβλημα του οποίου η ακριβής λύση είναι γνωστή και η μη γραμμικότητα οφείλεται στον συντελεστή του ελλειπτικού όρου ο οποίος είναι πολυώνυμο με μεταβλητή την εξαρτημένη μεταβλητή.

### 4.2 Μέθοδος υπολογισμού τάξης σφάλματος ως προς το χώρο και το χρόνο

Σε αυτή την παράγραφο θα περιγράψουμε τον τρόπο με τον οποίο θα υπολογίσουμε την τάξη του σφάλματος ως προς την χωρική παράμετρο  $h$  και

το χρονικό βήμα  $k$ . Έστω λοιπόν  $u$  η ακριβής λύση ενός παραβολικού προβλήματος αρχικών/συνοριακών τιμών και  $U^n$  η προσεγγιστική λύση του αντίστοιχου πλήρους διακριτού προβλήματος. Υποθέτουμε τώρα ότι χρησιμοποιούμε την έμμεση Euler Galerkin μέθοδο. Παρατηρούμε ότι στα προηγούμενα κεφάλαια δείξαμε ότι για το σφάλμα

$$e(h, k) = \max_n \|U^n - u(t_n)\|, \quad (4.1)$$

ισχύει η εκτίμηση

$$e(h, k) \leq C(u)(h^r + k), \quad (4.2)$$

όπου  $r \geq 2$  θετικός ακέραιος. Υποθέτουμε τώρα ότι

$$e(h, k) \approx C(u)(h^r + k). \quad (4.3)$$

Αν θέσουμε  $k = h^r$  τότε έχουμε ότι

$$e(h, h^r) \approx C(u)h^r. \quad (4.4)$$

Έστω τώρα  $h_1, h_2$  τέτοια ώστε  $h_1 \neq h_2$ . Τότε έχουμε

$$\begin{aligned} \frac{e(h_1, h_1^r)}{e(h_2, h_2^r)} &\approx \left(\frac{h_1}{h_2}\right)^r \Rightarrow \ln\left(\frac{e(h_1, h_1^r)}{e(h_2, h_2^r)}\right) \approx \ln\left(\frac{h_1}{h_2}\right)^r \\ &\Rightarrow r \approx \frac{\ln(e(h_1, h_1^r)/e(h_2, h_2^r))}{\ln(h_1/h_2)}. \end{aligned}$$

Έτσι ορίζουμε τη χωρική πειραματική τάξη σύγκλισης ως

$$O_s(h_1, h_2) = \frac{\ln(e(h_1, h_1^r)/e(h_2, h_2^r))}{\ln(h_1/h_2)}, \quad (4.5)$$

και περιμένουμε να είναι περίπου ίση με  $r$ . Έστω τώρα  $k_1, k_2$  τέτοια ώστε  $k_1 \neq k_2$ . Θέτοντας  $h = k^{1/r}$  καταλήγουμε με ανάλογο τρόπο να ορίσουμε την χρονική πειραματική τάξη σύγκλισης ως

$$O_t(k_1, k_2) = \frac{\ln(e(k_1^{1/r}, k_1)/e(k_2^{1/r}, k_2))}{\ln(k_1/k_2)}, \quad (4.6)$$

και περιμένουμε να είναι περίπου ίση με 1. Συνοψίζοντας, ο στόχος μας στα αριθμητικά μας πειράματα είναι να έχουμε χωρική τάξη σύγκλισης περίπου ίση με  $r$  και χρονική τάξη σύγκλισης περίπου ίση με 1.

### 4.3 Πρόβλημα αρχικών/συνοριακών τιμών για την εξίσωση της θερμότητας

Όπως είπαμε παραπάνω το πρώτο πρόβλημα που θα μελετήσουμε είναι το πρόβλημα αρχικών/συνοριακών τιμών για την εξίσωση της θερμότητας, το οποίο έχει γνωστή σε εμάς πραγματική ακριβή λύση. Έχουμε επίσης μηδενικές συνοριακές συνθήκες και συνεχή αρχικά δεδομένα. Έστω  $\Omega = [0, 1] \times [0, 1]$  και  $J = (0, \bar{t}]$ . Το πρόβλημα μας λοιπόν είναι το εξής

$$\begin{aligned} u_t - \Delta u &= f \text{ στο } \Omega, t \in J, \\ u &= 0 \text{ στο } \partial\Omega, t \in J, \\ u(\cdot, 0) &= (x - 1)xy(y - 1) \text{ στο } \Omega, \end{aligned} \quad (4.7)$$

όπου η συνάρτηση  $u$  είναι η λύση του προβλήματος που όπως αναφέραμε και προηγουμένως είναι γνωστή. Αυτή η συνάρτηση λοιπόν είναι η εξής

$$u(x, t) = e^{-t}(x - 1)xy(y - 1)(x^2t + y^2t - xt - yt + 1), \quad (4.8)$$

για  $(x, y) \in \Omega$  και  $t \in J$ .

Η συνάρτηση του δεξιού μέλους  $f$  όπως είπαμε και προηγουμένως είναι άγνωστη. Η τακτική που θα χρησιμοποιήσουμε για να υπολογίσουμε την  $f$  είναι να υπολογίσουμε τις μερικές παραγώγους της  $u$ , δηλαδή τις  $u_x$ ,  $u_y$  και  $u_t$ , και χρησιμοποιώντας αυτές να υπολογίσουμε την  $f$  από την ΜΔΕ  $f = u_t - \Delta u$ . Για λόγους ευκολίας αυτή την διαδικασία θα την κάνουμε μέσω της γλώσσας προγραμματισμού Python και της βιβλιοθήκης NGSolve, τα οποία θα χρησιμοποιήσουμε γενικά για την αριθμητική επίλυση του προβλήματος μας. Να σημειώσουμε ότι χρησιμοποιήσαμε κατά τμήματα γραμμικές συναρτήσεις για να προσεγγίσουμε την ακριβή λύση του προβλήματος, δηλαδή θέσαμε  $r = 2$ .

Επίσης χρησιμοποιήσαμε 5 διαφορετικά πλέγματα τα οποία ορίσαμε σύμφωνα με την παράμετρο  $h = \max_{\tau \in \mathcal{T}_h} \text{diam}(\tau)$ . Οι τιμές τις οποίες χρησιμοποιήσαμε για την παράμετρο  $h$  είναι  $1/4$ ,  $1/8$ ,  $1/16$ ,  $1/32$  και  $1/64$ .

