ΓΕΩΡΓΙΟΣ Δ. ΑΚΡΙΒΗΣ
Τμήμα Πληροφορικής
Πανεπιστήμιο Ιωαννίνων
e-mail: akrivis@cs.uoi.gr

ΒΑΣΙΛΕΙΟΣ Α. ΔΟΥΓΑΛΗΣ
Τμήμα Μαθηματικών
Πανεπιστήμιο Αθηνών
e-mail: doug@math.uoa.gr

ΑΡΙΘΜΗΤΙΚΕΣ ΜΕΘΟΔΟΙ ΓΙΑ ΜΕΡΙΚΕΣ ΔΙΑΦΟΡΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

ΙΩΑΝΝΙΝΑ, 2005

Πρόλογος

Η μελέτη των Διαφορικών Εξισώσεων, Συνήθων και Μερικών, αποτελεί κεντρικό αντικείμενο των Εφαρμοσμένων Μαθηματικών. Διαφορικές εξισώσεις εμφανίζονται πολύ συχνά σε μαθηματικά μοντέλα που περιγράφουν προβλήματα των φυσικών, των τεχνολογικών, των βιοϊατρικών, αλλά και των οικονομικών επιστημών. Συνήθως οι εξισώσεις αυτές δεν είναι δυνατόν να επιλυθούν με αναλυτικές μεθόδους· συνεπώς, η επίλυσή τους με προσεγγιστικές μεθόδους στον υπολογιστή αποτελεί ένα σημαντικό κεφάλαιο της Αριθμητικής Ανάλυσης.

Στις διδακτικές αυτές σημειώσεις επιχειρούμε μία εισαγωγή στο θέμα, δίνοντας ιδιαίτερη έμφαση, στο πρώτο μέρος, στην αριθμητική επίλυση του προβλήματος αρχικών τιμών για συνήθεις διαφορικές εξισώσεις. Στο δεύτερο μέρος δίνουμε μία σύντομη εισαγωγή σε μεθόδους διαφορών και πεπερασμένων στοιχείων για προβλήματα συνοριακών τιμών δύο σημείων για γραμμικές διαφορικές εξισώσεις δεύτερης τάξης. Τέλος, από το εκτενέστατο πεδίο της αριθμητικής ανάλυσης των Μερικών Διαφορικών Εξισώσεων, περιοριστήκαμε προς το παρόν, στο τρίτο μέρος, να αναλύσουμε μόνο ορισμένες βασικές μεθόδους διαφορών και πεπερασμένων στοιχείων για την επίλυση προβλημάτων αρχικών και συνοριακών τιμών για την εξίσωση της θερμότητας και την εξίσωση του κύματος.

Προηγούμενες μορφές διαφόρων κεφαλαίων των σημειώσεων αυτών έχουμε χρησιμοποιήσει κατά καιρούς τα τελευταία είκοσι χρόνια σε προπτυχιακά και μεταπτυχιακά μαθήματα Αριθμητικής Ανάλυσης στο Πανεπιστήμιο Κρήτης, στο Ε.Μ.Π., και στα Πανεπιστήμια Αθηνών και Ιωαννίνων.

Ιούνιος 2005

Γ. Δ. Ακρίβης και Β. Α. Δουγαλής

Περιεχόμενα

Μέρος I

Το πρόβλημα δύο σημείων

Η μαθηματική μοντελοποίηση των περισσότερων προβλημάτων της Φυσικής, της Μηχανικής και άλλων εφαρμοσμένων επιστημών οδηγεί σε Μερικές Διαφορικές Εξισώσεις. Οι εξισώσεις αυτές συμπληρώνονται με κατάλληλες συνοριακές ή αρχικές και συνοριακές συνθήκες έτσι ώστε να περιγράφουν κατά μοναδικό τρόπο τη λύση των προβλημάτων που μοντελοποιούν. Τέτοια προβλήματα μόνο σε σπάνιες και πολύ απλές περιπτώσεις μπορούν να λυθούν αναλυτικά, να πάρουμε δηλαδή τη λύση τους σε κλειστή μορφή. Αναγκαστικά λοιπόν προσφεύγουμε στην προσεγγιστική επίλυσή τους με μέσα της Αριθμητικής Ανάλυσης. Λόγω της μεγάλης σημασίας που έχουν αυτά τα προβλήματα για τις εφαρμογές αλλά και των πολλών ιδιομορφιών που παρουσιάζονται, η πλειονότητα των ερευνητών στην Αριθμητική Ανάλυση ασχολείται με τέτοια θέματα.

Οι μεγαλύτερες από τις πολλές κατηγορίες μεθόδων της Αριθμητικής Ανάλυσης που υπάρχουν για την αντιμετώπιση τέτοιων προβλημάτων είναι δύο: Οι μέθοδοι των πεπερασμένων διαφορών και οι μέθοδοι των πεπερασμένων στοιχείων. Η μέθοδος των πεπερασμένων διαφορών αναπτύχθηκε πάρα πολύ στις δεκαετίες του 50 και του 60, η μέθοδος των πεπερασμένων στοιχείων κυρίως μετά το 1970. Η μέθοδος των πεπερασμένων στοιχείων εφαρμόζεται σε περισσότερα προβλήματα, η ανάλυσή της είναι πιο συστηματικοποιημένη και γίνεται σε πολύπλοκους χώρους. Η ανάλυση των μεθόδων πεπερασμένων διαφορών γίνεται με στοιχειώδη μέσα.

Με την πάροδο του χρόνου όλο και περισσότεροι ερευνητές και χρήστες στρέφονται προς τη μέθοδο των πεπερασμένων στοιχείων και φυσιολογικά οι μέθοδοι των πεπερασμένων διαφορών έχουν χάσει κάτι από την παλιά τους αίγλη. Παρά ταύτα υπάρχουν πολλοί που πιστεύουν ότι οι μέθοδοι των πεπερασμένων διαφορών είναι αποτελεσματικότερες για κάποιες κατηγορίες προβλημάτων, όπως είναι τα υπερβολικά προβλήματα. Υπάρχουν πάντως πολλές ομοιότητες μεταξύ των δύο αυτών κατηγοριών μεθόδων και έτσι όταν γνωρίζει κανείς και τις δύο κατανοεί καλύτερα το θέμα.

Σ' αυτό το μέρος θα ασχοληθούμε με μεθόδους πεπερασμένων διαφορών και πεπερασμένων στοιχείων για ένα απλό πρόβλημα, το πρόβλημα δύο σημείων, με στόχο να εξοικειωθεί ο αναγνώστης με τις βασικές τεχνικές και έννοιες.

Το δεύτερο μέρος αποτελείται από τρία κεφάλαια. Στο κεφάλαιο 5 θα παρουσιάσουμε συνοπτικά μερικές ιδιότητες της λύσης του προβλήματος δύο σημείων σε τέτοια μορφή ώστε να οδηγηθούμε εύκολα στα κεφάλαια 6 και 7 σε ορισμένα διακριτά ανάλογά τους. Στο κεφάλαιο 6 θα γνωρίσουμε μία μέθοδο πεπερασμένων διαφορών και θα εκτιμήσουμε το σφάλμα της. Οι εκτιμήσεις θα γίνουν με τη μέθοδο της ενέργειας, αφ' ενός γιατί δίνει καλά αποτελέσματα και σε περιπτώσεις στις οποίες άλλες μέθοδοι αποτυγχάνουν, και αφ' ετέρου γιατί θα χρησιμοποιηθεί στο κεφάλαιο 7 όταν θα μελετήσουμε τη μέθοδο των πεπερασμένων στοιχείων.

Στο Μέρος III θα ασχοληθούμε ξανά με μεθόδους πεπερασμένων διαφορών και πεπερασμένων στοιχείων, εκεί για δυναμικές Μερικές Διαφορικές Εξισώσεις.

1. Το πρόβλημα δύο σημείων

Θεωρούμε το εξής πρόβλημα δύο σημείων για μία συνήθη διαφορική εξίσωση (Σ.Δ.Ε.) δεύτερης τάξης: Ζητείται μία συνάρτηση u, $u \in C^2[a,b]$, τέτοια ώστε

$$\begin{cases} -u'' + qu = f & \text{στο } [a,b] \\ u(a) = u(b) = 0, \end{cases} \tag{5.1}$$

όπου $a, b \in \mathbb{R}, a < b$ και q, f συνεχείς συναρτήσεις στο διάστημα $[a,b]$, $q, f \in C[a,b]$, και η q λαμβάνει μη αρνητικές τιμές, $q(x) \geq 0$ για κάθε $x \in [a,b]$. Είναι γνωστό ότι υπό τις ανωτέρω συνθήκες το πρόβλημα (5.1) έχει ακριβώς μία λύση.

Παρατήρηση 5.1 Αν η q παίρνει και αρνητικές τιμές, το πρόβλημα (5.1) μπορεί να έχει πολλές λύσεις, φερ' ειπείν το πρόβλημα

$$\begin{cases} -u'' - \pi^2 u = 0 & \text{στο } [0,1] \\ u(0) = u(1) = 0 \end{cases}$$

έχει τις λύσεις $u(x) = \alpha \sin(\pi x)$ με $\alpha \in \mathbb{R}$. Ακόμη είναι δυνατόν σε αυτή την περίπτωση να μην υπάρχει λύση, όπως, επί παραδείγματι, στην περίπτωση του προβλήματος

$$\begin{cases} -u'' - \pi^2 u = \pi^2 & \text{στο } [0,1] \\ u(0) = u(1) = 0. \end{cases}$$

Αυτό μπορούμε να το δούμε ως εξής: Η γενική λύση της Διαφορικής Εξίσωσης είναι $u(x) = \alpha \sin(\pi x) + \beta \cos(\pi x) - 1$, με $\alpha, \beta \in \mathbb{R}$. Αυτό είναι γνωστό από τη Θεωρία των Διαφορικών Εξισώσεων, επειδή όμως εδώ δεν προϋποτίθενται γνώσεις Σ.Δ.Ε. θα το αποδείξουμε στη συνέχεια με στοιχειώδη μέσα. Αν το δεχθούμε προς στιγμήν, τότε διαπιστώνουμε πολύ εύκολα ότι για καμμία επιλογή των α και β δεν πληρούνται συγχρόνως και οι δύο συνοριακές συνθήκες, δηλ. το πρόβλημά μας δεν επιδέχεται λύση. Επιστρέφουμε τώρα στη γενική λύση της Δ.Ε. Έστω w μία άλλη λύση. Τότε για την $v := u - w$ έχουμε

$$-v'' - \pi^2 v = 0 \tag{5.2}$$

και με κατάλληλη επιλογή των α και β μπορούμε να επιτύχουμε

$$v(0) = v'(0) = 0. \tag{5.2$'$}$$

Πολλαπλασιάζοντας την (5.2) με v' τη γράφουμε στη μορφή

$$[(v')^2]' + \pi^2(v^2)' = 0$$

και συμπεραίνουμε ότι

$$(v')^2 + \pi^2 v^2 = C, \qquad C\text{σταθερά.}$$

Σύμφωνα με την (5.2′) το C είναι μηδέν, και συνεπώς $v = 0$. Αυτό σημαίνει ότι και η w γράφεται στη μορφή

$$w(x) = \tilde{\alpha}\sin(\pi x) + \tilde{\beta}\cos(\pi x) - 1.$$

με $\tilde{\alpha}, \tilde{\beta} \in \mathbb{R}$. □

Θα δούμε τώρα μερικές βασικές ιδιότητες των λύσεων του προβλήματος (5.1). Ορισμένες από αυτές θα μας οδηγήσουν στην απόδειξη διακριτών αναλόγων για την προσεγγιστική λύση.

Συμβολίζουμε, ως συνήθως, με $(\cdot,\cdot)$ το L^2 εσωτερικό γινόμενο στο $[a, b]$ και με $\|\cdot\|$ την παραγόμενη από αυτό νόρμα, δηλ.

$$(v, w) := \int_a^b v(x)w(x)dx, \qquad \|v\| := (v, v)^{1/2}, \qquad v, w \in C[a, b].$$

Έστω $C_0^m[a, b] := \{v \in C^m[a, b] : v(a) = v(b) = 0\}$ και $v \in C_0^1[a, b]$. Λαμβάνοντας το εσωτερικό γινόμενο και των δύο μελών της εξίσωσης του προβλήματος (5.1) με το v έχουμε

$$-(u'', v) + (qu, v) = (f, v),$$

και, συνεπώς, με ολοκλήρωση κατά μέρη

$$(u', v') + (qu, v) = (f, v) \qquad \forall v \in C_0^1[a, b]. \tag{5.3}$$

Για $v := u$ η (5.3) δίνει αμέσως

$$\|u'\|^2 + (qu, u) = (f, u),$$

οπότε, επειδή η q είναι μη αρνητική, λαμβάνουμε

$$\|u'\|^2 \le (f, u).$$

Εκτιμώντας το δεξιό μέλος αυτής της σχέσης με την ανισότητα των Cauchy–Schwarz παίρνουμε

$$\|u'\|^2 \le \|f\| \, \|u\|. \tag{5.4}$$

Χρησιμοποιώντας τώρα την *ανισότητα των Poincaré–Friedrichs*

$$\|v\| \le (b - a)\|v'\| \qquad \forall v \in C_0^1[a, b], \tag{5.5}$$

η (5.4) δίνει

$$\|u'\| \le (b - a)\|f\|. \tag{5.6}$$

Η (5.5) αποδεικνύεται ως εξής: κατ' αρχάς

$$v(x) = \int_a^x v'(s)ds \qquad x \in [a, b],$$

συνεπώς

$$|v(x)|^2 = (\int_a^x v'(s)ds)^2 \le \int_a^x ds \int_a^x [v'(s)]^2 ds,$$

από την οποία παίρνουμε αμέσως

$$|v(x)|^2 \le (b - a)\|v'\|^2 \qquad x \in [a, b]. \tag{5.7}$$

Ολοκληρώνοντας και τα δύο μέλη της (5.7) στο $[a, b]$ λαμβάνουμε την (5.5). Άμεση συνέπεια της (5.7) είναι και η

$$\max_{a \le x \le b} |v(x)| \le \sqrt{b - a}\, \|v'\| \qquad v \in C_0^1[a, b], \tag{5.8}$$

η οποία λέγεται *ανισότητα του Sobolev*. Από τις (5.5) και (5.6) έπεται αμέσως

$$\|u\| \le (b-a)^2\|f\|. \tag{5.9}$$

Επίσης έχουμε

$$\|u''\| = \|qu - f\| \le \|qu\| + \|f\| \le \max_{a\le x\le b} q(x)\|u\| + \|f\|,$$

συνεπώς, σύμφωνα με την (5.9),

$$\|u''\| \le \left[(b-a)^2 \max_{a\le x\le b} q(x) + 1\right] \|f\|. \tag{5.10}$$

Από τις (5.6), (5.9) και (5.10) έπεται τώρα

$$\|u\| + \|u'\| + \|u''\| \le C\|f\| \tag{5.11}$$

με μία σταθερά C εξαρτώμενη μόνο από τα a, b και τη συνάρτηση q. Η (5.11) αναφέρεται ως *ελλειπτική ομαλότητα* (στην απλούστερη μορφή της).

Από την (5.9) έπεται ότι για $f = 0$ ισχύει $u = 0$, γεγονός το οποίο δείχνει ότι το πρόβλημα (5.1) έχει το πολύ μία λύση. Πραγματικά, αν v και w είναι λύσεις του (5.1), τότε για την $u := v - w$ έχουμε, όπως διαπιστώνει κανείς αμέσως,

$$\begin{cases} -u'' + qu = 0 \quad \text{στο } [a,b] \\ u(a) = u(b) = 0, \end{cases}$$

οπότε, σύμφωνα με την (5.9), $u = 0$, δηλαδή $v = w$.

Παρατήρηση 5.2

(i) Όταν προκαθορίζουμε τις τιμές της λύσης μιας Δ.Ε. στα άκρα του διαστήματος, τότε αυτές οι συνθήκες λέγονται *συνθήκες Dirichlet*. Στην περίπτωση του προβλήματος (5.1) έχουμε *ομογενείς* συνοριακές συνθήκες Dirichlet, γιατί οι προκαθορισμένες τιμές είναι μηδέν. Θεωρούμε τώρα γενικότερα το ακόλουθο πρόβλημα δύο σημείων με μη ομογενείς συνοριακές συνθήκες Dirichlet

$$\begin{cases} -u'' + qu = f \quad \text{στο } [a,b] \\ u(a) = c \\ u(b) = d \end{cases} \tag{5.12}$$

με $c, d \in \mathbb{R}$. Το πρόβλημα αυτό ανάγεται εύκολα σε ένα πρόβλημα της μορφής (5.1). Πραγματικά, κατ' αρχάς η συνάρτηση $w, w(x) := c\frac{b-x}{b-a} + d\frac{a-x}{a-b}$, ικανοποιεί τις συνοριακές συνθήκες του προβλήματος (5.12), $w(a) = c$ και $w(b) = d$. Για τη συνάρτηση $v,\ v := u - w$, έχουμε τώρα προφανώς $v(a) = v(b) = 0$, και επιπλέον

$$-v'' + qv = (-u'' + qu) - (-w'' + qw) = f - qw.$$

Συνεπώς, με $g := f - qw$, η v λύνει το πρόβλημα

$$\begin{cases} -v'' + qv = g & \text{στο } [a, b] \\ v(a) = v(b) = 0 \end{cases} \tag{5.13}$$

που είναι της ίδιας μορφής με το (5.1). Επίσης εύκολα διαπιστώνει κανείς ότι αν η v λύνει το πρόβλημα (5.13), τότε η $u := v + w$ λύνει το πρόβλημα (5.12). Μπορούμε λοιπόν χωρίς περιορισμό της γενικότητας να ασχοληθούμε με την περίπτωση ομογενών συνοριακών συνθηκών.

(ii) Θεωρούμε το ακόλουθο πρόβλημα δύο σημείων με ομογενείς συνοριακές συνθήκες Dirichlet για μία γενικότερη Δ.Ε. δεύτερης τάξης

$$\begin{cases} -(pu')' + qu = f & \text{στο } [a, b] \\ u(a) = u(b) = 0, \end{cases} \tag{5.14}$$

όπου p μία συνεχώς παραγωγίσιμη συνάρτηση με θετικές τιμές, $p \in C^1[a, b], p(x) > 0$ για κάθε $x \in [a, b]$, και q, f όπως στο πρόβλημα (5.1). Και αυτό το πρόβλημα έχει ακριβώς μία λύση. Η απόδειξη της μοναδικότητας της λύσης και της σχέσης (5.11) γι' αυτή την περίπτωση (με σταθερά C που εξαρτάται τώρα και από το p) είναι παρόμοιες των αντιστοίχων για το πρόβλημα (5.1), βλ. την Άσκηση 5.3.

(iii) Τροποποιούμε τώρα τις συνοριακές συνθήκες στο πρόβλημα (5.1) και θεωρούμε το εξής πρόβλημα δύο σημείων με ομογενείς συνοριακές *συνθήκες Neumann*

$$\begin{cases} -u'' + qu = f & \text{στο } [a, b] \\ u'(a) = u'(b) = 0 \end{cases} \tag{5.15}$$

όπου $q, f \in C[a, b]$ και η q λαμβάνει μη αρνητικές τιμές, $q(x) \geq 0$ για κάθε $x \in [a, b]$.

Η συνθήκη $q \neq 0$ είναι αναγκαία στην προκειμένη περίπτωση για να έχουμε μοναδικότητα της λύσης, αφού για $q = 0$ τόσο στην εξίσωση όσο και στις συνοριακές συνθήκες εμφανίζονται μόνο παράγωγοι της u, οι οποίες φυσικά δεν αλλάζουν αν στη u προσθέσουμε κάποια σταθερή συνάρτηση, το πρόβλημα, φερ' ειπείν,

$$\begin{cases} u'' = 0 & \text{στο } [a,b] \\ u'(a) = u'(b) = 0 \end{cases}$$

έχει άπειρες λύσεις της μορφής $u(x) = c$, c σταθερά.

Η μοναδικότητα της λύσης είναι εύκολο να αποδειχθεί και σε αυτή την περίπτωση, βλ. την Άσκηση 5.4.

Γενικότερα μπορούμε να θεωρήσουμε το πρόβλημα

$$\begin{cases} -u'' + qu = f & \text{στο } [a,b] \\ -u'(a) + \sigma_1 u(a) = 0 \\ u'(b) + \sigma_2 u(b) = 0 \end{cases} \tag{5.16}$$

με $q(x) \geq 0$, για κάθε $x \in [a,b]$, και $\sigma_1, \sigma_2 \geq 0$. Στην περίπτωση $q = 0$ υποθέτουμε επιπλέον ότι $\sigma_1 + \sigma_2 > 0$, ούτως ώστε, πέραν των παραγώγων της, να εμφανίζεται και η u σε κάποιο σημείο του προβλήματος. Και σ' αυτή την περίπτωση έχουμε ακριβώς μία λύση, για τη μοναδικότητα βλ. την Άσκηση 5.5.

(iv) Οι συνοριακές συνθήκες Dirichlet λέγονται και *ουσιαστικές συνοριακές συνθήκες* ή *συνοριακές συνθήκες πρώτου είδους*, οι συνθήκες Neumann λέγονται *φυσικές συνοριακές συνθήκες* ή *συνοριακές συνθήκες δευτέρου είδους*, και οι συνοριακές συνθήκες του προβλήματος (5.16) λέγονται *συνοριακές συνθήκες Robin* ή *μεικτές* (*ή τρίτου είδους*) *συνοριακές συνθήκες*.

Φυσικά δεν είναι απαραίτητο να έχουμε και στα δύο άκρα συνοριακές συνθήκες της ίδιας μορφής, μπορούμε να έχουμε λόγου χάριν προβλήματα δύο σημείων όπως το ακόλουθο

$$\begin{cases} -u'' + qu = f & \text{στο } [a,b] \\ u(a) = u'(b) = 0, \end{cases}$$

με f και q όπως στο πρόβλημα (5.1).

(v) Ας υποθέσουμε ότι οι συναρτήσεις q και f μπορούν να επεκταθούν περιοδικά στο $\mathbb{R}$ ως συνεχείς συναρτήσεις με περίοδο $b - a$, ότι ισχύει δηλαδή $q(b) = q(a)$ και

$f(b) = f(a)$. Αν θέλουμε οι επεκτάσεις αυτές να είναι ομαλές, πρέπει να απαιτήσουμε ισότητα αντιστοίχων παραγώγων στα άκρα. Σ' αυτή την περίπτωση μπορούμε να θεωρήσουμε το ακόλουθο *περιοδικό* πρόβλημα: Ζητείται μία περιοδική συνάρτηση u, $u \in C^2(\mathbb{R})$, με περίοδο $b - a$, τέτοια ώστε

$$-u'' + qu = f \quad \text{στο } \mathbb{R}. \tag{5.17}$$

Αν η q λαμβάνει μη αρνητικές τιμές, $q(x) \geq 0$ για κάθε $x \in \mathbb{R}$, και $q \neq 0$, τότε το πρόβλημα αυτό έχει μία ακριβώς λύση. □

Ασκήσεις

5.1. Στην απόδειξη της ανισότητας (5.5) χρησιμοποιήσαμε μόνο το γεγονός ότι $v \in C^1[a, b]$ και $v(a) = 0$. Χρησιμοποιήστε τώρα και το γεγονός ότι $v(b) = 0$ για να βελτιώσετε τη σταθερά σ' αυτή την ανισότητα από $b - a$ σε $(b - a)/2$, αποδείξτε δηλαδή ότι

$$\|v\| \leq \frac{b-a}{2}\|v'\| \qquad \forall v \in C_0^1[a, b]. \tag{5.18}$$

[*Υπόδειξη*: Χρησιμοποιήστε τις σχέσεις

$$v(x) = \int_a^x v'(s)ds \qquad x \in [a, \frac{a+b}{2}]$$
$$v(x) = -\int_x^b v'(s)ds \qquad x \in [\frac{a+b}{2}, b].]$$

Παρατηρήστε ότι

$$|v(x)|^2 \leq (x-a)\int_a^{\frac{a+b}{2}} |v'(s)|^2 ds \qquad x \in [a, \frac{a+b}{2}],$$

και αντίστοιχα για το διάστημα $[\frac{a+b}{2}, b]$, για να αντικαταστήσετε το 2 στην (5.18) με $2\sqrt{2}$.

5.2. Εργαστείτε ανάλογα για να βελτιώσετε τη σταθερά στην ανισότητα (5.8) από $\sqrt{b-a}$ σε $\sqrt{\frac{b-a}{2}}$.

5.3. Αποδείξτε μία ανισότητα της μορφής (5.11) για το πρόβλημα (5.14). [*Υπόδειξη*:

Αποδείξτε κατ' αρχάς ότι υπάρχει θετική σταθερά c τέτοια ώστε $p(x) \geq c \quad \forall x \in [a,b]$.]

5.4. Πολλαπλασιάστε τη Δ.Ε. του προβλήματος (5.15) επί u, ολοκληρώστε στο $[a,b]$ και ολοκληρώστε κατά μέρη. Συμπεράντε ότι το πρόβλημα έχει μοναδική λύση. Υποθέτοντας τώρα επιπλέον ότι η q είναι γνήσια θετική στο $[a,b]$, αποδείξτε μία ανισότητα της μορφής (5.11).

5.5. Αποδείξτε ότι το πρόβλημα (5.16) έχει το πολύ μία λύση.

5.6. Θεωρούμε το ακόλουθο πρόβλημα δύο σημείων

$$\begin{cases} -u'' + ru' + qu = f \quad \text{στο } [a,b] \\ u(a) = u(b) = 0, \end{cases} \tag{5.19}$$

όπου $r, q, f \in C[a,b], \quad q(x) \geq 0 \quad \forall x \in [a,b]$. Αποδείξτε ότι πολλαπλασιάζοντας τη Δ.Ε. στο (5.19) με $\mathrm{e}^{-\int_a^x r(s)ds}$ μπορούμε να αναγάγουμε το πρόβλημα αυτό σε ένα πρόβλημα της μορφής (5.14).

5.7. Θεωρούμε το πρόβλημα δύο σημείων

$$\begin{cases} -u'' + ru' + qu = 0 \quad \text{στο } [a,b] \\ u(a) = c \\ u(b) = d \end{cases} \tag{5.20}$$

με $r, q \in C[a,b]$ και $q(x) > 0 \quad \forall x \in [a,b]$. Αποδείξτε την εξής *αρχή μεγίστου*

$$\max_{a \leq x \leq b} |u(x)| \leq \max(|c|, |d|).$$

[*Υπόδειξη*: Αν η u δεν είναι σταθερά, αποδείξτε ότι δεν μπορεί να έχει θετικό τοπικό μέγιστο ή αρνητικό τοπικό ελάχιστο στο (a,b).]

5.8. Αποδείξτε για το πρόβλημα (5.17) μία ανισότητα της μορφής (5.11) και οδηγηθείτε στο συμπέρασμα ότι αυτό έχει το πολύ μία λύση. [*Υπόδειξη*: Αποδείξτε πρώτα ότι για $v \in C^1[a,b]$ ισχύει

$$\|v\|^2 \leq c\{\min_{a \leq x \leq b} |v(x)|^2 + \|v'\|^2\}$$

με μία σταθερά c ανεξάρτητη του v.]

5.9. Αποδείξτε για το πρόβλημα (5.15) ότι ισχύει

$$\max_{a \leq x \leq b} |u(x)| \leq C \max_{a \leq x \leq b} |f(x)|$$

με μία σταθερά C εξαρτώμενη μόνο από τα a, b και q. [*Υπόδειξη*: Έστω $v \in C^1[a,b]$. Αποδείξτε τότε ότι

$$\max_{a \le x \le b} |v(x)| \le |v(y)| + \sqrt{b-a}\ \ \|v'\| \qquad \forall y \in [a,b].$$

Επιπλέον έχουμε προφανώς

$$(qv, v) \ge \min_{a \le x \le b} |v(x)|^2 \quad \int_a^b q(x)dx.\,]$$

5.10. Αποδείξτε για το πρόβλημα (5.15) μία ανισότητα της μορφής (5.11), γενικεύοντας έτσι το αποτέλεσμα της Άσκησης 5.4 και για την περίπτωση που $\min_{a \le x \le b} q(x) = 0$.

5.11. Πώς μπορεί το πρόβλημα

$$\begin{cases} -u'' + qu = f & \text{στο } [a,b] \\ u'(a) = c \\ u'(b) = d \end{cases}$$

να αναχθεί σε ένα αντίστοιχο πρόβλημα με ομογενείς συνοριακές συνθήκες Neumann; [*Υπόδειξη*: Τι ιδιότητες έχει η συνάρτηση $v(x) := c(x-a) + \frac{d-c}{2(b-a)}(x-a)^2, \quad x \in [a,b]$;]

5.12. Έστω $\rho > 0$ και $x^\star \in (a,b)$. Θεωρούμε το εξής πρόβλημα δύο σημείων με διεπιφάνεια: Ζητείται μία συνεχής συνάρτηση $u : [a,b] \to \mathbb{R}$, ομαλή στα υποδιαστήματα $[a, x^\star]$ και $[x^\star, b]$, και τέτοια ώστε

$$\begin{cases} -u'' + qu = f & \text{στο } [a, x^\star) \cup (x^\star, b] \\ u'(x^\star-) = \rho u'(x^\star+) \\ u(a) = u(b) = 0, \end{cases}$$

όπου q, f συνεχείς συναρτήσεις στα $[a, x^\star)$, $(x^\star, b]$ οι οποίες μπορούν να επεκταθούν συνεχώς στα $[a, x^\star]$ και $[x^\star, b]$, και $q(x) \ge 0$ για κάθε $x \in [a,b]$. Αποδείξτε ότι το πρόβλημα αυτό έχει το πολύ μία ομαλή λύση. [*Υπόδειξη*. Χρησιμοποιήστε το εσωτερικό γινόμενο $(\cdot,\cdot)$, $(v,w) := \int_a^{x^\star} v(x)w(x)dx + \rho \int_{x^\star}^b v(x)w(x)dx$.]

5.13. Έστω $f : \mathbb{R} \to \mathbb{R}$ μία συνεχώς παραγωγίσιμη συνάρτηση τέτοια ώστε $-\gamma \le f'(s) \le 0$ για κάθε $s \in \mathbb{R}$ με μία θετική σταθερά γ. Αποδείξτε ότι το *μη γραμμικό* πρόβλημα δύο σημείων

$$\begin{cases} -u'' = f(u) \qquad \text{στο} \quad [a,b] \\ u(a) = u(b) = 0 \end{cases}$$

έχει το πολύ μία ομαλή λύση.

5.14. Έστω $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ μία συνεχώς παραγωγίσιμη συνάρτηση τέτοια ώστε $-\gamma \le \partial_2 f(x,s) \le 0$ για κάθε $x \in [a,b], s \in \mathbb{R}$, με μία θετική σταθερά γ. Αποδείξτε ότι το *μη γραμμικό* πρόβλημα δύο σημείων

$$\begin{cases} -(pu')' = f(x,u) \qquad \text{στο} \quad [a,b], \\ u(a) = u(b) = 0, \end{cases}$$

με $p \in C^1[a,b], p(x) > 0$ για κάθε $x \in [a,b]$, έχει το πολύ μία ομαλή λύση.

5.15. Εισάγουμε στον $C[a,b]$ την "αρνητική" (δηλ. αρνητικής τάξης) νόρμα $\|\cdot\|_{-1}$ δια

$$\|f\|_{-1} := \{\int_a^b [\int_a^x f(s)ds]^2 dx\}^{1/2}.$$

Για $f \in C[a,b]$ και $v \in C_0[a,b]$, αποδείξτε την εξής ανισότητα "τύπου Cauchy–Schwarz"

$$|(f,v)| \le \|f\|_{-1}\,\|v'\|.$$

Αποδείξτε για τη λύση u του προβλήματος (5.1) μία εκτίμηση της μορφής

$$\|u\| + \|u'\| \le C\|f\|_{-1} \tag{5.21}$$

με μία σταθερά C εξαρτώμενη μόνο από τα a, b και τη συνάρτηση q.

5.16. Αποδείξτε ότι

$$\|v\| \le \frac{b-a}{\sqrt{2}}\|v'\| \qquad \forall v \in C_0^1[a,b].$$

[*Υπόδειξη*: Λόγω του ότι $v(a) = 0$, έχουμε

$$|v(x)|^2 = |\int_a^x v'(s)ds|^2 \le (x-a)\int_a^x |v'(s)|^2 ds,$$

δηλαδή

$$|v(x)|^2 \le (x-a)\|v'\|^2.$$

Ολοκληρώστε ως προς x στο διάστημα [a,b].]

5.17. Βελτιώστε την εκτίμηση στην προηγούμενη Άσκηση, εκμεταλευόμενοι και το γεγονός ότι $v(b) = 0$, αποδεικνύοντας ότι

$$\|v\| \leq \frac{b-a}{2\sqrt{2}}\|v'\| \qquad \forall v \in C_0^1[a,b].$$

[*Υπόδειξη*: Συνδυάστε τις Υποδείξεις στις Ασκήσεις 5.1 και 5.16.]

5.18. Θεωρούμε το πρόβλημα της Άσκησης 5.6, και υποθέτουμε ότι η συνάρτηση r είναι σταθερή. Αποδείξτε απ' ευθείας, χωρίς να γράψετε το πρόβλημα στη μορφή (5.14), ότι το πρόβλημα έχει το πολύ μία ομαλή λύση. [*Υπόδειξη*: Χρησιμοποιήστε το γεγονός ότι $(u', u) = 0$.]

5.19. Έστω $f : \mathbb{R} \to \mathbb{R}$ μία περιοδική συνάρτηση με περίοδο ίση με ένα. Θεωρούμε το ακόλουθο περιοδικό πρόβλημα: Ζητείται μία περιοδική συνάρτηση $u : \mathbb{R} \to \mathbb{R}$, με περίοδο ένα, τέσσερεις φορές συνεχώς παραγωγίσιμη, τέτοια ώστε

$$u^{(4)} + u = f \qquad \text{στο} \ \ \mathbb{R}.$$

Αποδείξτε ότι το πρόβλημα αυτό έχει το πολύ μία ομαλή λύση.

2. Πεπερασμένες διαφορές για το πρόβλημα δύο σημείων

Έστω $v \in C^4[a,b]$, $x \in (a,b)$ και $h > 0$ τέτοιο ώστε $x+h,\ x-h \in [a,b]$. Αναπτύσσοντας κατά Taylor έχουμε

$$\tag{6.1} \begin{cases} v(x+h) = v(x) + hv'(x) + \dfrac{h^2}{2}v''(x) + \dfrac{h^3}{6}v'''(x) + \dfrac{h^4}{24}v^{(4)}(\vartheta_1) \\ v(x-h) = v(x) - hv'(x) + \dfrac{h^2}{2}v''(x) - \dfrac{h^3}{6}v'''(x) + \dfrac{h^4}{24}v^{(4)}(\vartheta_2) \end{cases}$$

με $\vartheta_1, \vartheta_2 \in (x-h, x+h)$. Προσθέτοντας αυτές τις σχέσεις λαμβάνουμε

$$v(x-h) - 2v(x) + v(x+h) = h^2 v''(x) + \frac{h^4}{24}[v^{(4)}(\vartheta_1) + v^{(4)}(\vartheta_2)]$$

ή, χρησιμοποιώντας το θεώρημα της ενδιάμεσης τιμής,

$$v(x-h) - 2v(x) + v(x+h) = h^2 v''(x) + \frac{h^4}{12}v^{(4)}(\xi)$$

με $\xi \in (x-h, x+h)$, και συμπεραίνουμε ότι

$$\tag{6.2} |\frac{v(x-h) - 2v(x) + v(x+h)}{h^2} - v''(x)| \le h^2 \frac{1}{12} \max_{a \le s \le b} |v^{(4)}(s)|.$$

Συνεπώς, για αρκετά μικρό h, η τιμή της δευτέρας παραγώγου της v σε ένα σημείο x προσεγγίζεται με μεγάλη ακρίβεια από την πεπερασμένη διαφορά $\frac{1}{h^2}[v(x-h) - 2v(x) + v(x+h)]$. Αυτό μας οδηγεί στην ακόλουθη μέθοδο πεπερασμένων διαφορών για το πρόβλημα δύο σημείων (5.1): Έστω $J \in \mathbb{N}$, $h := \frac{b-a}{J+1}$, και $x_i := a + ih$, $i = 0, \dots, J+1$, ο ομοιόμορφος διαμερισμός του διαστήματος $[a,b]$ με βήμα h. Έστω $\mathbb{R}_0^{J+2} := \{w = (w_0, \dots, w_{J+1})^T \in \mathbb{R}^{J+2} : w_0 = w_{J+1} = 0\}$. Προσεγγίζουμε το διάνυσμα $(u(x_0), u(x_1), \dots, u(x_{J+1}))^T$ με ένα διάνυσμα $U \in \mathbb{R}_0^{J+2}$ για το οποίο ισχύει

$$\tag{6.3} -\frac{1}{h^2}(U_{i-1} - 2U_i + U_{i+1}) + q(x_i)U_i = f(x_i), \quad i = 1, \dots, J.$$

Κατ' αυτόν τον τρόπο ορίζουμε μία προσέγγιση U_i για την τιμή της λύσης u του προβλήματος (5.1) στο σημείο x_i, δηλαδή για το $u(x_i)$. Στο *σχήμα* (6.3) (δηλαδή στη μέθοδο (6.3)) καταλήξαμε προσεγγίζοντας την τιμή $u''(x_i)$ της δευτέρας παραγώγου της λύσης u του προβλήματος (5.1) με μία πεπερασμένη διαφορά, και συγκεκριμένα με $\frac{1}{h^2}(u(x_{i-1}) - 2u(x_i) + u(x_{i+1}))$, και γιαυτό μιλάμε για *μεθόδους πεπερασμένων διαφορών*. Σημειώνουμε ότι η προσεγγιστική λύση ικανοποιεί τις συνοριακές συνθήκες του προβλήματος (5.1)· πράγματι έχουμε $U_0 = U_{J+1} = 0$. Το (6.3) είναι ένα $J \times J$ γραμμικό σύστημα με αγνώστους $U_1, \dots, U_J$, και έχει την εξής μορφή

$$A \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_{J-1} \\ U_J \end{pmatrix} = h^2 \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_{J-1}) \\ f(x_J) \end{pmatrix} \tag{6.4}$$

με πίνακα A,

$$A := \begin{pmatrix} 2+h^2q(x_1) & -1 & & & 0 \\ -1 & 2+h^2q(x_2) & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2+h^2q(x_{J-1}) & -1 \\ 0 & & & -1 & 2+h^2q(x_J) \end{pmatrix}.$$

Κατ' αρχάς ο πίνακας A του γραμμικού συστήματος (6.4) είναι τριδιαγώνιος, δηλ. $a_{ij} = 0$ για $|i-j| > 1$. Είναι επίσης συμμετρικός, δηλ. $a_{ij} = a_{ji}$. Επιπλέον μπορεί να αποδειχθεί ότι ο A είναι αντιστρέψιμος, γεγονός που συνεπάγεται ύπαρξη και μοναδικότητα της προσεγγιστικής λύσης U. Μάλιστα ο A είναι θετικά ορισμένος, στην περίπτωση δε που η συνάρτηση q λαμβάνει μόνο θετικές τιμές, δηλαδή $q(x) > 0$ για κάθε $x \in [a,b]$, ο A έχει επιπλέον αυστηρά κυριαρχική διαγώνιο. Οι ιδιότητες αυτές του A δεν είναι μόνο για θεωρητικούς σκοπούς χρήσιμες, αλλά είναι και πρακτικά πολύ σημαντικές. Γραμμικά συστήματα αυτής της μορφής είναι εύκολο να επιλυθούν γρήγορα και με μεγάλη ακρίβεια με κατάλληλες αριθμητικές μεθόδους. Για έναν κατάλληλο αλγόριθμο για την αριθμητική επίλυση του γραμμικού συστήματος (6.4), τον τριδιαγώνιο αλγόριθμο, παραπέμπουμε στο [1, §3.4]. Σε πολλά από τα ζητήματα που απλώς εθίχθησαν εδώ θα επανέλθουμε με περισσότερες λεπτομέρειες είτε στη θεωρία είτε στις ασκήσεις.

