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

ΜΕΘΟΔΟΙ ΠΕΠΕΡΑΣΜΕΝΩΝ ΣΤΟΙΧΕΙΩΝ

(πανεπιστημιακές παραδόσεις)

ΙΩΑΝΝΙΝΑ, 2007

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 μια εισαγωγή στο θέμα της επίλυσης διαφορικών εξισώσεων με μερικές παραγώγους με τη μέθοδο των πεπερασμένων στοιχείων.

Η μέθοδος των πεπερασμένων στοιχείων είναι η πλέον δημοφιλής για την επίλυση ελλειπτικών και παραβολικών εξισώσεων, τόσο μεταξύ των μηχανικών που την εφαρμόζουν για να επιλύσουν πραγματικά προβλήματα όσο και μεταξύ των αριθμητικοαναλυστών που ασχολούνται με αριθμητικές μεθόδους για διαφορικές εξισώσεις με μερικές παραγώγους. Οι βασικοί λόγοι γι' αυτό είναι ότι αυτή η μέθοδος επιτρέπει μεγάλη ευελιξία, τόσο στη γεωμετρία των χωρίων, όσο και στις εξισώσεις και τα είδη των συνοριακών συνθηκών, και επίσης το γεγονός ότι η μελέτη της είναι συστηματική και γενική. Η μέθοδος των πεπερασμένων στοιχείων αποτελεί ένα από τα καταλληλότερα υπολογιστικά εργαλεία για την επίλυση διαφορικών εξισώσεων με μερικές παραγώγους, που εμφανίζονται στις εφαρμογές.

Στόχος των ανά χείρας σημειώσεων είναι η διατύπωση της μεθόδου καθώς και η μελέτη των σημαντικότερων ιδιοτήτων της, όπως η ευστάθεια και η σύγκλιση. Ένα φυσιολογικό πλαίσιο στο οποίο διατυπώνεται και μελετάται η μέθοδος των πεπερασμένων στοιχείων αποτελούν οι χώροι του Sobolev· ένα μέρος των σημειώσεων αφιερώνεται ως εκ τούτου σε μια συνοπτική εισαγωγή στη θεωρία των χώρων του Sobolev.

Το κυριώτερο μέρος των σημειώσεων αναφέρεται στην επίλυση προβλημάτων συνοριακών τιμών για ελλειπτικές εξισώσεις. Επίσης θα ασχοληθούμε συνοπτικά και με την επίλυση προβλημάτων αρχικών και συνοριακών συνθηκών για παραβολικές εξισώσεις.

Οι σημειώσεις αποτελούνται από έξι κεφάλαια. Στο πρώτο παραθέτουμε κάποια θεμελιώδη αποτελέσματα σε αφηρημένη μορφή. Στο δεύτερο δίνουμε μια εισαγωγή στη θεωρία των χώρων του Sobolev σε μία διάσταση. Το τρίτο αναφέρεται σε μεθόδους πεπερασμένων στοιχείων για το πρόβλημα δύο σημείων. Οι χώροι του Sobolev σε πολλές διαστάσεις καθώς και η μέθοδος των πεπερασμένων στοιχείων για προβλήματα συνοριακών τιμών για ελλειπτικές εξισώσεις σε πολλές διαστάσεις θα μας απασχολήσουν στο τέταρτο κεφάλαιο. Το πέμπτο κεφάλαιο αφορά προβλήματα αρχικών και συνοριακών τιμών για παραβολικές εξισώσεις. Τέλος το έκτο κεφάλαιο αναφέρεται στη διακριτοποίηση στον χρόνο προβλημάτων αρχικών τιμών για παραβολικές εξισώσεις σε αφηρημένη μορφή.

Οι σημειώσεις γράφτηκαν για τα αντίστοιχα μαθήματα για προπτυχιακούς και μεταπτυχιακούς φοιτητές του Τμήματος Μαθηματικών και Στατιστικής του Πανεπιστημίου Κύπρου, και ακολουθούν στο τρίτο και στο πέμπτο κεφάλαιο τις σημειώσεις [4], βλέπε τη βιβλιογραφία.