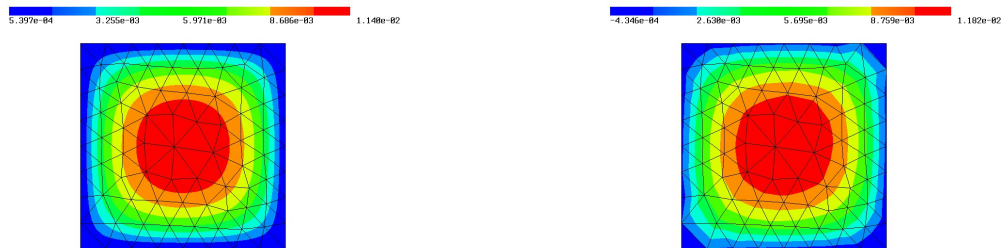
Ξεκινώντας από το πιο αραιό πλέγμα και καταλήγοντας στο πιο πυκνό παρουσιάζουμε παρακάτω τα γραφήματα για την ακριβή και την προσεγγιστική λύση του προβλήματος την χρονική στιγμή  $t = 1$  που πήραμε στο GUI του NGSolve τρέχοντας τον κώδικα μας.



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/4$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/4$ .

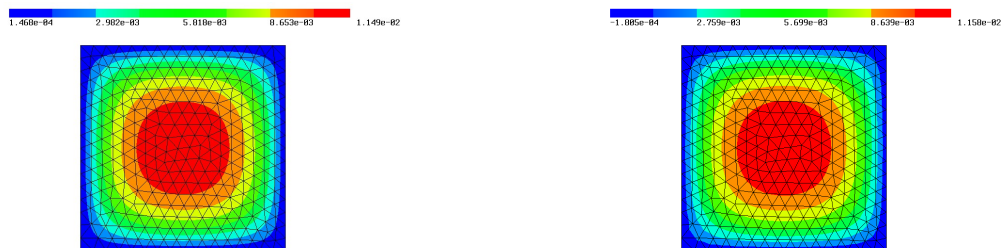
Σχήμα 4.1: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/4$ .



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/8$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/8$ .

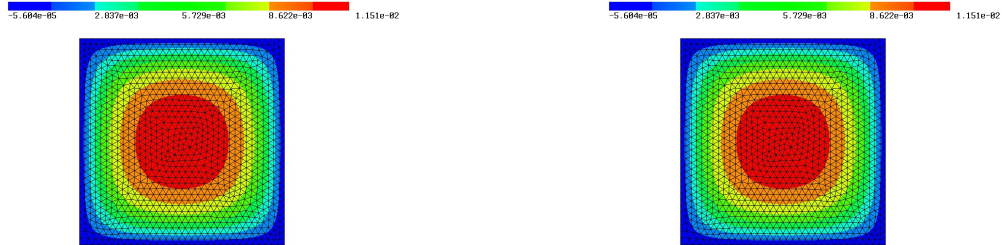
Σχήμα 4.2: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/8$ .



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/16$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/16$ .

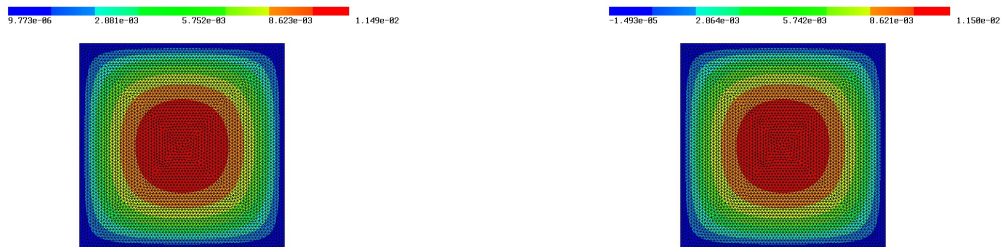
Σχήμα 4.3: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/16$ .



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/32$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/32$ .

Σχήμα 4.4: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/32$ .



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/64$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/64$ .

Σχήμα 4.5: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/64$ .

Όπως παρατηρούμε από τα παραπάνω γραφήματα όσο το  $h$  μειώνεται η προσεγγιστική λύση προσεγγίζει όλο και περισσότερο την ακριβή λύση. Αυτό το συμπεραίνουμε επίσης και από τα σφάλματα και τις τάξεις ακριβείας ως προς  $h$  και  $k$  που παρουσιάζουμε στον παρακάτω πίνακα.

$h$	$\max_n \ U^n - u(t_n)\ $	Τάξη ακρίβειας ως προς $h$	Τάξη ακρίβειας ως προς $k$
1/4	$2.92520464e - 03$	—	—
1/8	$4.63516291e - 04$	2.65784561	1.3289228
1/16	$9.18034727e - 05$	2.33599941	1.16799971
1/32	$2.16030717e - 05$	2.08731227	1.04365613
1/64	$5.39188388e - 06$	2.00237513	1.00118757

## 4.4 Πρόβλημα αρχικών/συνοριακών τιμών με γνωστή ακριβή λύση

Όπως είπαμε παραπάνω το δεύτερο πρόβλημα που θα μελετήσουμε έχει γνωστή σε εμάς πραγματική ακριβή λύση, η οποία είναι ομαλή, όπως και η συνάρτηση συντελεστών στον ελλειπτικό όρο. Έχουμε επίσης μηδενικές συνοριακές συνθήκες και συνεχή αρχικά δεδομένα. Έστω  $\Omega = [0, 1] \times [0, 1]$  και  $J = (0, \bar{t}]$ . Το πρόβλημα μας λοιπόν είναι το εξής

$$\begin{aligned} u_t - \nabla \cdot ((1 + e^{-t})\nabla u) &= f \text{ στο } \Omega, \quad t \in J, \\ u &= 0 \text{ στο } \partial\Omega, \quad t \in J, \\ u(\cdot, 0) &= (x - 1)xy(y - 1) \text{ στο } \Omega, \end{aligned} \quad (4.9)$$