Εισάγουμε τώρα στον $\mathbb{R}_0^{J+2}$ το διακριτό L^2 εσωτερικό γινόμενο $(\cdot,\cdot)_h$ δια

$$(v,w)_h := h\sum_{i=1}^{J} v_i w_i,$$

και συμβολίζουμε με $\|\cdot\|_h$ την παραγόμενη νόρμα, δηλ.

$$\|v\|_h = \big(h\sum_{i=1}^{J}|v_i|^2\big)^{1/2}.$$

Σχόλιο. Αν δεν υπήρχε ο παράγοντας h, που πολλαπλασιάζει το άθροισμα στον ορισμό του διακριτού εσωτερικού γινομένου, τότε αυτό θα συνέπιπτε με το Ευκλείδειο εσωτερικό γινόμενο στον $\mathbb{R}_0^{J+2}$. Έστω $v,w:[a,b]\to\mathbb{R}$ δύο ομαλές συναρτήσεις. Το L^2 εσωτερικό γινόμενό τους είναι

$$(v,w) := \int_a^b v(x)w(x)\,dx.$$

Προσεγγίζουμε τώρα το ολοκλήρωμα με τον σύνθετο τύπο του τραπεζίου με βήμα h και παίρνουμε

$$(v,w) \approx \frac{1}{2}hv(x_0)w(x_0) + h\sum_{i=1}^{J} v(x_i)w(x_i) + \frac{1}{2}hv(x_{J+1})w(x_{J+1}),$$

ιδιαίτερα λοιπόν, στην περίπτωση που οι συναρτήσεις μηδενίζονται στα άκρα του διαστήματος,

$$(v,w) \approx h\sum_{i=1}^{J} v(x_i)w(x_i).$$

Θεωρώντας τώρα τα διανύσματα $v=(v(x_0),v(x_1),\dots,v(x_{J+1}))^T$, $w=(w(x_0),w(x_1),\dots,w(x_{J+1}))^T\in\mathbb{R}_0^{J+2}$, το δεξιό μέλος αυτής της σχέσης λαμβάνει τη μορφή

$$h\sum_{i=1}^{J} v_i w_i$$

και αυτό ακριβώς είναι το κίνητρό μας για τον ορισμό του διακριτού εσωτερικού γινομένου $(\cdot,\cdot)_h$ κατά τον συγκεκριμένο τρόπο.

Επίσης θα χρησιμοποιήσουμε τη νόρμα $|\cdot|_{1,h}$ στον $\mathbb{R}_0^{J+2}$ που είναι το διακριτό ανάλογο της L^2 νόρμας της παραγώγου μίας συνεχώς παραγωγίσιμης συνάρτησης και δίνεται ως εξής

$$|v|_{1,h} := \Big(h\sum_{i=0}^{J}|\frac{v_{i+1}-v_i}{h}|^2\Big)^{1/2}.$$

Έστω $v\in\mathbb{R}_0^{J+2}$. Για $j\in\{1,\dots,J\}$ έχουμε, χρησιμοποιώντας την ανισότητα των Cauchy–Schwarz,

$$\begin{aligned}|v_j|^2 &= |\sum_{i=0}^{j-1}(v_{i+1}-v_i)|^2 \le j\sum_{i=0}^{j-1}|v_{i+1}-v_i|^2 \le (J+1)\sum_{i=0}^{J}|v_{i+1}-v_i|^2,\\ &= (J+1)h^2\sum_{i=0}^{J}|\frac{v_{i+1}-v_i}{h}|^2 = (b-a)h\sum_{i=0}^{J}|\frac{v_{i+1}-v_i}{h}|^2,\end{aligned}$$

δηλ.

(6.5) $$\max_{0\le i\le J+1}|v_i| \le \sqrt{b-a}\,|v|_{1,h} \qquad \forall v\in\mathbb{R}_0^{J+2},$$

που είναι μία *διακριτή ανισότητα του Sobolev*, το διακριτό αντίστοιχο της (5.8). Από την (6.5) έπεται αμέσως ότι

(6.6) $$\|v\|_h \le (b-a)|v|_{1,h} \qquad \forall v\in\mathbb{R}_0^{J+2},$$

που είναι μία *διακριτή ανισότητα των Poincaré–Friedrichs*, το διακριτό ανάλογο της (5.5).

Εισάγουμε τώρα τον τελεστή $\Delta_h:\mathbb{R}_0^{J+2}\to\mathbb{R}_0^{J+2}$ διά

$$\Delta_h v_i := (\Delta_h v)_i := \begin{cases}\dfrac{v_{i-1}-2v_i+v_{i+1}}{h^2}, & i=1,\dots,J,\\ 0, & i=0,\ J+1,\end{cases}$$

και γράφουμε την (6.3) στη μορφή

(6.7) $$-\Delta_h U_i + q(x_i)U_i = f(x_i),\quad i=1,\dots,J.$$

Τώρα, για $v, w \in \mathbb{R}_0^{J+2}$, έχουμε

$$\begin{aligned}(\Delta_h v, w)_h &= h\sum_{i=1}^{J}\frac{v_{i-1}-2v_i+v_{i+1}}{h^2}w_i\\ &= h\{\sum_{i=1}^{J}\frac{v_{i-1}-v_i}{h^2}w_i - \sum_{i=1}^{J}\frac{v_i-v_{i+1}}{h^2}w_i\} =\\ &= h\{\sum_{i=0}^{J-1}\frac{v_i-v_{i+1}}{h^2}w_{i+1} - \sum_{i=1}^{J}\frac{v_i-v_{i+1}}{h^2}w_i\} =\\ &= h\{\sum_{i=0}^{J}\frac{v_i-v_{i+1}}{h^2}w_{i+1} - \sum_{i=0}^{J}\frac{v_i-v_{i+1}}{h^2}w_i\},\end{aligned}$$

δηλ.

(6.8) $$-(\Delta_h v, w)_h = h\sum_{i=0}^{J}\frac{v_i-v_{i+1}}{h}\,\frac{w_i-w_{i+1}}{h} \qquad \forall v, w \in \mathbb{R}_0^{J+2},$$

και ιδιαίτερα

(6.8′) $$-(\Delta_h v, v)_h = |v|_{1,h}^2 \qquad \forall v \in \mathbb{R}_0^{J+2}.$$

Σημειώστε ότι η (6.8′) αποτελεί διακριτό ανάλογο της σχέσης

$$-(v'', v) = \|v'\|^2 \qquad \forall v \in C_0^2[a,b].$$

Σε αυτή τη σχέση οδηγούμαστε με ολοκλήρωση κατά μέρη, στην (6.8′) οδηγούμαστε αθροίζοντας κατά μέρη.

Ευστάθεια. Πολλαπλασιάζοντας την (6.7) επί hU_i, αθροίζοντας και χρησιμοποιώντας την (6.8) παίρνουμε

(6.9) $$|U|_{1,h}^2 + h\sum_{i=1}^{J}q(x_i)|U_i|^2 = h\sum_{i=1}^{J}f(x_i)U_i.$$

Επειδή ισχύει $q(x_i) \geq 0$, με την ανισότητα των Cauchy–Schwarz λαμβάνουμε

$$|U|_{1,h}^2 \leq \{h\sum_{i=1}^{J}|f(x_i)|^2\}^{1/2}\|U\|_h$$

και, συνεπώς, χρησιμοποιώντας την (6.6), συμπεραίνουμε

$$|U|_{1,h} \leq (b-a)\{h\sum_{i=1}^{J}|f(x_i)|^2\}^{1/2}. \tag{6.10}$$

Οι (6.5) και (6.10) δίνουν τώρα

$$\max_{0\leq i\leq J+1}|U_i| \leq (b-a)^{3/2}\{h\sum_{i=1}^{J}|f(x_i)|^2\}^{1/2}. \tag{6.11}$$

Η ανισότητα (6.11) αναφέρεται ως *ευστάθεια* της μεθόδου (6.3) και θα παίξει σημαντικό ρόλο στη συνέχεια στην απόδειξη της σύγκλισης της μεθόδου. Ένα άλλο συμπέρασμα που εξάγεται από την (6.11) είναι ότι η προσεγγιστική λύση U ορίζεται μοναδικά. Πραγματικά το πρόβλημά μας είναι ένα $J \times J$ γραμμικό σύστημα και το αντίστοιχο ομογενές έχει μόνο την τετριμμένη λύση, αφού αν υποθέσουμε ότι $f(x_i) = 0$, $i = 1,\dots,J$, συμπεραίνουμε, σύμφωνα με την (6.11), ότι $U_i = 0$, $i = 1,\dots,J$.

Γενικά λέμε ότι μία μέθοδος είναι *ευσταθής*, όταν η μεταβολή της αριθμητικής λύσης σε μία νόρμα φράσσεται από μία σταθερά επί τη μεταβολή των δεδομένων. Όταν αυτό ισχύει, τότε μικρές μεταβολές στα δεδομένα έχουν ως αποτέλεσμα μικρές μεταβολές στην αριθμητική λύση. Στην ειδική περίπτωση που το πρόβλημα είναι γραμμικό, ζητάμε η αριθμητική λύση να φράσσεται με μία σταθερά επί τα δεδομένα, όλα φυσικά μετρώμενα με νόρμες. Εύκολα διαπιστώνουμε τότε (πάντα στην περίπτωση γραμμικών προβλημάτων) ότι μικρές μεταβολές στα δεδομένα έχουν ως αποτέλεσμα μικρές μεταβολές στη λύση, βλ. την Άσκηση 6.1. Αυτό εξηγεί γιατί η (6.11) αναφέρεται ως ευστάθεια του σχήματος (6.3). Σημειώνουμε επίσης ότι η ευστάθεια είναι εσωτερική ιδιότητα του σχήματος, δεν έχει δηλαδή σχέση με το πρόβλημα που θέλουμε να προσεγγίσουμε. Φυσικά η (6.11) δεν είναι η μόνη ιδιότητα ευστάθειας του σχήματος (6.3)· ιδιότητα ευστάθειας είναι και η (6.10), στην οποία μετράμε τη μεταβολή της λύσης σε άλλη νόρμα. Μια άλλη βασική ιδιότητα αριθμητικών μεθόδων είναι η λεγόμενη *συνέπεια*. Σε αντίθεση με την ευστάθεια, η οποία όπως είπαμε είναι εσωτερική ιδιότητα της αριθμητικής μεθόδου, η συνέπεια αναφέρεται πάντα στη λύση ενός συγκεκριμένου προβλήματος, το ίδιο αριθμητικό σχήμα μπορεί να είναι συνεπές για ένα πρόβλημα και μη συνεπές για ένα άλλο.

Συνέπεια. Αν αντικαταστήσουμε την προσεγγιστική λύση U στην (6.3) με την ακριβή λύση u, τότε δεν θα έχουμε φυσικά ισότητα αλλά θα υπάρχει ένα σφάλμα. Αν το σφάλμα αυτό τείνει στο μηδέν, καθώς το h τείνει στο μηδέν, τότε η μέθοδος λέγεται *συνεπής* (για το πρόβλημα (5.1) εννοείται). Αυτό σημαίνει ότι, καθώς το h τείνει στο μηδέν, το σχήμα μας "τείνει" στη διαφορική εξίσωση. Εύκολα διαπιστώνουμε ότι

$$-\frac{u(x_{i-1}) - 2u(x_i) + u(x_{i+1})}{h^2} + q(x_i)u(x_i) = f(x_i) + r_i, \quad i = 1, \dots, J, \tag{6.12}$$

με

$$\max_{1 \le i \le J} |r_i| \le Ch^2, \tag{6.13}$$

βλ. την (6.2). Επομένως η μέθοδος (6.3) είναι συνεπής. Σημειώνουμε ότι, όπως και στην ευστάθεια, είναι δυνατόν να μετράμε και το σφάλμα συνέπειας σε διάφορες νόρμες, οπότε η μέθοδος μπορεί να είναι συνεπής ως προς κάποια νόρμα και μη συνεπής ως προς κάποια άλλη.

Μια προσεγγιστική ιδιότητα της U δίνεται στο επόμενο θεώρημα.

Θεώρημα 6.1 *Έστω ότι η λύση u του προβλήματος* (5.1) *είναι αρκετά ομαλή,* $u \in C^4[a, b]$. *Τότε υπάρχει μία σταθερά c ανεξάρτητη του h (η οποία εξαρτάται μόνο από τα a, b και το* $\max_x |u^{(4)}(x)|$) *τέτοια ώστε*

$$\max_{1 \le i \le J} |u(x_i) - U_i| \le ch^2, \tag{6.14}$$

όπου $U \in \mathbb{R}_0^{J+2}$ είναι η λύση του (6.3).

Απόδειξη. Συμβολίζουμε με $e \in \mathbb{R}_0^{J+2}$, $e_i := u(x_i) - U_i$, $i = 0, \dots, J+1$, το σφάλμα προσέγγισης. Αφαιρώντας κατά μέλη τις (6.12) και (6.3) έχουμε

$$-\Delta_h e_i + q(x_i)e_i = r_i, \quad i = 1, \dots, J. \tag{6.15}$$

Θέτουμε $r_0 := r_{J+1} := 0$, και $r := (r_0, \dots, r_{J+1})^T$. Πολλαπλασιάζοντας την (6.15) επί he_i, αθροίζοντας και χρησιμοποιώντας την (6.8) έχουμε

$$|e|_{1,h}^2 + h \sum_{i=1}^{J} q(x_i)|e_i|^2 = (r, e)_h,$$

δηλαδή

$$|e|_{1,h}^2 \leq (r,e)_h,$$

οπότε

$$|e|_{1,h}^2 \leq \|r\|_h \|e\|_h.$$

Χρησιμοποιώντας τώρα την (6.6) και στη συνέχεια την (6.5) λαμβάνουμε

$$\max_{0\leq i\leq J+1} |e_i| \leq (b-a)^{\frac{3}{2}} \|r\|_h. \tag{6.16}$$

Τώρα προφανώς

$$\|r\|_h \leq \sqrt{b-a}\, \max_{0\leq i\leq J} |r_i|,$$

οπότε, σύμφωνα με την (6.13),

$$\|r\|_h \leq Ch^2. \tag{6.17}$$

Οι (6.16) και (6.17) δίνουν τώρα αμέσως την (6.14). □

Το Θεώρημα 6.1 δεν μας λέει απλώς ότι η προσεγγιστική λύση συγκλίνει προς τη λύση καθώς το h τείνει στο μηδέν, αλλά μας δίνει και μία εκτίμηση του σφάλματος της τάξης του h^2. Λέμε ότι το σχήμα μας έχει *τάξη* δύο. Αυτό σημαίνει ότι, για αρκετά μικρό h, όταν το h υποδιπλασιάζεται, το σφάλμα περίπου υποτετραπλασιάζεται.

Όπως θα έχει ήδη διαπιστώσει ο προσεκτικός αναγνώστης, η απόδειξη της σύγκλισης, δηλ. του Θεωρήματος 6.1, είναι στην ουσία ένας συνδυασμός της ευστάθειας και της συνέπειας. Η συνέπεια της μεθόδου χρησιμοποιείται αυτούσια, ενώ από την ευστάθεια χρησιμοποιούμε τον τρόπο με τον οποίο αυτή αποδεικνύεται.

Παρατήρηση 6.1 Έστω προς στιγμήν ότι η q λαμβάνει μόνο θετικές τιμές, $q(x) > 0$ για κάθε $x \in [a,b]$. Τότε $q^\star := \min_x q(x) > 0$. Γράφοντας την (6.15) στη μορφή

$$[2 + h^2 q(x_i)]e_i = e_{i-1} + e_{i+1} + h^2 r_i, \quad i = 1,\dots,J,$$

και λαμβάνοντας υπ' όψιν το γεγονός ότι $e_0 = e_{J+1} = 0$, διαπιστώνουμε ότι

$$[2 + h^2 q(x_i)]\,|e_i| \leq 2 \max_{1\leq i\leq J} |e_i| + h^2 |r_i|, \quad i = 1,\dots,J,$$

και συμπεραίνουμε ότι

$$q^{\star} \max_{1\le i\le J} |e_i| \le \max_{1\le i\le J} |r_i|. \tag{6.18}$$

Οι (6.18) και (6.13) δίνουν τώρα πάλι την (6.14) και αυτό αποτελεί μία άλλη, απλούστερη απόδειξη του Θεωρήματος 6.1 στην περίπτωση που $\min_x q(x) > 0$. Η ουσία της απόδειξης αυτής έγκειται στο γεγονός ότι στην προκειμένη περίπτωση ο πίνακας του γραμμικού συστήματος (6.3) έχει αυστηρά κυριαρχική διαγώνιο. Χρησιμοποιώντας προχωρημένες τεχνικές της Αριθμητικής Γραμμικής Άλγεβρας, είναι δυνατόν να συμπληρώσουμε αυτή την απόδειξη για να καλύψει και την περίπτωση $\min_x q(x) = 0$. Η απόδειξη που δώσαμε στο Θεώρημα 6.1 μπορεί όμως να γενικευθεί εύκολα και να δώσει εκτιμήσεις βέλτιστης τάξης και σε άλλες περιπτώσεις, όπως π.χ. στην περίπτωση μη ομοιομόρφου διαμερισμού, σε αντίθεση με την απόδειξη που σχολιάσαμε σε αυτή την παρατήρηση. □

Παρατήρηση 6.2 Θεωρούμε τώρα το πρόβλημα (5.14). Το αντίστοιχο του σχήματος (6.3) είναι σ' αυτή την περίπτωση

$$-\frac{1}{h^2}[p(x_i+\frac{h}{2})(U_{i+1}-U_i) - p(x_i-\frac{h}{2})(U_i-U_{i-1})] + q(x_i)U_i = f(x_i), \quad i = 1,\dots,J, \tag{6.19}$$

όπου χρησιμοποιήσαμε τους ίδιους συμβολισμούς με το σχήμα (6.3). Εντελώς ανάλογα αποτελέσματα ισχύουν και σ' αυτή την περίπτωση, οι αποδείξεις είναι παρόμοιες γιαυτό και αφήνονται ως ασκήσεις. Σημειώνουμε μόνο ότι το $\frac{1}{h}p(x_i+\frac{h}{2})[u(x_{i+1})-u(x_i)]$ προσεγγίζει την $p(x_i+\frac{h}{2})u'(x_i+\frac{h}{2})$, το $\frac{1}{h}p(x_i-\frac{h}{2})[u(x_i)-u(x_{i-1})]$ προσεγγίζει την $p(x_i-\frac{h}{2})u'(x_i-\frac{h}{2})$ και η διαφορά των δύο διαιρούμενη δια h προσεγγίζει την τιμή της συνάρτησης $(pu')'$ στο σημείο x_i. □

Άλλες συνοριακές συνθήκες

Θεωρούμε κατ' αρχάς το πρόβλημα (5.15). Οι συνοριακές συνθήκες είναι δηλαδή τώρα συνθήκες Neumann. Έστω πάλι $J \in \mathbb{N}, h := \frac{b-a}{J+1}$, και $x_i = a+ih$, $i = 0,\dots,J+1$. Προσεγγίζουμε το διάνυσμα $(u(x_0), u(x_1),\dots,u(x_{J+1}))^T$ με ένα διάνυσμα $U \in \mathbb{R}^{J+2}$

για το οποίο ισχύει

$$
\begin{cases}
2\dfrac{U_0-U_1}{h^2}+q(x_0)U_0=f(x_0)\\
-\dfrac{1}{h^2}(U_{i-1}-2U_i+U_{i+1})+q(x_i)U_i=f(x_i), \quad i=1,\dots,J,\\
-2\dfrac{U_J-U_{J+1}}{h^2}+q(x_{J+1})U_{J+1}=f(x_{J+1}).
\end{cases}
\tag{6.20}
$$

Πριν προχωρήσουμε στην ανάλυση του σχήματος (6.20) θα σχολιάσουμε λίγο την πρώτη του εξίσωση, θέτοντας, χωρίς περιορισμό της γενικότητας, $a=0$. Αντίστοιχα σχόλια ισχύουν και για την τελευταία εξίσωση, ενώ οι υπόλοιπες είναι ίδιες με εκείνες του σχήματος (6.3). Είναι εύκολο να δούμε ότι η παράγωγος μίας ομαλής και άρτιας συνάρτησης μηδενίζεται στο μηδέν. Αν λοιπόν υποθέσουμε ότι η u επεκτείνεται άρτια αριστερά του μηδενός, τότε αυτόματα ικανοποιείται η συνοριακή συνθήκη $u'(0)=0$. Μένει μόνο να προσεγγίσουμε τη Δ.Ε., δηλαδή βασικά την $u''(0)$. Το $\frac{u(-h)-2u(0)+u(h)}{h^2}$ είναι κατά τα γνωστά μία καλή προσέγγιση της $u''(0)$, και επειδή η u είναι άρτια, αυτός ο λόγος γράφεται στη μορφή $-2\frac{u(0)-u(h)}{h^2}$ και διαφωτίζει κάπως την προέλευση της πρώτης εξίσωσης στο σχήμα (6.20). Όσον αφορά τη συνέπεια του σχήματος (6.20) αναπτύσσοντας κατάλληλα κατά Taylor διαπιστώνουμε εύκολα ότι

$$
\begin{cases}
2\dfrac{u(x_0)-u(x_1)}{h^2}+q(x_0)u(x_0)=f(x_0)+r_0,\\
-\dfrac{1}{h^2}[u(x_{i-1})-2u(x_i)+u(x_{i+1})]+q(x_i)u(x_i)=f(x_i)+r_i, \quad i=1,\dots,J,\\
-2\dfrac{u(x_J)-u(x_{J+1})}{h^2}+q(x_{J+1})u(x_{J+1})=f(x_{J+1})+r_{J+1},
\end{cases}
\tag{6.21}
$$

με

$$
\begin{cases}
\max\limits_{1\le i\le J}|r_i|\le Ch^2,\\
\max(|r_0|,|r_{J+1}|)\le Ch,
\end{cases}
\tag{6.22}
$$

όπου η σταθερά C εξαρτάται από τα $\max_{a\le x\le b}|u^{(3)}(x)|$ και $\max_{a\le x\le b}|u^{(4)}(x)|$. Προχωρούμε τώρα στη μελέτη της ευστάθειας του σχήματος (6.20). Για να απλοποιήσουμε λίγο τους συμβολισμούς εισάγουμε τον τελεστή $\Delta_h : \mathbb{R}^{J+2} \to \mathbb{R}^{J+2}$ δια

$$\Delta_h v_i := (\Delta_h v)_i := \begin{cases} 2\dfrac{-v_0 + v_1}{h^2}, & i = 0, \\ \dfrac{v_{i-1} - 2v_i + v_{i+1}}{h^2}, & i = 1, \dots, J, \\ 2\dfrac{v_J - v_{J+1}}{h^2}, & i = J+1. \end{cases}$$

Έτσι μπορούμε να συμπτύξουμε την (6.20) και να τη γράψουμε στη μορφή

$$-\Delta_h U_i + q(x_i) U_i = f(x_i), \quad i = 0, \dots, J+1. \tag{6.23}$$

Θα επεκτείνουμε επίσης τον ορισμό του εσωτερικού γινομένου $(\cdot,\cdot)_h$ και συνακόλουθα και της παραγόμενης από αυτό νόρμας στον $\mathbb{R}^{J+2}$ ως εξής

$$(v, w)_h := \frac{h}{2} v_0 w_0 + h \sum_{i=1}^{J} v_i w_i + \frac{h}{2} v_{J+1} w_{J+1} \quad v, w \in \mathbb{R}^{J+2}.$$

Σημειώνουμε ότι το εσωτερικό αυτό γινόμενο συμπίπτει στον $\mathbb{R}_0^{J+2}$ με εκείνο που χρησιμοποιούσαμε μέχρι τώρα. Τώρα, εύκολα αποδεικνύεται ότι και στην προκειμένη περίπτωση ισχύει το αντίστοιχο της (6.8), βλ. την Άσκηση 6.3,

$$-(\Delta_h v, v)_h = |v|_{1,h}^2 \qquad \forall v \in \mathbb{R}^{J+2}, \tag{6.24}$$

όπου η $|\cdot|_{1,h}$ ορίζεται στον $\mathbb{R}^{J+2}$ όπως ακριβώς και στον $\mathbb{R}_0^{J+2}$, δηλαδή δια

$$|v|_{1,h} = \Big\{ h \sum_{i=0}^{J} \Big| \frac{v_{i+1} - v_i}{h} \Big|^2 \Big\}^{\frac{1}{2}},$$

μόνο που στον $\mathbb{R}^{J+2}$ δεν είναι νόρμα αλλά ημινόρμα, δηλαδή έχει μεν τις υπόλοιπες ιδιότητες της νόρμας αλλά από $|v|_{1,h} = 0$ δεν έπεται $v = 0$.

Παίρνοντας τώρα το εσωτερικό γινόμενο με την U στην (6.23) και χρησιμοποιώντας την (6.24) έχουμε

$$|U|_{1,h}^2 + h {\sum_{i=0}^{J+1}}'' q(x_i) |U_i|^2 = h {\sum_{i=0}^{J+1}}'' f(x_i) U_i, \tag{6.25}$$

όπου οι δύο τόνοι στο σύμβολο της άθροισης υποδηλώνουν ότι ο πρώτος και ο τελευταίος όρος του αθροίσματος υποδιπλασιάζονται. Για να προχωρήσουμε χρειαζόμαστε κάποια ανισότητα που θα παίξει εδώ το ρόλο της (6.5), βλ. την Άσκηση 5.9.

Δεν μπορούμε βέβαια στον $\mathbb{R}^{J+2}$ να έχουμε ανισότητα της μορφής (6.5) γιατί υπάρχουν $v \in \mathbb{R}^{J+2}$ για τα οποία $|v|_{1,h} = 0$ και $\max_{0\le i\le J+1} |v_i| \neq 0$. Για $0 \le j \le \ell \le J+1$ και $v \in \mathbb{R}^{J+2}$ έχουμε

$$\begin{aligned}|v_\ell - v_j|^2 = |\sum_{i=j}^{\ell-1}(v_{i+1} - v_i)|^2 &\le (\ell - 1 - j)\sum_{i=j}^{\ell-1}|v_{i+1} - v_i|^2\\ &\le (J+1)\sum_{i=0}^{J}|v_{i+1} - v_i|^2,\end{aligned}$$

δηλαδή

$$|v_\ell - v_j|^2 \le (b-a)|v|_{1,h}^2 \qquad \forall v \in \mathbb{R}^{J+2} \quad \forall j, \ell \in \{0, \dots, J+1\}.$$

Από αυτή την ανισότητα συμπεραίνουμε εύκολα ότι

$$\max_{0\le i\le J+1} |v_i| \le \min_{0\le i\le J+1} |v_i| + \sqrt{b-a}\ |v|_{1,h} \quad \forall v \in \mathbb{R}^{J+2}. \tag{6.26}$$

Τώρα το $h{\sum}''^{J+1}_{i=0} q(x_i)$ είναι ένα άθροισμα Riemann (και συγκεκριμένα το αποτέλεσμα του σύνθετου τύπου του τραπεζίου) για το ολοκλήρωμα $\int_a^b q(x)dx$, και συνεπώς συγκλίνει προς το $\int_a^b q(x)dx$ καθώς το h τείνει στο μηδέν. Επομένως, επειδή $\int_a^b q(x)dx > 0$, για αρκετά μικρό h έχουμε

$$h{\sum_{i=0}^{J+1}}'' q(x_i) \ge \frac{1}{2}\int_a^b q(x)dx, \qquad h \text{ αρκετά μικρό.} \tag{6.27}$$

Οι (6.25) και (6.27) δίνουν, για αρκετά μικρό h,

$$|U|_{1,h}^2 + \frac{1}{2}\int_a^b q(x)dx \min_{0\le i\le J+1} |U_i|^2 \le h{\sum_{i=0}^{J+1}}'' f(x_i)U_i$$

και χρησιμοποιώντατην (6.26) λαμβάνουμε

$$\max_{0\le i\le J+1} |U_i|^2 \le ch{\sum_{i=0}^{J+1}}'' f(x_i)U_i \tag{6.28}$$

με κάποια σταθερά c. Από την (6.28) έπεται τώρα ευστάθεια της μεθόδου μας σε διάφορες νόρμες για τα δεδομένα, λόγου χάριν με την ανισότητα των Cauchy–Schwarz και την προφανή ανισότητα $\|v\|_h \le \sqrt{b-a}\ \max_i |v_i|$ παίρνουμε

$$\max_{0\le i\le J+1} |U_i| \le c\Big\{h{\sum_{i=0}^{J+1}}'' |f(x_i)|^2\Big\}^{1/2}, \qquad h \text{ αρκετά μικρό.} \tag{6.29}$$

Αυτού του είδους η ευστάθεια δεν μας δίνει όμως εκτιμήσεις δεύτερης τάξης για το σφάλμα της μεθόδου. Θα επανέλθουμε σε αυτό το σημείο μετά την απόδειξη της σύγκλισης.

Εκτιμώντας εξ άλλου το δεξιό μέλος της ανισότητας (6.28) δια

$$c\max_{0\le i\le J+1}|U_i|\big\{h\sum_{i=0}^{J+1}|f(x_i)|\big\}$$

λαμβάνουμε

$$\max_{0\le i\le J+1}|U_i|\le c\big\{h\sum_{i=0}^{J+1}|f(x_i)|\big\}\qquad h \text{ αρκετά μικρό.} \tag{6.30}$$

Αυτού του είδους η ευστάθεια θα μας δώσει σύγκλιση δεύτερης τάξης ως προς h. Πριν προχωρήσουμε στην απόδειξη της σύγκλισης σημειώνουμε ότι από την (6.30), ή και από την (6.29), συμπεραίνουμε ότι για αρκετά μικρό h η λύση U του γραμμικού συστήματος (6.20) ορίζεται μοναδικά.

Θεώρημα 6.2 *Έστω ότι η λύση u του προβλήματος* (5.15) *είναι αρκετά ομαλή, $u\in C^4[a,b]$. Τότε για αρκετά μικρό h η λύση U του* (6.20) *ορίζεται μοναδικά και ισχύει*

$$\max_{0\le i\le J+1}|u(x_i)-U_i|\le Ch^2 \tag{6.31}$$

με μία σταθερά C ανεξάρτητη του h.

Απόδειξη. Έχουμε ήδη αποδείξει ότι η U ορίζεται μοναδικά για αρκετά μικρό h, βλ. την (6.30). Μένει να αποδείξουμε την εκτίμηση (6.31). Έστω $e\in\mathbb{R}^{J+2}$, $e_i:=u(x_i)-U_i$, $i=0,\dots,J+1$, το σφάλμα προσέγγισης. Αφαιρώντας κατά μέλη τις (6.20) και (6.21) παίρνουμε

$$-\Delta_h e_i+q(x_i)e_i=r_i,\quad i=0,\dots,J+1,$$

συνεπώς

$$|e|_{1,h}^2+h\sum_{i=0}^{J+1}{}''q(x_i)|e_i|^2=(r,e)_h.$$

Χρησιμοποιώντας την (6.27) έχουμε λοιπόν, για αρκετά μικρό h,

$$|e|_{1,h}^2+\frac{1}{2}\int_a^b q(x)dx\quad\min_{0\le i\le J+1}|e_i|^2\le(r,e)_h$$

και επομένως, σύμφωνα με την (6.26), έχουμε

$$\max_{0\le i\le J+1} |e_i|^2 \le c(r,e)_h,$$

από την οποία έπεται αμέσως ότι

$$\max_{0\le i\le J+1} |e_i| \le ch\sum_{i=0}^{J+1}|r_i|. \tag{6.32}$$

Συνεπώς έχουμε

$$\begin{aligned}\max_{0\le i\le J+1} |e_i| &\le c\big\{h(|r_0|+|r_{J+1}|) + hJ \max_{1\le i\le J} |r_i|\big\}\\ &\le c\big\{h(|r_0|+|r_{J+1}|) + (b-a) \max_{1\le i\le J}|r_i|\big\}\end{aligned}$$

και χρησιμοποιώντας την (6.22) παίρνουμε την (6.31). □

Παρατήρηση 6.3 Αν στην απόδειξη του Θεωρήματος 6.2 δεν χρησιμοποιούσαμε την ευστάθεια της μεθόδου (6.20) στη μορφή (6.30) αλλά εκείνη στη μορφή (6.29), το αντίστοιχο της (6.32) θα ήταν

$$\max_{0\le i\le J+1} |e_i| \le c\|r\|_h, \tag{6.33}$$

η οποία συνδυαζόμενη με την (6.22) δίνει

$$\max_{0\le i\le J+1} |e_i| \le ch^{3/2}.$$

Η εκτίμηση του σφάλματος που παίρνουμε κατ' αυτόν τον τρόπο είναι τάξεως $3/2$, ενώ η βέλτιστη τάξη είναι δύο. □

Συνεχίζουμε τώρα με την περίπτωση συνοριακών συνθηκών τρίτου είδους. Θεωρούμε λοιπόν το πρόβλημα (5.16). Το αντίστοιχο του σχήματος (6.20) στην προκειμένη περίπτωση είναι

$$\begin{cases} \dfrac{2}{h^2}(U_0-U_1) + 2\dfrac{\sigma_1}{h}U_0 + q(x_0)U_0 = f(x_0) \\ -\dfrac{1}{h^2}(U_{i-1}-2U_i+U_{i+1}) + q(x_i)U_i = f(x_i), \qquad i=1,\dots,J \\ -\dfrac{2}{h^2}(U_J-U_{J+1}) + 2\dfrac{\sigma_2}{h}U_{J+1} + q(x_{J+1})U_{J+1} = f(x_{J+1}). \end{cases} \tag{6.34}$$

Δεν θα αποδείξουμε συνέπεια, ευστάθεια και σύγκλιση για το σχήμα (6.34), γιατί οι σχετικές αποδείξεις δεν διαφέρουν ουσιαστικά από τις αντίστοιχες στην περίπτωση συνοριακών συνθηκών Neumann. Θα περιοριστούμε μόνο σε δύο παρατηρήσεις σχετικά με το σχήμα (6.34): Στην περίπτωση $\sigma_1 = \sigma_2 = 0$ το πρόβλημα (5.16) είναι το ίδιο με το (5.15) και τα σχήματα (6.34) και (6.20) ταυτίζονται. Στην πρώτη εξίσωση του σχήματος (6.34) οδηγούμαστε ως εξής: Υποθέτουμε ότι τα δεδομένα μας q, f και η λύση u επεκτείνονται ομαλά και αριστερά του a κατά τρόπον ώστε να ισχύει και εκεί η Δ.Ε. Είναι εύκολο να δούμε ότι

$$\frac{u(a+h)-u(a-h)}{2h} = u'(a) + \frac{h^2}{6}u'''(\vartheta) \tag{6.35}$$

με $\vartheta \in [a-h, a+h]$. Αν το U_{-1} προσεγγίζει την $u(a-h)$, είναι λογικό να προσεγγίσουμε τη συνοριακή συνθήκη $-u'(a) + \sigma_1 u(a) = 0$ δια

$$-\frac{U_1 - U_{-1}}{2h} + \sigma_1 U_0 = 0 \tag{6.36}$$

Τη σχέση $-u''(a) + q(a)u(a) = f(a)$ την προσεγγίζουμε κατά τα γνωστά διά

$$-\frac{U_{-1} - 2U_0 + U_1}{h^2} + q(x_0)U_0 = f(x_0). \tag{6.37}$$

Λύνοντας την (6.36) ως προς U_{-1} και αντικαθιστώντας στην (6.37) οδηγούμαστε στην πρώτη εξίσωση του σχήματος (6.34). Η τελευταία εξίσωση σε αυτό το σχήμα προκύπτει εντελώς ανάλογα.

Θα κλείσουμε αυτή την ενότητα με μία σύντομη αναφορά στο πρόβλημα (5.17). Για να γράψουμε το σχήμα σε κομψή μορφή στην προκειμένη περίπτωση εισάγουμε πρώτα λίγο συμβολισμό. Έστω $J \in \mathbb{N}$, $\ h := \frac{b-a}{J+1}$, $\ x_i := a + ih$, $\ i \in \mathbb{Z}$. Έστω

$$\mathbb{R}^{J+1}_{\text{per}} := \big\{ v = (v_i)_{i\in\mathbb{Z}} : v_i \in \mathbb{R} \quad \text{και} \quad v_{i+J+1} = v_i, \ \ i \in \mathbb{Z} \big\}$$

και

$$\Delta_h : \mathbb{R}^{J+1}_{\text{per}} \longrightarrow \mathbb{R}^{J+1}_{\text{per}}, \quad \Delta_h v_i := \frac{v_{i-1} - 2v_i + v_{i+1}}{h^2}, \quad i \in \mathbb{Z}.$$

Προσεγγίζουμε την $(u(x_i))_{i\in\mathbb{Z}} \in \mathbb{R}^{J+1}_{\text{per}}$ με την $U \in \mathbb{R}^{J+1}_{\text{per}}$ που ορίζεται δια

$$-\Delta_h U_i + q(x_i)U_i = f(x_i), \quad i = 0, \dots, J. \tag{6.38}$$

Λόγω της περιοδικότητος των q και f και του ότι $U \in \mathbb{R}^{J+1}_{\text{per}}$, διαπιστώνουμε αμέσως ότι αν η ισότητα στην (6.38) ισχύει για $i = 0, \dots, J$, τότε ισχύει για κάθε ακέραιο i. Το (6.38) είναι ένα γραμμικό σύστημα $J+1$ εξισώσεων με $J+1$ αγνώστους. Οι άγνωστοι φαίνονται εκ πρώτης όψεως να είναι $J+3$, δηλαδή τα $U_{-1}, U_0, \dots, U_{J+1}$, στην ουσία όμως είναι $J+1$ γιατί $U_{-1} = U_J$ και $U_{J+1} = U_0$. Έτσι βλέπουμε εύκολα ότι η (6.38) γράφεται στη μορφή

$$\begin{cases} -\dfrac{1}{h^2}(U_J - 2U_0 + U_1) + q(x_0)U_0 = f(x_0) \\ -\Delta_h U_i + q(x_i)U_i = f(x_i), \qquad\qquad i = 1, \dots, J-1 \\ -\dfrac{1}{h^2}(U_{J-1} - 2U_J + U_0) + q(x_J)U_J = f(x_J). \end{cases} \tag{6.39}$$

Για μία αποτελεσματική μέθοδο για την αριθμητική επίλυση του γραμμικού συστήματος (6.39) παραπέμπουμε στην Άσκηση 6.8.

Ασκήσεις

6.1. Έστω $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ δύο γραμμικοί χώροι με νόρμα, και $L : X \longrightarrow Y$ μία γραμμική απεικόνιση. Υποθέτουμε ότι από τη σχέση $Lx = y$ έπεται $\|x\|_X \le C\|y\|_Y$ με μία σταθερά C ανεξάρτητη των x και y. Αποδείξτε τότε ότι από τις σχέσεις $Lx = y$ και $L\tilde{x} = \tilde{y}$ έπεται $\|x - \tilde{x}\|_X \le C\|y - \tilde{y}\|_Y$. Αν $y \in Y$, πόσες το πολύ λύσεις μπορεί να έχει το πρόβλημα "ζητείται $x \in X$ τέτοιο ώστε $Lx = y$" ;

6.2. Δείξτε ότι ο $J \times J$ πίνακας A του γραμμικού συστήματος (6.4) είναι θετικά ορισμένος, δηλαδή $y^T A y > 0$ για κάθε μη μηδενικό διάνυσμα του $\mathbb{R}^J$.

6.3. Αποδείξτε την (6.24).

6.4. Θεωρούμε το πρόβλημα (6.19). Ορίζουμε τη διγραμμική μορφή $a(\cdot,\cdot) : \mathbb{R}^{J+2}_0 \times \mathbb{R}^{J+2}_0 \longrightarrow \mathbb{R}$ δια

$$a(v, w) := h\sum_{i=1}^{J}\{-\frac{1}{h^2}[p(x_i + \frac{h}{2})(v_{i+1} - v_i) - p(x_i - \frac{h}{2})(v_i - v_{i-1})] + q(x_i)v_i\}w_i.$$

Αποδείξτε ότι:

i. Η a είναι συμμετρική, δηλ. $a(v, w) = a(w, v) \qquad \forall v, w \in \mathbb{R}^{J+2}_0$.

ii. Υπάρχουν θετικές σταθερές C_1 και C_2, ανεξάρτητες του h τέτοιες ώστε

$$|a(v,w)| \leq C_2|v|_{1,h}\,|w|_{1,h} \qquad v,w \in \mathbb{R}_0^{J+2}$$
$$a(v,v) \geq C_1|v|_{1,h}^2 \qquad v \in \mathbb{R}_0^{J+2}.$$

Η πρώτη από αυτές τις σχέσεις αναφέρεται ως *συνέχεια* και η δεύτερη ως *ελλειπτικότητα* της διγραμμικής μορφής.

iii. Υποθέτοντας ότι η λύση του προβλήματος (5.14) είναι αρκετά ομαλή, αποδείξτε μία εκτίμηση της μορφής (6.14) για το σφάλμα της μεθόδου (6.19).

6.5. Το αντίστοιχο του σχήματος (6.20) για το πρόβλημα

$$(6.40) \qquad \begin{cases} -(pu')' + qu = f \qquad \text{στο } [a,b] \\ u'(a) = u'(b) = 0 \end{cases}$$

είναι:

$$(6.41) \qquad \begin{cases} 2p(x_0+\frac{h}{2})\frac{U_0-U_1}{h^2} + q(x_0)U_0 = f(x_0) \\ -\frac{1}{h^2}[p(x_i+\frac{h}{2})(U_{i+1}-U_i) - p(x_i-\frac{h}{2})(U_i-U_{i-1})] + q(x_i)U_i = \\ \qquad\qquad = f(x_i), \quad 1,\dots,J, \\ -2p(x_{J+1}-\frac{h}{2})\frac{U_J-U_{J+1}}{h^2} + q(x_{J+1})U_{J+1} = f(x_{J+1}). \end{cases}$$

Αποδείξτε συνέπεια και ευστάθεια του σχήματος (6.41) και αποδείξτε μία εκτίμηση του σφάλματος της μορφής (6.31).

6.6. Αποδείξτε την (6.35).

6.7. Αποδείξτε ότι

$$\|v\|_h^2 \leq c\{\min_{0\leq i\leq J+1}|v_i|^2 + |v|_{1,h}^2\} \qquad \forall v \in \mathbb{R}^{J+2}$$

με μία σταθερά c ανεξάρτητη του h.