Σεπτέμβριος 2005

Γ. Δ. Ακρίβης

Περιεχόμενα

1. Βασικά αποτελέσματα σε αφηρημένη μορφή

Σε αυτό το κεφάλαιο παρουσιάζουμε ορισμένες βασικές ιδέες της θεωρίας των πεπερασμένων στοιχείων για ελλειπτικές εξισώσεις. Για να αποφύγουμε τεχνικές λεπτομέρειες, αλλά και για να αφήσουμε να αναδειχθούν καλύτερα οι ουσιαστικές υποθέσεις, δίνουμε τα αποτελέσματα σε αφηρημένη μορφή. Σε επόμενα κεφάλαια θα αναφερθούμε σε συγκεκριμένες εξισώσεις και θα εφαρμόσουμε και θα εμπλουτίσουμε τη θεωρία.

Έστω $\big(H, (\cdot,\cdot)\big)$ ένας πραγματικός χώρος Hilbert. Συμβολίζουμε με $\|\cdot\|$ την παραγόμενη από το εσωτερικό γινόμενο $(\cdot,\cdot)$ νόρμα. Θεωρούμε στον H μια *διγραμμική μορφή* a, $a(\cdot,\cdot) : H \times H \to \mathbb{R}$, δηλαδή μια μορφή που είναι γραμμική ως προς καθένα από τα δύο στοιχεία της ξεχωριστά. Υποθέτουμε ότι η διγραμμική μορφή a είναι *φραγμένη* στον H, δηλαδή ότι

$$\exists C \in \mathbb{R}\ \ \forall v, w \in H \quad |a(v,w)| \le C\|v\|\,\|w\|, \tag{1.1}$$

και *ελλειπτική* (coersiv) σε έναν κλειστό υπόχωρο V του H, δηλαδή ότι

$$\exists \alpha > 0\ \ \forall v \in V \quad a(v,v) \ge \alpha\|v\|^2. \tag{1.2}$$

Έστω ακόμα $F : V \to \mathbb{R}$ ένα φραγμένο γραμμικό συναρτησιακό. Θεωρούμε τότε το ακόλουθο *μεταβολικό* πρόβλημα: Ζητείται $u \in V$ τέτοιο ώστε

$$a(u,v) = F(v) \quad \forall v \in V. \tag{1.3}$$

Ουσιαστικά σε αυτό το κεφάλαιο θα ασχοληθούμε με τα θέματα της ύπαρξης και μοναδικότητας της λύσης u καθώς και με την προσέγγισή της. Εύκολα θα αποδείξουμε ύπαρξη και μοναδικότητα της u. Όταν όμως ο V είναι απειροδιάστατος, και αυτή είναι η ενδιαφέρουσα περίπτωση, τότε δεν μπορούμε να υπολογίσουμε τη u, οπότε αναγκαστικά καταφεύγουμε σε μεθόδους προσέγγισής της. Σημαντικό ρόλο εδώ παίζουν θέματα υπολογισμού των προσεγγίσεων και μελέτη ιδιοτήτων των προσεγγίσεων.

Αυτό το κεφάλαιο αποτελείται από τρεις ενότητες. Στην ενότητα 1.1, επί πλέον της συνέχειας και της ελλειπτικότητας, θα υποθέσουμε ότι η διγραμμική μορφή $a(\cdot,\cdot)$ είναι *συμμετρική*, δηλαδή ότι $a(v,w) = a(w,v)$, για κάθε $v,w \in H$, ενώ στην ενότητα 1.2 δεν θα υποθέσουμε συμμετρία. Στην ενότητα 1.3 εξετάζουμε γενικότερες περιπτώσεις όσον αφορά την προσεγγιστική λύση· τέτοιες καταστάσεις εμφανίζονται συχνά στις εφαρμογές. Συγκεκριμένα, επιτρέπουμε η προσεγγιστική λύση είτε να ορίζεται μέσω μιας προσεγγιστικής διγραμμικής μορφής και ενός προσεγγιστικού γραμμικού συναρτησιακού, είτε να ανήκει σε έναν χώρο που δεν είναι αναγκαστικά υπόχωρος του V.

1.1 Η συμμετρική περίπτωση

Σε αυτή την ενότητα υποθέτουμε ότι η διγραμμική μορφή $a(\cdot,\cdot)$ είναι *συμμετρική*, δηλαδή ότι $a(v,w) = a(w,v)$, για κάθε $v,w \in H$. Η μελέτη της συμμετρικής και της μη συμμετρικής περίπτωσης παρουσιάζεται ξεχωριστά για δύο λόγους: αφ' ενός γιατί η μελέτη στη συμμετρική περίπτωση είναι απλούστερη και αφ' ετέρου γιατί ορισμένα αποτελέσματα ισχύουν μόνο για συμμετρικές διγραμμικές μορφές.