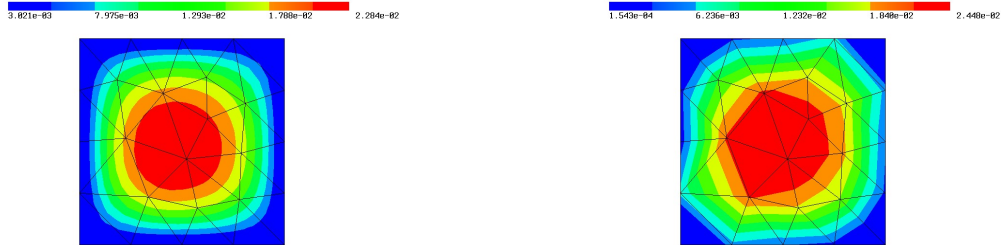
όπου η συνάρτηση  $u$  είναι η λύση του προβλήματος που όπως επανελημμένα αναφέραμε είναι γνωστή. Αυτή η συνάρτηση λοιπόν είναι η εξής

$$u(x, t) = e^{-t}(x - 1)xy(y - 1)(x^3t + y^3t - xt - yt + 1), \quad (4.10)$$

για  $(x, y) \in \Omega$  και  $t \in J$ . Η συνάρτηση του δεξιού μέλους  $f$  όπως είπαμε και προηγουμένως είναι άγνωστη. Η τακτική που θα χρησιμοποιήσουμε για να υπολογίσουμε την  $f$  είναι να υπολογίσουμε τις μερικές παραγώγους της  $u$ , δηλαδή τις  $u_x$ ,  $u_y$  και  $u_t$ , και χρησιμοποιώντας αυτές να υπολογίσουμε την  $f$  από την ΜΔΕ  $f = u_t - \nabla \cdot ((1 + e^{-t})\nabla u)$ . Για λόγους ευκολίας αυτή την διαδικασία θα την κάνουμε μέσω της γλώσσας προγραμματισμού Python και της βιβλιοθήκης NGSolve, όπως κάναμε και στο προηγούμενο πρόβλημα. Να σημειώσουμε ότι χρησιμοποιήσαμε κατά τμήματα γραμμικές συναρτήσεις για να προσεγγίσουμε την ακριβή λύση του προβλήματος, δηλαδή θέσαμε το  $r = 2$ .

Και πάλι χρησιμοποιήσαμε 5 διαφορετικά πλέγματα τα οποία ορίσαμε σύμφωνα με την παράμετρο  $h = \max_{\tau \in \mathcal{T}_h} \text{diam}(\tau)$ . Οι τιμές τις οποίες χρησιμοποιήσαμε για την παράμετρο  $h$  είναι  $1/4$ ,  $1/8$ ,  $1/16$ ,  $1/32$  και  $1/64$ , όπως και στο προηγούμενο αριθμητικό μας παράδειγμα.

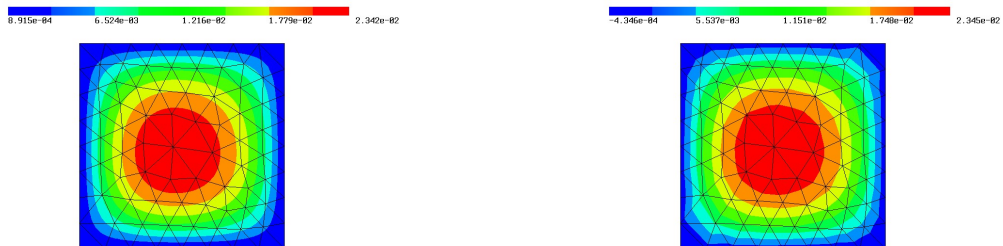
Ξεκινώντας από το πιο αραιό πλέγμα και καταλήγοντας στο πιο πυκνό παρουσιάζουμε παρακάτω τα γραφήματα για την ακριβή και την προσεγγιστική λύση του προβλήματος την χρονική στιγμή  $t = 0.5$  που πήραμε στο GUI του NGSolve τρέχοντας τον κώδικα μας.



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/4$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/4$ .

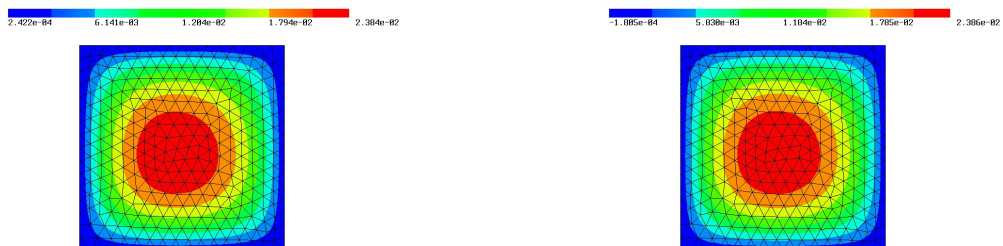
Σχήμα 4.6: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/4$ .



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/8$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/8$ .

Σχήμα 4.7: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/8$ .

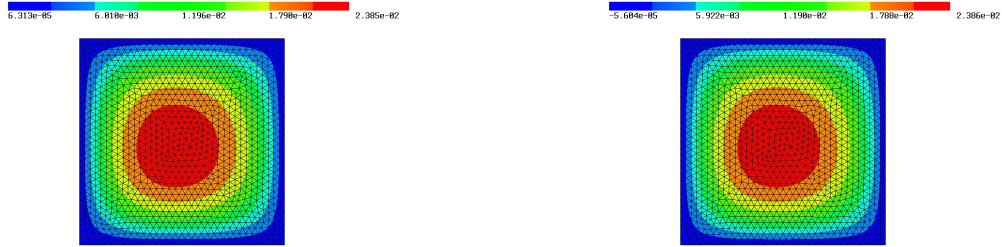


(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/16$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/16$ .

Σχήμα 4.8: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/16$ .

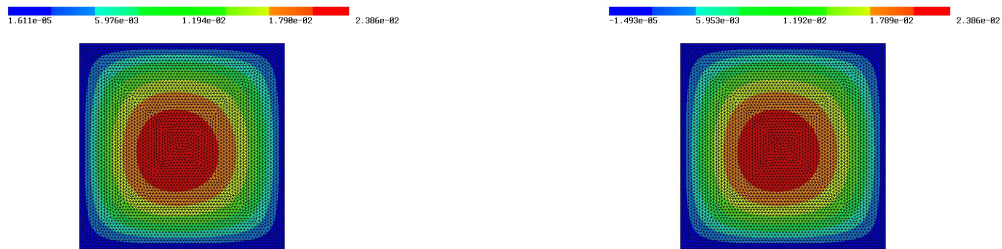




(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/32$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/32$ .