6.8. Έστω $A \in \mathbb{R}^{J,J}$ ένας συμμετρικός πίνακας τέτοιος ώστε $a_{ij} = 0$ για $1 < i-j < J-1$, βλ. την (6.39). Ένας λογικός τρόπος για να λύσουμε γραμμικά συστήματα της μορφής $Ax = y$ είναι ο εξής: Γράφουμε τον A ως άθροισμα $A = B + C$ όπου ο B είναι τριδιαγώνιος και ο C έχει μόνο δύο μη μηδενικά στοιχεία. Τότε έχουμε

$$Ax = y \iff Bx + Cx = y \iff x + B^{-1}Cx = B^{-1}y \iff$$

$$x + B^{-1}\begin{pmatrix} a_{1J}x_J \\ 0 \\ \vdots \\ 0 \\ a_{J1}x_1 \end{pmatrix} = B^{-1}y \iff$$

$$x + a_{1J}x_J b^1 + a_{J1}x_1 b^J = B^{-1}y, \tag{6.42}$$

όπου b^1 και b^J, η πρώτη και η τελευταία στήλη του B^{-1}, αντίστοιχα. Έστω τώρα $\tilde{x}$ η λύση του γραμμικού συστήματος

$$B\tilde{x} = y. \tag{6.43}$$

Οι (6.42) και (6.43) δίνουν τότε

$$\begin{aligned} x_1 + a_{1J}x_J b_1^1 + a_{J1}x_1 b_1^J &= \tilde{x}_1 \\ x_i + a_{1J}x_J b_i^1 + a_{J1}x_1 b_i^J &= \tilde{x}_i, \qquad i = 2, \dots, J-1, \\ x_J + a_{1J}x_J b_J^1 + a_{J1}x_1 b_J^J &= \tilde{x}_J. \end{aligned}$$

Πώς θα λύνατε λοιπόν ένα γραμμικό σύστημα αυτής της μορφής; Σημειώστε ότι ο υπολογισμός των $\tilde{x}, b^1$ και b^J απαιτεί την επίλυση τριών γραμμικών συστημάτων με τον ίδιο τριδιαγώνιο πίνακα .

6.9. Έστω ότι ισχύουν οι υποθέσεις του Θεωρήματος 6.1. Αποδείξτε ότι:

i. $|e|_{1,h} \le Ch^2$,

ii. $(h\sum_{i=0}^{J}|\frac{U_{i+1}-U_i}{h} - u'(x_i + \frac{h}{2})|^2)^{\frac{1}{2}} \le Ch^2$

με μία σταθερά C ανεξάρτητη του h.

6.10. Έστω q, f συνεχείς περιοδικές συναρτήσεις με περίοδο $b - a$, και $q(x) \ge 0$ $\forall x \in \mathbb{R}$, $q \ne 0$. Αν η λύση u του προβλήματος (5.17) είναι αρκετά ομαλή αποδείξτε για το σφάλμα του σχήματος (6.39) την εκτίμηση

$$\max_{0 \le i \le J} |u(x_i) - U_i| \le Ch^2$$

με μία σταθερά C ανεξάρτητη του h.

6.11. Έστω $J \in \mathbb{N}$, $\quad h := \frac{b-a}{J+1}$ και $x_i := a + ih$, $i = 0, \dots, J+1$. Προσεγγίζουμε το διάνυσμα $(u(x_0), u(x_1), \dots, u(x_{J+1}))^T \in \mathbb{R}_0^{J+2}$, όπου u η λύση του προβλήματος (5.1), με ένα διάνυσμα $U \in \mathbb{R}_0^{J+2}$, το οποίο ορίζεται δια

$$
\begin{aligned}
-\Delta_h U_i + \frac{1}{12}[q(x_{i-1})U_{i-1} + 10q(x_i)U_i + q(x_{i+1})U_{i+1}] = \\
= \frac{1}{12}[f(x_{i-1}) + 10f(x_i) + f(x_{i+1})], \quad i = 1, \dots, J,
\end{aligned}
\tag{6.44}
$$

με $\Delta_h : \mathbb{R}_0^{J+2} \to \mathbb{R}_0^{J+2}$ όπως στην (6.7).

(α) Αποδείξτε συνέπεια του σχήματος (6.44). Συγκεκριμένα, υποθέτοντας ότι $q, f \in C^4[a,b]$ και $u \in C^6[a,b]$, αναπτύσσοντας κατά Taylor ως προς το σημείο x_i και χρησιμοποιώντας το γεγονός ότι $-u^{(4)} + (qu)'' = f''$ στο $[a,b]$ διαπιστώστε ότι

$$
\begin{aligned}
-\frac{1}{h^2}[u(x_{i-1}) - 2u(x_i) + u(x_{i+1})] + \frac{1}{12}[(qu)(x_{i-1}) + 10(qu)(x_i) + (qu)(x_{i+1})] \\
= \frac{1}{12}[f(x_{i-1}) + 10f(x_i) + f(x_{i+1})] + r_i, \quad i = 1, \dots, J,
\end{aligned}
$$

όπου, με μια σταθερά C ανεξάρτητη του h,

$$\max_{1 \le i \le J} |r_i| \le Ch^4.$$

(β) Αποδείξτε ευστάθεια του σχήματος (6.44). Συγκεκριμένα, αποδείξτε κατ' αρχάς, για $v \in \mathbb{R}_0^{J+2}$, ότι

$$h\sum_{i=1}^{J}[q(x_{i-1})v_{i-1} + q(x_i)v_i + q(x_{i+1})v_{i+1}]v_i \ge -ch^2\|v\|_h^2.$$

Στη συνέχεια αποδείξτε ότι

$$|U|_{1,h}^2 - ch^2\|U\|_h^2 \le \Big\{h\sum_{i=0}^{J+1}|f(x_i)|^2\Big\}^{\frac{1}{2}}\|U\|_h$$

και, χρησιμοποιώντας την (6.6), αποδείξτε, για αρκετά μικρό h, ότι

$$|U|_{1,h} \le c\Big\{h\sum_{i=0}^{J+1}|f(x_i)|^2\Big\}^{\frac{1}{2}},$$

οπότε, σύμφωνα με την (6.5),

$$\max_{0\le i\le J+1}|U_i|\le c\Big\{h\sum_{i=0}^{J+1}|f(x_i)|^2\Big\}^{\frac12}.$$

(γ) Αποδείξτε σύγκλιση του σχήματος (6.44). Συγκεκριμένα, αν $u\in C^6[a,b],\ q,f\in C^4[a,b]$, αποδείξτε ότι για αρκετά μικρό h η λύση U του σχήματος (6.44) είναι καλά ορισμένη και ισχύει

$$\max_{0\le i\le J+1}|u(x_i)-U_i|\le Ch^4$$

με μια σταθερά C ανεξάρτητη του h.

6.12. Έστω $J\in\mathbb{N},\quad h:=\frac{b-a}{J+1}$ και $x_i:=a+ih,\ i=0,\dots,J+1$. Προσεγγίζουμε το διάνυσμα $(u(x_0),u(x_1),\dots,u(x_{J+1}))^T\in\mathbb{R}_0^{J+2}$, όπου u η λύση του προβλήματος της Άσκησης 5.13, με ένα διάνυσμα $U\in\mathbb{R}_0^{J+2}$, το οποίο ορίζεται δια

$$-\Delta_h U_i=f(U_i),\quad i=1,\dots,J, \tag{6.45}$$

με $\Delta_h:\mathbb{R}_0^{J+2}\to\mathbb{R}_0^{J+2}$ όπως στην (6.7). Μπορεί να αποδειχθεί ότι για αρκετά μικρό h η U είναι καλά ορισμένη.

(α) Αποδείξτε συνέπεια του σχήματος (6.45). Συγκεκριμένα, υποθέτοντας ότι $u\in C^4[a,b]$, αποδείξτε ότι

$$-\frac{1}{h^2}[u(x_{i-1})-2u(x_i)+u(x_{i+1})]=f(u(x_i))+r_i,\ \ i=1,\dots,J,$$

όπου, με μια σταθερά C ανεξάρτητη του h,

$$\max_{1\le i\le J}|r_i|\le Ch^2.$$

(β) Αποδείξτε ευστάθεια του σχήματος (6.45). Συγκεκριμένα, για $g,\tilde g\in\mathbb{R}_0^{J+2}$ και $v,w\in\mathbb{R}_0^{J+2}$ τέτοια ώστε

$$-\Delta_h v_i=f(v_i)+g_i,\quad i=1,\dots,J,$$
$$-\Delta_h w_i=f(w_i)+\tilde g_i,\quad i=1,\dots,J,$$

αποδείξτε ότι

$$\max_{0\le i\le J+1}|v_i-w_i|\le c\|g-\tilde g\|_h$$

με μία σταθερά C ανεξάρτητη των $g,\tilde g$.

(γ) Αποδείξτε σύγκλιση του σχήματος (6.45).

6.13. Θεωρούμε το πρόβλημα δύο σημείων με διεπιφάνεια που δίνεται στην Άσκηση 5.12. Διακριτοποιούμε το πρόβλημα με μια μέθοδο πεπερασμένων διαφορών: Έστω $J \in \mathbb{N},\ \ h := \frac{b-a}{J+1},\ \ x_i := a + ih,\ \ i = 0, \dots, J+1$. Υποθέτουμε ότι το $x^\star$ είναι κόμβος αυτού του διαμερισμού, έστω $x^\star = x_m$. Προσεγγίζουμε το $(u(x_0), \dots, u(x_{J+1}))^T \in \mathbb{R}_0^{J+2}$ με το $U \in \mathbb{R}_0^{J+2}$ που ορίζεται δια

$$
\left\{
\begin{aligned}
&-\frac{1}{h^2}(U_{i-1} - 2U_i + U_{i+1}) + q(x_i)U_i = f(x_i), \qquad i = 1, \dots, m-1,\\
&-\frac{2}{(1+\rho)h^2}[(U_{m-1} - U_m) - \rho(U_m - U_{m+1})]\\
&\qquad + \frac{1}{1+\rho}[q(x^\star-) + \rho q(x^\star+)]U_m = \frac{1}{1+\rho}[f(x^\star-) + \rho f(x^\star+)],\\
&-\frac{1}{h^2}(U_{i-1} - 2U_i + U_{i+1}) + q(x_i)U_i = f(x_i), \qquad i = m+1, \dots, J.
\end{aligned}
\right.
\tag{6.46}
$$

[Η εξίσωση που συνδέει τα U_{m-1}, U_m και U_{m+1} προκύπτει ως εξής: Έστω $v \in C[a,b]$ μια συνάρτηση που είναι τρεις φορές συνεχώς παραγωγίσιμη στα $[a, x^\star]$ και $[x^\star, b]$ και τέτοια ώστε $v'(x^\star-) = \rho v'(x^\star+)$. Τότε έχουμε

$$
\begin{aligned}
v(x^\star + h) &= v(x^\star) + hv'(x^\star+) + \frac{h^2}{2}v''(x^\star+) + \frac{h^3}{6}v'''(\vartheta_1)\\
v(x^\star - h) &= v(x^\star) - hv'(x^\star-) + \frac{h^2}{2}v''(x^\star-) - \frac{h^3}{6}v'''(\vartheta_2).
\end{aligned}
$$

Πολλαπλασιάζοντας την πρώτη σχέση επί ρ και προσθέτοντας στη δεύτερη έχουμε, σύμφωνα και με την $v'(x^\star-) = \rho v'(x^\star+)$,

$$
\begin{aligned}
&-\frac{2}{h^2}\big\{[v(x^\star - h) - v(x^\star)] - \rho[v(x^\star) - v(x^\star + h)]\big\} =\\
&\qquad = -[v''(x^\star-) + \rho v''(x^\star+)] + \frac{h}{3}[v'''(\vartheta_2) - pv'''(\vartheta_1)].
\end{aligned}
$$

Χρησιμοποιώντας αυτή τη σχέση για τη λύση u και το γεγονός ότι σύμφωνα με τη Δ.Ε.

$$
\begin{aligned}
u''(x^\star-) &= q(x^\star-)u(x^\star) - f(x^\star-)\\
u''(x^\star+) &= q(x^\star+)u(x^\star) - f(x^\star+)
\end{aligned}
$$

λαμβάνουμε

$$\begin{aligned}
&-\frac{2}{h^2}\big\{[u(x_{m-1})-u(x_m)]-\rho[u(x_m)-u(x_{m+1})]\big\}\\
&\qquad\qquad +[q(x_m-)+\rho q(x_m+)]u(x_m)\\
&\qquad\qquad =[f(x_m-)+\rho f(x_m+)]+\frac{h}{3}[u'''(\tilde\vartheta_2)-\rho u'''(\tilde\vartheta_1)].
\end{aligned}\tag{6.47}$$

Παραλείποντας τον τελευταίο όρο σ' αυτή τη σχέση, ο οποίος τείνει στο μηδέν καθώς το h τείνει στο μηδέν, οδηγούμεθα στην εξίσωση m του σχήματος (6.46). Σημειώστε ακόμη ότι για $\rho=1$ δεν υπάρχει διεπιφάνεια αφού η u' είναι συνεχής. Τότε το σχήμα (6.46) ταυτίζεται με το σχήμα (6.3), αν οι q και f είναι συνεχείς βέβαια.]

(α) Αποδείξτε συνέπεια του σχήματος (6.46). Ακριβέστερα αποδείξτε ότι

$$\begin{cases}
-\dfrac{1}{h^2}[u(x_{i-1})-2u(x_i)+u(x_{i+1})]+q(x_i)u(x_i)=f(x_i)+r_i,\\
\qquad\qquad\qquad\qquad i=1,\dots,m-1,m+1,\dots,J,\\
-\dfrac{2}{(1+\rho)h^2}\big\{[u(x_{m-1})-u(x_m)]-\rho[u(x_m)-u(x_{m+1})]\big\}\\
+\dfrac{1}{1+\rho}[q(x^\star-)+\rho q(x^\star+)]u(x_m)=\dfrac{1}{1+\rho}[f(x^\star-)+\rho f(x^\star+)]+r_m,
\end{cases}$$

με

$$\begin{cases}
|r_i|\le Ch^2, & i=1,\dots,m-1,\ m+1,\dots,J,\\
|r_m|\le Ch,
\end{cases}$$

όπου η σταθερά C εξαρτάται μόνο από τα a,b και τη u, βλ. την (6.47).

(β) Αποδείξτε ευστάθεια του σχήματος (6.46). Ακριβέστερα εισάγοντας την απεικόνιση $\Delta_h:\mathbb{R}_0^{J+2}\longrightarrow\mathbb{R}_0^{J+2}$,

$$\begin{cases}
\Delta_h v_i:=0, & i=0,J+1,\\
\Delta_h v_i:=\dfrac{1}{h^2}(v_{i-1}-2v_i+v_{i+1}), & i=1,\dots,m-1,m+1,\dots,J,\\
\Delta_h v_m:=\dfrac{2}{(1+\rho)h^2}[(v_{m-1}-v_m)-\rho(v_m-v_{m+1})],
\end{cases}$$

και συμβολίζοντας με $(\cdot,\cdot)_h$ το διακριτό L^2 εσωτερικό γινόμενο στον $\mathbb{R}_0^{J+2}$ με βάρος,

$$(v,w)_h:=h\Big\{\sum_{i=1}^{m-1}v_iw_i+\frac{1+\rho}{2}v_mw_m+\rho\sum_{i=m+1}^{J}v_iw_i\Big\},$$

με $\|\cdot\|_h$ την παραγόμενη από αυτό νόρμα, και με $|\cdot|_{1,h}$ τη διακριτή H^1-νόρμα με βάρος,

$$|v|_{1,h} := \big\{h\sum_{i=0}^{m-1}|\frac{v_{i+1}-v_i}{h}|^2 + \rho h\sum_{i=m}^{J}|\frac{v_{i+1}-v_i}{h}|^2\big\}^{\frac{1}{2}},$$

αποδείξτε κατ' αρχάς ότι

$$-(\Delta_h v, v)_h = |v|_{1,h}^2 \qquad \forall v \in \mathbb{R}_0^{J+2}. \tag{6.48}$$

Πολλαπλασιάζοντας τώρα στην (6.52) επί hU_i τις σχέσεις με $i = 1, \dots, m-1$, επί $\frac{1+\rho}{2}hU_m$ τη σχέση με $i = m$ και επί ρhU_i τις σχέσεις με $i = m+1, \dots, J$, προσθέτοντας και χρησιμοποιώντας την (6.48) αποδείξτε ότι

$$\begin{aligned}|U|_{1,h}^2 + h\big\{\sum_{i=1}^{m-1}q(x_i)|U_i|^2 + \frac{1}{2}[q(x^\star-) + \rho q(x^\star+)]|U_m|^2 + \rho\sum_{i=m+1}^{J}q(x_i)|U_i|^2\big\}\\ = h\big\{\sum_{i=1}^{m-1}f(x_i)U_i + \frac{1}{2}[f(x^\star-) + \rho f(x^\star+)]U_m + \rho\sum_{i=m+1}^{J}f(x_i)U_i\big\}.\end{aligned}$$

Από την (6.5) έπεται αμέσως ότι

$$\max_{0\le i\le J+1}|v_i| \le \frac{1}{\min(1,\sqrt{\rho})}\sqrt{b-a}\ \ |v|_{1,h} \qquad \forall v \in \mathbb{R}_0^{J+2}$$

(προσέξτε τη διαφορετική έννοια της $|\cdot|_{1,h}$ στην (6.5) και την προηγούμενη σχέση). Σύμφωνα με την (6.5) και την προηγούμενη σχέση, οδηγηθείτε στο συμπέρασμα ότι

$$\max_{0\le i\le J+1}|U_i| \le ch\big\{\sum_{i=1}^{m-1}|f(x_i)| + \frac{1}{2}|f(x^\star-) + \rho f(x^\star+)| + \rho\sum_{i=m+1}^{J}|f(x_i)|\big\}$$

με μια σταθερά $c = c(b-a, \rho)$.

(γ) Συνδυάστε ευστάθεια και συνέπεια για να αποδείξετε σύγκλιση. Συγκεκριμένα, υποθέτοντας ότι η λύση u είναι τέσσερεις φορές συνεχώς παραγωγίσιμη στα διαστήματα $[a, x^\star]$ και $[x^\star, b]$, αποδείξτε ότι

$$\max_{0\le i\le J+1}|u(x_i) - U_i| \le Ch^2$$

με μια σταθερά ανεξάρτητη του h.

6.14. Θεωρούμε τώρα το ακόλουθο πρόβλημα με διεπιφάνεια για την περίπτωση μίας γενικότερης Δ.Ε.

$$\begin{cases} -(pu')' + qu = f & \text{στο } [a, x^\star) \cup (x^\star, b] \\ p(x^\star-)u'(x^\star-) = \rho p(x^\star+)u'(x^\star+) \\ u(a) = u(b) = 0 \end{cases}$$

όπου p θετική συνάρτηση, συνεχώς παραγωγίσιμη στα $[a, x^\star]$ και $[x^\star, b]$ και πιθανώς ασυνεχής στο $x^\star$, και q, f όπως στην Άσκηση 5.12. Δώστε το αντίστοιχο του σχήματος (6.46) στην προκειμένη περίπτωση και αποδείξτε συνέπεια, ευστάθεια και σύγκλιση.

6.15. Στις Ασκήσεις 6.13 και 6.14 είναι ουσιαστική η υπόθεση ότι το $x^\star$ είναι κόμβος του διαμερισμού. Μερικές φορές αυτό είναι αδύνατον να επιτευχθεί με ομοιόμορφο διαμερισμό για οποιαδήποτε επιλογή του βήματος h (για να είναι αυτό εφικτό πρέπει ο αριθμός $\frac{b-x^\star}{x^\star-a}$ να είναι ρητός). Το πρόβλημα αυτό παρακάμπτεται εύκολα, αν εργασθούμε με έναν ομοιόμορφο διαμερισμό του $[a, x^\star]$ με βήμα h_-, ας πούμε, και έναν ομοιόμορφο διαμερισμό του $[x^\star, b]$ με βήμα h_+. Αν $x^\star = x_m$, τότε $x_i := a + ih_-$, $i = 0, \dots, m$, και $x_i := x_m + (i-m)h_+$, $i = m+1, \dots, J+1$. Το αντίστοιχο του σχήματος (6.52) είναι τώρα

$$\begin{cases} -\dfrac{1}{h_-^2}(U_{i-1} - 2U_i + U_{i+1}) + q(x_i)U_i = f(x_i), \quad i = 1, \dots, m-1, \\ -\dfrac{1}{\hat{h}}[\dfrac{1}{h_-}(U_{m-1} - U_m) - \rho\dfrac{1}{h_+}(U_m - U_{m+1})] + \\ \qquad + \dfrac{1}{2\hat{h}}[h_- q(x^\star-) + \rho h_+ q(x^\star+)]U_m = \dfrac{1}{2\hat{h}}[h_- f(x^\star-) + \rho h_+ f(x^\star+)], \\ -\dfrac{1}{h_+^2}(U_{i-1} - 2U_i + U_{i+1}) + q(x_i)U_i = f(x_i), \quad i = m+1, \dots, J, \end{cases}$$

όπου $\hat{h} := \frac{1}{2}(h_- + \rho h_+)$. Με τις προφανείς τροποποιήσεις στους ορισμούς των $(\cdot,\cdot)_h$, $\|\cdot\|_h$ και $|\cdot|_{1,h}$, το $(\cdot,\cdot)_h$ φερ' ειπείν ορίζεται τώρα δια

$$(v, w)_h := h_- \sum_{i=1}^{m-1} v_i w_i + \hat{h} v_m w_m + \rho h_+ \sum_{i=m+1}^{J} v_i w_i, \qquad v, w \in \mathbb{R}_0^{J+2},$$

αποδείξτε αποτελέσματα αντίστοιχα εκείνων στην Άσκηση 6.12, θέτοντας $h := \max(h_-, h_+)$.

6.16. (Μη ομοιόμορφος διαμερισμός.) Αρχίζουμε εισάγοντας λίγο συμβολισμό. Έστω $a = x_0 < x_1 < \cdots < x_J < x_{J+1} = b$ ένας διαμερισμός του $[a, b]$, $h_i := x_i - x_{i-1}$, $i = 1, \dots, J+1$, και $\hat{h}_i := \frac{1}{2}(h_i + h_{i+1})$, $i = 1, \dots, J$. Έστω u η λύση του προβλήματος (5.1). Προσεγγίζουμε τότε το $(u(x_0), u(x_1), \dots, u(x_{J+1}))^T \in \mathbb{R}_0^{J+2}$ με $U \in \mathbb{R}_0^{J+2}$ που δίνεται δια

$$-\frac{1}{\hat{h}_i}[\frac{U_{i-1} - U_i}{h_i} - \frac{U_i - U_{i+1}}{h_{i+1}}] + q(x_i)U_i = f(x_i), \quad i = 1, \dots, J. \tag{6.49}$$

(α) Αποδείξτε συνέπεια του σχήματος (6.49). Συγκεκριμένα, υποθέτοντας ότι $u \in C^4[a, b]$ και αναπτύσσοντας κατά Taylor, αποδείξτε κατ' αρχάς ότι

$$\begin{aligned} -\frac{1}{\hat{h}_i}[\frac{u(x_{i-1}) - u(x_i)}{h_i} - \frac{u(x_i) - u(x_{i+1})}{h_{i+1}}] + q(x_i)u(x_i) = \\ = f(x_i) + r_i, \quad i = 1, \dots, J, \end{aligned}$$

με $r \in \mathbb{R}_0^{J+2}$,

$$\begin{cases} r_i = -\frac{1}{3}(h_{i+1} - h_i)u'''(x_i) + \tilde{r}_i \\ |\tilde{r}_i| \le Ch^2 \end{cases}$$

με $h := \max_{1 \le i \le J+1} h_i$. Αποδείξτε εν συνεχεία ότι

$$\|r\|_{-1,H} \le Ch^2,$$

όπου $H = (h_1, \dots, h_{J+1})$ και η "αρνητική" νόρμα $\|\cdot\|_{-1,H}$ ορίζεται ως

$$\|v\|_{-1,H} := \Big\{\sum_{i=1}^{J} h_{i+1}(\sum_{j=1}^{i} \hat{h}_j v_j)^2\Big\}^{\frac{1}{2}}, \qquad v \in \mathbb{R}_0^{J+2},$$

βλ. την Άσκηση 5.15.

(β) Αποδείξτε ευστάθεια του σχήματος (6.49). Το φυσιολογικό διακριτό L^2 εσωτερικό γινόμενο $(\cdot, \cdot)_H$ ορίζεται τώρα ως εξής

$$(v, w)_H := \sum_{i=1}^{J} \hat{h}_i v_i w_i, \qquad v, w \in \mathbb{R}_0^{J+1}.$$

Συμβολίζουμε με $\|\cdot\|_H$ τη διακριτή L^2 νόρμα και ορίζουμε τη νόρμα $|\cdot|_{1,H}$ δια

$$|v|_{1,H} := \Big\{\sum_{i=1}^{J+1} h_i |\frac{v_i - v_{i-1}}{h_i}|^2\Big\}^{\frac{1}{2}}, \qquad v \in \mathbb{R}_0^{J+2}.$$

Αποδείξτε ότι

$$(v,w)_H = -\sum_{i=1}^{J} h_{i+1}\Big(\sum_{j=1}^{i} \hat{h}_j v_j\Big)\frac{w_{i+1}-w_i}{h_{i+1}}$$

και οδηγηθείτε στο συμπέρασμα ότι

$$(v,w)_H \le |v|_{-1,H}\ |w|_{1,H} \qquad \forall v,w \in \mathbb{R}_0^{J+2}.$$

Αποδείξτε τώρα ότι

$$|U|_{1,H} \le \Big\{\sum_{i=1}^{J} h_{i+1}\Big[\sum_{j=1}^{i} \hat{h}_j f(x_j)\Big]^2\Big\}^{\frac{1}{2}},$$

βλ. την Άσκηση 5.15.

(γ) Αποδείξτε τώρα σύγκλιση του σχήματος (6.49). Συγκεκριμένα, υποθέτοντας $u \in C^4[a,b]$, αποδείξτε ότι υπάρχει μία σταθερά c, ανεξάρτητη του διαμερισμού $a = x_0 < x_1 < x_2 < \cdots < x_{J+1} = b$ του $[a,b]$, τέτοια ώστε

$$\max_{1\le i\le J} |u(x_i) - U_i| \le ch^2.$$

6.17. Έστω $J \in \mathbb{N}, J \ge 5, h := \frac{1}{J+1}$, και $x_i := ih, i \in \mathbb{Z}$. Έστω

$$\mathbb{R}_{\text{per}}^{J+1} := \{v = (v_i)_{i\in\mathbb{Z}} : v_i \in \mathbb{R} \ \text{ και } \ v_{i+J+1} = v_i, i \in \mathbb{Z}\}.$$

Αν u είναι η λύση του προβλήματος της Άσκησης 5.19, προσεγγίζουμε την $\big(u(x_i)\big)_{i\in\mathbb{Z}}$ $\in \mathbb{R}_{\text{per}}^{J+1}$ με την $U \in \mathbb{R}_{\text{per}}^{J+1}$ που ορίζεται δια

$$\frac{1}{h^2}(U_{i-2} - 4U_{i-1} + 6U_i - 4U_{i+1} + U_{i+2}) + U_i = f(x_i), \quad i = 0,\dots,J.$$

Αποδείξτε ότι το διακριτό αυτό πρόβλημα έχει ακριβώς μία λύση, δηλαδή ότι η U είναι καλά ορισμένη.

3. Πεπερασμένα στοιχεία για το πρόβλημα δύο σημείων

Σ' αυτό το κεφάλαιο θα μελετήσουμε μεθόδους πεπερασμένων στοιχείων για το πρόβλημα δύο σημείων. Το κεφάλαιο αποτελείται από τρεις παραγράφους, οι δύο πρώτες εκ των οποίων αναφέρονται σε εκ των προτέρων εκτιμήσεις ενώ η τρίτη σε εκ των υστέρων εκτιμήσεις. Συγκεκριμένα, στην πρώτη παράγραφο θα ασχοληθούμε με το "ορισμένο" πρόβλημα, θα υποθέσουμε δηλαδή ότι η συνάρτηση q λαμβάνει μη αρνητικές τιμές, ενώ στη δεύτερη με το "μη ορισμένο" επιτρέποντας και αρνητικές τιμές της q, υποθέτοντας όμως πάντα ότι το πρόβλημα έχει μία ακριβώς λύση η οποία μάλιστα είναι αρκετά ομαλή, και θα δώσουμε εκ των προτέρων εκτιμήσεις, δηλαδή εκτιμήσεις στις οποίες υπεισέρχονται άγνωστες ποσότητες, όπως νόρμες της άγνωστης λύσης. Οι εκτιμήσεις στην τρίτη παράγραφο είναι διαφορετικής υφής, στα φράγματα υπεισέρχονται μόνο ποσότητες που μπορούν να υπολογισθούν, αρκεί να έχει πρώτα υπολογισθεί η προσεγγιστική λύση, πρόκειται λοιπόν για εκ των υστέρων εκτιμήσεις.

3.1 Το ορισμένο πρόβλημα

Θεωρούμε το πρόβλημα δύο σημείων (5.1). Σύμφωνα με την (5.3) ισχύει

$$(u', v') + (qu, v) = (f, v) \quad \forall v \in C_0^1[a, b]\,.$$

Είναι πολύ εύκολο να δει κανείς ότι αυτή η σχέση ισχύει και για συνεχείς και τμηματικά συνεχώς παραγωγίσιμες συναρτήσεις οι οποίες μηδενίζονται στα a και b, δηλ.

$$(u', v') + (qu, v) = (f, v) \quad \forall v \in V, \tag{7.1}$$

όπου $V := \{v \in C_0[a, b] : v$ τμηματικά συνεχώς παραγωγίσιμη$\}$.

Η (7.1) λέγεται *ασθενής* ή *γενικευμένη* ή *μεταβολική μορφή* του προβλήματος (5.1). Σε αυτή τη διατύπωση του προβλήματος βασίζεται η μέθοδος των πεπερασμένων στοιχείων.

Έστω τώρα $r \geq 2$ και $(S_h^r)_{0<h\leq b-a} \subset V$ μια οικογένεια υποχώρων του V, πεπερασμένης διάστασης, με την εξής *προσεγγιστική ιδιότητα*

$$\begin{aligned} &\exists c > 0 \quad \forall v \in C^s[a, b] \cap V \quad \exists \chi \in S_h^r \\ &\|v - \chi\| + h\|v' - \chi'\| \leq ch^s\|v\|_s, \quad s = 2,\ s = r^{\cdot} \end{aligned} \tag{7.2}$$

χρησιμοποιήσαμε εδώ τη νόρμα Sobolev $\|\cdot\|_s$ η οποία ορίζεται ως

$$\|v\|_s := \{\sum_{i=0}^{s} \|v^{(i)}\|^2\}^{1/2}$$

και είναι προφανώς ισοδύναμη με τη νόρμα $\|v\|^\star := \sum_{i=0}^{s} \|v^{(i)}\|$. Είναι προφανές ότι η νόρμα $\|\cdot\|_s$ παράγεται από ένα εσωτερικό γινόμενο.

Παράδειγμα 7.1 Έστω $S_h^r \subset V$ ο χώρος των splines ως προς τον ομοιόμορφο διαμερισμό του $[a, b]$ με βήμα h, οι οποίες σε κάθε υποδιάστημα αυτού του διαμερισμού είναι πολυώνυμα βαθμού το πολύ $r - 1$, και στο $[a, b]$ είναι $r - 2$ φορές συνεχώς παραγωγίσιμες συναρτήσεις. Είναι γνωστό ότι σ' αυτή την περίπτωση ισχύει η (7.2). Θα δώσουμε εδώ την απόδειξη για την περίπτωση $r = 2$. Έστω λοιπόν $v \in C^2[a, b] \cap V$ και $v_h \in S_h^2$ τέτοια ώστε $v_h(x_i) = v(x_i)$, $i = 0, \ldots, J + 1$, όπου $x_i = a + ih$, $i = 0, \ldots, J + 1, (J + 1)h = b - a$. Τότε, θέτοντας $w := v - v_h$ έχουμε, σύμφωνα με την (5.5),

$$\int_{x_i}^{x_{i+1}} [w(x)]^2 dx \leq h^2 \int_{x_i}^{x_{i+1}} [w'(x)]^2 dx, \quad i = 0, \ldots, J. \tag{7.3}$$

Αθροίζοντας από $i = 0$ έως $i = J$, παίρνουμε

$$\|w\| \leq h\|w'\|. \tag{7.4}$$

Σύμφωνα με το θεώρημα του Rolle υπάρχει $\xi \in (x_i, x_{i+1})$, τέτοιο ώστε $w'(\xi) = 0$. Τότε θα έχουμε

$$w'(x) = \int_\xi^x w''(s)ds = \int_\xi^x v''(s)ds, \quad x \in [x_i, x_{i+1}],$$

συνεπώς

$$|w'(x)|^2 \le h \int_{x_i}^{x_{i+1}} |v''(s)|^2 ds,$$

δηλ.

$$\int_{x_i}^{x_{i+1}} |w'(x)|^2 dx \le h^2 \int_{x_i}^{x_{i+1}} |v''(s)|^2 ds. \tag{7.5}$$

Αθροίζοντας από $i = 0$ έως $i = J$, λαμβάνουμε

$$\|w'\| \le h\|v''\|. \tag{7.6}$$

Από τις (7.4) και (7.6) έπεται η (7.2) για $s = r = 2$. □

Σε αυτή την παράγραφο υποθέτουμε ότι η συνάρτηση q λαμβάνει μόνο μη αρνητικές τιμές.

Θεωρούμε τώρα ένα σταθερό h. Προσεγγίζουμε τη λύση u του προβλήματος (5.1) με ένα στοιχείο $u_h \in S_h^r$ τέτοιο ώστε

$$(u_h', \chi') + (qu_h, \chi) = (f, \chi) \qquad \forall \chi \in S_h^r, \tag{7.7}$$

βλ. την (7.1). Αποδεικνύουμε κατ' αρχάς ότι η προσεγγιστική λύση u_h ορίζεται μονοσήμαντα. Έστω $N_h := \dim S_h^r$, και $\{\varphi_1, \dots, \varphi_{N_h}\}$ μια βάση του S_h^r. Η (7.7) γράφεται ισοδύναμα στη μορφή

$$(u_h', \varphi_i') + (qu_h, \varphi_i) = (f, \varphi_i), \qquad i = 1, \dots, N_h. \tag{7.8}$$

Παριστάνοντας την u_h στη μορφή $u_h = \alpha_1\varphi_1 + \cdots + \alpha_{N_h}\varphi_{N_h}$ διαπιστώνουμε εύκολα ότι το (7.8) είναι ένα γραμμικό σύστημα και ο αντίστοιχος πίνακας είναι συμμετρικός και θετικά ορισμένος, συνεπώς η u_h ορίζεται μονοσήμαντα. Ότι η u_h ορίζεται μονοσήμαντα, μπορεί να το δει κανείς επίσης μελετώντας το αντίστοιχο ομογενές σύστημα. Θέτοντας $f = 0$ και $\chi = u_h$ στην (7.7) (και χρησιμοποιώντας την (5.5), η οποία ισχύει και στον V) λαμβάνουμε αμέσως $u_h = 0$.

Το σύστημα (7.8) είναι της μορφής

$$A_h \alpha_h = f_h, \tag{7.9}$$

όπου $\alpha_h = (\alpha_1, \dots, \alpha_{N_h})^T, (f_h)_i = (f, \varphi_i)$, $i = 1, \dots, N_h$, και $(A_h)_{ij} = (\varphi_i', \varphi_j') + (q\varphi_i, \varphi_j)$, $i, j = 1, \dots, N_h$. Επιλέγοντας τη βάση του S_h^r κατά τρόπον ώστε τα στοιχεία της να έχουν *μικρό φορέα* (ιδιότητα στην οποία οφείλουν την ονομασία τους τα πεπερασμένα στοιχεία) επιτυγχάνουμε να είναι ο πίνακας A_h αραιός, και αυτό έχει ως συνέπεια η επίλυση του (7.9) να μπορεί να γίνει με χαμηλό υπολογιστικό κόστος. Επίσης η αρίθμηση των στοιχείων της βάσης επηρεάζει τη μορφή του πίνακα A_h (ακριβέστερα καθορίζει τις θέσεις όπου υπάρχουν μη μηδενικά στοιχεία).

Παράδειγμα 7.2 Έστω $h := \frac{b-a}{J+1}, x_i := a + ih, i = 0, \dots, J+1$, και S_h^2 ο χώρος των συνεχών συναρτήσεων οι οποίες μηδενίζονται στα σημεία a και b και είναι πολυώνυμα βαθμού το πολύ ένα σε καθένα των υποδιαστημάτων $(x_i, x_{i+1}), i = 0, \dots, J$. Επιλέγοντας ως βάση την $\{\varphi_1, \dots, \varphi_J\}$ με

$$\varphi_i(x) = \begin{cases} \dfrac{x - x_{i-1}}{h}, & x_{i-1} \le x \le x_i, \\ -\dfrac{x - x_{i+1}}{h}, & x_i \le x \le x_{i+1}, \\ \quad 0 \quad , & \text{διαφορετικά}, \end{cases}$$

διαπιστώνουμε αμέσως ότι ο πίνακας A_h είναι τριδιαγώνιος. □

Μια προσεγγιστική ιδιότητα της u_h δίνεται στο ακόλουθο θεώρημα.

Θεώρημα 7.1 *Έστω $r \ge 2$, και $u \in C^r[a, b]$ η λύση του προβλήματος* (5.1). *Έστω $S_h^r \subset V$ τέτοιος ώστε να ισχύει η* (7.2) *και $u_h \in S_h^r$ η λύση του* (7.7). *Τότε υπάρχει μία σταθερά C, ανεξάρτητη των u και h, τέτοια ώστε*

$$\|u' - u_h'\| \le Ch^{r-1}\|u\|_r, \tag{7.10}$$

και

$$\|u - u_h\| \le Ch^r\|u\|_r. \tag{7.11}$$

Απόδειξη. Εισάγουμε στον V τη *διγραμμική μορφή* $a(\cdot, \cdot)$, δηλ. γραμμική ως προς καθένα από τα στοιχεία της, δια

$$a(\cdot, \cdot) : V \times V \to \mathbb{R}, \quad a(v, w) := (v', w') + (qv, w). \tag{7.12}$$

Εύκολα διαπιστώνουμε ότι η διγραμμική μορφή a είναι στον $(V, \|\cdot\|_1)$ τόσο *συνεχής* όσο και *ελλειπτική*, δηλαδή ισχύουν

$$|a(v,w)| \le C\|v\|_1\,\|w\|_1 \qquad \forall v,w\in V \tag{7.13}$$

και

$$a(v,v) \ge c\|v\|_1^2 \qquad \forall v\in V, \tag{7.14}$$

αντίστοιχα, με δύο θετικές σταθερές c και C. Πράγματι, χρησιμοποιώντας την ανισότητα των Cauchy–Schwarz έχουμε

$$\begin{aligned} |a(v,w)| &\le \|v'\|\,\|w'\| + \max_x q(x)\|v\|\,\|w\| \\ &\le C[\|v'\|\,\|w'\| + \|v\|\,\|w\|] \\ &\le C\|v\|_1\,\|w\|_1 \end{aligned}$$

με $C := \max\{1, \max_x q(x)\}$. Εξ άλλου, επειδή υποθέσαμε ότι η q λαμβάνει μη αρνητικές τιμές,

$$a(v,v) \ge \|v'\|^2,$$

οπότε, βάσει της (5.5),

$$a(v,v) \ge \frac{1}{(b-a)^2}\|v\|^2,$$

και συνεπώς, αθροίζοντας τις δύο τελευταίες σχέσεις,

$$a(v,v) \ge \frac{1}{2}\min\Big(1, \frac{1}{(b-a)^2}\Big)\|v\|_1^2 \qquad \forall v\in V.$$

Οι (7.1) και (7.7) γράφονται τώρα στη μορφή

$$a(u,v) = (f,v) \quad \forall v\in V \tag{7.15}$$

$$a(u_h,\chi) = (f,\chi) \quad \forall \chi\in S_h^r. \tag{7.16}$$

Ιδιαίτερα, η (7.15) ισχύει για $v\in S_h^r$, δηλαδή

$$a(u,\chi) = (f,\chi) \quad \forall \chi\in S_h^r. \tag{7.17}$$

Αφαιρώντας κατά μέλη τις (7.16) και (7.17) λαμβάνουμε

$$a(u-u_h,\chi)=0 \quad \forall \chi \in S_h^r. \tag{7.18}$$

Συνεπώς, για $\chi \in S_h^r$ έχουμε

$$\begin{aligned} a(u-u_h,u-u_h) &= a(u-u_h,u)-a(u-u_h,u_h)=a(u-u_h,u)= \\ &= a(u-u_h,u)-a(u-u_h,\chi) \\ &= a(u-u_h,u-\chi), \end{aligned}$$

οπότε, σύμφωνα με τις (7.13) και (7.14),

$$\|u-u_h\|_1 \le \frac{C}{c}\min_{\chi\in S_h^r}\|u-\chi\|_1. \tag{7.19}$$

Χρησιμοποιώντας εδώ την προσεγγιστική ιδιότητα (7.2) οδηγούμαστε αμέσως στην (7.10).