Στη μέθοδο των πεπερασμένων στοιχείων προσεγγίζουμε τη λύση ενός προβλήματος με στοιχεία από κατάλληλους χώρους, κατά κανόνα υποχώρους πεπερασμένης διάστασης ενός χώρου Hilbert, στον οποίο ανήκει η ακριβής λύση.

Πριν προχωρήσουμε στο βασικό αντικείμενό μας σε αυτή την ενότητα, θα παραθέσουμε δύο χρήσιμα για τη συνέχεια αποτελέσματα από τη Συναρτησιακή Ανάλυση. Το πρώτο αφορά τις βέλτιστες προσεγγίσεις από κλειστούς υποχώρους χώρων Hilbert, και το δεύτερο είναι το θεώρημα αναπαράστασης του Riesz.

Πολλές φορές στις εφαρμογές προσεγγίζουμε στοιχεία ενός χώρου με στοιχεία υποσυνόλων του χώρου. Η περίπτωση που τα υποσύνολα είναι υπόχωροι του αρχικού χώρου είναι πολύ απλούστερη και εμφανίζεται κατά κόρον στις εφαρμογές, και γιαυτό θα περιοριστούμε σε αυτήν. Υπάρχουν γενικά πολλές "προσεγγίσεις", συνήθως μας ενδιαφέρει κάποια "βέλτιστη" υπό κάποια έννοια.

Ορισμός 1.1 (Βέλτιστη προσέγγιση.) Έστω $\big(X, \|\cdot\|\big)$ ένας χώρος με νόρμα, Y ένας υπόχωρός του, και $x \in X$. Ένα στοιχείο $y \in Y$ (αν υπάρχει, φυσικά), για το οποίο ισχύει

$$\|x-y\| = \inf_{z\in Y} \|x-z\| \tag{1.4}$$

λέγεται *βέλτιστη προσέγγιση* του x από τον Y.

Είναι πολύ εύκολο να πεισθεί κανείς ότι απαραίτητη προϋπόθεση για να υπάρχει μια βέλτιστη προσέγγιση, για κάθε $x \in X$, είναι ο υπόχωρος Y, από τον οποίο προσεγγίζουμε, να είναι κλειστός. Σε χώρους Hilbert αυτή η συνθήκη είναι και ικανή, όπως θα δούμε αμέσως στο ακόλουθο γενικότερο αποτέλεσμα, στο οποίο δίνεται και ένας χαρακτηρισμός της βέλτιστης προσέγγισης.

Πρόταση 1.1 (Βέλτιστης προσέγγισης.) *Έστω $\big(H, (\cdot,\cdot)\big)$ ένας χώρος Hilbert και $\tilde H$ ένας κλειστός υπόχωρός του. Τότε, για κάθε $x \in H$, υπάρχει ακριβώς μία βέλτιστη προσέγγιση $y \in \tilde H$ του x από τον $\tilde H$. Το y είναι το μόνο στοιχείο του $\tilde H$, για το οποίο η διαφορά $x-y$ είναι κάθετη στον $\tilde H$, δηλαδή*

$$\forall z \in \tilde H \quad (x-y, z) = 0. \tag{1.5}$$

Απόδειξη. Έστω $\delta := \inf_{z\in\tilde H} \|x-z\|$. Τότε υπάρχει μια ακολουθία $(y_n)_{n\in\mathbb{N}} \subset \tilde H$, τέτοια ώστε, με $\delta_n := \|x-y_n\|$, να ισχύει $\delta_n \to \delta, n \to \infty$. Θα αποδείξουμε πρώτα ότι η $(y_n)_{n\in\mathbb{N}}$ είναι ακολουθία Cauchy. Πράγματι, με $v_n := x - y_n$ έχουμε $\|v_n\| = \delta_n$ και

$$\|v_n + v_m\| = \|2x - y_n - y_m\| = 2\|x - \frac{1}{2}(y_n + y_m)\| \ge 2\delta,$$