Σχήμα 4.9: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/32$ .



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/64$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/64$ .

Σχήμα 4.10: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/64$ .

Όπως παρατηρούμε από τα παραπάνω γραφήματα όσο το  $h$  μειώνεται η προσεγγιστική λύση προσεγγίζει όλο και περισσότερο την ακριβή λύση. Αυτό το συμπεραίνουμε επίσης και από τα σφάλματα και τις τάξεις ακριβείας ως προς  $h$  και  $k$  που παρουσιάζουμε στον παρακάτω πίνακα.

$h$	$\max_n \ U^n - u(t_n)\ $	Τάξη ακρίβειας ως προς $h$	Τάξη ακρίβειας ως προς $k$
1/4	$2.32401493e - 03$	—	—
1/8	$4.57456774e - 04$	2.34491201	1.172456
1/16	$9.67632134e - 05$	2.24110484	1.12055242
1/32	$2.31353082e - 05$	2.06436236	1.03218118
1/64	$5.81266457e - 06$	1.99282476	0.99641238

## 4.5 Σχεδόν γραμμικό παραβολικό πρόβλημα αρχικών/συνοριακών τιμών

Αρχικά θα παρουσιάσουμε την μέθοδο σταθερού σημείου και έπειτα την μέθοδο Newton. Στην συνέχεια θα παρουσιάσουμε ένα αριθμητικό παράδειγμα σχεδόν γραμμικού παραβολικού προβλήματος αρχικών/συνοριακών τιμών.

### 4.5.1 Μέθοδος σταθερού σημείου

Η μέθοδος σταθερού σημείου είναι η απλούστερη μέθοδος επίλυσης μη γραμμικών εξισώσεων. Εφαρμόζεται σε εξισώσεις της μορφής

$$x = g(x), \quad (4.11)$$

όπου η συνάρτηση  $g$  είναι μη γραμμική συνάρτηση του  $x$ . Η βασική ιδέα είναι ότι θεωρούμε μια αρχική εκτίμηση  $x^{(0)}$  της λύσης  $\bar{x}$  και  $\varepsilon$  το μέγιστο σφάλμα που δεχόμαστε να έχουμε και έπειτα υπολογίζουμε επαναληπτικά για  $k = 0, 1, 2, \dots$

$$x^{(k+1)} = g(x^{(k)}), \quad (4.12)$$

μέχρι να έχουμε σύγκλιση, δηλαδή μέχρι  $\|x^{(k+1)} - x^{(k)}\| \leq \varepsilon$ . Η μέθοδος σταθερού σημείου συγκλίνει αν η συνάρτηση  $g$  είναι Lipschitz, δηλαδή αν υπάρχει σταθερά  $L < 1$  τέτοια ώστε  $\|g(x) - g(y)\| \leq L\|x - y\| \forall x, y$ . Η σύγκλιση της μεθόδου σταθερού σημείου αποδεικνύεται αφαιρώντας την παραπάνω σχέση από την (4.12)

$$\|x^{(k+1)} - \bar{x}\| = \|g(x^{(k)}) - g(\bar{x})\| \leq L\|x^{(k)} - \bar{x}\| \leq \dots \leq L^{k+1}\|x^{(0)} - \bar{x}\|.$$

Όπως αναφέραμε και προηγουμένως η μέθοδος σταθερού σημείου είναι η απλούστερη μέθοδος επίλυσης μη γραμμικών εξισώσεων όμως συνήθως η σύγκλιση της είναι αργή. Εφαρμόζοντας λοιπόν την μέθοδο σταθερού σημείου στην σχέση (3.43)

$$(\mathcal{B} + k\mathcal{A}(\alpha^n))\alpha^n = \mathcal{B}\alpha^{n-1} + k\tilde{f}(\alpha^n), \text{ για } t_n \in J,$$

παίρνουμε την επαναληπτική μέθοδο

$$\alpha^{n(k)} = (\mathcal{B} + k\mathcal{A}(\alpha^{n(k-1)}))^{-1}(\mathcal{B}\alpha^{n-1} + k\tilde{f}(\alpha^{n(k-1)})), \quad (4.13)$$

για  $n = 0, 1, 2, \dots$  και  $k = 0, 1, 2, \dots$ . Δηλαδή σε κάθε χρονική στιγμή λύνουμε μια μέθοδο σταθερού σημείου στην οποία η αρχική τιμή του διανύσματος  $\alpha^{n(0)}$  είναι η τιμή  $\alpha^{n-1}$  που υπολογίσαμε την προηγούμενη χρονική στιγμή.

### 4.5.2 Μέθοδος Newton

Μια εναλλακτική μέθοδος που χρησιμοποιείται για την επίλυση μη γραμμικών εξισώσεων είναι η μέθοδος Newton η οποία συγκλίνει πολύ γρηγορότερα από την μέθοδο σταθερού σημείου έχοντας όμως το αντίβαρο ότι είναι πιο περίπλοκη. Αυτή η μέθοδος εφαρμόζεται σε εξισώσεις της μορφής

$$g(x) = 0, \quad (4.14)$$

όπου η συνάρτηση  $g$  είναι μη γραμμική ως προς την μεταβλητή  $x$ .