Για να αποδείξουμε την (7.11) θα χρησιμοποιήσουμε το λεγόμενο *τέχνασμα του Nitsche*.

Θεωρούμε κατ' αρχάς το πρόβλημα

$$\begin{cases} -w''+qw=u-u_h & \text{στο } [a,b] \\ w(a)=w(b)=0. \end{cases} \tag{7.20}$$

Σύμφωνα με την (5.11) ισχύει τότε

$$\|w\|_2 \le c\|u-u_h\|. \tag{7.21}$$

Επιπλέον, σύμφωνα με την (7.15), έχουμε

$$a(w,v)=(u-u_h,v) \quad \forall v\in V.$$

Συνεπώς, για $v:=u-u_h$,

$$\|u-u_h\|^2=a(w,u-u_h). \tag{7.22}$$

Επομένως, σύμφωνα με την (7.18), για $\chi\in S_h^r$, έχουμε

$$\|u-u_h\|^2=a(w-\chi,u-u_h),$$

οπότε, λόγω της (7.13),

$$\|u-u_h\|^2 \leq C\|w-\chi\|_1\,\|u-u_h\|_1.$$

Σύμφωνα τώρα με τις (7.2) για $s=2$ και (7.21), αυτή η σχέση δίνει

$$\|u-u_h\| \leq Ch\|u-u_h\|_1\,, \tag{7.23}$$

η οποία συνδυαζόμενη με την (7.10) δίνει την (7.11). □

Παρατήρηση 7.1 Η μεταβολική διατύπωση (7.16) (ή ισοδύναμα η (7.7)) αναφέρεται ως *μέθοδος του Galerkin*. Το ίδιο πρόβλημα, στην περίπτωση που η διγραμμική μορφή a, πέραν της συνέχειας και ελλειπτικότητας που αναφέραμε, είναι και συμμετρική, όπως συμβαίνει στο πρόβλημά μας, μπορεί να διατυπωθεί και ως πρόβλημα ελαχιστοποίησης, και τότε αναφέρεται ως *μέθοδος του Ritz*. Συγκεκριμένα θα αποδείξουμε ότι η λύση u_h του προβλήματος (7.16) είναι το μόνο στοιχείο του S_h^r στο οποίο το συναρτησιακό J,

$$J(v) := a(v,v) - 2(f,v),$$

λαμβάνει την ελάχιστη τιμή του στον S_h^r. Πράγματι, για $v \in S_h^r$, έχουμε

$$\begin{aligned} J(u_h - v) &= a(u_h - v, u_h - v) - 2(f, u_h - v) \\ &= a(u_h, u_h) - 2(f, u_h) - 2a(u_h, v) + 2(f, v) + a(v, v), \end{aligned}$$

οπότε, επειδή $a(u_h, v) = (f, v)$,

$$J(u_h - v) = J(u_h) + a(v,v), \qquad \forall v \in S_h^r,$$

από την οποία προκύπτει αμέσως το αποτέλεσμα.

Σημειώνουμε πάντως ότι η διατύπωση Galerkin είναι γενικότερη γιατί εφαρμόζεται και σε περιπτώσεις όπου η αντίστοιχη διγραμμική μορφή δεν είναι συμμετρική, όπως, φερ’ ειπείν, σε εξισώσεις της μορφής $-u'' + p(x)u' + q(x)u = f$.

Παρατήρηση 7.2 Για να συγκρίνουμε κάπως την ομαλότητα που απαιτείται στις μεθόδους πεπερασμένων στοιχείων με την αντίστοιχη στις μεθόδους πεπερασμένων διαφορών, ας θεωρήσουμε την εκτίμηση (7.11) στην περίπτωση $r=2$,

$$\|u-u_h\| \leq Ch^2\|u\|_2.$$

Τόσο σε αυτή την εκτίμηση όσο και στην (6.14) το σφάλμα είναι της τάξεως $O(h^2)$, αλλά ενώ εδώ υπεισέρχεται μόνο η δεύτερη παράγωγος, στην (6.14) υπεισέρχεται η τέταρτη παράγωγος, απαιτείται δηλαδή περισσότερη ομαλότητα στη μέθοδο πεπερασμένων διαφορών. Αυτό οφείλεται στο γεγονός ότι στη μέθοδο πεπερασμένων στοιχείων χρησιμοποιούμε ολοκληρώματα των δεδομένων ενώ στις μεθόδους πεπερασμένων στοιχείων σημειακές τιμές τους. Αν προσεγγίσουμε τα ολοκληρώματα με τύπους αριθμητικής ολοκλήρωσης, θα αναγκασθούμε πάλι να απαιτήσουμε περισσότερη ομαλότητα.

3.2 Ένα μη ορισμένο πρόβλημα

Στη μέχρι τώρα μελέτη αριθμητικών μεθόδων για το πρόβλημα δύο σημείων (5.1) περιοριστήκαμε στην περίπτωση όπου η συνάρτηση q λαμβάνει μόνο μη αρνητικές τιμές. Ένα φυσιολογικό ερώτημα είναι πώς συμπεριφέρονται οι αριθμητικές μέθοδοι στην περίπτωση που το πρόβλημα (5.1) έχει μεν ακριβώς μία λύση, αρκετά ομαλή, αλλά η q επιτρέπεται να παίρνει και *αρνητικές* τιμές. Η μη αρνητικότητα της q χρησιμοποιήθηκε στην προηγούμενη παράγραφο σε δύο σημεία: για να αποδείξουμε ότι ο πίνακας του γραμμικού συστήματος (7.11) είναι θετικά ορισμένος, και ιδιαίτερα αντιστρέψιμος, και για να αποδείξουμε την ελλειπτικότητα της διγραμμικής μορφής a, βλ. την (7.14).

Η ανάλυση μεθόδων πεπερασμένων στοιχείων για το μη ορισμένο πρόβλημα γίνεται σχετικά εύκολα, και αποφασιστικό ρόλο παίζει το γεγονός ότι η σύγκλιση στην L^2-νόρμα είναι ταχύτερη της σύγκλισης στην H^1-νόρμα, βλ. τις (7.11) και (7.10). Υπάρχει ανάλυση και μεθόδων πεπερασμένων διαφορών για μη ορισμένα προβλήματα, παρά το ότι σ' αυτή την περίπτωση η τάξη σύγκλισης είναι η ίδια για τις διακριτές νόρμες L^2 και H^1, δεν θα ασχοληθούμε όμως εδώ με αυτό το θέμα.

Θεωρούμε λοιπόν το πρόβλημα (5.1) και υποθέτουμε ότι έχει μία ακριβώς λύση η οποία υποτίθεται αρκετά ομαλή. Επιτρέπουμε στην q να παίρνει και αρνητικές τιμές. Θεωρούμε μία οικογένεια $(S_h^r)_{0<h\leq b-a}$ υποχώρων πεπερασμένης διάστασης του V με την προσεγγιστική ιδιότητα (7.2). Έστω $u_h \in S_h^r$ μία προσεγγιστική λύση

πεπερασμένων στοιχείων τέτοια ώστε

$$(u_h', \chi') + (qu_h, \chi) = (f, \chi) \qquad \forall \chi \in S_h^r, \tag{7.24}$$

βλ. την (7.7). Εύκολα μπορεί κανείς να δώσει παραδείγματα στα οποία δεν εξασφαλίζεται ύπαρξη – μοναδικότητα της u_h, δηλαδή τέτοια ώστε ο πίνακας A_h του γραμμικού συστήματος (7.9) να μην είναι αντιστρέψιμος, όταν η q παίρνει και αρνητικές τιμές. Θα αποδείξουμε, όμως, στη συνέχεια, ότι, για αρκετά μικρό h, η u_h είναι καλώς ορισμένη και έχει προσεγγιστικές ιδιότητες ανάλογες των (7.10) και (7.11).

Στην συνέχεια θα χρησιμοποιήσουμε χωρίς να την αποδείξουμε την ελλειπτική ομαλότητα και για την περίπτωση που η συνάρτηση q παίρνει και αρνητικές τιμές. Είναι γνωστό ότι αν το πρόβλημα (5.1) έχει για κάποια δεδομένη συνάρτηση q μία ακριβώς ομαλή λύση για μία συγκεκριμένη ομαλή συνάρτηση f, τότε έχει μία ακριβώς λύση για κάθε ομαλή συνάρτηση f. Αυτό είναι ακριβώς ανάλογο της περιπτώσεως ενός γραμμικού συστήματος $Ax = b$ με τετραγωνικό πίνακα A. Μάλιστα στην περίπτωσή μας ισχύει και η εκτίμηση ελλειπτικής ομαλότητας (5.11). Τα αποτελέσματα που απλώς αναφέρθηκαν εδώ αποδεικνύονται στη συγκεκριμένη περίπτωση με μέσα που δεν έχουμε στη διάθεση μας σε αυτές τις σημειώσεις, π.χ., με το εναλλακτικό θεώρημα του Fredholm. Σημειώνουμε πάντως ότι εφόσον αποδειχθεί μία εκτίμηση της μορφής $\|u\| \le c\|f\|$, τότε η απόδειξη της εκτίμησης της ελλειπτικής ομαλότητας ολοκληρώνεται εύκολα με τη μέθοδο της ενέργειας, την οποία χρησιμοποιήσαμε επανειλημμένα εδώ.

Θεωρούμε τη διγραμμική μορφή a που ορίστηκε στην (7.12). Εύκολα διαπιστώνουμε ότι η a είναι συνεχής και στην περίπτωση που μελετάμε τώρα, δηλαδή ισχύει η (7.13), και μάλιστα με σταθερά $C = \max\{1, \max_x |q(x)|\}$, δεν είναι όμως ελλειπτική, δεν ισχύει δηλαδή η (7.14), αλλά αντ' αυτής έχουμε την *ανισότητα του Gårding*

$$a(v, v) \ge c_1\|v\|_1^2 - c_2\|v\|^2 \qquad \forall v \in V \tag{7.25}$$

με δύο θετικές σταθερές c_1 και c_2. Ακριβώς επειδή η a δεν είναι ελλειπτική στην παρούσα περίπτωση, αναφέρεται το πρόβλημα ως μη ορισμένο.

Θεώρημα 7.2 *Έστω $r \ge 2$, q μία δεδομένη συνάρτηση η οποία μπορεί να παίρνει και* *αρνητικές τιμές, είναι όμως τέτοια ώστε το πρόβλημα* (5.1) *να έχει μία ακριβώς ομαλή λύση* u, $u \in C^r[a,b]$. *Έστω* $S_h^r \subset V$ *χώροι πεπερασμένης διάστασης τέτοιοι ώστε να ισχύει η* (7.2). *Τότε υπάρχουν δύο θετικές σταθερές* h_0 *και* C, *τέτοιες ώστε για κάθε* $h \le h_0$ *το πρόβλημα* (7.24) *έχει μία ακριβώς λύση* u_h, *για την οποία μάλιστα ισχύουν οι εκτιμήσεις*

$$\|u' - u_h'\| \le Ch^{r-1}\|u\|_r, \tag{7.26}$$

και

$$\|u - u_h\| \le Ch^r\|u\|_r. \tag{7.27}$$

Απόδειξη. Θα αποδείξουμε πρώτα τις εκτιμήσεις (7.26) και (7.27) για κάθε προσεγγιστική λύση u_h της (7.24) που ενδεχομένως υπάρχει. Μετά θα αποδείξουμε ότι η προσεγγιστική λύση υπάρχει πράγματι για αρκετά μικρό h.

Θα χρησιμοποιήσουμε και εδώ τη διγραμμική μορφή a που εισαγάγαμε στην απόδειξη του Θεωρήματος 7.1, εδώ όμως φυσικά η συνάρτηση q επιτρέπεται να παίρνει και αρνητικές τιμές. Αφαιρώντας κατά μέλη τις (7.17) και (7.24) λαμβάνουμε

$$a(u - u_h, \chi) = 0 \quad \forall \chi \in S_h^r, \tag{7.28}$$

βλ. την (7.18). Σύμφωνα με τις (7.25) και (7.28) έχουμε, για $\chi \in S_h^r$,

$$\begin{aligned} c_1\|u - u_h\|_1^2 &\le a(u - u_h, u - u_h) + c_2\|u - u_h\|^2 \\ &= a(u - u_h, u - \chi) + c_2\|u - u_h\|^2, \end{aligned}$$

οπότε, χρησιμοποιώντας την (7.13),

$$c_1\|u - u_h\|_1^2 \le C\|u - u_h\|_1 \, \|u - \chi\|_1 + c_2\|u - u_h\|^2 \quad \forall \chi \in S_h^r.$$

Επομένως, χρησιμοποιώντας την προσεγγιστική ιδιότητα (7.2), έχουμε

$$c_1\|u - u_h\|_1^2 \le Ch^{r-1}\|u - u_h\|_1 \, \|u\|_r + c_2\|u - u_h\|^2. \tag{7.29}$$

Για να εκτιμήσουμε την L^2-νόρμα θα χρησιμοποιήσουμε πάλι το τέχνασμα του Nitsche. Θεωρούμε το πρόβλημα (7.20). Σύμφωνα με όσα αναφέραμε η w ορίζεται κατά μοναδικό τρόπο και ισχύει

$$\|w\|_2 \le c\|u - u_h\|. \tag{7.30}$$

Επιπλέον, εύκολα διαπιστώνουμε ότι

$$a(w, v) = (u - u_h, v) \quad \forall v \in V.$$

Συνεπώς, για $v := u - u_h$,

$$\|u - u_h\|^2 = a(w, u - u_h). \tag{7.31}$$

Επομένως, σύμφωνα με την (7.28), για $\chi \in S_h^r$, έχουμε

$$\|u - u_h\|^2 = a(w - \chi, u - u_h),$$

οπότε, λόγω της (7.12),

$$\|u - u_h\|^2 \le C\|w - \chi\|_1 \, \|u - u_h\|_1.$$

Σύμφωνα τώρα με τις (7.2) και (7.30), αυτή η σχέση δίνει

$$\|u - u_h\| \le Ch\|u - u_h\|_1 \, . \tag{7.32}$$

Από τις (7.29) και (7.32) λαμβάνουμε

$$c_1\|u - u_h\|_1 \le Ch^{r-1} \, \|u\|_r + \tilde{C}h^2\|u - u_h\|_1,$$

και επομένως, για αρκετά μικρό h,

$$\|u - u_h\|_1 \le Ch^{r-1}\|u\|_r. \tag{7.33}$$

Μέχρι τώρα υποθέσαμε ύπαρξη της προσεγγιστικής λύσης u_h. Τώρα μπορούμε να αποδείξουμε ύπαρξη και μοναδικότητα της u_h για αρκετά μικρό h. Το πρόβλημα (7.24) μπορεί να γραφεί ως γραμμικό σύστημα με τετραγωνικό πίνακα, βλ. την (7.9). Για $f = 0$ οδηγούμαστε στο αντίστοιχο ομογενές σύστημα. Τότε όμως $u = 0$, και, σύμφωνα με την (7.33), $u_h = 0$, δηλαδή το αντίστοιχο ομογενές σύστημα έχει μόνο την τετριμμένη λύση. Συνεπώς, η u_h είναι καλώς ορισμένη, για αρκετά μικρό h. Επιπλέον, η (7.26) έπεται αμέσως από την (7.33), και η (7.27) από τις (7.32) και (7.33).

□

3.3 Εκ των υστέρων εκτιμήσεις του σφάλματος

Εκτιμήσεις του σφάλματος της μορφής (7.10) και (7.11), καθώς και όλες οι άλλες που έχουμε δει μέχρι τώρα, λέγονται εκτιμήσεις *εκ των προτέρων* για το λόγο ότι τα φράγματα στο δεξιό μέλος δεν εξαρτώνται από την προσέγγιση u_h. Τέτοιες εκτιμήσεις είναι χρήσιμες γιατί μας δίνουν πληροφορίες σχετικά με τα ποιοτικά χαρακτηριστικά μιας μεθόδου, καθώς και για τις απαιτούμενες συνθήκες ομαλότητας της λύσης. Από την άλλη όμως πλευρά, αν θέλει κανείς, για κάποιο συγκεκριμένο h, να υπολογίσει το φράγμα στην (7.10), λόγου χάρη, δεν θα μπορέσει να το κάνει γιατί, αν και η τιμή της σταθεράς C είναι στην προκειμένη περίπτωση γνωστή, δεν γνωρίζει τη λύση u. (Θα μπορούσε κανείς να αντιπροτείνει να εκτιμήσουμε τη νόρμα της λύσης u με ποσότητες που εξαρτώνται από τα δεδομένα, όπως αυτές που είδαμε στο πέμπτο κεφάλαιο. Πράγματι κατ' αυτόν τον τρόπο οδηγείται κανείς σε μία ποσότητα που θα μπορούσε να υπολογισθεί, αυτή όμως δεν είναι πρακτικά χρήσιμη αφού είναι γενικά πολύ απαισιόδοξη, το φράγμα που προκύπτει είναι πολύ μεγαλύτερο από το πραγματικό σφάλμα. Μία αιτία για αυτό αποτελεί το γεγονός ότι στο φράγμα στην (7.10) χρησιμοποιείται μόνο το μέγιστο εύρος h αλλά όχι περισσότερες λεπτομέρειες για το διαμερισμό.)

Για να αντιμετωπισθούν τέτοια προβλήματα, πέραν των εκτιμήσεων εκ των προτέρων, έχουν επινοηθεί και διαφορετικής φύσεως εκτιμήσεις, οι λεγόμενες εκτιμήσεις *εκ των υστέρων*. Η αντίστοιχη της (7.10) εκ των υστέρων εκτίμηση είναι της μορφής

$$\|u' - u_h'\| \le \eta(u_h), \tag{7.34}$$

στην οποία το *συναρτησιακό εκτίμησης του σφάλματος* (error estimator), η, εξαρτάται από τα δεδομένα του προβλήματος, όπως οι συναρτήσεις q και f, και από ποσότητες που μπορούν να υπολογισθούν αν γνωρίζουμε την προσέγγιση u_h. Σε αυτού του είδους τις εκτιμήσεις χρησιμοποιούμε δηλαδή την προσέγγιση προκειμένου να αντλήσουμε πληροφορίες για την ποιότητά της.

Στη συνέχεια θα προσδιορίσουμε κατάλληλες εκφράσεις για το συναρτησιακό εκτίμησης του σφάλματος η. Για ευκολία υποθέτουμε ότι το πρόβλημα είναι ορισμένο, δηλαδή ότι η συνάρτηση q παίρνει μόνο μη αρνητικές τιμές.

Θεωρούμε έναν διαμερισμό $a = x_0 < x_1 < \cdots < x_J < x_{J+1} = b$ του $[a, b]$ και θέτουμε $h_i := x_{i+1} - x_i, i = 0, \dots, J$. Υποθέτουμε ότι ο χώρος S_h^r αποτελείται από συνεχείς συναρτήσεις, οι οποίες σε κάθε υποδιάστημα $[x_i, x_{i+1}]$ είναι πολυώνυμα βαθμού το πολύ $r - 1$, και μηδενίζονται στα άκρα a και b.

Έστω $u_h \in S_h^r$ η λύση πεπερασμένων στοιχείων για το πρόβλημα (5.1), βλ. την (7.16). Η u_h ικανοποιεί τις συνοριακές συνθήκες του προβλήματος (5.1). Ένα μέγεθος για το πόσο η προσεγγιστική λύση αποτυγχάνει να είναι ακριβής λύση του (5.1) αποτελεί το λεγόμενο *υπόλοιπο* r_h,

$$r_h(x) := -u_h''(x) + q(x)u_h(x) - f(x). \tag{7.35}$$

Το υπόλοιπο παίζει καθοριστικό ρόλο στις εκ των υστέρων εκτιμήσεις, ακριβώς αντίστοιχο του σφάλματος συνέπειας στις εκ των προτέρων εκτιμήσεις· υπενθυμίζουμε ότι το σφάλμα συνέπειας είναι ένα μέγεθος για το πόσο η ακριβής λύση αποτυγχάνει να είναι προσεγγιστική λύση, να ικανοποιεί δηλαδή την αριθμητική μέθοδο. Σημειώνουμε ακόμη ότι το υπόλοιπο δεν ορίζεται στους κόμβους x_i.

Το πρώτο αποτέλεσμα αυτής της παραγράφου αφορά την εκ των υστέρων εκτίμηση της $\|u' - u_h'\|$ και δίνεται στο ακόλουθο θεώρημα.

Θεώρημα 7.3 *Έστω $u_h \in S_h^r$ η λύση του* (7.16) *και r_h το υπόλοιπό της. Έστω $u \in C^2[a, b]$ η λύση του* (5.1). *Τότε, για το σφάλμα $u - u_h$ ισχύει η εκ των υστέρων εκτίμηση*

$$\|u' - u_h'\| \le \frac{1}{2}\Big(\sum_{i=0}^{J} h_i^2 \|r_h\|^2_{L^2(x_i, x_{i+1})}\Big)^{1/2}, \tag{7.36}$$

όπου $\|\cdot\|_{L^2(x_i,x_{i+1})}$ είναι η L^2 νόρμα στο διάστημα $[x_i, x_{i+1}]$,

$$\|v\|_{L^2(x_i,x_{i+1})} := \Big(\int_{x_i}^{x_{i+1}} |v(x)|^2 dx\Big)^{1/2}.$$

Απόδειξη. Συμβολίζουμε με e_h το σφάλμα, $e_h := u - u_h$. Προφανώς ισχύει

$$\|e_h'\|^2 \le a(e_h, e_h). \tag{7.37}$$

Για να οδηγηθούμε στην (7.36) αρκεί να εκτιμήσουμε κατάλληλα το δεξιό μέλος της (7.37). Υποθέτουμε κατ' αρχάς, για να γίνουν ευκολότερα οι υπολογισμοί, ότι έχουμε μία συνάρτηση v, συνεχή και τμηματικά δύο φορές συνεχώς παραγωγίσιμη, η οποία μηδενίζεται σε όλους τους κόμβους x_i, $i = 0, \dots, J+1$, και θα προσπαθήσουμε να προσδιορίσουμε μια κατάλληλη παράσταση του $a(e_h, v)$. Σημειώνουμε από τώρα ότι εμάς μας ενδιαφέρει το $a(e_h, e_h)$ και η e_h δεν μηδενίζεται κατ' ανάγκην στους κόμβους x_i, εκτός βέβαια αν $q = 0$, βλ. την Άσκηση 7.6, θα δούμε όμως αργότερα έναν τρόπο για να ξεπεράσουμε αυτό το εμπόδιο. Επειδή θέλουμε να ολοκληρώσουμε κατά μέρη και αυτό δεν μπορεί να γίνει κατ' ευθείαν στο διάστημα $[a, b]$, αφού η v δεν είναι ομαλή στους κόμβους, γράφουμε το $a(e_h, v)$ ως

$$a(e_h, v) = \sum_{i=0}^{J} \int_{x_i}^{x_{i+1}} e_h'(x) v'(x) dx + (q e_h, v).$$

Τώρα, ολοκληρώνοντας κατά μέρη και χρησιμοποιώντας την υπόθεση $v(x_i) = v(x_{i+1}) = 0$ έχουμε

$$\int_{x_i}^{x_{i+1}} e_h'(x) v'(x) dx = -\int_{x_i}^{x_{i+1}} e_h''(x) v(x) dx$$

και οδηγούμαστε εύκολα στη σχέση

$$a(e_h, v) = -(e_h'', v) + (q e_h, v) = -(u'' - u_h'', v) + (q(u - u_h), v),$$

δηλαδή

$$a(e_h, v) = (-u'' + qu, v) - (-u_h'' + q u_h, v).$$

Τώρα $-u'' + qu = f$ και χρησιμοποιώντας και τον ορισμό του r_h οδηγούμαστε στην επιθυμητή παράσταση

$$a(e_h, v) = -(r_h, v). \tag{7.38}$$

Τώρα, λόγω της (7.18) έχουμε $a(e_h, e_h) = a(e_h, e_h - \chi)$, για οποιοδήποτε $\chi \in S_h^r$. Επιλέγουμε ως χ μια συνεχή συνάρτηση, τμηματικά πολυώνυμο πρώτου βαθμού, τέτοια ώστε $e_h(x_i) - \chi(x_i) = 0$, $i = 0, \dots, J+1$, δηλαδή η χ είναι η τμηματικά γραμμική παρεμβάλλουσα της e_h στους κόμβους $x_i, i = 0, \dots, J+1$. Σημειώνουμε από τώρα ότι η χ δεν είναι κάποια εκ των υστέρων ποσότητα, δεν μπορούμε να την υπολογίσουμε δηλαδή, αυτό όμως δεν θα μας εμποδίσει να τη χρησιμοποιήσουμε στη θεωρία για να οδηγηθούμε στην εκτίμηση που θέλουμε. Πριν προχωρήσουμε σημειώνουμε ότι

$$\int_{x_i}^{x_{i+1}} \big(e_h'(x) - \chi'(x)\big) \chi'(x) dx = -\int_{x_i}^{x_{i+1}} \big(e_h(x) - \chi(x)\big) \chi''(x) dx,$$

και, αφού η χ είναι στο διάστημα $[x_i, x_{i+1}]$ πολυώνυμο πρώτου βαθμού, συμπεραίνουμε αμέσως ότι

$$\int_{x_i}^{x_{i+1}} \big(e_h'(x) - \chi'(x)\big)\chi'(x)dx = 0.$$

Επομένως ισχύει

$$\int_{x_i}^{x_{i+1}} |e_h'(x) - \chi'(x)|^2 dx = \int_{x_i}^{x_{i+1}} \big(e_h'(x) - \chi'(x)\big)e_h'(x)dx$$

και εκτιμώντας το δεξιό μέλος με την ανισότητα των Cauchy–Schwarz λαμβάνουμε

$$\|e_h' - \chi'\|_{L^2(x_i,x_{i+1})} \le \|e_h'\|_{L^2(x_i,x_{i+1})}. \tag{7.39}$$

Εξ άλλου, χρησιμοποιώντας την ανισότητα των Poincaré–Friedrichs (5.18) έχουμε

$$\|e_h - \chi\|_{L^2(x_i,x_{i+1})} \le \frac{1}{2} h_i \|e_h' - \chi'\|_{L^2(x_i,x_{i+1})}\,. \tag{7.40}$$

Συνδυάζοντας τις (7.39) και (7.40) οδηγούμαστε σε ένα χρήσιμο για τη συνέχεια αποτέλεσμα,

$$\|e_h - \chi\|_{L^2(x_i,x_{i+1})} \le \frac{1}{2} h_i \|e_h'\|_{L^2(x_i,x_{i+1})}\,. \tag{7.41}$$

Σύμφωνα με τα προηγούμενα έχουμε τώρα, χρησιμοποιώντας πρώτα την ανισότητα των Cauchy–Schwarz για ολοκληρώματα, μετά την (7.41) και τέλος πάλι την ανισότητα των Cauchy–Schwarz, αυτή τη φορά για αθροίσματα,

$$\begin{aligned}
a(e_h, e_h) &= a(e_h, e_h - \chi) = -(r_h, e_h - \chi) \\
&\le \sum_{i=0}^{J} \|r_h\|_{L^2(x_i,x_{i+1})} \, \|e_h - \chi\|_{L^2(x_i,x_{i+1})} \\
&\le \frac{1}{2} \sum_{i=0}^{J} \|r_h\|_{L^2(x_i,x_{i+1})} \, h_i \|e_h'\|_{L^2(x_i,x_{i+1})} \\
&\le \frac{1}{2} \Big(\sum_{i=0}^{J} h_i^2 \|r_h\|^2_{L^2(x_i,x_{i+1})} \Big)^{1/2} \Big(\sum_{i=0}^{J} \|e_h'\|^2_{L^2(x_i,x_{i+1})} \Big)^{1/2},
\end{aligned}$$

δηλαδή

$$a(e_h, e_h) \le \frac{1}{2} \Big(\sum_{i=0}^{J} h_i^2 \|r_h\|^2_{L^2(x_i,x_{i+1})} \Big)^{1/2} \|e_h'\|\,. \tag{7.42}$$

Από τις (7.37) και (7.42) έπεται αμέσως η εκτίμηση (7.36). $\square$

Θα προχωρήσουμε τώρα στην εκ των υστέρων εκτίμηση του σφάλματος $u - u_h$ στη νόρμα του L^2. Αρχίζουμε με κάποια προκαταρκτικά αποτελέσματα. Όπως και στην αντίστοιχη εκ των προτέρων εκτίμηση, το βασικό επιχείρημα βασίζεται στο τέχνασμα του Nitsche. Θεωρούμε λοιπόν το πρόβλημα (7.20). Κατ' αρχάς, σύμφωνα με την ανισότητα των Poincaré–Friedrichs (5.18), έχουμε

$$\|w\| \le \frac{b-a}{2}\|w'\|. \tag{7.43}$$

Παίρνοντας το εσωτερικό γινόμενο στη διαφορική εξίσωση του (7.20) με την w διαπιστώνουμε αμέσως ότι $\|w'\|^2 \le \|u-u_h\|\,\|w\|$, οπότε η (7.43) δίνει

$$\|w\| \le \frac{(b-a)^2}{4}\|u-u_h\|. \tag{7.44}$$

Χρησιμοποιώντας τώρα τη διαφορική εξίσωση του (7.20) παίρνουμε

$$\|w''\| \le \|q\|_\infty\,\|w\| + \|u-u_h\|,$$

όπου $\|\cdot\|_\infty$ είναι η νόρμα μεγίστου στο διάστημα $[a,b]$, επομένως, συνδυάζοντας με την (7.44) λαμβάνουμε τη σημαντική για τη συνέχεια εκτίμηση

$$\|w''\| \le \big[1+\frac{(b-a)^2}{4}\|q\|_\infty\big]\,\|u-u_h\|. \tag{7.45}$$

Θεώρημα 7.4 *Έστω $u_h \in S_h^r$ η λύση του* (7.16) *και r_h το υπόλοιπό της. Έστω $u \in C^2[a,b]$ η λύση του* (5.1). *Τότε, για το σφάλμα $u-u_h$ ισχύει η εκ των υστέρων εκτίμηση*

$$\|u-u_h\| \le \frac{1}{2}\big[1+\frac{(b-a)^2}{4}\|q\|_\infty\big]\Big(\sum_{i=0}^{J} h_i^4\|r_h\|^2_{L^2(x_i,x_{i+1})}\Big)^{1/2}. \tag{7.46}$$

Απόδειξη. Συμβολίζουμε πάλι με e_h το σφάλμα, $e_h := u-u_h$. Σύμφωνα με την (7.22) έχουμε

$$\|e_h\|^2 = a(e_h,w). \tag{7.47}$$

Έστω τώρα χ η τμηματικά γραμμική παρεμβάλουσα της w στους κόμβους $x_i, i = 0,\dots,J+1$, $w(x_i)-\chi(x_i)=0$, $i=0,\dots,J+1$. Τότε $\chi \in S_h^r$, και η (7.47) γράφεται στη μορφή

$$\|e_h\|^2 = a(e_h,w-\chi). \tag{7.48}$$

Τώρα

$$\|w-\chi\|_{L^2(x_i,x_{i+1})} \le \frac{1}{2}h_i\|w'-\chi'\|_{L^2(x_i,x_{i+1})}\,, \tag{7.49}$$

βλ. την (7.40). Εξ άλλου, σύμφωνα με το θεώρημα του Rolle, η $w'-\chi'$ μηδενίζεται σε κάποιο σημείο $\xi\in(x_i,x_{i+1})$, οπότε

$$w'(x)-\chi'(x)=\int_\xi^x[w''(s)-\chi''(s)]\,ds=\int_\xi^x w''(s)\,ds,\quad x\in(x_i,x_{i+1}),$$

συνεπώς, χρησιμοποιώντας την ανισότητα των Cauchy–Schwarz,

$$|w'(x)-\chi'(x)|^2\le h_i\int_{x_i}^{x_{i+1}}|w''(x)|^2\,dx,\quad x\in(x_i,x_{i+1}).$$

Από αυτή τη σχέση έπεται αμέσως ότι

$$\|w'-\chi'\|_{L^2(x_i,x_{i+1})}\le h_i\|w''\|_{L^2(x_i,x_{i+1})}\,. \tag{7.50}$$

Οι (7.49) και (7.50) δίνουν

$$\|w-\chi\|_{L^2(x_i,x_{i+1})}\le \frac{1}{2}h_i^2\|w''\|_{L^2(x_i,x_{i+1})}\,. \tag{7.51}$$

Προχωρώντας όπως στην απόδειξη της (7.42) και χρησιμοποιώντας τις (7.48), (7.38) και (7.51), έχουμε

$$\begin{aligned}\|e_h\|^2&=a(e_h,w-\chi)=-(r_h,w-\chi)\\&\le\sum_{i=0}^J\|r_h\|_{L^2(x_i,x_{i+1})}\,\|w-\chi\|_{L^2(x_i,x_{i+1})}\\&\le\frac{1}{2}\sum_{i=0}^J\|r_h\|_{L^2(x_i,x_{i+1})}\,h_i^2\|w''\|_{L^2(x_i,x_{i+1})}\\&\le\frac{1}{2}\Big(\sum_{i=0}^J h_i^4\|r_h\|^2_{L^2(x_i,x_{i+1})}\Big)^{1/2}\Big(\sum_{i=0}^J\|w''\|^2_{L^2(x_i,x_{i+1})}\Big)^{1/2},\end{aligned}$$

δηλαδή

$$\|e_h\|^2\le\frac{1}{2}\Big(\sum_{i=0}^J h_i^4\|r_h\|^2_{L^2(x_i,x_{i+1})}\Big)^{1/2}\|w''\|\,. \tag{7.52}$$

Από τις (7.45) και (7.52) έπεται αμέσως η εκτίμηση (7.46). □

Παρατήρηση 7.3

i. Χρησιμοποιώντας τη διαφορική εξίσωση του (5.1) στην (7.35), μπορούμε να γράψουμε το υπόλοιπο στη μορφή

$$r_h = (u - u_h)'' + q(u - u_h). \tag{7.53}$$

Αυτή η παράσταση είναι χρήσιμη για τη μελέτη της ασυμπτωτικής συμπεριφοράς του υπολοίπου, καθώς το h τείνει στο μηδέν.

ii. Στην απόδειξη του Θεωρήματος 7.3 είδαμε ότι ισχύει η σχέση $a(e_h, e_h) = -(r_h, e_h - \chi)$. Εκτιμώντας το δεξιό μέλος με την ανισότητα του τύπου της Cauchy–Schwarz της Άσκησης 5.15, λαμβάνουμε

$$a(e_h, e_h) \le \|r_h\|_{-1}\, \|e_h' - \chi'\|,$$

οπότε, με τη βοήθεια και της (7.39), έχουμε

$$a(e_h, e_h) \le \|r_h\|_{-1}\, \|e_h'\|.$$

Συνδυάζοντας αυτό το αποτέλεσμα με την (7.37) οδηγούμαστε στην εξής εκ των υστέρων εκτίμηση, βλ. την (7.36),

$$\|u' - u_h'\| \le \|r_h\|_{-1}. \tag{7.54}$$

Παρατήρηση 7.3 Οι εκ των υστέρων εκτιμήσεις είναι προφανώς χρήσιμες γιατί μας δίνουν ακριβείς πληροφορίες για την προσεγγιστική ιδιότητα της προσεγγιστικής λύσης. Αποτελούν επίσης τη βάση για μεθόδους αυτόματης επιλογής κατάλληλων διαμερισμών του διαστήματος $[a, b]$. Αυτό αποτελεί θέμα μεγάλης πρακτικής σημασίας, γιατί η λύση u προσεγγίζεται δύσκολα στα διαστήματα όπου ταλαντώνεται γρήγορα, μεταβάλλεται δηλαδή γρήγορα, και προσεγγίζεται πολύ ευκολότερα στα διαστήματα όπου μεταβάλλεται αργά. Δυστυχώς δεν γνωρίζουμε γενικά εκ των προτέρων σε ποιες περιοχές συμβαίνει είτε το ένα είτε το άλλο. Το να καταφύγει κανείς σε πολύ εκλεπτυσμένους διαμερισμούς προκειμένου να προσεγγίσει τη λύση παντού καλά, αποτελεί πολύ δαπανηρή διέξοδο. Στη συνέχεια θα δούμε έναν τρόπο αυτόματης επιλογής του διαμερισμού, ο οποίος βασίζεται σε εκ των υστέρων εκτιμήσεις, για να εντοπίσει τα διαστήματα στα οποία χρειάζονται πολλοί κόμβοι καθώς και εκείνα στα οποία λίγοι κόμβοι αρκούν.

Ας υποθέσουμε φερ' ειπείν ότι θέλουμε η νόρμα $\|u' - u_h'\|$ να μην ξεπερνά έναν προκαθορισμένο θετικό αριθμό ε. Σύμφωνα με την (7.36), εύκολα διαπιστώνουμε ότι αυτό εξασφαλίζεται από τις συνθήκες

$$h_i \|r_h\|^2_{L^2(x_i,x_{i+1})} \leq \frac{4\varepsilon^2}{b-a} \tag{7.55}$$

για όλα τα υποδιαστήματα $[x_i, x_{i+1}]$. Το μέγεθος $h_i \|r_h\|^2_{L^2(x_i,x_{i+1})}$ αποτελεί ένδειξη για την *τοπική* συμπεριφορά του σφάλματος, ακριβέστερα για τη συνεισφορά του διαστήματος $[x_i, x_{i+1}]$ στο συναρτησιακό εκτίμησης του (συνολικού) σφάλματος και λέγεται *συναρτησιακό ένδειξης του σφάλματος* (error indicator). Μεγάλο συναρτησιακό ένδειξης σφάλματος είναι ένδειξη μη καλής προσέγγισης της λύσης, δηλαδή ταχείας μεταβολής της λύσης, οπότε εκλεπτύνουμε τοπικά τον διαμερισμό, μικρό συναρτησιακό ένδειξης σφάλματος είναι ένδειξη καλής προσέγγισης της λύσης, δηλαδή βραδείας μεταβολής της λύσης, οπότε μπορούμε να αφαιρέσουμε τοπικά κάποιους από τους κόμβους του διαμερισμού, γεγονός που συνεπάγεται μείωση του υπολογιστικού κόστους.

Ένας τρόπος κατάλληλης επιλογής του διαμερισμού, ο οποίος βασίζεται στην πληροφορία που παίρνουμε από το συναρτησιακό ένδειξης του σφάλματος, είναι συνεπώς ο εξής, βλέπε την αντίστοιχη συζήτηση στην παράγραφο 3.5, η οποία αφορά την αυτόματη επιλογή του βήματος στην επίλυση προβλημάτων αρχικών τιμών με μεθόδους των Runge–Kutta: Αρχίζουμε με έναν διαμερισμό με λίγους κόμβους, ας πούμε ομοιόμορφο. Θέτουμε $\varepsilon_1 := \frac{4\varepsilon^2}{b-a}$. Διακρίνουμε τώρα τρεις περιπτώσεις:

I. Αν το συναρτησιακό ένδειξης του σφάλματος στο διάστημα $[x_i, x_{i+1}]$ δεν είναι ούτε υπερβολικά μεγάλο ούτε υπερβολικά μικρό, π.χ. αν ισχύει

$$\frac{1}{10}\varepsilon_1 \leq h_i \|r_h\|^2_{L^2(x_i,x_{i+1})} \leq \varepsilon_1,$$

τότε αφήνουμε το διάστημα $[x_i, x_{i+1}]$ όπως είναι.

II. Αν το συναρτησιακό ένδειξης του σφάλματος στο διάστημα $[x_i, x_{i+1}]$ είναι υπερβολικά μεγάλο, δηλαδή αν ισχύει

$$h_i \|r_h\|^2_{L^2(x_i,x_{i+1})} > \varepsilon_1,$$

τότε εκλεπτύνουμε τοπικά τον διαμερισμό, συμπεριλαμβάνοντας στο διαμερισμό έναν, φερ' ειπείν, επί πλέον κόμβο στο διάστημα $[x_i, x_{i+1}]$, το μέσον του διαστήματος παραδείγματος χάριν.

III. Αν το συναρτησιακό ένδειξης του σφάλματος στα διαστήματα $[x_{i-1}, x_i]$ και $[x_i, x_{i+1}]$ είναι υπερβολικά μικρό, π.χ. αν ισχύει

$$h_{i-1}\|r_h\|^2_{L^2(x_{i-1},x_i)} < \frac{1}{10}\varepsilon_1 \quad \text{και} \quad h_i\|r_h\|^2_{L^2(x_i,x_{i+1})} < \frac{1}{10}\varepsilon_1,$$

τότε αυτό είναι ένδειξη ότι ο κόμβος x_i μπορεί να μην είναι απαραίτητος στον διαμερισμό και τον αφαιρούμε. Αν για τη νέα προσέγγιση για το νέο διάστημα οδηγηθούμε στην περίπτωση II, δηλαδή προκύψουν υπερβολικά μεγάλα συναρτησιακά σφάλματος, τότε συμπεριλαμβάνουμε πάλι το x_i στο διαμερισμό και δεν το αφαιρούμε ποτέ στο μέλλον.

Φυσικά, δεν μπορούμε να έχουμε κόμβους με πολύ μικρή απόσταση μεταξύ τους, επομένως στην περίπτωση II, αν το h_i είναι πολύ μικρό, μικρότερο από έναν προκαθορισμένο θετικό αριθμό δ, αφήνουμε το διάστημα $[x_i, x_{i+1}]$ όπως είναι, και στο τέλος του υπολογισμού δίνουμε σχετικό μήνυμα στο χρήστη.