οπότε, με τη βοήθεια και της ισότητας του παραλληλογράμμου, βλ. την Άσκηση 1.1,

$$\|y_n - y_m\|^2 = \|v_n - v_m\|^2 = -\|v_n + v_m\|^2 + 2\big(\|v_n\|^2 + \|v_m\|^2\big) \le -(2\delta)^2 + 2(\delta_n^2 + \delta_m^2).$$

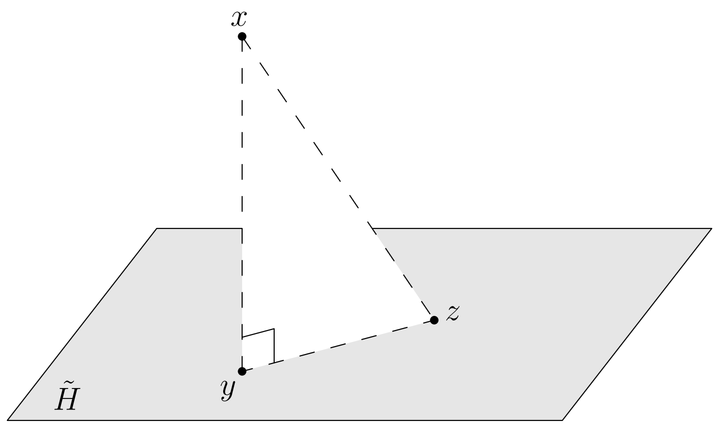
Επομένως, η $(y_n)_{n\in\mathbb{N}}$ είναι ακολουθία Cauchy και, συνεπώς, συγκλίνουσα, $y_n \to y, n \to \infty$. Λόγω της κλειστότητας του $\tilde H$ ισχύει φυσικά $y \in \tilde H$. Προφανώς $\|x-y\| \ge \delta$. Επί πλέον έχουμε

$$\|x-y\| \le \|x-y_n\| + \|y_n - y\| = \delta_n + \|y_n - y\| \to \delta, \;\; n\to\infty,$$

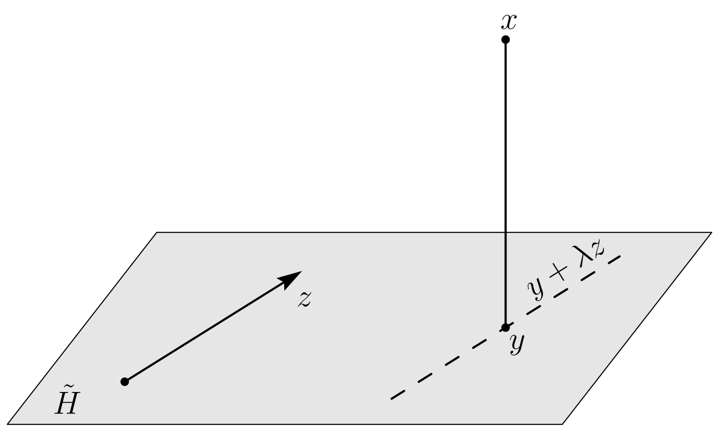
οπότε $\|x-y\| = \delta$.

Θα αποδείξουμε τώρα ότι ισχύει η (1.5) με απαγωγή σε άτοπο. Έστω λοιπόν, κατ' αρχάς, ότι έχουμε ένα $y \in \tilde H$, το οποίο ικανοποιεί την (1.5). Τότε, για τυχαίο $z \in \tilde H$, θα έχουμε, σύμφωνα με το Πυθαγόρειο θεώρημα, βλ. την Άσκηση 1.1,

$$\|x-z\|^2 = \|(x-y)+(y-z)\|^2 = \|x-y\|^2 + \|y-z\|^2,$$



Σχήμα 1.1: (1.5) $\implies$ y είναι βέλτιστη προσέγγιση



Σχήμα 1.2: y είναι βέλτιστη προσέγγιση $\implies$ (1.5)

συνεπώς το y είναι βέλτιστη προσέγγιση, και μάλιστα μοναδική, του x από τον $\tilde H$, βλ. το Σχήμα 1.1.