Το πρώτο βήμα που πρέπει να κάνουμε για να κατασκευάσουμε την μέθοδο Newton είναι να υποθέσουμε ότι η λύση  $\bar{x}$  της παραπάνω εξίσωσης μπορεί να γραφτεί στην μορφή

$$\bar{x} = x^{(0)} + \delta x, \quad (4.15)$$

όπου  $x^{(0)}$  είναι μια γνωστή προσέγγιση της λύσης  $\bar{x}$  και το  $\delta x$  καλείται διόρθωση. Υποθέτουμε ότι η αρχική προσέγγιση είναι αρκετά κοντά στη λύση  $\bar{x}$  έτσι ώστε το  $\delta x$  να είναι μικρό και χρησιμοποιώντας το ανάπτυγμα Taylor στην συνάρτηση  $g(x)$  γύρω από το σημείο  $\bar{x}$  έχουμε

$$g(\bar{x}) = g(x^{(0)} + \delta x) = g(x^{(0)}) + g'(x^{(0)})\delta x + \mathcal{O}(\delta x^2) \quad (4.16)$$

Πετώντας τους όρους δεύτερης τάξης και άνω και χρησιμοποιώντας την σχέση (4.14) έχουμε ότι

$$0 \approx g(x^{(0)}) + g'(x^{(0)})\delta x. \quad (4.17)$$

Παρατηρούμε ότι η παραπάνω σχέση είναι γραμμική ως προς  $\delta x$  σε αντίθεση με την (4.14) η οποία είναι μη γραμμική ως προς  $x$ . Επομένως θέτοντας

$$\delta x = -\frac{g(x^{(0)})}{g'(x^{(0)})}, \quad (4.18)$$

προσεγγίζουμε την λύση  $\bar{x}$  από την σχέση (4.15) και έτσι παίρνουμε μια καλύτερη προσέγγιση από την  $x^{(0)}$ . Η μέθοδος Newton υλοποιείται επαναλαμβάνοντας την παραπάνω διαδικασία μέχρι για ένα αυθαίρετα μικρό  $\varepsilon > 0$  να έχουμε  $|\delta x| < \varepsilon$  ή μέχρι ένα μέγιστο αριθμό επαναλήψεων τον οποίο θέτουμε επίσης αυθαίρετα.

Η μέθοδος Newton είναι δημοφιλής επειδή συνήθως συγκλίνει γρήγορα. Μπορεί ναδειχθεί ότι

$$\|x^{(k+1)} - \bar{x}\| \leq C\|x^{(k)} - \bar{x}\|^2, \quad (4.19)$$

όπου το  $x^{(k)}$  είναι αρκούντως κοντά στη λύση  $\bar{x}$ . Παρατηρούμε λοιπόν ότι υπό την προϋπόθεση ότι η συνάρτηση  $g$  είναι διαφορίσιμη και η  $g' \neq 0$  κοντά στο  $\bar{x}$  η μέθοδος Newton πετυχαίνει τετραγωνική τάξη σύγκλισης.

Πάμε λοιπόν να εφαρμόσουμε την μέθοδο Newton στην σχέση (3.43)

$$(\mathcal{B} + k\mathcal{A}(\alpha^n))\alpha^n = \mathcal{B}\alpha^{n-1} + k\tilde{f}(\alpha^n), \text{ για } t_n \in J.$$

Αυτό σημαίνει ότι πρέπει να λύσουμε μια σχέση της μορφής

$$r(\alpha^n) = 0, \quad (4.20)$$

όπου  $r(\alpha^n)$  είναι το διάνυσμα υπολοίπων

$$(\mathcal{B} + k\mathcal{A}(\alpha^n))\alpha^n - \mathcal{B}\alpha^{n-1} - k\tilde{f}(\alpha^n). \quad (4.21)$$

Ο Ιακωβιανός πίνακας  $J$  έχει ως στοιχεία τις μερικές παραγώγους  $J_{ij} = \partial r_i / \partial (\alpha^n)_j$  για  $i, j = 1, 2, \dots, N_x$ . Εμείς στο υπολογιστικό μας παράδειγμα θα χρησιμοποιήσουμε την μέθοδο Newton. Δεν χρειάζεται όμως να υπολογίσουμε οι ίδιοι τον παραπάνω Ιακωβιανό πίνακα μιας και θα το κάνει αυτόματα για εμάς το λογισμικό NGSolve.

### 4.5.3 Υπολογιστικό παράδειγμα

Το πρόβλημα που θα μελετήσουμε έχει γνωστή σε εμάς πραγματική ακριβή λύση, η οποία είναι ομαλή, και η συνάρτηση  $a(u)$  στον ελλειπτικό όρο είναι ένα πολυώνυμο δευτέρου βαθμού. Έχουμε επίσης μηδενικές συνοριακές συνθήκες και συνεχή αρχικά δεδομένα. Έστω  $\Omega = [0, 1] \times [0, 1]$  και  $J = (0, \bar{t}]$ . Το πρόβλημα μας λοιπόν είναι το εξής

$$\begin{aligned} u_t - \nabla \cdot ((u^2 + 1)\nabla u) &= f \text{ στο } \Omega, \quad t \in J, \\ u &= 0 \text{ στο } \partial\Omega, \quad t \in J, \\ u(\cdot, 0) &= (x - 1)xy(y - 1) \text{ στο } \Omega, \end{aligned} \quad (4.22)$$

όπου η συνάρτηση  $u$  είναι η λύση του προβλήματος που όπως αναφέραμε και προηγουμένως είναι γνωστή. Αυτή η συνάρτηση λοιπόν είναι η εξής

$$u(x, t) = e^{-t}(x - 1)xy(y - 1)(x^2t + y^2t - xt - yt + 1), \quad (4.23)$$

για  $(x, y) \in \Omega$  και  $t \in J$ . Η συνάρτηση του δεξιού μέλους  $f$  είναι άγνωστη. Για να υπολογίσουμε την  $f$  είναι να υπολογίσουμε τις μερικές παραγώγους της

$u$ , δηλαδή τις  $u_x$ ,  $u_y$  και  $u_t$ , και χρησιμοποιώντας αυτές να υπολογίσουμε την  $f$  από την ΜΔΕ  $f = u_t - \nabla \cdot ((u^2 + 1)\nabla u)$ . Για λόγους ευκολίας αυτή την διαδικασία θα την κάνουμε μέσω της γλώσσας προγραμματισμού Python και της βιβλιοθήκης NGSolve, όπως κάναμε και στα προηγούμενα προβλήματα. Να σημειώσουμε ότι χρησιμοποιήσαμε κατά τμήματα γραμμικές συναρτήσεις για να προσεγγίσουμε την ακριβή λύση του προβλήματος, δηλαδή θέσαμε  $r = 2$ .