Η διαδικασία αυτή γίνεται για όλα τα υποδιαστήματα του αρχικού διαμερισμού και οδηγεί σε έναν νέο διαμερισμό. Εν συνεχεία αν δεν ισχύει ακόμη η (7.55) για όλα τα υποδιαστήματα με μήκος μεγαλύτερο του δ, επαναλαμβάνεται η όλη διαδικασία για τον νέο διαμερισμό.

Ασκήσεις

7.1. Πώς διατυπώνεται η μέθοδος των πεπερασμένων στοιχείων για το πρόβλημα (5.14); Αποδείξτε για αυτή την περίπτωση ένα αποτέλεσμα αντίστοιχο του Θεωρήματος 7.1.

7.2. Θεωρούμε το πρόβλημα (5.15). Θεωρούμε μια οικογένεια $(S_h^r)_{0<h\le b-a}$ γραμμικών χώρων πεπερασμένης διάστασης συνεχών και τμηματικά συνεχώς παραγωγίσιμων συναρτήσεων, με προσεγγιστική ιδιότητα ανάλογη της (7.2) αλλά για $v \in C^s[a,b]$ (χωρίς να απαιτείται δηλαδή $v(a) = v(b) = 0$). Ορίστε την προσεγγιστική

λύση $u_h \in S_h^r$ για αυτό το πρόβλημα και αποδείξτε εκτιμήσεις αντίστοιχες των (7.10) και (7.11) γι' αυτή την περίπτωση.

7.3. Τι χώρους πεπερασμένων στοιχείων θα επιλέγατε για το πρόβλημα δύο σημείων

$$\begin{cases} -u'' + qu = f & \text{στο } [a,b] \\ u(a) = u'(b) = 0\,; \end{cases}$$

7.4. Έστω ότι η συνάρτηση q λαμβάνει μόνο μη αρνητικές τιμές. Αποδείξτε ότι η λύση u του προβλήματος (5.1) είναι το μόνο στοιχείο του χώρου V στο οποίο το συναρτησιακό J, $J(v) := a(v,v) - 2(f,v)$, λαμβάνει την ελάχιστη τιμή του στον V.

7.5. Θεωρούμε το πρόβλημα δύο σημείων με διεπιφάνεια της Άσκησης 5.12. Δώστε μία μέθοδο πεπερασμένων στοιχείων για αυτό το πρόβλημα, αντίστοιχη της (7.7), και αποδείξτε εκτιμήσεις αντίστοιχες των (7.10) και (7.11).

7.6. Θεωρούμε το πρόβλημα δύο σημείων

$$\begin{cases} -u'' = f & \text{στο } [a,b] \\ u(a) = u(b) = 0\,. \end{cases}$$

Έστω S_h^r ένας υπόχωρος του V αποτελούμενος από συνεχείς συναρτήσεις οι οποίες σε καθένα των υποδιαστημάτων $[x_i, x_{i+1}]$, $i = 0, \dots, J$, ενός διαμερισμού $a = x_0 < x_1 < \cdots < x_{J+1} = b$ του $[a,b]$ είναι πολυώνυμα βαθμού το πολύ r. Αποδείξτε ότι

$$u_h(x_i) = u(x_i), \quad i = 0, \dots, J+1.$$

[*Υπόδειξη*: Για $i = 1, \dots, J$ η συνάρτηση χ,

$$\chi(x) = \begin{cases} \dfrac{x-a}{x_i - a}, & a \le x \le x_i, \\ \dfrac{x-b}{x_i - b}, & x_i < x \le b, \end{cases}$$

είναι στοιχείο του S_h^r.]

7.7. Θεωρούμε το πρόβλημα (5.1) και την προσεγγιστική λύση πεπερασμένων στοιχείων $u_h \in S_h^2$, με S_h^2 όπως στο Παράδειγμα 7.2. Αποδείξτε ότι

$$-\frac{1}{h}[u_h(x_{i-1}) - 2u_h(x_i) + u_h(x_{i+1})] + \int_{x_{i-1}}^{x_{i+1}} q(x) u_h(x) \varphi_i(x)\,dx =$$
$$= \int_{x_{i-1}}^{x_{i+1}} f(x) \varphi_i(x)\,dx, \ i = 1, \dots, J.$$

Προσεγγίζοντας τα ολοκληρώματα με τον σύνθετο τύπο του τραπεζίου, δηλ. το $\int_{x_{i-1}}^{x_{i+1}} g(x)dx$ με $h[g(x_{i-1})/2 + g(x_i) + g(x_{i+1})/2]$, οδηγούμαστε στο σχήμα

$$-\frac{1}{h}[U_h(x_{i-1}) - 2U_h(x_i) + U_h(x_{i+1})] + hq(x_i)U_h(x_i) = hf(x_i), \quad i = 1, \dots, J.$$

Ποια σχέση έχει αυτό το σχήμα με τη μέθοδο πεπερασμένων διαφορών (6.3);

7.8. Θεωρούμε το πρόβλημα δύο σημείων (5.1). Έστω S_h^2 ένας υπόχωρος του V αποτελούμενος από συνεχείς συναρτήσεις οι οποίες σε καθένα των υποδιαστημάτων $[x_i, x_{i+1}]$, $i = 0, \dots, J$, ενός διαμερισμού $a = x_0 < x_1 < \cdots < x_{J+1} = b$ του $[a, b]$ είναι πολυώνυμα βαθμού το πολύ ένα. Χρησιμοποιήστε τις εκ των υστέρων εκτιμήσεις (7.36) και (7.46) για να οδηγηθείτε στις εκ των προτέρων εκτιμήσεις (7.10) και (7.11), με $r = 2$. Αυτή η διαδικασία επιβεβαιώνει ότι οι εκ των υστέρων εκτιμήσεις, τουλάχιστον στην εν λόγω περίπτωση, είναι βέλτιστης τάξεως. [*Υπόδειξη*: Από την (7.36), φερ' ειπείν, παίρνουμε

$$\|u' - u_h'\| \le \frac{1}{2}h\|r_h\|.$$

Αρκεί συνεπώς να αποδείξουμε ότι η $\|r_h\|$ είναι φραγμένη, με σταθερά ανεξάρτητη του διαμερισμού. Η (7.53) τώρα γράφεται στη μορφή $r_h = u'' + q(u - u_h)$ και αν χρησιμοποιήσουμε το γεγονός ότι $\|u_h\| \le \frac{b-a}{2}\|f\|$ οδηγούμεθα εύκολα στο ζητούμενο.]

Μέρος ΙΙ

Δυναμικές μερικές διαφορικές εξισώσεις

Στο δεύτερο μέρος γνωρίσαμε μεθόδους πεπερασμένων διαφορών καθώς και πεπερασμένων στοιχείων για το πρόβλημα δύο σημείων. Στο μέρος αυτό θα ασχοληθούμε με μεθόδους πεπερασμένων διαφορών και πεπερασμένων στοιχείων για απλά προβλήματα δυναμικών Μερικών Διαφορικών Εξισώσεων, και συγκεκριμένα για προβλήματα αρχικών και συνοριακών συνθηκών για μία παραβολική καθώς και για μία υπερβολική εξίσωση σε μία χωρική διάσταση.

Το Μέρος III αποτελείται από πέντε κεφάλαια. Στο κεφάλαιο 8 παρουσιάζουμε συνοπτικά ένα πρόβλημα αρχικών και συνοριακών συνθηκών για μία παραβολική εξίσωση. Στα κεφάλαια 9 και 10 μελετούμε μεθόδους πεπερασμένων διαφορών και μεθόδους πεπερασμένων στοιχείων, αντίστοιχα, για παραβολικά προβλήματα. Τέλος στο κεφάλαιο 11 ασχολούμαστε με ένα πρόβλημα αρχικών και συνοριακών συνθηκών για μία υπερβολική εξίσωση και στο κεφάλαιο 12 με τη διακριτοποίησή του με μεθόδους πεπερασμένων διαφορών.

4. Μία παραβολική εξίσωση

Έστω $T > 0$. Θεωρούμε το εξής πρόβλημα αρχικών και συνοριακών συνθηκών για μία παραβολική εξίσωση: Ζητείται μία συνάρτηση $u : [a, b] \times [0, T] \to \mathbb{R}$, δύο φορές συνεχώς παραγωγίσιμη ως προς x και μία φορά ως προς t τέτοια ώστε

$$\begin{cases} u_t = (pu_x)_x - qu + f & \text{στο } [a, b] \times [0, T] \\ u(a, \cdot) = u(b, \cdot) = 0 & \text{στο } [0, T] \\ u(\cdot, 0) = u_0 & \text{στο } [a, b], \end{cases} \tag{8.1}$$

όπου $q, f \in C([a, b] \times [0, T])$ και $p \in C^1([a, b] \times [0, T]), \quad p(x, t) > 0$ για $x \in [a, b]$ και $t \in [0, T]$. Στη συνέχεια θα υποθέτουμε επιπλέον συχνά ότι

$$q(x, t) \geq 0 \qquad \forall (x, t) \in [a, b] \times [0, T]. \tag{8.2}$$

Σε αντίθεση με το πρόβλημα δύο σημείων, στην προκειμένη περίπτωση η συνθήκη (8.2) δεν είναι ουσιαστική, με την έννοια ότι δεν επηρεάζει την ύπαρξη και τη μοναδικότητα της λύσης του προβλήματος. Πραγματικά θέτοντας $\tilde{u}(x, t) := u(x, t)\mathrm{e}^{-\lambda t}$ με μια θετική σταθερά λ, το πρόβλημα (8.1) γράφεται ισοδύναμα στη μορφή

$$\begin{cases} \tilde{u}_t = (p\tilde{u}_x)_x - (\lambda + q)\tilde{u} + \mathrm{e}^{-\lambda t} f & \text{στο } [a, b] \times [0, T] \\ \tilde{u}(a, \cdot) = \tilde{u}(b, \cdot) = 0 & \text{στο } [0, T] \\ \tilde{u}(\cdot, 0) = u_0 & \text{στο } [a, b]. \end{cases} \tag{8.1$'$}$$

Αν το λ επιλεγεί αρκετά μεγάλο, τότε η αντίστοιχη της συνθήκης (8.2) για το πρόβλημα (8.1$'$) προφανώς ικανοποιείται. Στο εξής θα υποθέτουμε συχνά ότι έχει ήδη γίνει ένας τέτοιος μετασχηματισμός, θα θεωρούμε το πρόβλημα στη μορφή (8.1), και θα υποθέτουμε ότι ισχύει η (8.2).

Είναι γνωστό ότι αν τα δεδομένα είναι αρκετά ομαλά και συμβατά, τότε το πρόβλημα (8.1) έχει ακριβώς μια λύση. Αναγκαία συνθήκη για τη συνέχεια της u είναι η $u_0(a) = u_0(b) = 0$, ενώ για τη συνέχεια της u_t απαιτείται η παράσταση

$$(pu_0')' - qu_0 + f$$

να μηδενίζεται στα σημεία $(a, 0)$ και $(b, 0)$.

Συνεχής εξάρτηση της λύσης από τα αρχικά δεδομένα. Έστω u_1 μια λύση του προβλήματος (8.1), και u_2 μια λύση του αντίστοιχου προβλήματος με την ίδια Δ.Ε. και τις ίδιες συνοριακές συνθήκες αλλά με αρχική συνθήκη $\tilde{u}_0$ αντί για u_0. Τότε για την $u := u_1 - u_2$ έχουμε

$$\begin{cases} u_t = (pu_x)_x - qu & \text{στο } [a,b] \times [0,T] \\ u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T] \\ u(\cdot,0) = u_0 - \tilde{u}_0 & \text{στο } [a,b]. \end{cases} \tag{8.3}$$

Πολλαπλασιάζοντας τη Δ.Ε. στο (8.3) επί u, ολοκληρώνοντας στο $[a, b]$, ολοκληρώνοντας κατά μέρη και χρησιμοποιώντας τις συνοριακές συνθήκες παίρνουμε, λαμβάνοντας υπ' όψιν και την (8.2),

$$\int_a^b u_t(x,t)u(x,t)dx \le -\int_a^b p(x,t)[u_x(x,t)]^2 dx.$$

Τώρα

$$\int_a^b u_t(x,t)u(x,t)dx = \frac{1}{2}\frac{d}{dt}\int_a^b [u(x,t)]^2 dx = \frac{1}{2}\frac{d}{dt}\|u(\cdot,t)\|^2,$$

οπότε η προηγούμενη σχέση γράφεται στη μορφή

$$\frac{d}{dt}\|u(\cdot,t)\|^2 \le -2\int_a^b p(x,t)[u_x(x,t)]^2 dx. \tag{8.4}$$

Από εδώ συμπεραίνουμε αμέσως ότι η $\|u(\cdot,t)\|$ είναι φθίνουσα συνάρτηση του t, ιδιαίτερα λοιπόν

$$\|u(\cdot,t)\| \le \|u(\cdot,0)\| \qquad \forall t \in [0,T] \tag{8.5}$$

ή

$$\max_{0\le t\le T} \|u_1(\cdot,t) - u_2(\cdot,t)\| \le \|u_0 - \tilde{u}_0\|, \tag{8.6}$$

μικρή δηλαδή μεταβολή της αρχικής συνθήκης συνεπάγεται μικρή μεταβολή της λύσης ή, όπως συνηθίζεται να λέγεται, η λύση *εξαρτάται συνεχώς* από τα αρχικά δεδομένα.

Μπορούμε εύκολα να βελτιώσουμε την εκτίμηση (8.5), αν εκμεταλλευθούμε καλύτερα το δεξιό μέλος της (8.4). Σύμφωνα με την (5.5), από την (8.4) λαμβάνουμε

$$\begin{aligned}\frac{d}{dt}\|u(\cdot,t)\|^2 &\le -2\,\min_{x,t} p(x,t)\,\|u_x(\cdot,t)\|^2\\ &\le -2\mu\,\|u(\cdot,t)\|^2,\end{aligned}$$

όπου $\mu := \frac{1}{(b-a)^2}\min_{x,t} p(x,t)$. Συνεπώς

$$\frac{d}{dt}[\mathrm{e}^{2\mu t}\|u(\cdot,t)\|^2] \le 0,$$

οπότε

$$\|u(\cdot,t)\| \le \mathrm{e}^{-\mu t}\|u(\cdot,0)\| \qquad \forall t\in[0,T]. \tag{8.7}$$

Συνεχής εξάρτηση από τον μη ομογενή όρο. Πολλαπλασιάζοντας τη Δ.Ε. στο (8.1) επί u, ολοκληρώνοντας στο $[a,b]$, ολοκληρώνοντας κατά μέρη και χρησιμοποιώντας τις συνοριακές συνθήκες και την (8.2) λαμβάνουμε

$$\frac{1}{2}\frac{d}{dt}\|u(\cdot,t)\|^2 \le -\min_{x,t} p(x,t)\|u_x(\cdot,t)\|^2 + (f(\cdot,t),u(\cdot,t)). \tag{8.8}$$

Συνεπώς, σύμφωνα με την (5.5),

$$\frac{d}{dt}\|u(\cdot,t)\|^2 \le -2\mu\|u(\cdot,t)\|^2 + 2\|f(\cdot,t)\|\,\|u(\cdot,t)\|$$

ή

$$\frac{d}{dt}\|u(\cdot,t)\|^2 \le \frac{1}{2\mu}\|f(\cdot,t)\|^2$$

οπότε

$$\|u(\cdot,t)\|^2 \le \|u(\cdot,0)\|^2 + \frac{1}{2\mu}\int_0^t \|f(\cdot,s)\|^2 ds \qquad \forall t\in[0,T], \tag{8.9}$$

όπου $\mu = \frac{1}{(b-a)^2}\min_{x,t} p(x,t)$. Αν u_1 και u_2 είναι λύσεις των προβλημάτων της μορφής (8.1) με τις ίδιες αρχικές και συνοριακές συνθήκες αλλά με Δ.Ε. με μη ομογενείς όρους f_1 και f_2 αντίστοιχα, τότε εύκολα συμπεραίνουμε από τη (8.9) ότι

$$\|u_1(\cdot,t) - u_2(\cdot,t)\|^2 \leq \frac{1}{2\mu}\int_0^t \|f_1(\cdot,s) - f_2(\cdot,s)\|^2 ds \qquad \forall t \in [0,T]. \tag{8.10}$$

Παρατηρήσεις 8.1.

(i). Τόσο από την (8.6) όσο και από την (8.10) έπεται ότι το πρόβλημα (8.1) έχει το πολύ μια ομαλή λύση.

(ii). Μελετήσαμε ξεχωριστά τη συνεχή εξάρτηση της λύσεως από τα αρχικά δεδομένα και από τον μη ομογενή όρο για καθαρά παιδαγωγικούς λόγους. Από την (8.9) συμπεραίνουμε ότι η λύση εξαρτάται συνεχώς τόσο από τα αρχικά δεδομένα όσο και από τον μη ομογενή όρο.

(iii). Οι συνοριακές συνθήκες Dirichlet στο πρόβλημα (8.1) μπορούν να αντικατασταθούν με άλλες συνοριακές συνθήκες, όπως ακριβώς και στο πρόβλημα δύο σημείων, φερ' ειπείν με συνοριακές συνθήκες Neumann, με μεικτές συνοριακές συνθήκες ή με περιοδικές συνοριακές συνθήκες. Θα ασχοληθούμε εν συντομία με την περίπτωση συνοριακών συνθηκών Neumann. Θεωρούμε λοιπόν το πρόβλημα αρχικών και συνοριακών συνθηκών

$$\begin{cases} u_t = (pu_x)_x - qu + f & \text{στο } [a,b]\times[0,T] \\ u_x(a,\cdot) = u_x(b,\cdot) = 0 & \text{στο } [0,T] \\ u(\cdot,0) = u_0 & \text{στο } [a,b]. \end{cases} \tag{8.11}$$

Οι συνθήκες ομαλότητος των p, q και f καθώς και του προσήμου της p είναι οι ίδιες με εκείνες για το πρόβλημα (8.1), για την q όμως επιπλέον της (8.2) υποθέτουμε τώρα ότι

$$q^\star := \min_{0\leq t\leq T}\int_a^b q(x,t)dx > 0. \tag{8.12}$$

Όπως στην περίπτωση του προβλήματος (8.1) έτσι και τώρα ο περιορισμός αυτός στην q δεν είναι ουσιαστικός, βλ. την (8.1′). Πολλαπλασιάζοντας τη Δ.Ε. στο (8.11) επί u, ολοκληρώνοντας στο $[a,b]$ και ολοκληρώνοντας κατά μέρη λαμβάνουμε

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|u(\cdot,t)\|^2 \leq &-\min_{x,t} p(x,t)\|u_x(\cdot,t)\|^2 - q^\star \min_x |u(x,t)|^2\\ &+ (f(\cdot,t),u(\cdot,t)).\end{aligned}$$

Σύμφωνα με την υπόδειξη στην Άσκηση 5.9 έχουμε λοιπόν

$$\frac{1}{2}\frac{d}{dt}\|u(\cdot,t)\|^2 \leq -c\|u(\cdot,t)\|^2 + \|f(\cdot,t)\|\,\|u(\cdot,t)\|$$

με μια θετική σταθερά c, και συμπεραίνουμε εύκολα ότι

$$\|u(\cdot,t)\|^2 \leq \|u(\cdot,0)\|^2 + c_1\int_0^t \|f(\cdot,s)\|^2 ds \qquad \forall t\in[0,T]. \tag{8.13}$$

Από την (8.13) έπεται η συνεχής εξάρτηση της λύσεως του προβλήματος (8.11), τόσο από τα αρχικά δεδομένα όσο και από τον μη ομογενή όρο.
(iv). Θεωρούμε τώρα το αντίστοιχο του προβλήματος (8.1) για μια πιο γενική Δ.Ε.

$$\begin{cases} u_t = (pu_x)_x - ru_x - qu + f & \text{στο } [a,b]\times[0,T]\\ u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T]\\ u(\cdot,0) = u_0 & \text{στο } [a,b]\end{cases} \tag{8.14}$$

με τις ίδιες συνθήκες ομαλότητος για τις p,q,f και $p(x,t)>0$ και υποθέτουμε τώρα ότι η r είναι συνεχώς παραγωγίσιμη ως προς την πρώτη μεταβλητή και

$$q(x,t) - \frac{1}{2}r_x(x,t) \geq 0 \quad \forall (x,t)\in[a,b]\times[0,T]. \tag{8.15}$$

Χρησιμοποιώντας το γεγονός ότι

$$\begin{aligned}\int_a^b r(x,t)u_x(x,t)u(x,t)dx &= \frac{1}{2}\int_a^b r(x,t)\{[u(x,t)]^2\}_x dx\\ &= -\frac{1}{2}\int_a^b r_x(x,t)[u(x,t)]^2 dx\end{aligned}$$

μπορούμε εύκολα να αποδείξουμε και γι' αυτό το πρόβλημα εκτιμήσεις της μορφής (8.6), (8.7) και (8.10).

Ασκήσεις

8.1. (*Το Λήμμα του Gronwall*) Έστω $\alpha, \beta \in \mathbb{R}, \quad \beta > 0$, και $\varphi \in C[0,T]$ τέτοια ώστε

$$\varphi(t) \leq \alpha + \beta \int_0^t \varphi(s)ds \qquad \forall t \in [0,T].$$

Αποδείξτε ότι

$$\varphi(t) \leq \alpha \mathrm{e}^{\beta t} \qquad \forall t \in [0,T].$$

[*Υπόδειξη*: Έστω $\varepsilon > 0$ και $\psi(t) := (\alpha+\varepsilon)\mathrm{e}^{\beta t}, t \in [0,T]$. Τότε $\psi' = \beta\psi$ και συνεπώς

$$\psi(t) = \alpha + \varepsilon + \beta \int_0^t \psi(s)ds \qquad \forall t \in [0,T].$$

Προφανώς $\varphi(0) < \psi(0)$. Δεχθείτε ότι υπάρχει $t_0 \in (0,T]$ τέτοιο ώστε $\varphi(t) < \psi(t) \quad \forall t \in [0,t_0)$ και $\varphi(t_0) = \psi(t_0)$ και οδηγηθείτε σε άτοπο.]

8.2. Έστω $\varphi \in C^1[0,T], \quad \beta > 0$ και $\varphi'(t) \leq \beta\varphi(t), t \in [0,T]$. Αποδείξτε ότι

$$\varphi(t) \leq \varphi(0)\mathrm{e}^{\beta t} \qquad \forall t \in [0,T].$$

8.3. (*Το γενικευμένο Λήμμα του Gronwall*) Έστω $\varphi, f \in C[0,T], \quad f(t) \geq 0, t \in [0,T]$, και

$$\varphi(t) \leq \alpha + \int_0^t f(s)\varphi(s)ds \qquad \forall t \in [0,T],$$

όπου $\alpha \in \mathbb{R}$. Αποδείξτε ότι

$$\varphi(t) \leq \alpha \mathrm{e}^{\int_0^t f(s)ds} \qquad \forall t \in [0,T].$$

[*Υπόδειξη*: Για $\varepsilon > 0$ έστω $\psi(t) := (\alpha+\varepsilon)\mathrm{e}^{\int_0^t f(s)ds}$. Τότε $\psi(t) = \alpha + \varepsilon + \int\limits_0^t f(s)\psi(s)ds$. Εργασθείτε όπως στην Άσκηση 8.1.]

8.4. Έστω $\varphi \in C^1[0,T]$ και $f \in C[0,T], \quad f(t) \geq 0 \quad \forall t \in [0,T]$. Αν

$$\varphi'(t) \leq f(t)\varphi(t) \quad \forall t \in [0,T],$$

αποδείξτε ότι

$$\varphi(t) \leq \varphi(0)\mathrm{e}^{\int_0^t f(s)ds} \qquad \forall t \in [0,T].$$

8.5. Θεωρούμε το πρόβλημα

$$(8.16)\qquad \begin{cases} u_t = (pu_x)_x - ru_x - qu & \text{στο } [a,b]\times[0,T] \\ u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T] \\ u(\cdot,0) = u_0 & \text{στο } [a,b], \end{cases}$$

όπου $p \in C^1([a,b]\times[0,T]),\ \ p(x,t) > 0,\quad q, r \in C([a,b]\times[0,T]), q(x,t) > 0$. Αν

$$(8.17)\qquad \max_{x,t} |r(x,t)|^2 \le 4 \min_{x,t} p(x,t)\ \ \min_{x,t} q(x,t),$$

αποδείξτε ότι

$$\|u(\cdot,t)\| \le \|u_0\| \qquad \forall t \in [0,T].$$

8.6. Θεωρούμε το πρόβλημα (8.16) χωρίς όμως τώρα τη συνθήκη (8.17). Αποδείξτε ότι

$$\frac{1}{2}\frac{d}{dt}\|u(\cdot,t)\|^2 \le c\|u(\cdot,t)\|^2 \qquad \forall t \in [0,T]$$

και οδηγηθείτε στο συμπέρασμα ότι

$$\|u(\cdot,t)\| \le \mathrm{e}^{ct}\,\|u_0\| \qquad \forall t \in [0,T]$$

με μια θετική σταθερά c.

8.7. Θεωρούμε το πρόβλημα (8.16) χωρίς τη συνθήκη (8.17) και πάλι. Δείξτε ότι

$$\frac{1}{2}\frac{d}{dt}\int_a^b p(u_x)^2 dx \le c\|u_x(\cdot,t)\|^2 \qquad \forall t \in [0,T]$$

και

$$\|u_x(\cdot,t)\| \le C\mathrm{e}^{ct}\|u_0'\| \qquad \forall t \in [0,T]$$

με σταθερές C και c. [*Υπόδειξη*: Πολλαπλασιάστε τη Δ.Ε. στο (8.16) επί u_t, ολοκληρώστε στο $[a,b]$, ολοκληρώστε κατά μέρη, χρησιμοποιήστε την αριθμητική–γεωμετρική ανισότητα και την (5.5).]

8.8. Θεωρούμε το ακόλουθο πρόβλημα αρχικών και συνοριακών συνθηκών για την εξίσωση του Schrödinger

$$\begin{cases} u_t = \mathrm{i}\,\alpha u_{xx} + \mathrm{i}\,\beta(x,t)u & \text{στο } [a,b]\times[0,T] \\ u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T] \\ u(\cdot,0) = u_0 & \text{στο } [a,b], \end{cases}$$

όπου i η φανταστική μονάδα, α ένας μη μηδενικός πραγματικός αριθμός και β μια ομαλή συνάρτηση η οποία παίρνει πραγματικές τιμές. Αποδείξτε ότι η L^2 νόρμα της $u(\cdot, t)$ είναι ανεξάρτητη του t,

$$\|u(\cdot, t)\| = \|u_0\|.$$

[*Υπόδειξη*: Πολλαπλασιάστε τη Δ.Ε. επί $\bar{u}$ (τη συζυγή της u), ολοκληρώστε στο διάστημα $[a, b]$ και πάρτε πραγματικά μέρη.]

5. Μέθοδοι πεπερασμένων διαφορών για παραβολικές εξισώσεις

Σ' αυτό το κεφάλαιο θα μελετήσουμε τρείς μεθόδους πεπερασμένων διαφορών για το πρόβλημα (8.1), την πεπλεγμένη μέθοδο του Euler, την άμεση μέθοδο του Euler, και τη μέθοδο των Crank–Nicolson. Η άμεση μέθοδος του Euler είναι ευσταθής υπό περιοριστικές συνθήκες μεταξύ του χρονικού και του χωρικού βήματος, οι άλλες δύο είναι απεριόριστα ευσταθείς. Η άμεση μέθοδος του Euler μελετάται εδώ για παιδαγωγικούς λόγους, δεν είναι χρήσιμη μέθοδος για το εν λόγω πρόβλημα. Η πεπλεγμένη μέθοδος του Euler είναι πρώτης τάξεως ως προς το χρονικό βήμα, η μέθοδος των Crank–Nicolson δεύτερης τάξης.

5.1 Η πεπλεγμένη μέθοδος του Euler

Για να απλοποιήσουμε κάπως τους συμβολισμούς θεωρούμε το πρόβλημα (8.1) στην ειδική περίπτωση όπου $p(x,t) = 1$, $(x,t) \in [a,b] \times [0,T]$, θεωρούμε δηλαδή το πρόβλημα

$$(9.1) \qquad \begin{cases} u_t = u_{xx} - qu + f & \text{στο } [a,b] \times [0,T] \\ u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T] \\ u(\cdot,0) = u_0 & \text{στο } [a,b], \end{cases}$$

όπου $q, f \in C([a,b] \times [0,T])$ και $q(x,t) \geq 0$.
Έστω $J, N \in \mathbb{N}$, $h := \frac{b-a}{J+1}$, $k := \frac{T}{N}$, $x_i := a + ih$, $i = 0, \dots, J+1$, και $t^n := nk$, $n = 0, \dots, N$. Θα προσεγγίσουμε τις τιμές της λύσεως στα σημεία (x_i, t^n), $i = 0, \dots, J+1$, $n = 0, \dots, N$. Έστω $u^n \in \mathbb{R}_0^{J+2}$, $u_i^n := u(x_i, t^n)$. Προσεγγίζουμε τα u^n, $n = 1, \dots, N$,

με διανύσματα $U^1, \dots, U^N \in \mathbb{R}_0^{J+2}$ τέτοια ώστε με $U^0 := u^0$ να ισχύει, για $n = 0, \dots, N-1$,

$$\partial U_i^n = \Delta_h U_i^{n+1} - q(x_i, t^{n+1}) U_i^{n+1} + f(x_i, t^{n+1}), \quad i = 1, \dots, J. \tag{9.2}$$

Στο σχήμα αυτό χρησιμοποιήσαμε τους εξής συμβολισμούς: Για $v^0, \dots, v^N \in \mathbb{R}_0^{J+2}$ θέτουμε $\partial v^n := \frac{1}{k}(v^{n+1} - v^n)$, $\partial v_i^n := (\partial v^n)_i$, και ορίζουμε τον τελεστή $\Delta_h : \mathbb{R}_0^{J+2} \to \mathbb{R}_0^{J+2}$ κατά τα γνωστά δια

$$\Delta_h v_i := (\Delta_h v)_i := \frac{1}{h^2}(v_{i-1} - 2v_i + v_{i+1}), \quad i = 1, \dots, J.$$

Στο σχήμα (9.2) οδηγούμαστε προσεγγίζοντας τις παραγώγους στη Δ.Ε. στο σημείο (x_i, t^{n+1}) με πεπερασμένες διαφορές: Η $u_{xx}(x_i, t^{n+1})$ προσεγγίζεται κατά τα γνωστά δια $\Delta_h u_i^{n+1}$ και η $u_t(x_i, t^{n+1})$ δια ∂u_i^n. Η μέθοδος που περιγράφεται από το σχήμα (9.2) λέγεται *πεπλεγμένη μέθοδος του Euler*. Οι υπολογισμοί των προσεγγίσεων $U^n, n = 1, \dots, N$, γίνονται βήμα προς βήμα. Γνωρίζοντας ήδη το U^n μπορούμε από την (9.2) να υπολογίσουμε το U^{n+1}. Ο υπολογισμός αυτός απαιτεί την επίλυση ενός τριδιαγώνιου $J \times J$ γραμμικού συστήματος και σ' αυτό ακριβώς το γεγονός οφείλεται ο χαρακτηρισμός της μεθόδου ως πεπλεγμένης. Με την άμεση μέθοδο του Euler, η οποία συχνά αναφέρεται απλώς ως μέθοδος του Euler, θα ασχοληθούμε στην επόμενη παράγραφο. Σημειώνουμε ακόμη ότι η μέθοδος (9.2) ανήκει στην κατηγορία των λεγόμενων μονοβηματικών μεθόδων, για τον υπολογισμό του U^{n+1} απαιτείται μόνο γνώση του U^n αλλά όχι προσεγγίσεων σε προηγούμενες χρονικές στιγμές. Συνήθως δεν μας ενδιαφέρει η προσέγγιση της λύσης του προβλήματος (9.1) για όλα τα t παρά μόνο για $t = T$, δηλαδή μας ενδιαφέρει βασικά μόνο ο υπολογισμός του U^N. Έτσι στα μονοβηματικά σχήματα όταν υπολογίσουμε το U^n δεν χρειάζεται πια να κρατάμε αποθηκευμένο στον υπολογιστή το U^{n-1}, γεγονός που συνεπάγεται εξοικονόμηση χώρου μνήμης.

Συνέπεια. Έστω ότι η λύση u του προβλήματος (9.1) είναι αρκετά ομαλή, ας πούμε τέσσερεις φορές συνεχώς παραγωγίσιμη ως προς x και δύο ως προς t. Αναπτύσσοντας τότε κατά Taylor ως προς το σημείο (x_i, t^{n+1}) διαπιστώνουμε εύκολα ότι, για $n = 0, \dots, N-1$,

$$\partial u_i^n = \Delta_h u_i^{n+1} - q(x_i, t^{n+1}) u_i^{n+1} + f(x_i, t^{n+1}) + r_i^n, \quad i = 1, \dots, J, \tag{9.3}$$

όπου

$$\max_{i,n} |r_i^n| \leq C(k+h^2) \tag{9.4}$$

με μια σταθερά C ανεξάρτητη των k και h. Η εκτίμηση είναι πρώτης τάξης ως προς k γιατί $\partial u_i^n = u_t(x_i, t^{n+1}) - \frac{k}{2} u_{tt}(x_i, \vartheta)$ με $t^n < \vartheta < t^{n+1}$.

Στη συνέχεια θα δώσουμε εκτιμήσεις του σφάλματος στη διακριτή L^2 νόρμα.

Σύγκλιση στη διακριτή L^2 νόρμα. Πριν προχωρήσουμε στην εκτίμηση του σφάλματος θα αποδείξουμε ευστάθεια του σχήματος (9.2) στη διακριτή L^2 νόρμα. Πολλαπλασιάζοντας την (9.2) επί hU_i^{n+1}, αθροίζοντας από $i=1$ έως $i=J$ και χρησιμοποιώντας τις (6.8′) και (8.2) λαμβάνουμε

$$\|U^{n+1}\|_h^2 - (U^{n+1}, U^n)_h \leq kh \sum_{i=1}^J f(x_i, t^{n+1}) U_i^{n+1}$$

από την οποία έπεται αμέσως ότι

$$\|U^{n+1}\|_h \leq \|U^n\|_h + k\big\{h \sum_{i=1}^J [f(x_i, t^{n+1})]^2\big\}^{\frac{1}{2}}, \quad n = 0, \dots, N-1. \tag{9.5}$$

Άμεση συνέπεια της (9.5) είναι το γεγονός ότι τα γραμμικά συστήματα που παριστούν οι (9.2) έχουν ακριβώς μια λύση, οι προσεγγιστικές λύσεις ορίζονται συνεπώς καλώς. Επίσης από την (9.5) συμπεραίνουμε εύκολα ότι

$$\|U^n\|_h \leq \|U^0\|_h + t^n \max_{1\leq m\leq n} \big\{h \sum_{i=1}^J [f(x_i, t^m)]^2\big\}^{\frac{1}{2}}, \qquad n = 1, \dots, N, \tag{9.6}$$

και ιδιαίτερα ότι

$$\max_{1\leq n\leq N} \|U^n\|_h \leq \|U^0\|_h + T \max_{1\leq n\leq N} \big\{h \sum_{i=1}^J [f(x_i, t^n)]^2\big\}^{\frac{1}{2}}. \tag{9.7}$$

Θεώρημα 9.1 *Έστω ότι η λύση u του προβλήματος* (9.1) *είναι αρκετά ομαλή έτσι ώστε να ισχύει η* (9.4). *Τότε, για αρκετά μικρό k,*

$$\max_{1\leq n\leq N} \|u^n - U^n\|_h \leq C(k+h^2), \tag{9.8}$$

με μια σταθερά C ανεξάρτητη των k και h.

Απόδειξη. Έστω $e^n \in \mathbb{R}_0^{J+2}$, $e^n := u^n - U^n$, $n = 0, \dots, N$. Αφαιρώντας κατά μέλη τις (9.3) και (9.2) λαμβάνουμε

$$\partial e_i^n = \Delta_h e_i^{n+1} - q(x_i, t^{n+1}) e_i^{n+1} + r_i^n, \quad i = 1, \dots, J.$$

Παίρνοντας το εσωτερικό γινόμενο με e^{n+1} λαμβάνουμε

$$\|e^{n+1}\|_h^2 - (e^{n+1}, e^n)_h \le k(r^n, e^{n+1})_h,$$

οπότε έχουμε

$$\|e^{n+1}\|_h \le \|e^n\|_h + k\|r^n\|_h, \;\; n = 0, \dots, N-1.$$

Επομένως, επειδή $e^0 = 0$,

$$\max_{1 \le n \le N} \|e^n\|_h \le T \max_{0 \le n \le N-1} \|r^n\|_h$$

και με την (9.4) παίρνουμε αμέσως την (9.8). □

5.2 Η άμεση μέθοδος του Euler

Στην προηγούμενη παράγραφο μελετήσαμε την πεπλεγμένη μέθοδο του Euler για το πρόβλημα (9.1). Στην παρούσα παράγραφο θα προσεγγίσουμε τη λύση του (9.1) με την άμεση μέθοδο του Euler.

Έστω $J, N \in \mathbb{N}$, $h := \frac{b-a}{J+1}$, $k := \frac{T}{N}$, $x_i := a + ih$, $i = 0, \dots, J+1$, και $t^n := nk$, $n = 0, \dots, N$. Θα προσεγγίσουμε πάλι τις τιμές της λύσεως στα σημεία (x_i, t^n), $i = 0, \dots, J+1$, $n = 0, \dots, N$. Έστω $u^n \in \mathbb{R}_0^{J+2}$, $u_i^n := u(x_i, t^n)$. Προσεγγίζουμε τα u^n, $n = 1, \dots, N$, με διανύσματα $U^1, \dots, U^N \in \mathbb{R}_0^{J+2}$ που δίνονται δια $U^0 := u^0$ και, για $n = 0, \dots, N-1$,

$$\partial U_i^n = \Delta_h U_i^n - q(x_i, t^n) U_i^n + f(x_i, t^n), \;\; i = 1, \dots, J, \tag{9.9}$$

συγκρίνετε με το σχήμα (9.2). Τα U^n υπολογίζονται και εδώ βήμα προς βήμα, γνωρίζοντας το U^n υπολογίζουμε το U^{n+1}. Το σχήμα (9.9) είναι άμεσο, το U_i^{n+1} υπολογίζεται αμέσως από το (9.9) χρησιμοποιώντας τα U_{i-1}^n, U_i^n και U_{i+1}^n, δεν απαιτείται επίλυση κάποιου γραμμικού συστήματος.

Στο σχήμα (9.9) οδηγούμαστε προσεγγίζοντας τις παραγώγους στη Δ.Ε. στο σημείο (x_i, t^n) με πεπερασμένες διαφορές όπως και στην πεπλεγμένη μέθοδο του Euler.

Συνέπεια. Έστω ότι η λύση u του προβλήματος (9.1) είναι αρκετά ομαλή. Αναπτύσσοντας τότε κατά Taylor ως προς το σημείο (x_i, t^n) διαπιστώνουμε εύκολα ότι, για $n = 0, \dots, N-1$,

$$\partial u_i^n = \Delta_h u_i^n - q(x_i, t^n) u_i^n + f(x_i, t^n) + r_i^n, \quad i = 1, \dots, J, \tag{9.10}$$

όπου

$$\max_{i,n} |r_i^n| \le C(k + h^2) \tag{9.11}$$

με μια σταθερά C ανεξάρτητη των k και h.