Έστω τώρα, αντίστροφα, ότι $y \in \tilde H$ είναι η βέλτιστη προσέγγιση του x και ότι υπάρχει $z \in \tilde H$ τέτοιο ώστε $(x-y,z) \neq 0$. Τότε, για το τριώνυμο φ, $\varphi(\lambda) := \|x-(y+\lambda z)\|^2$, θα είχαμε

$$\varphi(\lambda) = \|z\|^2\lambda^2 - 2(x-y,z)\lambda + \|x-y\|^2,$$

οπότε

$$\min_{\lambda\in\mathbb{R}} \varphi(\lambda) = \varphi\Big(\frac{(x-y,z)}{\|z\|^2}\Big) < \varphi(0) = \|x-y\|^2,$$

το οποίο αντίκειται στην υπόθεση ότι το y είναι βέλτιστη προσέγγιση του x από τον

$\tilde{H}$. Γεωμετρικά, αυτό σημαίνει ότι, αν το $x-y$ δεν είναι κάθετο στο z, τότε η κάθετος από το x στην ευθεία $y+\lambda z$ δεν είναι η xy, βλ. το Σχήμα 1.2. □

Στην περίπτωση που προσεγγίζουμε από έναν υπόχωρο πεπερασμένης διάστασης, το πρόβλημα του προσδιορισμού της βέλτιστης προσέγγισης ανάγεται στην επίλυση ενός γραμμικού συστήματος εξισώσεων, του λεγόμενου συστήματος *κανονικών εξισώσεων*, βλ. την Άσκηση 1.2.

Ένα βασικότατο και ιδιαίτερα χρήσιμο αποτέλεσμα για χώρους Hilbert αποτελεί το λεγόμενο θεώρημα αναπαράστασης του Riesz. Για $z \in H$, το συναρτησιακό $F: H \to \mathbb{R}$, $F(x) := (z,x)$, είναι προφανώς γραμμικό και φραγμένο και μάλιστα, όπως θα δούμε και σε λίγο, για τη νόρμα του F, $\|F\| := \sup_{x\in H,\, x\neq 0} |F(x)|/\|x\|$, ισχύει $\|F\| = \|z\|$. Το θεώρημα αναπαράστασης του Riesz είναι το αντίστροφο αυτού του αποτελέσματος, λέει δηλαδή ότι κάθε *φραγμένο* γραμμικό συναρτησιακό μπορεί να παρασταθεί ως εσωτερικό γινόμενο με ένα κατάλληλο στοιχείο του χώρου.

Θεώρημα 1.1 (Θεώρημα αναπαράστασης του Riesz.) *Έστω $\big(H,(\cdot,\cdot)\big)$ ένας χώρος Hilbert και $F: H \to \mathbb{R}$ ένα φραγμένο γραμμικό συναρτησιακό. Τότε υπάρχει $z \in H$, τέτοιο ώστε*

$$\forall x \in H \quad F(x) = (z,x). \tag{1.6}$$

Επί πλέον, το z ορίζεται μονοσήμαντα και ισχύει $\|F\| = \|z\|$.

Απόδειξη. Αρχίζουμε με την απόδειξη της μοναδικότητας. Έστω λοιπόν $y,z \in H$, τέτοια ώστε

$$\forall x \in H \quad F(x) = (y,x) = (z,x).$$

Τότε, έχουμε $(y-z,x) = 0$, για κάθε $x \in H$, και με $x := y-z$ οδηγούμαστε στο συμπέρασμα $y = z$, αποδείξαμε δηλαδή τη μοναδικότητα.

Έστω τώρα $z \in H$ τέτοιο ώστε $F(x) = (z,x)$, για κάθε $x \in H$. Θα αποδείξουμε ότι $\|F\| = \|z\|$. Πράγματι έχουμε αφ' ενός

$$|F(x)| = |(z,x)| \leq \|z\|\,\|x\|,$$

δηλαδή $\|F\| \leq \|z\|$ και έχουμε αφ' ετέρου $|F(z)| = \|z\|\,\|z\|$, δηλαδή $\|F\| \geq \|z\|$. Επομένως $\|F\| = \|z\|$.