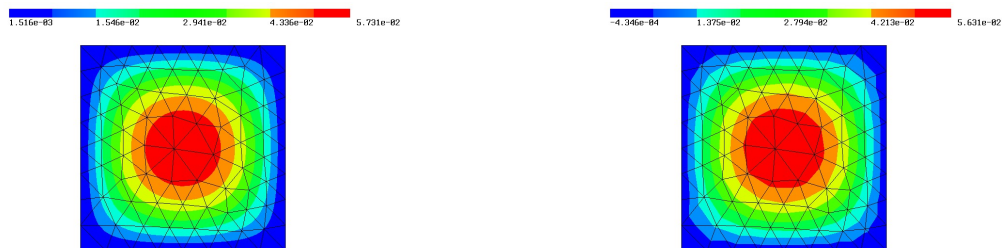
Χρησιμοποιήσαμε και πάλι τα ίδια πλέγματα όπως και στα προηγούμενα παραδείγματα. Επίσης θέσαμε  $\varepsilon = 10^{-13}$  και μέγιστο αριθμό επαναλήψεων 25. Ξεκινώντας από το πιο αραιό πλέγμα και καταλήγοντας στο πιο πυκνό παρουσιάζουμε παρακάτω τα γραφήματα για την ακριβή και την προσεγγιστική λύση του προβλήματος την χρονική στιγμή  $t = 0.05$  που πήραμε στο GUI του NGSolve τρέχοντας τον κώδικα μας.



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/4$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/4$ .

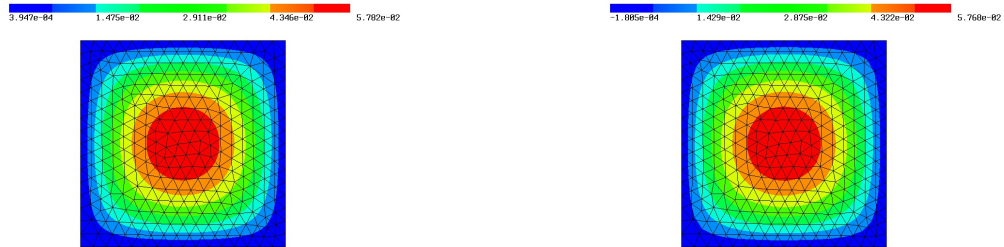
Σχήμα 4.11: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/4$ .



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/8$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/8$ .

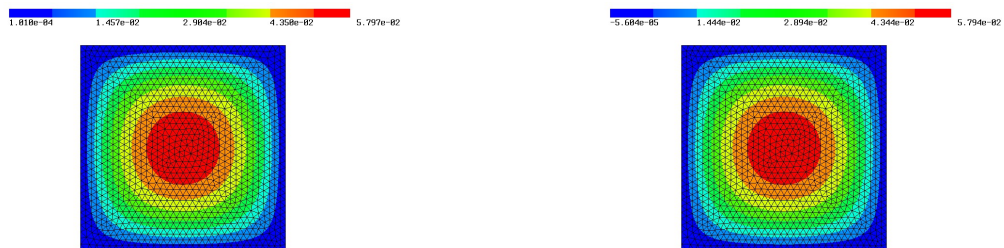
Σχήμα 4.12: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/8$ .



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/16$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/16$ .

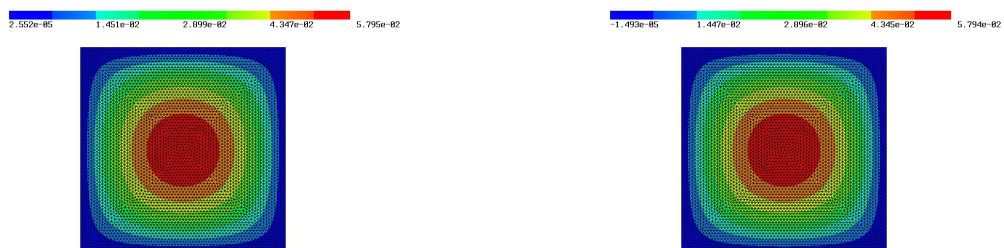
Σχήμα 4.13: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/16$ .



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/32$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/32$ .

Σχήμα 4.14: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/32$ .



(a) Ακριβής λύση στο πλέγμα που ορίζεται για  $h = 1/64$ .

(b) Προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/64$ .

Σχήμα 4.15: Ακριβής και προσεγγιστική λύση στο πλέγμα που ορίζεται για  $h = 1/64$ .

Όπως παρατηρούμε και πάλι από τα παραπάνω γραφήματα όσο το  $h$  μειώνεται η προσεγγιστική λύση προσεγγίζει όλο και περισσότερο την ακριβή λύση. Αυτό το συμπεραίνουμε επίσης και από τα σφάλματα και τις τάξεις ακριβείας ως προς  $h$  και  $k$  που παρουσιάζουμε στον παρακάτω πίνακα.

$h$	$\max_n \ U^n - u(t_n)\ $	Τάξη ακρίβειας ως προς $h$	Τάξη ακρίβειας ως προς $k$
1/4	$2.13118653e - 03$	—	—
1/8	$4.47129120e - 04$	2.25289346	1.12644673
1/16	$9.16048800e - 05$	2.28719514	1.14359757
1/32	$2.13563947e - 05$	2.10075634	1.05037817
1/64	$5.29609208e - 06$	2.011668	1.005834