Ευστάθεια στη διακριτή L^2 νόρμα. Πολλαπλασιάζουμε την (9.9) επί hU_i^n και αθροίζουμε από $i = 1$ έως $i = J$, οπότε παίρνουμε

$$(U^{n+1}, U^n)_h - \|U^n\|_h^2 + k|U^n|_{1,h}^2 \le kh \sum_{i=1}^{J} f(x_i, t^n) U_i^n$$

ή

$$\frac{1}{2}\big\{\|U^{n+1}\|_h^2 - \|U^n\|_h^2 - \|U^{n+1} - U^n\|_h^2\big\} + k|U^n|_{1,h}^2 \le kh \sum_{i=1}^{J} f(x_i, t^n) U_i^n. \tag{9.12}$$

Πολλαπλασιάζοντας εξ άλλου την (9.9) επί $h(U^{n+1} - U^n)_i$ και αθροίζοντας από $i = 1$ έως $i = J$ παίρνουμε

$$\begin{aligned}
\|U^{n+1} - U^n\|_h^2 &= k(\Delta_h U^n, U^{n+1} - U^n)_h - kh \sum_{i=1}^{J} q(x_i, t^n) U_i^n (U_i^{n+1} - U_i^n) \\
&\quad + kh \sum_{i=1}^{J} f(x_i, t^n)(U_i^{n+1} - U_i^n) \\
&\le k|U^n|_{1,h}\,|U^{n+1} - U^n|_{1,h} + Ck\|U^n\|_h\,\|U^{n+1} - U^n\|_h \\
&\quad + kh \sum_{i=1}^{J} f(x_i, t^n)(U_i^{n+1} - U_i^n),
\end{aligned}$$

όπου χρησιμοποιήσαμε μια προφανή συνέπεια της (6.8), βλ. την Άσκηση 9.1. Τώρα για $v \in \mathbb{R}_0^{J+2}$ έχουμε

$$|v|_{1,h}^2 = h\sum_{i=0}^{J} |\frac{v_{i+1}-v_i}{h}|^2 \le \frac{2}{h}\sum_{i=0}^{J}(v_{i+1}^2+v_i^2) \le \frac{4}{h^2}\|v\|_h^2$$

και η ανωτέρω σχέση δίνει

$$\begin{aligned}\|U^{n+1}-U^n\|_h^2 \le \frac{2k}{h}|U^n|_{1,h}\,\|U^{n+1}-U^n\|_h + Ck\|U^n\|_h\,\|U^{n+1}-U^n\|_h\\ + k\big\{h\sum_{i=1}^{J}[f(x_i,t^n)]^2\big\}^{\frac12}\|U^{n+1}-U^n\|_h,\end{aligned}$$

συνεπώς

$$\|U^{n+1}-U^n\|_h \le 2\frac{k}{h}|U^n|_{1,h} + Ck\|U^n\|_h + k\big\{h\sum_{i=1}^{J}[f(x_i,t^n)]^2\big\}^{\frac12}\ .$$

ή

$$\begin{aligned}\|U^{n+1}-U^n\|_h^2 \le 4\frac{k^2}{h^2}|U^n|_{1,h}^2 + k^2\big[C\|U^n\|_h + \big\{h\sum_{i=1}^{J}[f(x_i,t^n)]^2\big\}^{\frac12}\big]^2\\ + 4\frac{k^2}{h}|U^n|_{1,h}\big[C\|U^n\|_h + \big\{h\sum_{i=1}^{J}[f(x_i,t^n)]^2\big\}^{\frac12}\big]\ .\end{aligned} \tag{9.13}$$

Από τις (9.12) και (9.13) έπεται τώρα, χρησιμοποιώντας πάλι την εκτίμηση $|v|_{1,h} \le 2\|v\|_h/h$,

$$\begin{aligned}\frac12\big\{\|U^{n+1}\|_h^2 - \|U^n\|_h^2\big\} + k|U^n|_{1,h}^2 \le 2\frac{k^2}{h^2}|U^n|_{1,h}^2\\ + \frac{k^2}{2}\big[C\|U^n\|_h + \big\{h\sum_{i=1}^{J}[f(x_i,t^n)]^2\big\}^{\frac12}\big]^2\\ + 4\frac{k^2}{h^2}\|U^n\|_h\big[C\|U^n\|_h + \big\{h\sum_{i=1}^{J}[f(x_i,t^n)]^2\big\}^{\frac12}\big] + k\big\{h\sum_{i=1}^{J} f(x_i,t^n)U_i^n\big\}\ .\end{aligned}$$

Υποθέτοντας τώρα ότι

$$2\frac{k}{h^2} \le 1 \tag{9.14}$$

λαμβάνουμε από την προηγούμενη ανισότητα

$$\|U^{n+1}\|_h^2 \le (1+Ck)\|U^n\|_h^2 + Ck\big\{h\sum_{i=1}^{J}[f(x_i,t^n)]^2\big\} \tag{9.15}$$

και συμπεραίνουμε ότι

$$\max_{0\le n\le N}\|U^n\|_h \le C\|U^0\| + C\max_{0\le n\le N}\big\{h\sum_{i=1}^{J}[f(x_i,t^n)]^2\big\}^{\frac{1}{2}}, \tag{9.16}$$

βλ. την Άσκηση 9.2.

Συνδυάζοντας την ευστάθεια και τη συνέπεια κατά τα γνωστά μπορούμε να αποδείξουμε το ακόλουθο αποτέλεσμα:

Θεώρημα 9.2 *Έστω ότι η λύση u του προβλήματος* (9.1) *είναι αρκετά ομαλή και ότι ικανοποιείται η* (9.14). *Τότε ισχύει*

$$\max_{0\le n\le N}\|u^n - U^n\|_h \le C(k+h^2) \tag{9.17}$$

με μια σταθερά C ανεξάρτητη των k και h. □

Παρατήρηση 9.1. Η συνθήκη (9.14) είναι πολύ περιοριστική για το χρονικό βήμα k στην άμεση μέθοδο του Euler και στην ουσία καθιστά αυτή τη μέθοδο μη ενδιαφέρουσα για τις εφαρμογές. Η άμεση μέθοδος του Euler είναι όπως λέμε *ευσταθής υπό συνθήκες*, σε αντίθεση με την πεπλεγμένη μέθοδο του Euler, η οποία είναι ευσταθής χωρίς κανέναν περιορισμό στη σχέση μεταξύ χρονικού και χωρικού βήματος. Μια συνθήκη της μορφής (9.14) είναι αναμενόμενη για τον εξής λόγο: Είναι γνωστό ότι η τιμή της λύσης u του προβλήματος (9.1) στο σημείο (x_i,t^{n+1}) εξαρτάται από τις τιμές της λύσης στα σημεία $(x,t^n),\ \ a\le x\le b$, και όχι μόνο από τις τιμές της λύσης σε σημεία (x,t^n) με x “κοντά” στο x_i, όσο μικρή και αν είναι η διαφορά $t^{n+1}-t^n$. Όπως λέμε, η θερμότητα διαδίδεται με “άπειρη” ταχύτητα. Στο διακριτό ανάλογο για να πάρει κανείς λογική προσέγγιση U_i^{n+1} της $u(x_i,t^{n+1})$ από τις U_{i-1}^n, U_i^n και U_{i+1}^n και μόνο, πρέπει η διαφορά $t^{n+1}-t^n=k$ να είναι μικρή συγκρινόμενη με το h. □

5.3 Η μέθοδος των Crank–Nicolson

Η μέθοδος των Crank–Nicolson χρησιμοποιείται ευρέως για την προσεγγιστική επίλυση του προβλήματος (8.1). Η μέθοδος των Crank–Nicolson είναι πεπλεγμένη και όπως και η πεπλεγμένη μέθοδος του Euler είναι ευσταθής χωρίς περιορισμούς στη σχέση μεταξύ χρονικού και χωρικού βήματος. Το βασικό πλεονέκτημα της μεθόδου των Crank–Nicolson έναντι της πεπλεγμένης μεθόδου του Euler είναι ότι είναι δεύτερης τάξης και ως προς το χρονικό και ως προς το χωρικό βήμα.

Έστω $J, N \in \mathbb{N}, h := \frac{b-a}{J+1}$, $k := \frac{T}{N}$, $x_i := a + ih$, $i = 0, \dots, J+1$, και $t^n := nk$, $n = 0, \dots, N$. Θα προσεγγίσουμε τις τιμές της λύσης u του προβλήματος (9.1) στα σημεία (x_i, t^n), $i = 0, \dots, J+1$, $n = 0, \dots, N$. Έστω $u^n \in \mathbb{R}_0^{J+2}$, $u_i^n := u(x_i, t^n)$. Προσεγγίζουμε τα $u^n,\ n = 1, \dots, N$, με διανύσματα $U^1, \dots, U^N \in \mathbb{R}_0^{J+2}$, τέτοια ώστε με $U^0 := u^0$ να ισχύει, για $n = 0, \dots, N$,

$$\partial U_i^n = \Delta_h U_i^{n+\frac{1}{2}} - q(x_i, t^{n+\frac{1}{2}}) U_i^{n+\frac{1}{2}} + f(x_i, t^{n+\frac{1}{2}}), \quad i = 1, \dots, J. \tag{9.18}$$

Επιπλέον των συμβολισμών που χρησιμοποιήσαμε μέχρι τώρα, στο σχήμα (9.18) χρησιμοποιήσαμε και τους συμβολισμούς $t^{n+\frac{1}{2}} := t^n + \frac{k}{2} = \frac{1}{2}(t^n + t^{n+1})$, και, για $v^0, \dots, v^N \in \mathbb{R}_0^{J+2}$, $v^{n+\frac{1}{2}} := \frac{1}{2}(v^n + v^{n+1})$. Στη μέθοδο των Crank–Nicolson οδηγούμαστε προσεγγίζοντας τη λύση και τις παραγώγους της στη Δ.Ε. του (9.1) στο σημείο $(x_i, t^{n+\frac{1}{2}})$ με τον προφανή από το σχήμα (9.18) τρόπο. Η ∂u_i^n προσεγγίζει την u_t στο σημείο $(x_i, t^{n+\frac{1}{2}})$ με δεύτερη τάξη, ενώ στα σημεία (x_i, t^{n+1}) και (x_i, t^n) την προσεγγίζει μόνο με πρώτη τάξη.

Υποθέτοντας ότι έχουμε ήδη προσδιορίσει το U^n, μπορούμε να υπολογίσουμε το U^{n+1} στο σχήμα (9.18) λύνοντας ένα τριδιαγώνιο $J \times J$ γραμμικό σύστημα.

Συνέπεια. Έστω ότι η λύση u του προβλήματος (9.1) είναι αρκετά ομαλή. Αναπτύσσοντας κατά Taylor ως προς τα σημεία (x_i, t^{n+1}) και (x_i, t^n) και στη συνέχεια ως προς το σημείο $(x_i, t^{n+\frac{1}{2}})$ διαπιστώνουμε εύκολα ότι

$$\partial u_i^n = \Delta_h u_i^{n+\frac{1}{2}} - q(x_i, t^{n+\frac{1}{2}}) u_i^{n+\frac{1}{2}} + f(x_i, t^{n+\frac{1}{2}}) + r_i^n, \quad i = 1, \dots, J, \tag{9.19}$$

όπου

$$\max_{i,n} |r_i^n| \le C(k^2 + h^2). \tag{9.20}$$

Σύγκλιση στη διακριτή L^2 νόρμα. Πιστεύοντας ότι ο αναγνώστης έχει ήδη εξοικειωθεί με τον τρόπο που γίνεται η απόδειξη σύγκλισης με συνδυασμό της ευστάθειας με τη συνέπεια των μεθόδων πεπερασμένων διαφορών, σε αυτή την παράγραφο θα προχωρήσουμε κατ' ευθείαν στην εκτίμηση του σφάλματος της μεθόδου των Crank–Nicolson χωρίς να μελετήσουμε ξεχωριστά την ευστάθεια της μεθόδου.

Θεώρημα 9.3 *Έστω ότι η λύση u του προβλήματος* (9.1) *είναι αρκετά ομαλή έτσι ώστε να ικανοποιείται η* (9.20). *Τότε ισχύει*

$$\max_{0\le n\le N}\|u^n-U^n\|_h\le C(k^2+h^2) \tag{9.21}$$

με μια σταθερά C ανεξάρτητη των k και h.

Απόδειξη. Έστω $e^n\in\mathbb{R}_0^{J+2},\ e^n:=u^n-U^n,\ \ n=0,\dots,N$. Αφαιρώντας κατά μέλη τις (9.19) και (9.18) λαμβάνουμε, για $n=0,\dots,N-1$,

$$\partial e_i^n=\Delta_h e_i^{n+\frac12}-q(x_i,t^{n+\frac12})e_i^{n+\frac12}+r_i^n,\ \ i=1,\dots,J. \tag{9.22}$$

Παίρνοντας στην (9.22) το εσωτερικό γινόμενο με $e^{n+\frac12}$ και λαμβάνοντας υπ' όψιν την (6.8′) και την (8.2) διαπιστώνουμε εύκολα ότι

$$\frac{1}{2k}(\|e^{n+1}\|_h^2-\|e^n\|_h^2)\le(r^n,e^{n+\frac12})_h$$

ή

$$\|e^{n+1}\|_h\le\|e^n\|_h+k\|r^n\|_h,\quad n=0,\dots,N-1. \tag{9.23}$$

Από την (9.23) έπεται εύκολα κατά τα γνωστά, επειδή $e^0=0$, ότι

$$\max_{0\le n\le N}\|e^n\|_h\le C\max_{0\le n\le N-1}\|r^n\|_h$$

και με την (9.20) λαμβάνουμε την (9.21). □

Σύγκλιση στη διακριτή νόρμα μεγίστου.

Θεώρημα 9.4 *Έστω ότι η λύση u του προβλήματος* (9.1) *είναι αρκετά ομαλή έτσι ώστε να ισχύει η* (9.20). *Τότε, για αρκετά μικρό k,*

$$\max_{0\le n\le N}|u^n-U^n|_{1,h}\le C(k^2+h^2) \tag{9.24}$$

και

$$\max_{0\le n\le N}\max_{0\le i\le J+1}|u_i^n-U_i^n|\le C(k^2+h^2) \tag{9.25}$$

με μια σταθερά C ανεξάρτητη των k και h.

Απόδειξη. Χρησιμοποιούμε τους ίδιους συμβολισμούς όπως στην προηγούμενη απόδειξη. Κατ' αρχάς από την (6.8) συμπεραίνουμε αμέσως ότι για $v, w \in \mathbb{R}_0^{J+2}$ ισχύει

$$(\Delta_h v, w)_h = (v, \Delta_h w)_h. \tag{9.26}$$

Παίρνοντας στην (9.22) το εσωτερικό γινόμενο με ∂e^n και χρησιμοποιώντας τις (6.8′) και (9.26) λαμβάνουμε

$$\|\partial e^n\|_h^2 + \frac{1}{2k}|e^{n+1}|_{1,h}^2 = \frac{1}{2k}|e^n|_{1,h}^2 - h\sum_{i=1}^{J} q(x_i, t^{n+\frac{1}{2}})e_i^{n+\frac{1}{2}}\partial e_i^n + (r^n, \partial e^n)_h$$

ή

$$\|\partial e^n\|_h^2 + \frac{1}{2k}|e^{n+1}|_{1,h}^2 \le \frac{1}{2k}|e^n|_{1,h}^2 + C\|e^{n+\frac{1}{2}}\|^2 + \frac{1}{2}\|r^n\|_h^2 + \|\partial e^n\|_h^2$$

ή

$$(1 - 2Ck)|e^{n+1}|_{1,h}^2 \le (1 + 2Ck)|e^n|_{1,h}^2 + k\|r^n\|_h^2,$$

οπότε, για αρκετά μικρό k,

$$|e^{n+1}|_{1,h}^2 \le (1 + ck)|e^n|_{1,h}^2 + ck\|r^n\|_h^2, \quad n = 0, \dots, N-1. \tag{9.27}$$

Επειδή $e^0 = 0$, από την (9.27) έπεται ότι

$$\max_{0\le n\le N} |e^n|_{1,h} \le C \max_{0\le n\le N-1} \|r^n\|_h,$$

βλ. την Άσκηση 9.2, και σε συνδυασμό με την (9.20) λαμβάνουμε την (9.24). Η (9.25) είναι άμεση συνέπεια των (9.24) και (6.5). □

Ασκήσεις

9.1. Χρησιμοποιήστε την (6.8) για να αποδείξετε ότι

$$|(\Delta_h v, w)_h| \le |v|_{1,h}\,|w|_{1,h} \qquad \forall v, w \in \mathbb{R}_0^{J+2}.$$

9.2. Έστω $T > 0,\ N \in \mathbb{N}$ και $k := \frac{T}{N}$. Αν $\alpha_0, \dots, \alpha_N$ και $\varepsilon_0, \dots, \varepsilon_{N-1}$ μη αρνητικοί αριθμοί τέτοιοι ώστε με $\gamma > 0$

$$\alpha_{n+1} \le (1 + \gamma k)\alpha_n + k\varepsilon_n, \quad n = 0, \dots, N-1,$$

αποδείξτε ότι

$$\max_{1\le n\le N}\alpha_n \le \mathrm{e}^{\gamma T}\alpha_0 + \frac{1}{\gamma}\mathrm{e}^{\gamma T}\max_{0\le n\le N-1}\varepsilon_n. \tag{9.28}$$

9.3. Έστω $T>0$, $0=t^0<t^1<\cdots<t^N=T$ ένας διαμερισμός του $[0,T]$, $k_n:=t^{n+1}-t^n$, $n=0,\dots,N-1$. Αν $\alpha_0,\dots,\alpha_N$ και $\varepsilon_0,\dots,\varepsilon_{N-1}$ μη αρνητικοί αριθμοί τέτοιοι ώστε με $\gamma>0$

$$\alpha_{n+1}\le(1+\gamma k_n)\alpha_n+k_n\varepsilon_n,\quad n=0,\dots,N-1,$$

αποδείξτε ότι

$$\max_{1\le n\le N}\alpha_n\le \mathrm{e}^{\gamma T}\alpha_0+\int_0^T \mathrm{e}^{\gamma(T-s)}ds\quad \max_{0\le n\le N-1}\varepsilon_n.$$

[*Υπόδειξη*: Αποδείξτε ότι

$$\begin{aligned}\alpha_{n+1}&\le(1+\gamma k_n)(1+\gamma k_{n-1})\cdots(1+\gamma k_0)\alpha_0\\&\quad+\big\{k_n+(1+\gamma k_n)k_{n-1}+(1+\gamma k_n)(1+\gamma k_{n-1})k_{n-2}+\cdots\\&\quad+(1+\gamma k_n)(1+\gamma k_{n-1})\cdots(1+\gamma k_1)k_0\big\}\max_{0\le m\le n}\varepsilon_m\\&\le \mathrm{e}^{\gamma t^{n+1}}\alpha_0+\big\{k_n+\mathrm{e}^{\gamma(t^{n+1}-t^n)}(t^n-t^{n-1})+\mathrm{e}^{\gamma(t^{n+1}-t^{n-1})}(t^{n-1}-t^{n-2})\\&\quad+\cdots+\mathrm{e}^{\gamma(t^{n+1}-t^1)}(t^1-t^0)\big\}\max_{0\le m\le n}\varepsilon_m\end{aligned}$$

και ότι

$$\mathrm{e}^{\gamma(t^{n+1}-t^\kappa)}(t^\kappa-t^{\kappa-1})\le\int_{t^{\kappa-1}}^{t^\kappa}\mathrm{e}^{\gamma(t^{n+1}-s)}ds.]$$

9.4. Έστω $J\in\mathbb{N}$, $h:=\frac{b-a}{J+1}$, $x_i:=a+ih$, $i=0,\dots,J+1$, ένας ομοιόμορφος διαμερισμός του $[a,b]$, και $0=t^0<t^1<\cdots<t^N=T$ ένας διαμερισμός του $[0,T]$. Προσεγγίζουμε τα $u^n\in\mathbb{R}_0^{J+2}$, όπου $u_i^n:=u(x_i,t^n)$ και u η λύση του (8.1), με διανύσματα $U^n\in\mathbb{R}_0^{J+2}$ τέτοια ώστε με $U^0:=u^0$ να ισχύει, για $n=0,\dots,N-1$,

$$\frac{1}{k_n}(U_i^{n+1}-U_i^n)=\Delta_h U_i^{n+1}-q(x_i,t^{n+1})U_i^{n+1}+f(x_i,t^{n+1}),\ \ i=1,\dots,J, \tag{9.29}$$

όπου $k_n:=t^{n+1}-t^n$. Αν η λύση u του προβλήματος (8.1) είναι αρκετά ομαλή, αποδείξτε για τις προσεγγίσεις του σχήματος (9.29), δηλ. για την πεπλεγμένη μέθοδο του Euler με μη ομοιόμορφο διαμερισμό στο χρόνο, μια εκτίμηση της μορφής (9.8) με $k := \max_n k_n$ και σταθερά C ανεξάρτητη των διαμερισμών. [*Υπόδειξη*: Χρησιμοποιήστε την Άσκηση 9.3.]

9.5. Θεωρούμε την πεπλεγμένη μέθοδο του Euler (9.2). Αποδείξτε εκτιμήσεις του σφάλματος στη διακριτή νόρμα μεγίστου, καθώς και στη νόρμα $|\cdot|_{1,h}$, πρώτης τάξης ως προς k και δεύτερης τάξης ως προς h.

9.6. Θεωρούμε το πρόβλημα (8.1), και ομοιόμορφους διαμερισμούς των $[a,b]$ και $[0,T]$ με βήμα h και k αντίστοιχα. Προσεγγίζουμε τα $u^n \in \mathbb{R}_0^{J+2}$ με διανύσματα $U^n \in \mathbb{R}_0^{J+2}$ τέτοια ώστε $U^0 = u^0$ και, για $n = 0, \dots, N-1$,

$$
\begin{aligned}
\partial U_i^n = \frac{1}{h^2}\big[p(x_i + \frac{h}{2}, t^{n+1})(U_{i+1}^{n+1} - U_i^{n+1})& \\
- p(x_i - \frac{h}{2}, t^{n+1})(U_i^{n+1} - U_{i-1}^{n+1})\big]& \\
- q(x_i, t^{n+1})U_i^{n+1} + f(x_i, t^{n+1}), \quad i = 1, \dots, J.&
\end{aligned}
\tag{9.30}
$$

Αν η λύση u του προβλήματος (8.1) είναι αρκετά ομαλή, αποδείξτε για τις προσεγγίσεις του σχήματος (9.30) μία εκτίμηση της μορφής (9.8).

9.7. Θεωρούμε το εξής πρόβλημα αρχικών και συνοριακών συνθηκών: Ζητείται μια ομαλή συνάρτηση u τέτοια ώστε

$$
\begin{cases}
u_t = u_{xx} - r u_x - q u & \text{στο } [a,b] \times [0,T] \\
u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T] \\
u(\cdot,0) = u_0 & \text{στο } [a,b],
\end{cases}
\tag{9.31}
$$

όπου $q, r \in C([a,b] \times [0,T])$. Θεωρούμε ομοιόμορφους διαμερισμούς των $[a,b]$ και $[0,T]$ με βήμα h και k, αντίστοιχα. Προσεγγίζουμε τα $u^n \in \mathbb{R}_0^{J+2}$ με διανύσματα $U^n \in \mathbb{R}_0^{J+2}$ τέτοια ώστε $U^0 := u^0$ και, για $n = 0, \dots, N-1$,

$$
\begin{aligned}
\partial U_i^n = \Delta_h U_i^{n+1} - r(x_i, t^{n+1})\frac{1}{2h}(U_{i+1}^{n+1} - U_{i-1}^{n+1}) - q(x_i, t^{n+1})U_i^{n+1},& \\
i = 1, \dots, J.&
\end{aligned}
\tag{9.32}
$$

Αποδείξτε κατ' αρχάς ότι

$$\|U^n\|_h \le C\|U^0\|_h, \quad n = 1, \dots, N,$$

και στη συνέχεια, υποθέτοντας ότι η λύση u του προβλήματος (9.31) είναι αρκετά ομαλή, αποδείξτε για τις προσεγγίσεις του σχήματος (9.32) μια εκτίμηση της μορφής (9.8).

9.8. Θεωρούμε το πρόβλημα (8.11) με $p(x,t) = 1$ και ομοιόμορφους διαμερισμούς των $[a,b]$ και $[0,T]$ με βήμα h και k, αντίστοιχα. Προσεγγίζουμε τα $u^n \in \mathbb{R}^{J+2}$, $u_i^n = u(x_i, t^n)$, με διανύσματα $U^n \in \mathbb{R}^{J+2}$ τέτοια ώστε $U^0 := u^0$ και, για $n = 0, \dots, N-1$,

$$\partial U_i^n = \Delta_h U_i^{n+1} - q(x_i, t^{n+1})U_i^{n+1} + f(x_i, t^{n+1}), \quad i = 0, \dots, J+1, \tag{9.33}$$

όπου $\Delta_h : \mathbb{R}^{J+2} \to \mathbb{R}^{J+2}$ όπως ορίστηκε μεταξύ των σχέσεων (6.22) και (6.23). Έστω ότι η λύση u του προβλήματος (8.11) είναι αρκετά ομαλή. Αποδείξτε κατ' αρχάς ότι

$$\partial u_i^n = \Delta_h u_i^{n+1} - q(x_i, t^{n+1})u_i^{n+1} + f(x_i, t^{n+1}) + r_i^n, \quad i = 0, \dots, J+1$$

όπου

$$\begin{cases} \max\limits_{\substack{1 \le i \le J \\ n}} |r_i^n| \le C(k + h^2) \\ \max\limits_{n} |r_i^n| \le C(k + h) \quad i = 0, J+1. \end{cases}$$

Στη συνέχεια αποδείξτε για τις προσεγγίσεις του σχήματος (9.33) μία εκτίμηση της μορφής (9.8) με κατάλληλη νόρμα $\|\cdot\|_h$.

9.9. Θεωρούμε το πρόβλημα (8.1), και ομοιομόρφους διαμερισμούς των $[a,b]$ και $[0,T]$ με βήμα h και k, αντίστοιχα. Προσεγγίζουμε τα $u^n \in \mathbb{R}_0^{J+2}$ με διανύσματα $U^n \in \mathbb{R}_0^{J+2}$ τέτοια ώστε $U^0 := u^0$ και, για $n = 0, \dots, N-1$,

$$\begin{aligned} \partial U_i^n = \frac{1}{h^2}[p(x_i + \frac{h}{2}, t^{n+\frac12})(U_{i+1}^{n+\frac12} - U_i^{n+\frac12}) \\ - p(x_i - \frac{h}{2}, t^{n+\frac12})(U_i^{n+\frac12} - U_{i-1}^{n+\frac12})] \\ - q(x_i, t^{n+\frac12})U_i^{n+\frac12} + f(x_i, t^{n+\frac12}), \ \ i = 1, \dots, J. \end{aligned} \tag{9.34}$$

Αν η λύση u του προβλήματος (8.1) είναι αρκετά ομαλή, αποδείξτε για τις προσεγγίσεις του σχήματος (9.34) εκτιμήσεις της μορφής (9.21), (9.24) και (9.25).

9.10. Θεωρούμε το πρόβλημα (9.31), και ομοιόμορφους διαμερισμούς των $[a,b]$ και $[0,T]$ με βήμα h και k, αντίστοιχα. Προσεγγίζουμε τα $u^n \in \mathbb{R}_0^{J+2}$ με διανύσματα

$U^n \in \mathbb{R}_0^{J+2}$ τέτοια ώστε $U^0 := u^0$ και, για $n = 0, \dots, N-1$,

$$\partial U_i^n = \Delta_h U_i^{n+\frac{1}{2}} - r(x_i, t^{n+\frac{1}{2}}) \frac{1}{2h} (U_{i+1}^{n+\frac{1}{2}} - U_{i-1}^{n+\frac{1}{2}}) - q(x_i, t^{n+\frac{1}{2}}) U_i^{n+\frac{1}{2}}, \quad i = 1, \dots, J. \tag{9.35}$$

Αν η λύση του προβλήματος (9.31) είναι αρκετά ομαλή, αποδείξτε για τις προσεγγίσεις του σχήματος (9.35) εκτιμήσεις της μορφής (9.21), (9.24) και (9.25).

9.11. Θεωρούμε το πρόβλημα (8.11) με $p(x,t) = 1$ και ομοιόμορφους διαμερισμούς των $[a,b]$ και $[0,T]$ με βήμα h και k, αντίστοιχα. Προσεγγίζουμε τα $u^n \in \mathbb{R}^{J+2}$, $u_i^n = u(x_i, t^n)$, με διανύσματα $U^n \in \mathbb{R}^{J+2}$ τέτοια ώστε $U^0 := u^0$ και για $n = 0, \dots, N-1$

$$\partial U_i^n = \Delta_h U_i^{n+\frac{1}{2}} - q(x_i, t^{n+\frac{1}{2}}) U_i^{n+\frac{1}{2}} + f(x_i, t^{n+\frac{1}{2}}), \quad i = 0, \dots, J+1, \tag{9.36}$$

όπου $\Delta_h : \mathbb{R}^{J+2} \to \mathbb{R}^{J+2}$ όπως ορίστηκε μεταξύ των σχέσεων (6.22) και (6.23). Έστω ότι η λύση u του προβλήματος (8.11) είναι αρκετά ομαλή. Αποδείξτε κατ' αρχάς ότι

$$\partial u_i^n = \Delta_h u_i^{n+\frac{1}{2}} - q(x_i, t^{n+\frac{1}{2}}) u_i^{n+\frac{1}{2}} + f(x_i, t^{n+\frac{1}{2}}) + r_i^n, \quad i = 0, \dots, J+1,$$

όπου

$$\begin{cases} \max\limits_{\substack{1 \le i \le J \\ n}} |r_i^n| \le C(k^2 + h^2), \\ \max\limits_{n} |r_i^n| \le C(k^2 + h), \qquad i = 0, J+1. \end{cases}$$

Στη συνέχεια αποδείξτε για τις προσεγγίσεις του σχήματος (9.36) μία εκτίμηση της μορφής (9.21) με κατάλληλη νόρμα $\| \cdot \|_h$.

6. Μέθοδοι πεπερασμένων στοιχείων για παραβολικές εξισώσεις

Σ' αυτό το κεφάλαιο θα μελετήσουμε μεθόδους πεπερασμένων στοιχείων για την εξίσωση της θερμότητας σε μία διάσταση. Πρώτα θα ασχοληθούμε με το ημιδιακριτό πρόβλημα, το οποίο είναι διακριτό ως προς το χώρο και συνεχές ως προς το χρόνο. Για να οδηγηθούμε σε μεθόδους οι οποίες δίνουν προσεγγίσεις που μπορούν να υπολογισθούν, πρέπει να διακριτοποιήσουμε και ως προς το χρόνο, παίρνοντας έτσι πλήρως διακριτά σχήματα. Η διακριτοποίηση στο χρόνο θα γίνει με τρείς μεθόδους, τις μεθόδους του Euler, άμεση και πεπλεγμένη, και με τη μέθοδο των Crank–Nicolson. Όπως και στο προηγούμενο κεφάλαιο, η άμεση μέθοδος του Euler είναι ευσταθής υπό συνθήκες, ενώ οι άλλες δύο μέθοδοι είναι απεριόριστα ευσταθείς. Οι μέθοδοι του Euler είναι πρώτης τάξης ως προς το χρονικό βήμα, η μέθοδος των Crank–Nicolson είναι δεύτερης τάξης.

6.1 Εισαγωγή

Έστω $[a, b]$ ένα φραγμένο διάστημα και $T > 0$. Θεωρούμε το εξής πρόβλημα αρχικών και συνοριακών συνθηκών

$$\begin{cases} u_t = u_{xx} + f & \text{στο } [a,b] \times [0,T] \\ u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T] \\ u(\cdot,0) = u^0 & \text{στο } [a,b]. \end{cases} \tag{10.1}$$

Εδώ $f : [a, b] \times [0, T] \to \mathbb{R}$ είναι μια δεδομένη συνάρτηση και $u^0 : [a, b] \to \mathbb{R}$ είναι μια δεδομένη αρχική τιμή. Υποθέτουμε ότι τα δεδομένα είναι ομαλά και συμβατά, ώστε το πρόβλημα (10.1) να έχει μια λύση, η οποία μάλιστα είναι αρκετά ομαλή ώστε να έχουν έννοια οι νόρμες της u οι οποίες θα χρησιμοποιηθούν στην συνέχεια.

Έστω $v : [a, b] \to \mathbb{R}$ μια συνεχής και κατά τμήματα συνεχώς παραγωγίσιμη συνάρτηση η οποία μηδενίζεται στα σημεία a και b. Παίρνοντας στη διαφορική εξίσωση του (10.1) το L^2 εσωτερικό γινόμενο και των δύο μελών με την v, ολοκληρώνοντας στον πρώτο όρο στο δεξιό μέλος κατά μέρη και χρησιμοποιώντας τις συνοριακές συνθήκες, γράφουμε το πρόβλημα (10.1) σε μεταβολική μορφή ως εξής: Ζητείται $u(\cdot, t), t \in [0, T]$, τέτοια ώστε

$$\begin{cases} (u_t(\cdot, t), v) + (u_x(\cdot, t), v') = (f(\cdot, t), v) \quad \forall v \in V \quad \forall t \in [0, T] \\ u(\cdot, 0) = u^0 , \end{cases} \tag{10.2}$$

όπου V ο χώρος που ορίστηκε αμέσως μετά την (7.1). Η διακριτοποίηση του προβλήματος (10.1) με τη μέθοδο των πεπερασμένων στοιχείων βασίζεται στη *μεταβολική* διατύπωση (10.2).

Στην επόμενη παράγραφο θα μελετήσουμε τη διακριτοποίηση στο χώρο του προβλήματος (10.1) με τη μέθοδο των πεπερασμένων στοιχείων, τη λεγόμενη *ημιδιακριτοποίηση*. Η ημιδιακριτοποίηση έχει μόνο θεωρητική σημασία, αποτελεί ένα βήμα για τον υπολογισμό προσεγγίσεων της λύσης του (10.1). Η ημιδιακριτή λύση είναι λύση ενός προβλήματος αρχικών τιμών για ένα σύστημα συνήθων διαφορικών εξισώσεων. Θα μελετήσουμε προσεγγιστικές ιδιότητες της ημιδιακριτής λύσης.

Για να μπορέσουμε να υπολογίσουμε προσεγγίσεις της λύσης του προβλήματος (10.1), πρέπει να διακριτοποιήσουμε ως προς τον χρόνο το ημιδιακριτό πρόβλημα. Αυτό γίνεται με τα λεγόμενα *πλήρως διακριτά σχήματα*. Στην τρίτη παράγραφο θα μελετήσουμε τρία πλήρως διακριτά σχήματα για το πρόβλημα (10.1): την πεπλεγμένη μέθοδο του Euler, τη μέθοδο των Crank–Nicolson, και τη μέθοδο του Euler. Θα δούμε πλεονεκτήματα και μειονεκτήματα αυτών των μεθόδων και θα μελετήσουμε προσεγγιστικές ιδιότητες πλήρως διακριτών προσεγγίσεων.

6.2 Ημιδιακριτοποίηση

Έστω S_h^r μια οικογένεια υποχώρων του $V, N_h := \dim S_h^r < \infty$, με την προσεγγιστική ιδιότητα (7.2). Κατ' αναλογίαν προς το πρόβλημα (10.2) ορίζουμε την *ημιδιακριτή* *λύση* $u_h, u_h(\cdot,t) \in S_h^r, t \in [0,T]$, δια

$$(10.3) \qquad \begin{cases} (u_{ht}(\cdot,t),\chi) + (u_{hx}(\cdot,t),\chi') = (f(\cdot,t),\chi) \quad \forall \chi \in S_h^r \quad \forall t \in [0,T] \\ u_h(\cdot,0) = u_h^0, \end{cases}$$

όπου $u_h^0 \in S_h^r$ μια προσέγγιση της αρχικής τιμής u^0.

Θα αποδείξουμε κατ' αρχάς ύπαρξη και μοναδικότητα της ημιδιακριτής λύσης u_h. Έστω λοιπόν $\{\varphi_1,\dots,\varphi_{N_h}\}$ μια βάση του S_h^r. Επειδή $u_h(\cdot,t) \in S_h^r, t \in [0,T]$, η u_h γράφεται στη μορφή

$$u_h(x,t) = \sum_{j=1}^{N_h} \alpha_j(t)\varphi_j(x)$$

και το πρόβλημα (10.3) γράφεται ως

$$(10.4) \qquad \begin{cases} \displaystyle\sum_{j=1}^{N_h} (\varphi_j,\varphi_i)\alpha_j'(t) + \sum_{j=1}^{N_h} (\varphi_j',\varphi_i')\alpha_j(t) = (f(\cdot,t),\varphi_i), \quad i = 1,\dots,N_h, \\ \alpha_j(0) = \gamma_j, \quad j = 1,\dots,N_h, \end{cases}$$

όπου τα γ_j είναι τέτοια ώστε $\quad u_h^0 = \sum_{j=1}^{N_h} \gamma_j \varphi_j$.

Το (10.4) είναι ένα πρόβλημα αρχικών τιμών για ένα σύστημα συνήθων διαφορικών εξισώσεων. Επειδή ο πίνακας μάζης $(\varphi_j,\varphi_i)_{j,i=1,\dots,N_h}$ είναι αντιστρέψιμος, ως πίνακας του Gram, το (10.4) λύνεται μονοσήμαντα.

Πριν προχωρήσουμε στη μελέτη προσεγγιστικών ιδιοτήτων της ημιδιακριτής λύσης, δίνουμε ένα βοηθητικό αποτέλεσμα, το οποίο θα παίξει σημαντικό ρόλο στη συνέχεια. Το αποτέλεσμα αναφέρεται στη λεγόμενη *ελλειπτική προβολή*. Ο λόγος για τον οποίον εισάγουμε την ελλειπτική προβολή θα καταστεί σαφής στη συνέχεια. Προς το παρόν αναφέρουμε απλώς το κίνητρο. Αν αφαιρέσουμε κατά μέλη της δύο πρώτες σχέσεις των (10.2) και (10.3) καταλήγουμε στην

$$((u-u_h)_t(\cdot,t),\chi) + ((u-u_h)_x(\cdot,t),\chi') = 0 \quad \forall \chi \in S_h^r \quad \forall t \in [0,T].$$

Τονίζουμε ότι αυτή η σχέση ισχύει για $\chi \in S_h^r$, αφού ο S_h^r είναι υπόχωρος του V. Για να εκτιμήσουμε το σφάλμα $u-u_h$ θα θέλαμε να επιλέξουμε $\chi = u(\cdot,t) - u_h(\cdot,t)$, δυστυχώς όμως αυτό δεν επιτρέπεται, γιατί η $u(\cdot,t)$ δεν είναι γενικά στοιχείο του S_h^r. Αντί λοιπόν να συγκρίνουμε την u_h απ' ευθείας με τη u, θα τη συγκρίνουμε με μία “κατάλληλη” προβολή της u στον S_h^r. Το κατάλληλη εδώ αναφέρεται σε δύο πράγματα: Αφ' ενός η προβολή πρέπει να διαφέρει από τη u τάξη $O(h^r)$ στη νόρμα L^2, αφ' ετέρου δε πρέπει να ικανοποιεί μια σχέση αντίστοιχη της (10.3) με ένα σφάλμα συνέπειας της τάξεως $O(h^r)$. Για να επιτευχθεί ο τελευταίος στόχος πρέπει να αναιρεθούν κατά κάποιον τρόπο οι όροι που περιέχουν χωρικές παραγώγους γιατί για κάθε χωρική παράγωγο η τάξη προσέγγισης μειώνεται κατά ένα. Και οι δύο στόχοι επιτυγχάνονται με την ελλειπτική προβολή, την οποία εισάγουμε στη συνέχεια. Σημειώνουμε ακόμη ότι η ελλειπτική προβολή εξυπηρετεί καθαρά λόγους θεωρητικής μελέτης, ο υπολογισμός της ούτε εφικτός ούτε αναγκαίος είναι.

Λήμμα 2.1. *Έστω $w \in C_0^r[a,b]$ και $R_h w$ η ελλειπτική προβολή του w στον S_h^r, η οποία ορίζεται δια*

$$((R_h w)', \chi') = (w', \chi') \quad \forall \chi \in S_h^r. \tag{10.5}$$

Η $R_h w$ ορίζεται μονοσήμαντα και ισχύει

$$\|w - R_h w\| + h\|w' - (R_h w)'\| \le Ch^r \|w\|_r. \tag{10.6}$$

Απόδειξη. Θέτουμε τώρα $\varphi := -w''$, οπότε η w είναι λύση του ελλειπτικού προβλήματος

$$\begin{cases} -w'' = \varphi & \text{στο } [a,b] \\ w(a) = w(b) = 0. \end{cases}$$

Εύκολα βλέπουμε ότι η $R_h w$ είναι η προσεγγιστική λύση της w, όπως αυτή ορίστηκε στη μελέτη της μεθόδου των πεπερασμένων στοιχείων για ελλειπτικά προβλήματα στο Κεφάλαιο 7. Από αυτή τη μελέτη γνωρίζουμε ότι η $R_h w$ είναι καλώς ορισμένη και ότι ισχύει η (10.6). □

Ευστάθεια του ημιδιακριτού προβλήματος. Έστω u_h η λύση του προβλήματος (10.3) και $\tilde{u}_h$ αντί της u_h η αντίστοιχη λύση με αρχική τιμή $\tilde{u}_h^0$ αντί της u_h^0. Τότε για τη διαφορά $v_h := u_h - \tilde{u}_h$ έχουμε

$$(v_{ht}(\cdot,t), \chi) + (v_{hx}(\cdot,t), \chi') = 0. \tag{10.7}$$

Ευστάθεια στην L^2 νόρμα. Θέτοντας $\chi := v_h$ στην (10.7) λαμβάνουμε

$$\frac{1}{2}\frac{d}{dt}\|v_h(\cdot,t)\|^2 + \|v_{hx}(\cdot,t)\|^2 = 0, \quad t \in [0,T],$$

ή

$$\frac{d}{dt}\|v_h(\cdot,t)\|^2 \le 0$$

ή

$$\|v_h(\cdot,t)\| \le \|v_h(\cdot,s)\| \quad 0 \le s \le t \le T.$$

Ιδιαίτερα λοιπόν έχουμε

$$\|v_h(\cdot,t)\| \le \|v_h(\cdot,0)\|, \quad t \in [0,T]. \tag{10.8}$$

Ευστάθεια στην H^1 νόρμα. Θέτοντας $\chi := v_{ht}$ στην (10.7) λαμβάνουμε

$$\|v_{ht}(\cdot,t)\|^2 + \frac{1}{2}\frac{d}{dt}\|v_{hx}(\cdot,t)\|^2 = 0 \quad t \in [0,T],$$

ή

$$\frac{d}{dt}\|v_{hx}(\cdot,t)\|^2 \le 0$$

ή

$$\|v_{hx}(\cdot,t)\| \le \|v_{hx}(\cdot,s)\| \quad 0 \le s \le t \le T.$$

Ιδιαίτερα λοιπόν έχουμε

$$\|v_{hx}(\cdot,t)\| \le \|v_{hx}(\cdot,0)\|, \quad t \in [0,T]. \tag{10.9}$$

Συνέπεια. Έστω $W(\cdot,t) := R_h u(\cdot,t)$ η ελλειπτική προβολή της λύσης του προβλήματος (10.1). Θα αποδείξουμε συνέπεια του ημιδιακριτού σχήματος (10.2) για την W. Έχουμε, για $\chi \in S_h^r$,

$$\begin{aligned}
&(W_t(\cdot,t),\chi) + (W_x(\cdot,t),\chi') - (f(\cdot,t),\chi) = \\
&= (W_t(\cdot,t),\chi) + (u_x(\cdot,t),\chi') - (f(\cdot,t),\chi) = \\
&= (W_t(\cdot,t),\chi) - (u_t(\cdot,t),\chi),
\end{aligned}$$

δηλαδή

$$(W_t(\cdot,t),\chi)+(W_x(\cdot,t),\chi')=(f(\cdot,t),\chi)-(u_t(\cdot,t)-W_t(\cdot,t),\chi)\quad \forall\chi\in S_h^r. \tag{10.10}$$

Τώρα ο τελεστής της ελλειπτικής προβολής και η παραγώγιση ως προς t αντιμετατίθενται. Πράγματι έχουμε, για $\chi\in S_h^r$,

$$\begin{aligned}
&((R_hu_t)_x,\chi')=((u_t)_x,\chi')=(\frac{\partial}{\partial t}u_x,\chi')=\\
&=\frac{d}{dt}(u_x,\chi')=\frac{d}{dt}((R_hu)_x,\chi')=(\frac{\partial}{\partial t}((R_hu)_x),\chi')=\\
&=(((R_hu)_t)_x,\chi').
\end{aligned}$$

Συνεπώς, σύμφωνα με την (10.6), έχουμε

$$\|u_t(\cdot,t)-W_t(\cdot,t)\|\le Ch^r,\quad t\in[0,T]. \tag{10.11}$$

Σύγκλιση. Συνδυάζοντας τώρα ευστάθεια και συνέπεια αποδεικνύουμε σύγκλιση της ημιδιακριτής λύσης προς την ακριβή λύση του προβλήματος (10.1).