Απομένει τώρα να αποδείξουμε ύπαρξη του z. Αν το F είναι το μηδενικό συναρτησιακό, θέτουμε απλούστατα $z = 0$. Έστω λοιπόν ότι $F \neq 0$. Για να ισχύει η (1.6) πρέπει, προφανώς, το z να είναι κάθετο στον πυρήνα $\operatorname{Ker} F$ του F, $\operatorname{Ker} F := \{x \in H : F(x) = 0\}$. Εύκολα διαπιστώνουμε ότι ο $\operatorname{Ker} F$ είναι κλειστός υπόχωρος του H, αφού το F είναι φραγμένο και, συνακόλουθα, συνεχές γραμμικό συναρτησιακό. Επίσης προφανώς ισχύει $\operatorname{Ker} F \neq H$, αφού υποθέσαμε ότι $F \neq 0$. Επομένως ο $(\operatorname{Ker} F)^{\perp}$ είναι μη μηδενικός υπόχωρος του H. Έστω λοιπόν z_0 ένα μη μηδενικό στοιχείο του, $z_0 \in (\operatorname{Ker} F)^{\perp}$. Για $x \in H$ θέτουμε

$$v := F(x)z_0 - F(z_0)x.$$

Τότε, προφανώς,

$$F(v) = F(x)F(z_0) - F(z_0)F(x) = 0,$$

συνεπώς $v \in \operatorname{Ker} F$. Επειδή όμως $z_0 \in (\operatorname{Ker} F)^{\perp}$, θα έχουμε τότε

$$0 = (v, z_0) = \big(F(x)z_0 - F(z_0)x, z_0\big) = F(x)\|z_0\|^2 - F(z_0)(x, z_0),$$

συνεπώς

$$F(x) = \frac{F(z_0)}{\|z_0\|^2}(x, z_0),$$

οπότε, με

$$z := \frac{F(z_0)}{\|z_0\|^2} z_0,$$

έχουμε

$$\forall x \in H \quad F(x) = (z, x). \qquad \square$$

Παρατήρηση 1.1 (Ο δυϊκός ενός χώρου Hilbert.) Συμβολίζουμε με H' τον (τοπολογικό) δυϊκό του χώρου H, δηλαδή τον χώρο των *φραγμένων* γραμμικών συναρτησιακών στον H. Τη νόρμα του H' τη συμβολίζουμε πάλι με $\|\cdot\|$,

$$\|F\| := \sup_{v \in H,\ v \neq 0} \frac{|F(v)|}{\|v\|}.$$

Σύμφωνα με το θεώρημα αναπαράστασης του Riesz, η απεικόνιση $\tau : H' \to H, F \mapsto z$, είναι ισομετρία. Συνεπώς, οι χώροι H' και H έχουν την ίδια δομή· συνηθίζεται ως εκ τούτου, ως αντικείμενα μελέτης της Συναρτησιακής Ανάλυσης, να τους θεωρούμε ως τον ‘ίδιο’ χώρο, να τους ‘ταυτίζουμε’ όπως λέμε. Τονίζουμε πάντως ότι αυτή η ‘ταύτιση’ έχει νόημα μόνο όταν μας απασχολεί η αφηρημένη δομή των χώρων, οι H' και H μπορεί να περιέχουν εντελώς διαφορετικά στοιχεία· επί πλέον η ισομετρία, δηλαδή ο προσδιορισμός του z, όταν μας έχει δοθεί το F, μπορεί να παριστά ένα πολύπλοκο μαθηματικό πρόβλημα. Στα θέματα αυτά θα επανέλθουμε αργότερα με λεπτομέρειες. □

Σημειώνουμε ότι συνήθως γράφουμε Fx αντί για $F(x)$, όταν ο τελεστής (δηλαδή η απεικόνιση) F είναι γραμμικός· ιδιαίτερα, συνεπώς, αυτός ο συμβολισμός χρησιμοποιείται όταν το F είναι γραμμικό συναρτησιακό.

Υποθέτουμε τώρα ότι η διγραμμική μορφή a ικανοποιεί τις (1.1) και (1.2) και είναι συμμετρική. Αμέσως διαπιστώνουμε ότι η a είναι εσωτερικό γινόμενο στον V. Μάλιστα, θα δούμε αμέσως τώρα ότι ο $\big(V, a(\cdot,\cdot)\big)$ είναι χώρος Hilbert· ιδιαίτερα, θα μπορέσουμε αργότερα να εφαρμόσουμε σε αυτόν τον χώρο το θεώρημα αναπαράστασης του Riesz.