# Παράρτημα

## Κώδικες Python

### Κώδικας για την εξίσωση της θερμότητας

```
from netgen import gui
from math import pi
from ngsolve import *
from netgen.geom2d import SplineGeometry
import numpy as np

geo = SplineGeometry()
geo.AddRectangle( (0, 0), (1, 1),
                 bcs = ("bottom", "right", "top", "left"))

def BackwardEuler(h,r):
    mesh = Mesh( geo.GenerateMesh(maxh=h) )
    Draw(mesh)
    fes = H1(mesh, order=r-1, dirichlet="bottom|right|left|
top")

    u,v = fes.TnT()

    time = 0.0
    dt = h**r

    a = BilinearForm(fes, symmetric=False)
    a += grad(u)*grad(v)*dx
    a.Assemble()

    b = BilinearForm(fes, symmetric=False)
```



```

b += u*v*dx
b.Assemble()

mstar = b.mat.CreateMatrix()
mstar.AsVector().data = b.mat.AsVector() + dt * a.mat.
AsVector()
invmstar = mstar.Inverse(freedofs=fes.FreeDofs())

t = Parameter(0.0)
u_exact = exp(-t)*(x-1)*x*y*(y-1)*(x*x *t + y*y * t -x *
t - y * t + 1)
u_t = u_exact.Diff(t)
grad_u = CoefficientFunction( (u_exact.Diff(x), u_exact.
Diff(y)))
u_x = grad_u[0]; u_y = grad_u[1]
Delta = u_x.Diff(x) + u_y.Diff(y)
f = u_t - Delta

Draw(f,mesh,"ft",sd=4)
Draw(u_exact,mesh,"u_exact")

ft = LinearForm(fes)
ft += f*v*dx
time = 0.0
gfu = GridFunction(fes)
gfu.Set((x-1)*x*y*(y-1))
Draw(gfu,mesh,"u")

tstep = 1 # time that we want to step over within one
block-run
t_intermediate=0 # time counter within one block-run

L2err = []
times = []
L2err_time_square = Integrate((gfu-u_exact)*(gfu-u_exact)
,mesh)
L2err_time = sqrt(L2err_time_square.real)
L2err.append(L2err_time)
times.append(t_intermediate)

res = gfu.vec.CreateVector()
while t_intermediate < tstep - 0.5 * dt:
    t.Set(time+t_intermediate+dt)
    ft.Assemble()
    # res.data = b.mat * gfu.vec + dt * ft.vec
    res.data = dt * ft.vec - dt * a.mat * gfu.vec
    gfu.vec.data += invmstar * res
    L2err_time_square = Integrate((gfu-u_exact)*(gfu-

```

```

    u_exact), mesh)
        L2err_time = sqrt(L2err_time_square.real)
        L2err.append(L2err_time)
        times.append(t_intermediate)
        t_intermediate += dt
        print("\r", time+t_intermediate, end="")
        Redraw(blocking=True)
    print("")
    time+=t_intermediate

    L2error = max(L2err)
    return L2error

h = [1/4, 1/8, 1/16, 1/32, 1/64]
L2errors = np.zeros(len(h))

for i in range(len(h)):
    L2errors[i] = BackwardEuler(h[i], 2)
    print(L2errors)

L2orders_space = np.zeros(len(h)-1)
L2orders_time = np.zeros(len(h)-1)

for i in range(len(h)-1):
    L2orders_space[i] = log(L2errors[i+1]/L2errors[i])/log(h[
i+1]/h[i])
    L2orders_time[i] = L2orders_space[i]/2
print(L2orders_space)
print(L2orders_time)

```

Κώδικας για την εξίσωση με τον χρονικά εξαρτούμενο ελλειπτικό όρο και την γνωστή ακριβή λύση

```

from netgen import gui
from math import pi
from ngsolve import *
from netgen.geom2d import SplineGeometry
import matplotlib.pyplot as plt
import numpy as np

pt = [0.01*i for i in range(1000)]
vals = [(exp(-p)+1) for p in pt]
plt.plot(pt, vals)

```

```
plt.xlabel('t')
plt.xlabel('a(t)')
plt.show()

geo = SplineGeometry()
geo.AddRectangle( (0, 0), (1, 1),
                 bcs = ("bottom", "right", "top", "left"))

def BackwardEuler(h,r):
    mesh = Mesh( geo.GenerateMesh(maxh=h))
    Draw(mesh)
    fes = H1(mesh, order=r-1, dirichlet="bottom|right|left|
top")

    u,v = fes.TnT()

    time = 0.0
    dt = h**r
    t = Parameter(0.0)
    at = exp(-t)+1

    a = BilinearForm(fes, symmetric=False)
    a += at*grad(u)*grad(v)*dx

    b = BilinearForm(fes, symmetric=False)
    b += u*v*dx
    b.Assemble()

    mstar = b.mat.CreateMatrix()

    t = Parameter(0.0)
    u_exact = exp(-t)*(x-1)*x*y*(y-1)*(x*x*x *t + y*y*y * t -
x * t - y * t + 1)
    u_t = u_exact.Diff(t)
    grad_u = CoefficientFunction(( u_exact.Diff(x), u_exact.
Diff(y)))
    au_x = at*grad_u[0]; au_y = at*grad_u[1]
    Div = CoefficientFunction(au_x.Diff(x) + au_y.Diff(y))
    f = u_t - Div

    Draw(f,mesh,"ft",sd=4)
    Draw(u_exact,mesh,"u_exact")

    ft = LinearForm(fes)
    ft += f*v*dx
    time = 0.0
    gfu = GridFunction(fes)
    gfu.Set((x-1)*x*y*(y-1))
```

```

Draw(gfu, mesh, "u")

tstep = 0.5 # time that we want to step over within one
block-run
t_intermediate=0 # time counter within one block-run

L2err = []
times = []
L2err_time_square = Integrate((gfu-u_exact)*(gfu-u_exact)
, mesh)
L2err_time = sqrt(L2err_time_square.real)
L2err.append(L2err_time)
times.append(t_intermediate)

res = gfu.vec.CreateVector()
while t_intermediate < tstep - 0.5 * dt:
    a.Assemble()
    mstar.AsVector().data = b.mat.AsVector() + dt * a.mat
.AsVector()
    invmstar = mstar.Inverse(freedofs=fes.FreeDofs())
    t.Set(time+t_intermediate+dt)
    # res.data = b.mat * gfu.vec + dt * ft.vec
    a.Assemble()
    ft.Assemble()
    res.data = dt * ft.vec - dt * a.mat * gfu.vec
    gfu.vec.data += invmstar * res
    L2err_time_square = Integrate((gfu-u_exact)*(gfu-
u_exact), mesh)
    L2err_time = sqrt(L2err_time_square.real)
    L2err.append(L2err_time)
    times.append(t_intermediate)
    t_intermediate += dt
    print("\r", time+t_intermediate, end=" ")
    Redraw(blocking=True)
print("")
time+=t_intermediate

L2error = max(L2err)
return L2error

h = [1/4, 1/8, 1/16, 1/32, 1/64]
L2errors = np.zeros(len(h))

for i in range(len(h)):
    L2errors[i] = BackwardEuler(h[i], 2)
    print(L2errors)

L2orders_space = np.zeros(len(h)-1)
L2orders_time = np.zeros(len(h)-1)

```

```
for i in range(len(h)-1):
    L2orders_space[i] = log(L2errors[i+1]/L2errors[i])/log(h[
    i+1]/h[i])
    L2orders_time[i] = L2orders_space[i]/2
print(L2orders_space)
print(L2orders_time)
```