Σύγκλιση στην L^2 νόρμα.

Θεώρημα 10.1. *Έστω ότι η λύση u του προβλήματος* (10.1) *είναι αρκετά ομαλή. Τότε ισχύει*

$$\max_{0\le t\le T}\|u(\cdot,t)-u_h(\cdot,t)\|\le C\{\|u^0-u_h^0\|+h^r\}. \tag{10.12}$$

Απόδειξη. Με $\rho(\cdot,t)=u(\cdot,t)-W(\cdot,t)$ και $\vartheta(\cdot,t):=W(\cdot,t)-u_h(\cdot,t)$ έχουμε $u-u_h=\rho+\vartheta$. Τώρα, σύμφωνα με την (10.6),

$$\max_{0\le t\le T}\|\rho(\cdot,t)\|\le Ch^r. \tag{10.13}$$

Μένει συνεπώς να εκτιμήσουμε την $\|\vartheta(\cdot,t)\|$. Αφαιρώντας την πρώτη σχέση της (10.3) από την (10.10) έχουμε

$$(\vartheta_t(\cdot,t),\chi)+(\vartheta_x(\cdot,t),\chi')=-(\rho_t(\cdot,t),\chi)\quad\forall\chi\in S_h^r\quad\forall t\in[0,T]. \tag{10.14}$$

Θέτοντας $\chi:=\vartheta(\cdot,t)$ στην (10.14) λαμβάνουμε

$$\frac{1}{2}\frac{d}{dt}\|\vartheta(\cdot,t)\|^2+\|\vartheta_x(\cdot,t)\|^2=-(\rho_t(\cdot,t),\vartheta(\cdot,t))$$

ή

$$\frac{d}{dt}\|\vartheta(\cdot,t)\|^2 \leq \|\rho_t(\cdot,t)\|^2 + \|\vartheta(\cdot,t)\|^2$$

ή

$$\frac{d}{dt}\{\mathrm{e}^{-t}\|\vartheta(\cdot,t)\|^2\} \leq \mathrm{e}^{-t}\|\rho_t(\cdot,t)\|^2$$

ή

$$\mathrm{e}^{-t}\|\vartheta(\cdot,t)\|^2 \leq \|\vartheta(\cdot,0)\|^2 + \int_0^t \mathrm{e}^{-s}\|\rho_t(\cdot,s)\|^2 ds$$

ή, σύμφωνα με την (10.11),

$$\|\vartheta(\cdot,t)\|^2 \leq C\{\|\vartheta(\cdot,0)\|^2 + h^{2r}\}.$$

Επομένως

$$\|\vartheta(\cdot,t)\| \leq C\{\|\vartheta(\cdot,0)\| + h^r\}. \tag{10.15}$$

Τώρα $\vartheta(\cdot,0) = W(\cdot,0) - u_h^0 = (u^0 - u_h^0) - (u^0 - W(\cdot,0))$, οπότε σύμφωνα με την (10.6),

$$\|\vartheta(\cdot,0)\| \leq C\|u^0 - u_h^0\| + Ch^r.\, tag 10.16 \tag{6.1}$$

Από τις (10.15) και (10.16) έπεται ότι

$$\max_{0\leq t\leq T}\|\vartheta(\cdot,t)\| \leq C\{\|u^0 - u_h^0\| + Ch^r\}. \tag{10.17}$$

Οι (10.13) και (10.17) δίνουν αμέσως την (10.12). □

Σύγκλιση στην H^1 νόρμα.

Θεώρημα 10.2. *Έστω ότι η λύση u του προβλήματος* (10.1) *είναι αρκετά ομαλή. Τότε ισχύει*

$$\max_{0\leq t\leq T}\|u_x(\cdot,t) - u_{hx}(\cdot,t)\| \leq \|(u^0 - u_h^0)'\| + Ch^{r-1}. \tag{10.18}$$

Απόδειξη. Χρησιμοποιούμε τους συμβολισμούς που χρησιμοποιήσαμε στην απόδειξη του Θεωρήματος 10.1. Σύμφωνα με την (10.6) έχουμε κατ' αρχάς

$$\max_{0\leq t\leq T}\|\rho_x(\cdot,t)\| \leq Ch^{r-1}. \tag{10.19}$$

Θέτοντας τώρα $\chi := \vartheta_t(\cdot,t)$ στην (10.14) λαμβάνουμε

$$\|\vartheta_t(\cdot,t)\|^2 + \frac{1}{2}\frac{d}{dt}\|\vartheta_x(\cdot,t)\|^2 = -(\rho_t(\cdot,t),\vartheta_t(\cdot,t))$$

ή

$$\|\vartheta_t(\cdot,t)\|^2 + \frac{1}{2}\frac{d}{dt}\|\vartheta_x(\cdot,t)\|^2 \le \frac{1}{2}\|\rho_t(\cdot,t)\|^2 + \frac{1}{2}\|\vartheta_t(\cdot,t)\|^2$$

ή

$$\frac{d}{dt}\|\vartheta_x(\cdot,t)\|^2 \le \|\rho_t(\cdot,t)\|^2$$

ή

$$\|\vartheta_x(\cdot,t)\|^2 \le \|\vartheta_x(\cdot,0)\|^2 + \int_0^t \|\rho_t(\cdot,s)\|^2 ds.$$

Χρησιμοποιώντας εδώ την (10.11) έχουμε

$$\|\vartheta_x(\cdot,t)\|^2 \le \|\vartheta_x(\cdot,0)\|^2 + C^2h^{2r}$$

ή

$$\|\vartheta_x(\cdot,t)\| \le \|\vartheta_x(\cdot,0)\| + Ch^r. \tag{10.20}$$

Τώρα,

$$\|\vartheta_x(\cdot,0)\| \le \|(u^0-u_h^0)'\| + \|u_x^0 - W_x(\cdot,0)\|,$$

οπότε, σύμφωνα με την (10.6),

$$\|\vartheta_x(\cdot,0)\| \le \|(u^0-u_h^0)'\| + Ch^{r-1}. \tag{10.21}$$

Από τις (10.20) και (10.21) έπεται ότι

$$\max_{0\le t\le T}\|\vartheta_x(\cdot,t)\| \le \|(u^0-u_h^0)'\| + Ch^{r-1}. \tag{10.22}$$

Οι (10.19) και (10.22) δίνουν αμέσως την (10.18). □

Παρατήρηση 10.1 Για $u_h^0 := W(\cdot,0)$ έχουμε $\vartheta(\cdot,0) = 0$, και η (2.19) δίνει το εξής αποτέλεσμα *“υπερσύγκλισης”*

$$\max_{0\le t\le T}\|\vartheta_x(\cdot,t)\| \le Ch^r. \tag{10.23}$$

6.3 Πλήρως διακριτά σχήματα

Σε αυτή την παράγραφο θα μελετήσουμε τρία πλήρως διακριτά σχήματα για το πρόβλημα (10.1): την πεπλεγμένη μέθοδο του Euler, τη μέθοδο των Crank–Nicolson και τη μέθοδο του Euler.

6.3.1 Η πεπλεγμένη μέθοδος του Euler

Έστω $N \in \mathbb{N}, k := \frac{T}{N}$, και $t^n := nk, n = 0, \dots, N$. Για $v^0, \dots, v^N \in H_0^1$ θα χρησιμοποιήσουμε τον συμβολισμό

$$\partial v^n := \frac{v^{n+1} - v^n}{k}.$$

Διακριτοποιούμε το ημιδιακριτό πρόβλημα (10.3) ως προς το χρόνο με την πεπλεγμένη μέθοδο του Euler και οδηγούμαστε σε προσεγγίσεις $U^n \in S_h^r$ των $u^n := u(\cdot, t^n)$, οι οποίες δίνονται αναδρομικά από το σχήμα

$$\begin{cases} (\partial U^n, \chi) + (U_x^{n+1}, \chi') = (f(\cdot, t^{n+1}), \chi) \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1, \\ U^0 := u_h^0. \end{cases} \tag{10.24}$$

Έστω ότι έχουμε υπολογίσει το U^n. Αν χρησιμοποιήσουμε μια βάση του S_h^r, τότε βλέπουμε αμέσως ότι για να υπολογίσουμε το U^{n+1} πρέπει να λύσουμε ένα $\dim S_h^r \times \dim S_h^r$ γραμμικό σύστημα. Ο αντίστοιχος πίνακας είναι ανεξάρτητος του n, είναι συμμετρικός και θετικά ορισμένος. Ιδιαίτερα λοιπόν οι προσεγγίσεις $U^0, \dots, U^n$ ορίζονται καλώς δια της (10.24).

Ευστάθεια. Έστω $U^0, \dots, U^N \in S_h^r$ και $\tilde U^0, \dots, \tilde U^N \in S_h^r$ δύο ακολουθίες που ικανοποιούν την πρώτη σχέση της (3.1), δηλαδή προσεγγίσεις Euler που αντιστοιχούν σε αρχικές τιμές U^0 και $\tilde U^0$. Τότε για τις διαφορές $V^n := U^n - \tilde U^n, n = 0, \dots, N$, έχουμε

$$(\partial V^n, \chi) + (V_x^{n+1}, \chi') = 0 \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1. \tag{10.25}$$

Ευστάθεια στην L^2 νόρμα. Θέτοντας $\chi := V^{n+1}$ στην (10.25) λαμβάνουμε

$$(\partial V^n, V^{n+1}) + \|V_x^{n+1}\|^2 = 0$$

ή

$$(\partial V^n, V^{n+1}) \le 0$$

ή

$$\|V^{n+1}\|^2 \le (V^n, V^{n+1})$$

ή

$$\|V^{n+1}\|^2 \le \|V^n\|\,\|V^{n+1}\|,$$

οπότε

$$\|V^{n+1}\| \le \|V^n\|, \quad n = 0, \dots, N-1. \tag{10.26}$$

Ιδιαίτερα λοιπόν έχουμε

$$\max_{1\le n\le N} \|V^n\| \le \|V^0\|. \tag{10.27}$$

Παρατήρηση 10.2 Για δεδομένο $n,$ αν θεωρήσουμε το αντίστοιχο ομογενές γραμμικό σύστημα του (10.24), θέσουμε δηλαδή $U^n = 0$ και $f = 0,$ τότε σύμφωνα με την (10.26) θα έχουμε $U^{n+1} = 0.$ Έτσι διαπιστώνουμε ξανά ότι τα $U^1, \dots, U^N$ ορίζονται καλώς.

Ευστάθεια στην H^1 νόρμα. Θέτοντας $\chi := \partial V^n$ στην (10.25) λαμβάνουμε

$$\|\partial V^n\|^2 + (V_x^{n+1}, \partial V_x^n) = 0$$

και συμπεραίνουμε εύκολα ότι

$$\|V_x^{n+1}\| \le \|V_x^n\|, \quad n = 0, \dots, N-1. \tag{10.28}$$

Ιδιαίτερα λοιπόν εχουμε

$$\max_{1\le n\le N} \|V_x^n\| \le \|V_x^0\|. \tag{10.29}$$

Συνέπεια. Όπως και στην περίπτωση του ημιδιακριτού προβλήματος, θα αποδείξουμε συνέπεια του σχήματος (10.24) για την ελλειπτική προβολή W της λύσης u του προβλήματος (10.1), $W(\cdot,t) := R_h u(\cdot,t).$ Θέτουμε $W^n := W(\cdot,t^n), n = 0, \dots, N,$ και έχουμε, για $\chi \in S_h^r,$ σύμφωνα με την (10.5),

$$\begin{aligned}(\partial W^n,\chi) &+ (W_x^{n+1},\chi') - (f(\cdot,t^{n+1}),\chi)\\ &= (\partial W^n,\chi) + (u_x^{n+1},\chi') - (f(\cdot,t^{n+1}),\chi).\end{aligned}$$

Λαμβάνοντας τώρα υπ' όψιν την (10.2) έχουμε

$$(10.30) \qquad \begin{aligned}(\partial W^n, \chi) + (W_x^{n+1}, \chi') &= (f(\cdot, t^{n+1}), \chi) \\ &+ (\partial W^n - u_t(\cdot, t^{n+1}), \chi) \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1.\end{aligned}$$

Θα πρέπει λοιπόν να εκτιμήσουμε την διαφορά $\partial W^n - u_t(\cdot, t^{n+1})$. Θέτουμε $E^n := \partial W^n - u_t(\cdot, t^{n+1}), E_1^n := \partial W^n - \partial u^n$ και $E_2^n := \partial u^n - u_t(\cdot, t^{n+1})$. Προφανώς ισχύει $E^n = E_1^n + E_2^n$. Τώρα

$$E_1^n = \frac{1}{k} \int_{t^n}^{t^{n+1}} [W_t(\cdot, t) - u_t(\cdot, t)] dt,$$

συνεπώς

$$\|E_1^n\| \le \frac{1}{k} \int_{t^n}^{t^{n+1}} \|W_t(\cdot, t) - u_t(\cdot, t)\| dt,$$

οπότε, σύμφωνα με την (10.11), έχουμε

$$(10.31i) \qquad \max_{0 \le n \le N-1} \|E_1^n\| \le C h^r.$$

Επίσης, αναπτύσσοντας κατά Taylor έχουμε

$$E_2^n = \frac{1}{k} \int_{t^n}^{t^{n+1}} (t^n - s) u_{tt}(\cdot, s) ds,$$

συνεπώς

$$(10.31ii) \qquad \max_{0 \le n \le N-1} \|E_2^n\| \le C k.$$

Από τις (10.31i) και (10.31ii) συμπεραίνουμε αμέσως ότι ισχύει

$$(10.32) \qquad \max_{0 \le n \le N-1} \|E^n\| \le C(k + h^r).$$

Σύγκλιση. Συνδυάζοντας τώρα ευστάθεια και συνέπεια αποδεικνύουμε σύγκλιση των προσεγγίσεων, που δίνει το πεπλεγμένο σχήμα του Euler, προς την ακριβή λύση του προβλήματος (10.1).

Θεώρημα 10.3 *Έστω ότι η λύση u του προβλήματος* (10.1) *είναι αρκετά ομαλή. Τότε για τις προσεγγίσεις $U^0, \dots, U^N$, που δίνονται από την* (10.24), *ισχύει*

$$(10.33) \qquad \max_{0 \le n \le N} \|u^n - U^n\| \le \|u^0 - u_h^0\| + C(k + h^r).$$

Απόδειξη. Με $\rho^n := u^n - W^n$ και $\vartheta^n := W^n - U^n, n = 0, \dots, N$, έχουμε $u^n - U^n = \rho^n + \vartheta^n$. Τώρα, σύμφωνα με την (10.6),

$$\max_{0 \le n \le N} \|\rho^n\| \le Ch^r. \tag{10.34}$$

Μένει συνεπώς να εκτιμήσουμε την $\|\vartheta^n\|$. Αφαιρώντας την πρώτη σχέση της (10.24) από την (10.30) λαμβάνουμε

$$(\partial\vartheta^n, \chi) + (\vartheta_x^{n+1}, \chi') = (E^n, \chi) \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1. \tag{10.35}$$

Θέτοντας $\chi = \vartheta^{n+1}$ στην (10.35) και πολλαπλασιάζοντας επί k έχουμε

$$\|\vartheta^{n+1}\|^2 - (\vartheta^n, \vartheta^{n+1}) + k\|\vartheta_x^{n+1}\|^2 = k(E^n, \vartheta^{n+1})$$

ή

$$\|\vartheta^{n+1}\|^2 \le \{\|\vartheta^n\| + k\|E^n\|\}\|\vartheta^{n+1}\|$$

ή

$$\|\vartheta^{n+1}\| \le \|\vartheta^n\| + k\|E^n\|.$$

Επομένως

$$\|\vartheta^n\| \le \|\vartheta^0\| + k\sum_{j=0}^{n-1} \|E^j\|$$

ή

$$\|\vartheta^n\| \le \|\vartheta^0\| + Nk \max_{0 \le j \le N-1} \|E^j\|,$$

οπότε, σύμφωνα με την (10.32),

$$\max_{0 \le n \le N} \|\vartheta^n\| \le \|\vartheta^0\| + C(k + h^r). \tag{10.36}$$

Τώρα

$$\|\vartheta^0\| \le \|u^0 - u_h^0\| + \|u^0 - W^0\| \le \|u^0 - u_h^0\| + Ch^r,$$

και η (10.36) δίνει

$$\max_{0 \le n \le N} \|\vartheta^n\| \le \|u^0 - u_h^0\| + C(k + h^r). \tag{10.37}$$

Από τις (10.34) και (10.37) έπεται αμέσως η (10.33). □

6.3.2 Η μέθοδος των Crank–Nicolson

Όπως βλέπουμε από την εκτίμηση (10.33), η πεπλεγμένη μέθοδος του Euler έχει τάξη ένα ως προς την παράμετρο διακριτοποίησης στο χρόνο k. Η μέθοδος των Crank–Nicolson, την οποία θα μελετήσουμε σε αυτό το εδάφιο, είναι δεύτερης τάξης ως προς k. Σημειώνουμε πάντως ότι η πεπλεγμένη μέθοδος του Euler είναι πιο "φυσιολογική" από την μέθοδο των Crank–Nicolson για παραβολικά προβλήματα με την έννοια ότι οι προσεγγίσεις που δίνει η μέθοδος του Euler μιμούνται καλύτερα τις ιδιότητες ομαλοποίησης των λύσεων παραβολικών εξισώσεων. Η εξήγηση για αυτό το γεγονός ξεφεύγει από τους σκοπούς της παρούσης εισαγωγής στο θέμα.

Πέραν των συμβολισμών του προηγουμένου εδαφίου θα χρησιμοποιήσουμε εδώ και τον εξής συμβολισμό
$$v^{n+\frac{1}{2}} := \frac{1}{2}(v^n + v^{n+1})$$
για $v^0, \dots, v^N \in V$, και, ανάλογα, $t^{n+\frac{1}{2}} := \frac{1}{2}(t^n + t^{n+1})$.

Οι προσεγγίσεις των Crank–Nicolson $U^n \in S_h^r$ των $u^n = u(\cdot, t^n)$ δίνονται αναδρομικά από το εξής σχήμα

$$\text{(10.38)} \qquad \begin{cases} (\partial U^n, \chi) + (U_x^{n+\frac{1}{2}}, \chi') = (f(\cdot, t^{n+\frac{1}{2}}), \chi) \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1, \\ U^0 := u_h^0. \end{cases}$$

Ευστάθεια. Κατ' αναλογίαν προς την (10.25) οδηγούμαστε στη μελέτη του σχήματος

$$\text{(10.39)} \qquad (\partial V^n, \chi) + (V_x^{n+\frac{1}{2}}, \chi') = 0 \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1,$$

με $V^n \in S_h^r, \quad n = 0, \dots, N.$

Ευστάθεια στην L^2 νόρμα. Θέτοντας $\chi := V^{n+\frac{1}{2}}$ στην (10.39) λαμβάνουμε

$$(\partial V^n, V^{n+\frac{1}{2}}) + \|V_x^{n+\frac{1}{2}}\|^2 = 0$$

ή

$$(V^{n+1} - V^n, V^{n+1} + V^n) \leq 0$$

ή

$$\|V^{n+1}\|^2 \leq \|V^n\|^2,$$

οπότε

$$\|V^{n+1}\| \le \|V^n\|, \quad n = 0, \dots, N-1. \tag{10.40}$$

Ιδιαίτερα λοιπόν έχουμε

$$\max_{1\le n\le N} \|V^n\| \le \|V^0\|. \tag{10.41}$$

Από την (10.40) έπεται εύκολα ότι οι προσεγγίσεις των Crank–Nicolson $U^1, \dots, U^N$ ορίζονται καλώς.

Ευστάθεια στην H^1 νόρμα. Θέτοντας $\chi := \partial V^n$ στην (10.39) λαμβάνουμε

$$\|\partial V^n\|^2 + (V_x^{n+\frac12}, \partial V_x^n) = 0$$

ή

$$((V^n + V^{n+1})_x, (V^{n+1} - V^n)_x) \le 0$$

ή

$$\|V_x^{n+1}\| \le \|V_x^n\|, \quad n = 0, \dots, N-1. \tag{10.42}$$

Ιδιαίτερα λοιπόν έχουμε

$$\max_{1\le n\le N} \|V_x^n\| \le \|V_x^0\|. \tag{10.43}$$

Συνέπεια. Όπως και στην περίπτωση της πεπλεγμένης μεθόδου του Euler, θα αποδείξουμε συνέπεια του σχήματος των Crank–Nicolson για την ελλειπτική προβολή W της λύσης u του προβλήματος (10.1). Χρησιμοποιώντας την (10.5), για $\chi \in S_h^r$, έχουμε

$$\begin{aligned}(\partial W^n, \chi) + (W_x^{n+\frac12}, \chi') - (f(\cdot, t^{n+\frac12}), \chi) =\\ = (\partial W^n, \chi) + (u_x^{n+\frac12}, \chi') - (f(\cdot, t^{n+\frac12}), \chi),\end{aligned}$$

οπότε λαμβάνοντας υπ' όψιν την (10.2) έχουμε, για $n = 0, \dots, N-1$,

$$(\partial W^n, \chi) + (W_x^{n+\frac12}, \chi') = (f(\cdot, t^{n+\frac12}), \chi) + (E^n, \chi) \quad \forall \chi \in S_h^r, \tag{10.44}$$

με $E^n := (\partial W^n - \partial u^n) + [\partial u^n - u_t(\cdot, t^{n+\frac{1}{2}})] + [u_{xx}(\cdot, t^{n+\frac{1}{2}}) - u_{xx}^{n+\frac{1}{2}}]$. Θέτουμε

$$\begin{aligned} E_1^n &:= \partial W^n - \partial u^n \\ E_2^n &:= \partial u^n - u_t(\cdot, t^{n+\frac{1}{2}}) \\ E_3^n &:= u_{xx}(\cdot, t^{n+\frac{1}{2}}) - u_{xx}^{n+\frac{1}{2}} \end{aligned}$$

και προφανώς έχουμε $E^n = E_1^n + E_2^n + E_3^n$. Το E_1^n είναι το ίδιο όπως και στην πεπλεγμένη μέθοδο του Euler, συνεπώς κατά την (10.31i) έχουμε

$$\max_{1\le n\le N-1} \|E_1^n\| \le Ch^r. \tag{10.45i}$$

Επίσης, αναπτύσσοντας κατά Taylor ως προς το $t^{n+\frac{1}{2}}$, έχουμε

$$E_2^n = \frac{1}{2k}[\int_{t^n}^{t^{n+\frac{1}{2}}} (s-t^n)^2 u_{ttt}(\cdot, s)ds + \int_{t^{n+\frac{1}{2}}}^{t^{n+1}} (s-t^{n+1})^2 u_{ttt}(\cdot, s)ds],$$

συνεπώς

$$\max_{1\le n\le N-1} \|E_2^n\| \le Ck^2. \tag{10.45ii}$$

Επιπλέον έχουμε

$$2[u(\cdot, t^{n+\frac{1}{2}}) - u^{n+\frac{1}{2}}] = \int_{t^{n+\frac{1}{2}}}^{t^{n+1}} (s-t^{n+1}) u_{tt}(\cdot, s)ds - \int_{t^n}^{t^{n+\frac{1}{2}}} (s-t^n) u_{tt}(\cdot, s)ds,$$

συνεπώς

$$\max_{1\le n\le N-1} \|E_3^n\| \le Ck^2. \tag{10.45iii}$$

Από τις (10.45i), (10.45ii) και (10.45iii) συμπεραίνουμε αμέσως ότι ισχύει

$$\max_{1\le n\le N-1} \|E^n\| \le C(k^2 + h^r). \tag{10.46}$$

Σύγκλιση. Συνδυάζοντας τώρα ευστάθεια και συνέπεια αποδεικνύουμε σύγκλιση.

Θεώρημα 10.4 *Έστω ότι η λύση u του προβλήματος* (10.1) *είναι αρκετά ομαλή. Τότε για τις προσεγγίσεις $U^0, \dots, U^N$, που δίνονται από την* (10.38), *ισχύει*

$$\max_{1\le n\le N} \|u^n - U^n\| \le \|u^0 - u_h^0\| + C(k^2 + h^r). \tag{10.47}$$

Απόδειξη. Με $\rho^n := u^n - W^n$ και $\vartheta^n := W^n - U^n, \quad n = 0, \dots, N$, έχουμε $u^n - U^n = \rho^n + \vartheta^n$. Τώρα, σύμφωνα με την (10.6),

$$\max_{1 \le n \le N} \|\rho^n\| \le Ch^r. \tag{10.48}$$

Μένει συνεπώς να εκτιμήσουμε την $\|\vartheta^n\|$. Αφαιρώντας την πρώτη σχέση της (10.38) από την (10.44) λαμβάνουμε

$$(\partial \vartheta^n, \chi) + (\vartheta_x^{n+\frac{1}{2}}, \chi') = (E^n, \chi) \quad \forall \chi \in S_h^r \quad n = 0, \dots, N-1. \tag{10.49}$$

Θέτοντας $\chi = \vartheta^{n+\frac{1}{2}}$ στην (10.49) και πολλαπλασιάζοντας επί $2k$ έχουμε

$$\|\vartheta^{n+1}\|^2 - \|\vartheta^n\|^2 + 2k\|\vartheta_x^{n+\frac{1}{2}}\|^2 = k(E^n, \vartheta^n + \vartheta^{n+1})$$

ή

$$\|\vartheta^{n+1}\|^2 - \|\vartheta^n\|^2 \le k\|E^n\|(\|\vartheta^n\| + \|\vartheta^{n+1}\|)$$

ή

$$\|\vartheta^{n+1}\| \le \|\vartheta^n\| + k\|E^n\|, \quad n = 0, \dots, N-1. \tag{10.50}$$

Προχωρώντας τώρα όπως στην απόδειξη του Θεωρήματος 10.3 και λαμβάνοντας υπ' όψιν την (10.46) αποδεικνύουμε ότι

$$\max_{1 \le n \le N} \|\vartheta^n\| \le \|u^0 - u_h^0\| + C(k^2 + h^r). \tag{10.51}$$

Από τις (10.48) και (10.51) έπεται αμέσως η (10.47). □

6.3.3 Η μέθοδος του Euler

Οι δύο μέθοδοι που μελετήσαμε μέχρι τώρα, δηλ. η πεπλεγμένη μέθοδος του Euler και η μέθοδος των Crank–Nicolson, είναι πεπλεγμένες. Δείξαμε, μεταξύ των άλλων, ότι αμφότερες είναι *απεριόριστα ευσταθείς*, δηλ. ευσταθείς χωρίς περιορισμούς στις παραμέτρους διακριτοποίησης. Η μέθοδος του Euler, την οποία θα μελετήσουμε τώρα, είναι πρώτης τάξης ως προς k, *ευσταθής υπό συνθήκες*, και όχι χρήσιμη για το συγκεκριμένο πρόβλημα. Δίνεται εδώ περισσότερο για παιδαγωγικούς λόγους.

Διακριτοποιούμε το ημιδιακριτό πρόβλημα (10.2) ως προς το χρόνο με τη μέθοδο του Euler και οδηγούμαστε σε προσεγγίσεις $U^n \in S_h^r$ των $u^n = u(\cdot, t^n)$, οι οποίες δίνονται αναδρομικά από το σχήμα

$$\begin{cases} (\partial U^n, \chi) + (U_x^n, \chi') = (f(\cdot, t^n), \chi) \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1, \\ U^0 := u_h^0. \end{cases} \tag{10.52}$$

Ευστάθεια. Κατ' αναλογίαν προς την (10.25) οδηγούμαστε στη μελέτη του σχήματος

$$(\partial V^n, \chi) + (V_x^n, \chi') = 0 \qquad \forall \chi \in S_h^r, \qquad n = 0, \dots, N-1, \tag{10.53}$$

με $V^n \in S_h^r, \quad n = 0, \dots, N.$

Πέραν της προσεγγιστικής ιδιότητας (7.2) υποθέτουμε στο εξής και την ακόλουθη αντίστροφη ανισότητα για την οικογένεια $\{S_h^r\}_{0<h\le 1}$

$$\|\chi'\| \le C_1 h^{-1} \|\chi\| \qquad \forall \chi \in S_h^r \tag{10.54}$$

με μια σταθερά C_1 ανεξάρτητη του h, βλ. την Άσκηση 10.3.

Θέτοντας $\chi := V^n$ στην (10.53) λαμβάνουμε

$$(V^{n+1}, V^n) - \|V^n\|^2 + k\|V_x^n\|^2 = 0$$

ή

$$\frac{1}{2}\{\|V^{n+1}\|^2 - \|V^n\|^2 - \|V^{n+1} - V^n\|^2\} + k\|V_x^n\|^2 = 0. \tag{10.55i}$$

Θέτοντας εξ άλλου $\chi := \partial V^n$ στην (10.53) λαμβάνουμε

$$\|\partial V^n\|^2 + (V_x^n, \partial V_x^n) = 0$$

ή

$$\|V^{n+1} - V^n\|^2 + k(V_x^n, V_x^{n+1} - V_x^n) = 0$$

ή

$$\|V^{n+1} - V^n\|^2 \le k\|V_x^n\| \|V_x^{n+1} - V_x^n\|.$$

Χρησιμοποιώντας συνεπώς την (10.54) έχουμε

$$\|V^{n+1} - V^n\|^2 \le C_1 \frac{k}{h} \|V_x^n\| \|V^{n+1} - V^n\|,$$

οπότε

$$\|V^{n+1} - V^n\| \le C_1 \frac{k}{h} \|V_x^n\|. \tag{10.55ii}$$

Από τις (10.55i) και (10.55ii) λαμβάνουμε τώρα

$$\|V^{n+1}\|^2 - \|V^n\|^2 \le (C_1^2 \frac{k}{h^2} - 2)k\|V_x^n\|^2. \tag{10.56}$$

Υποθέτουμε τώρα ότι ικανοποιείται η εξής συνθήκη

$$\frac{k}{h^2} \le \frac{2}{C_1^2}. \tag{10.57}$$

Από τις (10.56) και (10.57) λαμβάνουμε

$$\|V^{n+1}\| \le \|V^n\|, \qquad n = 0, \dots, N-1. \tag{10.58}$$

Ιδιαίτερα επομένως έχουμε

$$\max_{1\le n\le N} \|V^n\| \le \|V^0\|. \tag{10.59}$$

Παρατήρηση 10.3 Η συνθήκη (10.57), υπό την οποία αποδείξαμε ευστάθεια της μεθόδου του Euler, είναι πολύ περιοριστική, γεγονός που καθιστά τη μέθοδο σχεδόν άχρηστη για το συγκεκριμένο πρόβλημα.

Συνέπεια. Εργαζόμενοι όπως στην περίπτωση της πεπλεγμένης μεθόδου του Euler, αποδεικνύουμε εύκολα ότι ισχύει

$$\begin{aligned} &(\partial W^n, \chi) + (W_x^n, \chi') = (f(\cdot, t^n), \chi) + (E^n, \chi) \quad \forall \chi \in S_h^r, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad n = 0, \dots, N-1, \end{aligned} \tag{10.60}$$

με E^n τέτοια ώστε

$$\max_{0\le n\le N-1} \|E^n\| \le C(k + h^r). \tag{10.61}$$

Σύγκλιση. Συνδυάζοντας τώρα ευστάθεια και συνέπεια αποδεικνύουμε σύγκλιση.

Θεώρημα 10.5 *Έστω ότι η λύση u του προβλήματος* (10.1) *είναι αρκετά ομαλή, και ότι ισχύουν οι* (10.54) *και* (10.57). *Τότε για τις προσεγγίσεις $U^0, \dots, U^N$, που δίνονται από την* (10.52), *ισχύει*

$$\max_{0\le n\le N} \|u^n - U^n\| \le C\|u^0 - u_h^0\| + (k + h^r). \tag{10.62}$$

Απόδειξη. Με $\rho^n := u^n - W^n$ και $\vartheta^n := W^n - U^n$, $n = 0, \dots, N$, έχουμε $u^n - U^n = \rho^n + \vartheta^n$. Τώρα, σύμφωνα με την (10.6),

$$\max_{0\le n\le N} \|\rho^n\| \le Ch^r. \tag{10.63}$$

Μένει συνεπώς να εκτιμήσουμε την $\|\vartheta^n\|$. Αφαιρώντας την πρώτη σχέση της (10.52) από την (10.60) λαμβάνουμε

$$(\partial\vartheta^n, \chi) + (\vartheta^n_x, \chi') = (E^n, \chi) \qquad \forall \chi \in S^r_h, \qquad n = 0, \dots, N-1. \tag{10.64}$$

Θέτοντας $\chi := \vartheta^n$ στην (10.64) λαμβάνουμε

$$(\vartheta^{n+1}, \vartheta^n) - \|\vartheta^n\|^2 + k\|\vartheta^n_x\|^2 = k(E^n, \vartheta^n)$$

ή

$$\frac{1}{2}\{\|\vartheta^{n+1}\|^2 - \|\vartheta^n\|^2 - \|\vartheta^{n+1} - \vartheta^n\|^2\} + k\|\vartheta^n_x\|^2 = k(E^n, \vartheta^n). \tag{10.65i}$$

Θέτοντας εξ άλλου $\chi := \partial\vartheta^n$ στην (10.64) λαμβάνουμε

$$\|\vartheta^{n+1} - \vartheta^n\|^2 + k(\vartheta^n_x, \vartheta^{n+1}_x - \vartheta^n_x) = k(E^n, \vartheta^{n+1} - \vartheta^n),$$

συνεπώς

$$\|\vartheta^{n+1} - \vartheta^n\|^2 \le k\|\vartheta^n_x\|\,\|\vartheta^{n+1}_x - \vartheta^n_x\| + k\|E^n\|\,\|\vartheta^{n+1} - \vartheta^n\|,$$

οπότε, σύμφωνα με την (10.54),

$$\|\vartheta^{n+1} - \vartheta^n\|^2 \le C_1\frac{k}{h}\|\vartheta^n_x\|\,\|\vartheta^{n+1} - \vartheta^n\| + k\|E^n\|\,\|\vartheta^{n+1} - \vartheta^n\|$$

ή

$$\|\vartheta^{n+1} - \vartheta^n\| \le C_1\frac{k}{h}\|\vartheta^n_x\| + k\|E^n\|$$

ή

$$\|\vartheta^{n+1} - \vartheta^n\|^2 \le C_1^2\frac{k^2}{h^2}\|\vartheta^n_x\|^2 + k^2\|E^n\|^2 + 2C_1\frac{k}{h}\|\vartheta^n_x\|\,\|E^n\|. \tag{10.65ii}$$

Από τις (10.65i) και (10.65ii), χρησιμοποιώντας πάλι την (10.54), λαμβάνουμε

$$\begin{aligned}\|\vartheta^{n+1}\|^2 - \|\vartheta^n\|^2 \le{}& (C_1^2\frac{k^2}{h^2} - 2k)\|\vartheta^n_x\|^2 + k^2\|E^n\|^2 + 2k(E^n, \vartheta^n)\\ &+ 2C_1^2\frac{k^2}{h^2}\|\vartheta^n\|\,\|E^n\|\end{aligned}$$

ή

$$\|\vartheta^{n+1}\|^2 \le (1+ck)\|\vartheta^n\|^2 + Ck\|E^n\|^2.$$

Επομένως

$$\begin{aligned}
\|\vartheta^n\|^2 &\le (1+ck)^n\|\vartheta^0\|^2 + Ck\sum_{j=0}^{n-1}(1+ck)^j\|E^j\|^2\\
&\le \mathrm{e}^{nck}\|\vartheta^0\|^2 + \max_{0\le j\le N-1}\|E^j\|^2\cdot Ck\sum_{j=0}^{N-1}\mathrm{e}^{jck}\\
&= \mathrm{e}^{nck}\|\vartheta^0\|^2 + \max_{0\le j\le N-1}\|E^j\|^2\cdot Ck\frac{\mathrm{e}^{Nck}-1}{\mathrm{e}^{ck}-1}\\
&\le \mathrm{e}^{cT}\|\vartheta^0\|^2 + \frac{C}{c}\max_{0\le j\le N-1}\|E^j\|^2\,(\mathrm{e}^{cT}-1).
\end{aligned}$$

Χρησιμοποιώντας εδώ την (10.61) λαμβάνουμε

$$\max_{0\le n\le N}\|\vartheta^n\|^2 \le C\|\vartheta^0\|^2 + C(k+h^r)^2$$

ή

$$\max_{0\le n\le N}\|\vartheta^n\| \le C\|\vartheta^0\| + C(k+h^r). \tag{10.66}$$

Από τις (10.63) και (10.66) έπεται εύκολα η (10.62). □

Ασκήσεις

10.1. Θεωρούμε το ακόλουθο πρόβλημα αρχικών και συνοριακών συνθηκών

$$\begin{cases}
u_t = u_{xx} - u + f & \text{στο } [a,b]\times[0,T]\\
u_x(a,\cdot) = u_x(b,\cdot) = 0 & \text{στο } [0,T]\\
u(\cdot,0) = u^0 & \text{στο } [a,b].
\end{cases}$$

Θεωρήστε έναν κατάλληλο χώρο πεπερασμένων στοιχείων για αυτό το πρόβλημα, βλ. την Άσκηση 7.1. Ποια μορφή παίρνουν για αυτό το πρόβλημα οι μέθοδοι πεπερασμένων στοιχείων Euler, πεπλεγμένη Euler, και Crank–Nicolson. Αποδείξτε εκτιμήσεις αντίστοιχες αυτών που δώθηκαν σε αυτό το Κεφάλαιο.

10.2. Θεωρούμε το πρόβλημα της Άσκησης 8.8. Μελετήστε τη μέθοδο πεπερασμένων στοιχείων/ Crank–Nicolson για αυτό το πρόβλημα. Ιδιαίτερα, αποδείξτε για τις προσεγγίσεις U^n ότι έχουν την ίδια L^2 νόρμα για όλα τα n (για κάθε συγκεκριμένο h), το σχήμα είναι όπως λέμε *συντηρητικό*.

10.3. (Η αντίστροφη ανισότητα (10.54).)

α) Έστω $J \in \mathbb{N}, h := \frac{b-a}{J+1}$ και $x_i := a + ih, i = 0, \dots, J+1$, ένας ομοιόμορφος διαμερισμός του διαστήματος $[a, b]$. Θεωρούμε τη συνάρτηση $v : [a, b] \to \mathbb{R}$,

$$v(x) := \begin{cases} \dfrac{x - x_0}{h}, & x_0 \le x \le x_1, \\ \dfrac{x_2 - x}{h}, & x_1 < x \le x_2, \\ 0, & \text{διαφορετικά}. \end{cases}$$

Αποδείξτε ότι $\|v'\| = \sqrt{3}\, h^{-1}\|v\|$, και οδηγηθείτε στο συμπέρασμα ότι ο εκθέτης του h στο δεξιό μέλος της αντίστροφης ανισότητας (10.54) είναι ο καλύτερος δυνατός.

β) Έστω $a = x_0 < x_1 < \cdots < x_J = b$ ένας διαμερισμός του διαστήματος $[a, b]$, $h_i := x_i - x_{i-1}, i = 1, \dots, J$, και

$$S_h^r := \{\chi \in C_0[a, b] : \chi|_{[x_{i-1}, x_i]} \in \mathbb{P}_{r-1}, i = 1 \dots, J\}.$$

Έστω $v \in S_h^r$ και $i \in \{0, \dots, J-1\}$. Ορίζουμε τη συνάρτηση $\hat{v} : [0, 1] \to \mathbb{R}, \hat{v}(s) := v\big(x_i + (x_{i+1} - x_i)s\big)$. Αποδείξτε ότι

$$\|v\|^2_{L^2(x_i, x_{i+1})} = h_i^2 \int_0^1 |\hat{v}(s)|^2 \, ds$$

και

$$\|v'\|^2_{L^2(x_i, x_{i+1})} = \int_0^1 |\hat{v}'(s)|^2 \, ds.$$

Τώρα $\hat{v} \in \mathbb{P}_{r-1}$, επομένως, επειδή είμαστε σε έναν χώρο πεπερασμένης διάστασης, υπάρχει σταθερά C τέτοια ώστε

$$\int_0^1 |\hat{v}'(s)|^2 \, ds \le C \int_0^1 |\hat{v}(s)|^2 \, ds.$$

Συνδυάστε τις τρεις αυτές σχέσεις για να διαπιστώσετε ότι

$$\|v'\|^2_{L^2(x_i, x_{i+1})} \le C \frac{1}{(h_i)^2} \|v\|^2_{L^2(x_i, x_{i+1})},$$

οπότε, αθροίζοντας ως προς i,

$$\|v'\|_{L^2}^2 \leq C(\min_i h_i)^{-2}\|v\|_{L^2}^2.$$

Οδηγηθείτε στο συμπέρασμα ότι η αντίστροφη ανισότητα (10.54) ισχύει, αν η οικογένεια των διαμερισμών είναι ημιομοιόμορφη, δηλαδή αν

$$\frac{\max_i(x_{i+1}-x_i)}{\min_i(x_{i+1}-x_i)} \leq c,$$

με την ίδια σταθερά c για όλους τους διαμερισμούς.

10.4. Έστω $\vartheta \in [0,1]$. Θεωρούμε την εξής μέθοδο για τη διακριτοποίηση του ημιδιακριτού προβλήματος (10.3) στο χρόνο

$$\begin{cases} (\partial U^n, \chi) + (\vartheta U_x^{n+1} + (1-\vartheta)U_x^n, \chi') = \\ \qquad\qquad = (\vartheta f(\cdot, t^{n+1}) + (1-\vartheta) f(\cdot, t^{n+1}), \chi) \quad \forall \chi \in S_h^r, \\ \qquad\qquad\qquad\qquad\qquad\qquad n = 0, \dots, N-1, \\ U^0 := u_h^0. \end{cases}$$

Ποιες μέθοδοι προκύπτουν στις περιπτώσεις $\vartheta = 0$ και $\vartheta = 1$, αντίστοιχα; Ποια είναι η διαφορά της μεθόδου για $\vartheta = 1/2$ από τη μέθοδο των Crank–Nicolson; Σημειώνουμε ότι η κλάση των εν λόγω μεθόδων αναφέρεται στη βιβλιογραφία ως *μέθοδοι θήτα*.

7. Μία υπερβολική εξίσωση

Έστω $T > 0$. Θεωρούμε το εξής πρόβλημα αρχικών και συνοριακών συνθηκών για μια υπερβολική εξίσωση δευτέρας τάξεως: Ζητείται μια συνάρτηση $u : [a, b] \times [0, T] \longrightarrow \mathbb{R}$, δύο φορές συνεχώς παραγωγίσιμη, τέτοια ώστε

$$\begin{cases} u_{tt} = (pu_x)_x - qu + f & \text{στο } [a,b] \times [0,T] \\ u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T] \\ u(\cdot,0) = u_0 & \text{στο } [a,b] \\ u_t(\cdot,0) = u_1 & \text{στο } [a,b] \end{cases} \tag{11.1}$$

όπου $q, f \in C([a, b] \times [0, T])$ και $p \in C^1([a, b] \times [0, T])$, $\quad p(x, t) > 0$ για $x \in [a, b]$ και $t \in [0, T]$.

Πολλαπλασιάζουμε τη Δ.Ε. του (11.1) επί u_t και ολοκληρώνουμε στο $[a, b]$ οπότε παίρνουμε

$$\frac{1}{2}\frac{d}{dt}\|u_t(\cdot,t)\|^2 = -\int_a^b pu_xu_{xt}dx - \int_a^b quu_tdx + \int_a^b fu_tdx$$

ή

$$\frac{1}{2}\frac{d}{dt}\big\{\|u_t(\cdot,t)\|^2 + \|\sqrt{p(\cdot,t)}u_x(\cdot,t)\|^2\big\} = \frac{1}{2}\int_a^b p_t(u_x)^2dx - \int_a^b quu_tdx + \int_a^b fu_tdx$$

και συμπεραίνουμε εύκολα ότι

$$\begin{aligned} &\frac{d}{dt}\big\{\|u_t(\cdot,t)\|^2 + \|\sqrt{p(\cdot,t)}u_x(\cdot,t)\|^2\big\} \\ &\qquad\qquad \le C\big\{\|u_t(\cdot,t)\|^2 + \|\sqrt{p(\cdot,t)}u_x(\cdot,t)\|^2\big\} + 2\int_a^b fu_tdx. \end{aligned} \tag{11.2}$$

Εύκολα βλέπουμε ότι ο τελευταίος όρος στην (11.2) μπορεί να αντικατασταθεί από $\|f(\cdot,t)\|^2$ αν μεγαλώσουμε λίγο τη σταθερά C. Έτσι παίρνουμε αμέσως

$$\frac{d}{dt}\big\{\mathrm{e}^{-Ct}[\|u_t(\cdot,t)\|^2+\|\sqrt{p(\cdot,t)}u_x(\cdot,t)\|^2]\big\}\le \mathrm{e}^{-Ct}\|f(\cdot,t)\|^2.$$

Ολοκληρώνοντας αυτή τη σχέση από 0 έως t λαμβάνουμε

$$\|u_t(\cdot,t)\|^2+\|\sqrt{p(\cdot,t)}u_x(\cdot,t)\|^2\le \mathrm{e}^{Ct}\big\{\|u_1\|^2+\|\sqrt{p(\cdot,0)}u_0'\|^2\big\}+\int_0^t \mathrm{e}^{C(t-s)}\|f(\cdot,s)\|^2ds$$

και συμπεραίνουμε εύκολα ότι

$$\max_{0\le t\le T}\big\{\|u_t(\cdot,t)\|+\|u_x(\cdot,t)\|\big\}\le C(\|u_1\|+\|u_0'\|)+C\max_{0\le t\le T}\|f(\cdot,t)\|. \tag{11.3}$$

Θα προσπαθήσουμε τώρα να αποδείξουμε μια ανισότητα παρόμοια με την (11.3) αλλά με την $\|\cdot\|_{-1}$ νόρμα της $f, f_t,\dots,$ στη θέση της $\|f(\cdot,t)\|$. Αυτό είναι χρήσιμο, γιατί διακριτά ανάλογα μας δίνουν εκτιμήσεις βέλτιστης τάξεως και στην περίπτωση μη ομοιόμορφου διαμερισμού στο χώρο, θέμα με το οποίο δεν θα ασχοληθούμε πάντως στη συνέχεια. Επιστρέφουμε λοιπόν στην (11.2) και τη γράφουμε στη μορφή

$$\frac{d}{dt}\big\{\mathrm{e}^{-Ct}[\|u_t(\cdot,t)\|^2+\|\sqrt{p(\cdot,t)}u_x(\cdot,t)\|^2]\big\}\le 2\mathrm{e}^{-Ct}\int_a^b fu_tdx.$$

Ολοκληρώνοντας από 0 έως t λαμβάνουμε, υποθέτοντας ότι η f_t είναι συνεχής,

$$\begin{aligned}&\frac{1}{2}\mathrm{e}^{-Ct}[\|u_t(\cdot,t)\|^2+\|\sqrt{p(\cdot,t)}u_x(\cdot,t)\|^2]\le\int_a^b\int_0^t \mathrm{e}^{-Ct}fu_tdtdx=\\&\int_a^b \mathrm{e}^{-Ct}f(x,t)u(x,t)dx-\int_a^b f(x,0)u(x,0)dx-\int_a^b\int_0^t(\mathrm{e}^{-Ct}f)_t(x,t)u(x,t)dxdt.\end{aligned}$$

Με τη βοήθεια της ανισότητας $(f,u)\le\|f\|_{-1}\|u_x\|$ και της αριθμητικής–γεωμετρικής ανισότητας συμπεραίνουμε εύκολα ότι

$$\begin{aligned}\max_{0\le t\le T}\big\{\|u_t(\cdot,t)\|+\|u_x(\cdot,t)\|_{-1}\big\}\le C(\|u_1\|+\|u_0'\|)&+C\max_{0\le t\le T}\|f(\cdot,t)\|_{-1}\\&+C\max_{0\le t\le T}\|f_t(\cdot,t)\|_{-1}.\end{aligned}\tag{11.4}$$

Από τις (11.3) ή (11.4) έπεται ότι η λύση u του προβλήματος (11.1) εξαρτάται συνεχώς από τα αρχικά δεδομένα και τον μη ομογενή όρο, και ιδιαίτερα ότι το πρόβλημα (11.1) έχει το πολύ μια λύση. Είναι επίσης γνωστό ότι, αν τα δεδομένα του προβλήματος (11.1) είναι ομαλά και συμβατά, τότε αυτό έχει λύση.

Ασκήσεις

11.1. Θεωρούμε το ακόλουθο πρόβλημα αρχικών και συνοριακών συνθηκών: Ζητείται μια συνάρτηση $u : [a,b] \times [0,T] \longrightarrow \mathbb{R}$, δύο φορές συνεχώς παραγωγίσιμη, τέτοια ώστε

$$\begin{cases} u_{tt} = (pu_x)_x - ru_x - qu + f & \text{στο } [a,b] \times [0,T] \\ u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T] \\ u(\cdot,0) = u_0 & \text{στο } [a,b] \\ u_t(\cdot,0) = u_1 & \text{στο } [a,b], \end{cases} \tag{11.5}$$

όπου $u_0, u_1 : [a,b] \longrightarrow \mathbb{R}$ δεδομένες αρχικές τιμές και $r,\ q,\ f$ συνεχείς στο $[a,b] \times [0,T]$ και p συνεχώς παραγωγίσιμη και θετική στο $[a,b] \times [0,T]$. Αποδείξτε γι' αυτό το πρόβλημα μια εκτίμηση της μορφής (11.3) και υποθέτοντας ότι η f είναι συνεχώς παραγωγίσιμη ως προς t αποδείξτε και μια εκτίμηση της μορφής (11.4).

11.2. Θεωρούμε το ακόλουθο πρόβλημα αρχικών και συνοριακών συνθηκών: Ζητείται μια συνάρτηση $u : [a,b] \times [0,T] \longrightarrow \mathbb{R}$, δύο φορές συνεχώς παραγωγίσιμη τέτοια ώστε

$$\begin{cases} u_{tt} = (pu_x)_x - qu + f & \text{στο } [a,b] \times [0,T] \\ u_x(a,\cdot) = u_x(b,\cdot) = 0 & \text{στο } [0,T] \\ u(\cdot,0) = u_0 & \text{στο } [a,b] \\ u_t(\cdot,0) = u_1 & \text{στο } [a,b], \end{cases} \tag{11.6}$$

όπου u_0, u_1, p, q και f όπως στην προηγούμενη Άσκηση. Αποδείξτε γι' αυτό το πρόβλημα εκτιμήσεις της μορφής (11.3) και (11.4).

8. Μέθοδοι πεπερασμένων διαφορών για υπερβολικές εξισώσεις

Σ' αυτό το κεφάλαιο θα ασχοληθούμε με μία μέθοδο πεπερασμένων διαφορών για υπερβολικές εξισώσεις, η οποία είναι γνωστή ως μέθοδος των Courant, Friedrichs και Lewy.

Για να απλοποιήσουμε κάπως τους συμβολισμούς θεωρούμε το πρόβλημα (11.1) στην ειδική περίπτωση όπου $p(x,t)=1,\ \ (x,t)\in[a,b]\times[0,T]$, θεωρούμε δηλαδή το πρόβλημα

$$
(12.1)\qquad \begin{cases} u_{tt}=u_{xx}-qu+f & \text{στο } [a,b]\times[0,T]\\ u(a,\cdot)=u(b,\cdot)=0 & \text{στο } [0,T]\\ u(\cdot,0)=u_0 & \text{στο } [a,b]\\ u_t(\cdot,0)=u_1 & \text{στο } [a,b]. \end{cases}
$$

Επιπλέον σ' αυτή την παράγραφο θα υποθέσουμε ότι

$$
(12.2)\qquad q(x,t)\geq 0 \qquad \forall (x,t)\in[a,b]\times[0,T].
$$

Έστω $J,N\in\mathbb{N},\quad h:=\frac{b-a}{J+1},\quad k:=\frac{T}{N}\quad x_i:=a+ih,\quad i=0,\dots,J+1$, και $t^n:=nk,\quad n=0,\dots,N$. Θα προσεγγίσουμε τις τιμές της λύσεως u στα σημεία $(x_i,t^n),\ \ i=0,\dots,J+1,\ \ n=0,\dots,N$. Έστω $u^n\in\mathbb{R}_0^{J+2},\ \ u_i^n:=u(x_i,t^n)$. Θα προσεγγίσουμε τα $u^n,\ \ n=0,\dots,N$, με διανύσματα $U^0,\dots,U^N\in\mathbb{R}_0^{J+2}$. Η μέθοδος την οποία θα μελετήσουμε εδώ είναι διβηματική, για να μπορέσουμε να την εφαρμόσουμε απαιτούνται δύο αρχικές τιμές. Κατ' αρχάς θέτουμε $U^0:=u^0$. Για να ορίσουμε το U^1 παρατηρούμε ότι, για $i=1,\dots,J$,

$$
u(x_i,t^1)=u(x_i,k)=u(x_i,0)+ku_t(x_i,0)+\frac{k^2}{2}u_{tt}(x_i,0)+\frac{k^3}{6}u_{ttt}(x_i,\vartheta_i),
$$

συνεπώς, χρησιμοποιώντας τα δεδομένα και τη Δ.Ε. του (12.1),

$$
\begin{aligned}
u(x_i,t^1) &= u_0(x_i)+ku_1(x_i)+\frac{k^2}{2}[u_0''(x_i)-q(x_i,0)u_0(x_i)+f(x_i,0)] \\
&\quad +\frac{k^3}{6}u_{ttt}(x_i,\vartheta_i), \qquad i=1,\dots,J.
\end{aligned} \tag{12.3}
$$

Ορίζουμε λοιπόν το $U^1\in\mathbb{R}_0^{J+2}$ δια

$$
U_i^1 := u_0(x_i)+ku_1(x_i)+\frac{k^2}{2}[\Delta_h U_i^0-q(x_i,0)U_i^0+f(x_i,0)], \quad i=1,\dots,J. \tag{12.4}
$$

Τα $U^2,\dots,U^N\in\mathbb{R}_0^{J+2}$ είναι τέτοια ώστε για $n=1,\dots,N-1$

$$
\partial^2 U_i^{n-1}=\Delta_h U_i^n-q(x_i,t^n)U_i^n+f(x_i,t^n),\ \ i=1,\dots,J. \tag{12.5}
$$

Εκτός των συνηθισμένων συμβολισμών χρησιμοποιήσαμε εδώ και το ∂^2 που σημαίνει ότι εφαρμόζουμε δύο φορές το ∂. Για $v^0,\dots,v^N\in\mathbb{R}_0^{J+2}$ έχουμε δηλαδή, για $n=1,\dots,N-1$,

$$
\partial^2 v^{n-1}=\partial\partial v^{n-1}=\partial\frac{v^n-v^{n-1}}{k}=\frac{1}{k}\partial v^n-\frac{1}{k}\partial v^{n-1}=\frac{v^{n+1}-2v^n+v^{n-1}}{k^2}.
$$

Στο σχήμα (12.5) οδηγούμαστε προσεγγίζοντας τις παραγώγους στη Δ.Ε. στο σημείο (x_i,t^n) με πεπερασμένες διαφορές: Η $u_{xx}(x_i,t^n)$ προσεγγίζεται κατά τα γνωστά δια $\Delta_h u_i^n$ και η $u_{tt}(x_i,t^n)$ προσεγγίζεται κατά εντελώς ανάλογο τρόπο δια $\partial^2 u_i^{n-1}$.

Η μέθοδος (12.5) αναφέρεται ως *μέθοδος των Courant, Friedrichs και Lewy* ή, για προφανείς λόγους, ως *μέθοδος των πέντε σημείων*. Το σχήμα (12.5) είναι άμεσο, το U_i^{n+1} υπολογίζεται αν γνωρίζουμε τα $U_{i-1}^n,U_i^n,U_{i+1}^n$ και U_i^{n-1}.

Συνέπεια. Έστω ότι η λύση u του προβλήματος (12.1) είναι αρκετά ομαλή, ας πούμε τέσσερεις φορές συνεχώς παραγωγίσιμη ως προς x και ως προς t. Αναπτύσσοντας κατά Taylor ως προς το σημείο (x_i,t^n) διαπιστώνουμε εύκολα ότι για $n=1,\dots,N-1$

$$
\partial^2 u_i^{n-1}=\Delta_h u_i^n-q(x_i,t^n)u_i^n+f(x_i,t^n)+r_i^n,\ \ i=1,\dots,J\,, \tag{12.6}
$$

όπου

$$
\max_{1\le n\le N-1}\max_{1\le i\le J}|r_i^n|\le C(k^2+h^2). \tag{12.7}
$$

Σύγκλιση στη διακριτή L^2 νόρμα. Δεν θα αποδείξουμε ξεχωριστά ευστάθεια του σχήματος των Courant, Friedrichs και Lewy, αλλά θα προχωρήσουμε κατ' ευθείαν στην εκτίμηση του σφάλματος.

Θεώρημα 12.1 *Έστω ότι η λύση u του προβλήματος* (12.1) *είναι τέσσερεις φορές συνεχώς παραγωγίσιμη ως προς x και ως προς t. Τότε για k αρκετά μικρό και $\frac{k}{h} \le r < 1$ ισχύει*

$$\max_{0\le n\le N} \|u^n - U^n\|_h \le C(k^2 + h^2) \tag{12.8}$$

με μια σταθερά C ανεξάρτητη των k και h.

Απόδειξη. Έστω $e^n \in \mathbb{R}_0^{J+2}$, $\;e^n := u^n - U^n$, $\;n = 0, \dots, N$. Αφαιρώντας κατά μέλη τις (12.6) και (12.5) λαμβάνουμε για $n = 1, \dots, N-1$

$$\partial^2 e_i^{n-1} = \Delta_h e_i^n - q(x_i, t^n) e_i^n + r_i^n, \quad i = 1, \dots, J. \tag{12.9}$$

Πολλαπλασιάζουμε αυτή τη σχέση επί $h(e_i^{n+1} - e_i^{n-1}) = hk(\partial e_i^n + \partial e_i^{n-1})$ και αθροίζουμε από $i = 1$ έως $i = J$, οπότε παίρνουμε

$$\begin{aligned}\|\partial e^n\|_h^2 - \|\partial e^{n-1}\|_h^2 = (\Delta_h e^n, e^{n+1} - e^{n-1})_h - h \sum_{i=1}^J q(x_i, t^n) e_i^n (e_i^{n+1} - e_i^{n-1})\\ + (r^n, e^{n+1} - e^{n-1})_h.\end{aligned}$$

Χρησιμοποιώντας εδώ την (9.26) έχουμε

$$\begin{aligned}\|\partial e^n\|_h^2 - \|\partial e^{n-1}\|_h^2 = (\Delta_h e^{n+1}, e^n)_h - (\Delta_h e^n, e^{n-1})_h - h \sum_{i=1}^J q(x_i, t^n) e_i^{n+1} e_i^n\\ + h \sum_{i=1}^J q(x_i, t^n) e_i^n e_i^{n-1} + (r^n, e^{n+1} - e^{n-1})_h,\end{aligned}$$

οπότε αθροίζοντας από $n = 1$ έως $n = \ell \le N-1$ λαμβάνουμε

$$\begin{aligned}\|\partial e^\ell\|_h^2 - \|\partial e^0\|_h^2 = (\Delta_h e^{\ell+1}, e^\ell)_h - \sum_{n=1}^\ell h \sum_{i=1}^J q(x_i, t^n) e_i^{n+1} e_i^n\\ + \sum_{n=1}^\ell h \sum_{i=1}^J q(x_i, t^n) e_i^n e_i^{n-1} + k \sum_{n=1}^\ell (r^n, \partial e^n)_h + k \sum_{n=1}^\ell (r^n, \partial e^{n-1})_h.\end{aligned} \tag{12.10}$$

Τώρα

$$
\begin{aligned}
&-\sum_{n=1}^{\ell} q(x_i,t^n)e_i^{n+1}e_i^n + \sum_{n=1}^{\ell} q(x_i,t^n)e_i^n e_i^{n-1} = \\
&= -\sum_{n=1}^{\ell} q(x_i,t^n)e_i^{n+1}e_i^n + \sum_{n=0}^{\ell-1} q(x_i,t^{n+1})e_i^{n+1}e_i^n \\
&= -\sum_{n=1}^{\ell} q(x_i,t^n)e_i^{n+1}e_i^n + \sum_{n=1}^{\ell-1} q(x_i,t^{n+1})e_i^{n+1}e_i^n \\
&= -q(x_i,t^\ell)e_i^{\ell+1}e_i^\ell + \sum_{n=1}^{\ell-1}[q(x_i,t^{n+1}) - q(x_i,t^n)]e_i^{n+1}e_i^n \\
&= -q(x_i,t^\ell)e_i^{\ell+1}e_i^\ell + k\sum_{n=1}^{\ell-1} q_t(x_i,\vartheta_{n,i})e_i^{n+1}e_i^n
\end{aligned}
$$

και έτσι παίρνουμε εύκολα από την (12.10) την εκτίμηση

$$
\begin{aligned}
\|\partial e^\ell\|_h^2 - \|\partial e^0\|_h^2 &\le (\Delta_h e^{\ell+1}, e^\ell)_h - h\sum_{i=1}^{J} q(x_i,t^\ell)e_i^{\ell+1}e_i^\ell \\
&\quad + Ck\sum_{n=1}^{\ell-1} \|e^{n+1}\|_h\,\|e^n\|_h + k\sum_{n=1}^{\ell}(r^n,\partial e^n)_h + k\sum_{n=1}^{\ell}(r^n,\partial e^{n-1})_h.
\end{aligned} \tag{12.11}
$$

Σε χώρους με εσωτερικό γινόμενο ισχύουν τώρα γενικά, όπως διαπιστώνει κανείς εύκολα, οι ταυτότητες

$$(x,y) = \frac{1}{4}\big\{(x+y,x+y) - (x-y,x-y)\big\} \tag{12.12i}$$

$$2(x,x) + 2(y,y) = (x+y,x+y) + (x-y,x-y), \tag{12.12ii}$$

η δεύτερη των οποίων αναφέρεται ως *ισότητα του παραλληλογράμμου*. Χρησιμοποιώντας τις (12.12) και (6.8′), παίρνουμε από την (12.11)

$$
\begin{aligned}
\|\partial e^\ell\|_h^2 - \|\partial e^0\|_h^2 + \frac{1}{4}|e^{\ell+1}+e^\ell|_{1,h}^2 + \frac{h}{4}\sum_{i=1}^{J} q(x_i,t^\ell)(e_i^{\ell+1}+e_i^\ell)^2 \le \\
\le \frac{1}{4}|e^{\ell+1}-e^\ell|_{1,h}^2 + \frac{h}{4}\sum_{i=1}^{J} q(x_i,t^\ell)(e_i^{\ell+1}-e_i^\ell)^2 \\
+ Ck\sum_{n=1}^{\ell-1}\left\{\|e^{n+1}+e^n\|_h^2 + \|e^{n+1}-e^n\|_h^2\right\} + k\sum_{n=1}^{\ell}(r^n,\partial e^n)_h \\
+ k\sum_{n=1}^{\ell}(r^n,\partial e^{n-1})_h.
\end{aligned}
\tag{12.13}
$$

Λαμβάνοντας υπ' όψιν την (12.2) έχουμε συνεπώς

$$
\begin{aligned}
\|\partial e^\ell\|_h^2 - \frac{k^2}{4}|\partial e^\ell|_{1,h}^2 - Ck^2\|\partial e^\ell\|_h^2 + |e^{\ell+\frac{1}{2}}|_{1,h}^2 \le \|\partial e^0\|_h^2 + Ck\sum_{n=1}^{\ell-1}\|e^{n+\frac{1}{2}}\|_h^2 \\
+ Ck^3\sum_{n=1}^{\ell-1}\|\partial e^n\|_h^2 + k\sum_{n=1}^{\ell}(r^n,\partial e^n)_h + k\sum_{n=1}^{\ell}(r^n,\partial e^{n-1})_h.
\end{aligned}
\tag{12.14}
$$

Χρησιμοποιώντας τώρα την ανισότητα

$$
|v|_{1,h}^2 \le \frac{4}{h^2}\|v\|_h^2 \qquad \forall v \in \mathbb{R}_0^{J+2},
$$

η οποία αποδείχθηκε μεταξύ των σχέσεων (9.12) και (9.13), και τη συνθήκη $\frac{k}{h} \le r < 1$, βλέπουμε ότι για αρκετά μικρό k ισχύει

$$
\begin{aligned}
\|\partial e^\ell\|_h^2 + |e^{\ell+\frac{1}{2}}|_{1,h}^2 \le Ck\sum_{n=1}^{\ell-1}\|e^{n+\frac{1}{2}}\|_h^2 + Ck^3\sum_{n=1}^{\ell-1}\|\partial e^n\|_h^2 + C\|\partial e^0\|_h^2 \\
+ ck\sum_{n=1}^{\ell}(r^n,\partial e^n)_h + ck\sum_{n=1}^{\ell}(r^n,\partial e^{n-1})_h.
\end{aligned}
$$

Χρησιμοποιώντας λοιπόν τις (6.6) και (12.7) λαμβάνουμε εύκολα

$$
\begin{aligned}
\|\partial e^\ell\|_h^2 + |e^{\ell+\frac{1}{2}}|_{1,h}^2 \le C\|\partial e^0\|_h^2 + Ck\sum_{n=1}^{\ell-1}(\|\partial e^n\|_h^2 + |e^{n+\frac{1}{2}}|_{1,h}^2) \\
+ C(k^2+h^2)^2, \quad \ell = 1,\dots,N-1.
\end{aligned}
\tag{12.15}
$$

Χρησιμοποιώντας το διακριτό Λήμμα του Gronwall (βλ. την Άσκηση 12.1) συμπεραίνουμε από εδώ εύκολα ότι

$$\max_{1\le\ell\le N-1}\{\|\partial e^\ell\|_h^2+|e^{\ell+\frac12}|_{1,h}^2\}\le C\|\partial e^0\|_h^2+C(k^2+h^2)^2, \tag{12.15$'$}$$

οπότε έχουμε

$$\max_{1\le\ell\le N-1}\|\partial e^\ell\|_h\le C\|\partial e^0\|_h+C(k^2+h^2). \tag{12.16}$$

Από τις (12.3) και (12.4) έπεται εξ άλλου

$$\max_i|e_i^1|\le k^2(h^2+k),$$

οπότε, επειδή $e^0=0$, έχουμε

$$\|\partial e^0\|_h\le C(k^2+h^2). \tag{12.17}$$

Οι (12.16) και (12.17) δίνουν

$$\max_{0\le n\le N-1}\|\partial e^n\|_h\le C(k^2+h^2) \tag{12.18}$$

ή

$$\|e^{n+1}\|_h\le\|e^n\|_h+Ck(k^2+h^2)\qquad n=0,\dots,N-1,$$

από την οποία έπεται αμέσως η (12.8). □

Παρατήρηση 12.1. Από την (12.15′) έπεται εύκολα ότι

$$\max_{1\le\ell\le N-1}|e^{\ell+\frac12}|_{1,h}\le C(k^2+h^2).$$

Εύκολα διαπιστώνουμε επίσης ότι το $|e^{\frac12}|_{1,h}$ φράσσεται με τον ίδιο τρόπο, συνεπώς έχουμε

$$\max_{0\le n\le N-1}|e^{n+\frac12}|_{1,h}\le C(k^2+h^2). \tag{12.19}$$

Κατά τα γνωστά από την (12.19) έπεται

$$\max_{0\le n\le N-1}\ \max_{0\le i\le J+1}|e_i^{n+\frac12}|\le C(k^2+h^2), \tag{12.20}$$

οπότε λαμβάνουμε

$$\max_{0\leq n\leq N-1}\max_{0\leq i\leq J+1}|\frac{U_i^n+U_i^{n+1}}{2}-u(x_i,t^{n+\frac{1}{2}})|\leq C(k^2+h^2). \tag{12.21}$$ □

Παρατήρηση 12.2. Το σχήμα (12.5) ανήκει στην ακόλουθη οικογένεια διβηματικών μεθόδων

$$\begin{aligned}\partial^2U_i^{n-1}&=\Delta_h[\vartheta U^{n+1}+(1-2\vartheta)U^n+\vartheta U^{n-1}]_i\\&-q(x_i,t^n)[\vartheta U^{n+1}+(1-2\vartheta)U^n+\vartheta U^{n-1}]_i+f(x_i,t^n),\quad i=1,\dots,J,\end{aligned}\tag{12.22}$$

όπου $\vartheta\geq 0$ είναι μια σταθερά παράμετρος. Οι μέθοδοι (12.22) είναι γνωστές ως *μέθοδοι του von Neumann*. Για $\vartheta=0$ προκύπτει η μέθοδος πέντε σημείων των Courant, Friedrichs και Lewy (12.5), για $\vartheta=\frac{1}{12}$ προκύπτει η *μέθοδος των Størmer–Numerov*. Για $\vartheta=0$ παίρνουμε όπως είδαμε μια άμεση μέθοδο. Για $\vartheta>0$ προκύπτουν πεπλεγμένες μέθοδοι, ο υπολογισμός του U^{n+1}, όταν γνωρίζουμε τα U^n και U^{n-1}, απαιτεί την επίλυση ενός τριδιαγώνιου γραμμικού συστήματος. Για $\vartheta\geq\frac{1}{4}$ προκύπτουν απεριόριστα ευσταθείς μέθοδοι, για $0\leq\vartheta<\frac{1}{4}$ οι μέθοδοι είναι ευσταθείς υπό συνθήκες της μορφής $\frac{k}{h}<\varphi(\vartheta)<\infty$. Όλες οι μέθοδοι είναι δεύτερης τάξης ως προς h, για $\vartheta\neq\frac{1}{12}$ επίσης δεύτερης τάξης ως προς k, ενώ για $\vartheta=\frac{1}{12}$ προκύπτει μια μέθοδος τετάρτης τάξεως ως προς k. Από τις μεθόδους της οικογένειας (12.22) στην πράξη χρησιμοποιείται συχνότερα η μέθοδος των Courant, Friedrichs και Lewy. □

Ασκήσεις

12.1. (*Διακριτή ανισότητα του Gronwall*) Έστω $T>0$, $N\in\mathbb{N}$ και $k:=\frac{T}{N}$. Αν $\alpha_0,\dots,\alpha_N$ και E μη αρνητικοί αριθμοί τέτοιοι ώστε με $\gamma>0$ να ισχύει

$$\alpha_{n+1}\leq E+\gamma k\sum_{\ell=0}^{n}\alpha_\ell\qquad n=0,\dots,N-1,$$

αποδείξτε ότι

$$\max_{1\leq n\leq N}\alpha_n\leq C(k\alpha_0+E)\tag{12.23}$$

με μια σταθερά C ανεξάρτητη του k. [*Υπόδειξη*: Έστω $\varphi_n:=\gamma k\sum_{\ell=0}^n\alpha_\ell$. Τότε $\varphi_{n+1}-\varphi_n=\gamma k\alpha_{n+1}$, συνεπώς

$$\varphi_{n+1}\leq\gamma kE+(1+\gamma k)\varphi_n,\quad n=0,\dots,N-1.$$

Χρησιμοποιήστε την Άσκηση 9.2 για να εκτιμήσετε τα φ_n, και την

$$\alpha_{n+1} \leq E + \varphi_n$$

για να εκτιμήσετε τα α_n.]

12.2. Δώστε το αντίστοιχο του σχήματος (12.5) (και της (12.4)) για την περίπτωση του προβλήματος (11.5) και αποδείξτε μια εκτίμηση της μορφής (12.8).

12.3. Δώστε το αντίστοιχο του σχήματος (12.5) (και της (12.4)) για την περίπτωση του προβλήματος (11.6) και αποδείξτε μια εκτίμηση της μορφής (12.8).

Βιβλιογραφία

1. R. A. Adams, J. J. F. Fournier: *Sobolev Spaces*. 2nd edition, Academic Press, New York, 2003.
2. Γ. Δ. Ακρίβης, Β. Α. Δουγαλής: *Εισαγωγή στην Αριθμητική Ανάλυση*. Πανεπιστημιακές Εκδόσεις Κρήτης, Ηράκλειο, 1997. (β′, αναθεωρημένη έκδοση, 2004, α′ ανατύπωση, 2005.)
3. Γ. Δ. Ακρίβης, Β. Α. Δουγαλής: *Αριθμητικές Μέθοδοι για Συνήθεις Διαφορικές Εξισώσεις*. Πανεπιστημιακές Εκδόσεις Κρήτης, Ηράκλειο, 2005.
4. Ν. Δ. Αλικάκος, Γ. Η. Καλογερόπουλος: *Συνήθεις Διαφορικές Εξισώσεις*. Σύγχρονη Εκδοτική, Αθήνα, 2003.
5. S. C. Brenner, L. R. Scott: *The Mathematical Theory of Finite Element Methods*. Springer–Verlag, New York, 1994. (2nd ed., 2002.)
6. J. Butcher: *Implicit Runge–Kutta processes*. Mathematics of Computation v. **18** (1964), pp. 50–64.
7. J. Butcher: *The Numerical Analysis of Ordinary Differential Equations – Runge–Kutta and General Linear Methods*. Wiley, Chichester, 1987.
8. W. Cheney: *Analysis for Applied Mathematics*. Springer–Verlag, New York, 2001.
9. P. G. Ciarlet: *The Finite Element Method for Elliptic Problems*. North Holland, Amsterdam, 1978. (Ανατύπωση SIAM, Classics in Applied Mathematics v. 40, Philadelphia, 2002.)
10. P. G. Ciarlet, J. L. Lions (editors): *Handbook of Numerical Analysis, vol. I, Finite Difference Methods* (*Part 1*), *Solution of Equations in* $\mathbb{R}^n$ (*Part 1*). North Holland, Amsterdam, 1990.

11. P. G. Ciarlet, J. L. Lions (editors): *Handbook of Numerical Analysis, vol. II, Finite Element Methods* (*Part 1*). North Holland, Amsterdam, 1991.

12. P. G. Ciarlet, J. L. Lions (editors): *Handbook of Numerical Analysis, vol. III, Techniques of Scientific Computing* (*Part 1*), *Numerical Methods for Solids* (*Part 1*), *Solution of Equations in* $\mathbb{R}^n$ (*Part 2*). North Holland, Amsterdam, 1994.

13. P. G. Ciarlet, J. L. Lions (editors): *Handbook of Numerical Analysis, vol. IV, Finite Element Methods* (*Part 2*), *Numerical Methods for Solids* (*Part 2*). North Holland, Amsterdam, 1995.

14. P. G. Ciarlet, J. L. Lions (editors): *Handbook of Numerical Analysis, vol. V, Techniques of Scientific Computing* (*Part 2*). North Holland, Amsterdam, 1997.

15. P. G. Ciarlet, J. L. Lions (editors): *Handbook of Numerical Analysis, vol. VII, Solution of Equations in* $\mathbb{R}^n$ (*Part 3*), *Techniques of Scientific Computing* (*Part 3*). North Holland, Amsterdam, 2000.

16. P. G. Ciarlet, J. L. Lions (editors): *Handbook of Numerical Analysis, vol. VIII, Solution of Equations in* $\mathbb{R}^n$ (*Part 4*), *Techniques of Scientific Computing* (*Part 4*), *Numerical Methods for Fluids* (*Part 2*). North Holland, Amsterdam, 2001.

17. M. Crouzeix: *Sur l'approximation des équations différentielles operationelles linéaires par des méthodes de Runge–Kutta*. Thèse d'état, Univ. Paris VI, 1975.

18. M. Crouzeix, A. L. Mignot: *Analyse Numérique des Équations Différentielles*. 2η έκδοση. Masson, Paris, 1989.

19. G. Dahlquist: *Convergence and stability in the numerical integration of ordinary differential equations*. Math. Scand. v. **4** (1956), pp. 33–53.

20. K. Dekker, J. G. Verwer: *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations*. North Holland, Amsterdam, 1984.

21. K. Eriksson, D. Estep, C. Johnson: *Applied Mathematics: Body and Soul*. Vol. I: Derivatives and Geometry in $\mathbb{R}^n$. Vol. II: Integrals and Geometry in $\mathbb{R}^n$. Vol. III: Calculus in Several Dimensions. Springer–Verlag, Berlin, 2004.

22. L. C. Evans: *Partial Differential Equations*. Graduate Studies in Mathemetics, vol. 19, AMS, Providence, Rhode Island, 1998.

23. G. E. Forsythe, M. A. Malcolm, C. B. Moler: *Computer Methods for Mathematical Computations*. Prentice–Hall, Englewood Cliffs, N.J., 1977. (Ελληνική μετάφραση με τίτλο *Αριθμητικές μέθοδοι και προγράμματα για μαθηματικούς υπολογισμούς*. γ′ έκδοση, Πανεπιστημιακές Εκδόσεις Κρήτης, Ηράκλειο, 1998.)
24. H. Fujita, N. Saito, T. Suzuki: *Operator Theory and Numerical Methods*. Elsevier, Amsterdam, 2001.
25. C. W. Gear: *Numerical Initial Value Problems in Ordinary Differential Equations*. Prentice–Hall, Englewood Cliffs, N.J., 1971.
26. R. D. Grigorieff: *Numerik gewöhnlicher Differentialgleichungen*. Bd. I, II, B. G. Teubner, Stuttgart, Bd. I: 1972, Bd. II: 1977.
27. E. Hairer, Ch. Lubich, G. Wanner: *Geometric Numerical Integration: Structure–Preserving Algorithms for Ordinary Differential Equations*. Springer–Verlag, Berlin, 2002.
28. E. Hairer, S. P. Nørsett, G. Wanner: *Solving Ordinary Differential Equations I – Nonstiff Problems*. Springer–Verlag, Berlin, 2nd revised ed., 1993, corr. 2nd printing, 2000.
29. E. Hairer, G. Wanner: *Solving Ordinary Differential Equations II – Stiff and Differential–Algebraic Problems*. Springer–Verlag, Berlin, 2nd revised ed., 2002.
30. P. Henrici: *Discrete Variable Methods in Ordinary Differential Equations*. Wiley, New York, 1962.
31. W. Hundsdorfer, J. G. Verwer: *Numerical Solution of Time–Dependent Advection–Diffusion–Reaction Equations*. Springer–Verlag, Berlin, 2003.
32. A. Iserles: *A First Course in the Numerical Analysis of Differential Equations*. Cambridge U. P., Cambridge, 1996.
33. C. Johnson: *Numerical Solutions of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, Cambridge, 1987.
34. H. B. Keller: *Numerical Methods for Two–Point Boundary Value Problems*. Blaisdell, Waltham, Mass., 1968. (Ανατύπωση Dover, New York, 1992.)

35. J. D. Lambert: *Numerical Methods for Ordinary Differential Systems*. Wiley, Chichester, 1991.

36. S. Larsson, V. Thomée: *Partial Differential Equations with Numerical Methods*. Springer–Verlag, Berlin, 2003.

37. C. Moler: *Numerical Computing with MATLAB*. SIAM, Philadelphia, 2004.

38. R. D. Richtmyer, K. W. Morton: *Difference Methods for Initial–Value Problems*. 2nd ed., Wiley, New York, 1967.

39. W. Rudin: *Principles of Mathematical Analysis*. 3rd ed., McGraw–Hill, Singapore, 1976. (Ελληνική μετάφραση με τίτλο *Αρχές Μαθηματικής Αναλύσεως*, Εκδόσεις Leader Books, Αθήνα, 2000.)

40. A. A. Samarskij: *Theorie der Differenzenverfahren*. Akademische Verlagsgesellschaft Geest und Portig, Leipzig, 1984.

41. L. L. Schumaker: *Spline Functions: Basic Theory*. Wiley, New York, 1981.

42. J. C. Strikwerda: *Finite Difference Schemes and Partial Differential Equations*. Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989.

43. V. Thomée: *Galerkin Finite Element Methods for Parabolic Problems*. Lecture Notes in Mathematics, v. 1054, Springer–Verlag, Berlin, 1984.

44. V. Thomée: *Galerkin Finite Element Methods for Parabolic Problems*. Springer–Verlag, Berlin, 1997.

45. V. Thomée: *Finite difference methods for linear parabolic equations*. Στο: P. G. Ciarlet, J. L. Lions, εκδότες, Handbook of Numerical Analysis, Vol. I, pp. 5–196. North–Holland, Amsterdam, 1990.

Ευρετήριο