Κώδικας για την εξίσωση με τον χρονικά εξαρτούμενο ελλειπτικό όρο και την γνωστή ακριβή λύση

```
from netgen import gui
from math import pi
from ngsolve import *
from netgen.geom2d import SplineGeometry
import matplotlib.pyplot as plt
import numpy as np

geo = SplineGeometry()
geo.AddRectangle( (0, 0), (1, 1),
                 bcs = ("bottom", "right", "top", "left"))

def BackwardEuler(h,r):
    mesh = Mesh( geo.GenerateMesh(maxh=h))
    Draw(mesh)
    fes = H1(mesh, order=r-1, dirichlet="bottom|right|left|
    top")

    u,v = fes.TnT()

    time = 0.0
    dt = h**r

    t = Parameter(0.0)
    u_exact = exp(-t)*(x-1)*x*y*(y-1)*(x*x *t + y*y * t -x *
    t - y * t + 1)
    u_t = u_exact.Diff(t)
    grad_u = CoefficientFunction( (u_exact.Diff(x), u_exact.
    Diff(y)))
    au_x = (u_exact**2+1)*grad_u[0]; au_y = (u_exact**2+1)*
    grad_u[1]
    Div = CoefficientFunction(au_x.Diff(x) + au_y.Diff(y))
```

```

f = u_t - Div

Draw(f,mesh,"ft",sd=4)
Draw(u_exact,mesh,"u_exact")

time = 0.0
gfu = GridFunction(fes)
gfu.Set((x-1)*x*y*(y-1))
gfu_old = GridFunction(fes)
gfu_old.Set((x-1)*x*y*(y-1))

Draw(gfu,mesh,"u")

a = BilinearForm(fes, symmetric=False)
a +=( u*v +dt*(u**2+1)*grad(u)*grad(v) - dt*f*v- gfu_old*
v) * dx

tstep = 0.05 # time that we want to step over within one
block-run
t_intermediate=0 # time counter within one block-run

L2err = []
times = []
L2err_time_square = Integrate((gfu-u_exact)*(gfu-u_exact)
,mesh)
L2err_time = sqrt(L2err_time_square.real)
L2err.append(L2err_time)
times.append(t_intermediate)
tol = 1e-13

res = gfu.vec.CreateVector()
du = gfu.vec.CreateVector()
while t_intermediate < tstep - 0.5 * dt:
    t.Set(time+t_intermediate+dt)
    for it in range(25):
        a.Apply(gfu.vec, res)
        a.AssembleLinearization(gfu.vec)
        inv = a.mat.Inverse(freedofs=fes.FreeDofs())
        du.data = inv * res
        #update iteration
        gfu.vec.data -= du
        #stopping criteria
        stopcritval = sqrt(abs(InnerProduct(du,res)))
        if stopcritval < tol:
            break
        gfu_old.vec.data = gfu.vec
        L2err_time_square = Integrate((gfu-u_exact)*(gfu-
u_exact),mesh)

```

```
        L2err_time = sqrt(L2err_time_square.real)
        L2err.append(L2err_time)
        times.append(t_intermediate)
        t_intermediate += dt
        print("\r",time+t_intermediate,end=" ")
        Redraw(blocking=True)
    print("")
    time+=t_intermediate

L2error = max(L2err)
return L2error

h = [1/4, 1/8, 1/16, 1/32, 1/64]
L2errors = np.zeros(len(h))

for i in range(len(h)):
    L2errors[i] = BackwardEuler(h[i],2)
    print(L2errors)

L2orders_space = np.zeros(len(h)-1)
L2orders_time = np.zeros(len(h)-1)

for i in range(len(h)-1):
    L2orders_space[i] = log(L2errors[i+1]/L2errors[i])/log(h[
i+1]/h[i])
    L2orders_time[i] = L2orders_space[i]/2
print(L2orders_space)
print(L2orders_time)
```

# Βιβλιογραφία

- [1] Susanne C. Brenner L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods*. Texts in Mathematics. Springer, 1994. ISBN: 978-0-387-75933-3.
- [2] Philippe G. Ciarlet. *The Finite Element Method for Elliptic Problems*. SIAM's Classics in Applied Mathematics. SIAM, 1978. ISBN: 0-89871-514-8.
- [3] Εμμανουήλ Γεωργούλης. *Σημειώσεις μαθήματος: Πεπερασμένα Στοιχεία και Πεπερασμένες Διαφορές για ΜΔΕ*. 2020.
- [4] Lawrence C. Evans. *Partial Differential Equation*. Graduate Studies in Mathematics. American Mathematical Society, 1998.
- [5] Fredrik Bengzon Mats G. Larson. *The Finite Element Method: Theory, Implementation, and Applications*. Texts in Computational Science and Engineering. Springer, 2013. ISBN: 978-3-642-33286-9.
- [6] Vidar Thomée. *Galerkin Finite Element Methods for Parabolic Problems*. Springer Series in Computational Mathematics. Springer, 2006. ISBN: 0179-3632.
- [7] P. E. Sobolevskii. *Equations of parabolic type in a Banach space*. Tr. Mosk. Mat. Obs., 10, pp. 297–350, 1961.
- [8] Σταυραράκης Νικόλαος. *Συνήθειες Διαφορικές Εξισώσεις, 2η Αναθεωρημένη Έκδοση: Γραμμική και μη γραμμική θεωρία με εφαρμογές από τη φύση και τη ζωή*. Εκδόσεις Παπασωτηρίου, 2011. ISBN: 9789607182074.
- [9] Netgen/NGSolve team. *Interactive NGSolve Tutorial*. 2017. URL: <https://ngsolve.org/docu/latest/i-tutorials/index.html>.



- 
- [10] Γ. Δ. Ακρίβης- Β. Α. Δουγάλης. *Εισαγωγή στην Αριθμητική Ανάλυση* 4η έκδοση. Πανεπιστημιακή Βιβλιοθήκη Θετικών Επιστημών/Μαθηματικά. Πανεπιστημιακές Εκδόσεις Κρήτης, 2015. ISBN: 978-960-524-022-6.
- [11] Robert Johansson. *Numerical Python: A Practical Techniques Approach for Industry*. Apress, 2015. ISBN: (pbk): 978-1-4842-0554-9 ISBN (electronic): 978-1-4842-0553-2.