Πρόταση 1.2 *Ο χώρος $\big(V, a(\cdot,\cdot)\big)$ είναι χώρος Hilbert.*

Απόδειξη. Συμβολίζουμε με $\|\cdot\|_E$ τη *νόρμα ενέργειας*, δηλαδή τη νόρμα που παράγεται από τη (συμμετρική) διγραμμική μορφή a, $\|v\|_E := a(v,v)^{1/2}$, $v \in V$. Θεωρούμε μια ακολουθία $(v_n)_{n\in\mathbb{N}} \subset V$, που είναι ακολουθία Cauchy στον $\big(V, \|\cdot\|_E\big)$. Από την ελλειπτικότητα (1.2) έπεται αμέσως ότι η $(v_n)_{n\in\mathbb{N}}$ είναι ακολουθία Cauchy και στον $\big(H, \|\cdot\|\big)$. Συνεπώς, υπάρχει $v \in H$ τέτοιο ώστε $\|v_n - v\| \to 0$, $n \to \infty$. Μάλιστα, ισχύει $v \in V$, αφού ο V είναι κλειστός. Αλλά, από τη συνέχεια (1.1) έπεται ότι

$$\|v_n - v\|_E \le \sqrt{C}\,\|v_n - v\|$$

και έτσι οδηγούμαστε στο συμπέρασμα ότι $\|v_n - v\|_E \to 0$, $n \to \infty$. □

Τονίζουμε ότι η ελλειπτικότητα της διγραμμικής μορφής $a(\cdot,\cdot)$ είναι αναγκαία για να είναι ο $\big(V, a(\cdot,\cdot)\big)$ χώρος Hilbert· στο θέμα αυτό θα επανέλθουμε αργότερα με παραδείγματα, όταν θα έχουμε εξοικειωθεί με τους χώρους του Sobolev.

Το συμμετρικό μεταβολικό πρόβλημα. Έστω a μια φραγμένη και συμμετρική διγραμμική μορφή, η οποία είναι ελλειπτική σε έναν κλειστό υπόχωρο V του H, και F ένα φραγμένο γραμμικό συναρτησιακό στον V, $F \in V'$. Θεωρούμε τότε το *συμμετρικό μεταβολικό πρόβλημα*: Ζητείται $u \in V$ τέτοιο ώστε

$$a(u,v) = F(v) \quad \forall v \in V. \tag{1.7}$$

Η ύπαρξη και η μοναδικότητα της λύσης u έπονται από το γεγονός ότι ο $\big(V, a(\cdot,\cdot)\big)$ είναι, όπως είδαμε, χώρος Hilbert και το θεώρημα αναπαράστασης του Riesz. □

Ένας χαρακτηρισμός της λύσης του προβλήματος (1.7) δίνεται στην ακόλουθη Πρόταση:

Πρόταση 1.3 *Υπό τις υποθέσεις που αναφέρθηκαν προηγουμένως, το συναρτησιακό* $J : V \to \mathbb{R}$,

$$J(v) := \frac{1}{2}a(v,v) - F(v),$$

λαμβάνει το ελάχιστό του ακριβώς σε ένα σημείο $u \in V$, *και μάλιστα στη λύση του προβλήματος* (1.7).

Απόδειξη. Σύμφωνα με όσα αναφέρθηκαν προηγουμένως, υπάρχει ακριβώς ένα $u \in V$ που ικανοποιεί την (1.7). Για οποιοδήποτε $w \in V$ έχουμε τώρα

$$\begin{aligned} J(u+w) &= \frac{1}{2}a(u+w,u+w) - F(u+w) \\ &= \Big[\frac{1}{2}a(u,u) - F(u)\Big] + \frac{1}{2}a(w,w) + \big[a(u,w) - F(w)\big]. \end{aligned}$$

Όμως, ο πρώτος όρος στο δεξιό μέλος ισούται με $J(u)$ και ο τελευταίος ισούται με μηδέν, επομένως

$$J(u+w) = J(u) + \frac{1}{2}a(w,w) \geq J(u) + \frac{\alpha}{2}\|w\|^2,$$

και το αποτέλεσμα έπεται αμέσως. □

Πρόταση 1.4 (Η μέθοδος του Ritz.) *Έστω V_h ένας υπόχωρος του V με πεπερασμένη διάσταση. Τότε το πρόβλημα προσέγγισης: Ζητείται $u_h \in V_h$ τέτοιο ώστε*

$$a(u_h, \chi) = F(\chi) \quad \forall \chi \in V_h, \tag{1.8}$$

έχει ακριβώς μία λύση, η οποία μάλιστα ελαχιστοποιεί το συναρτησιακό $J : V_h \to \mathbb{R}$, $J(v) := \frac{1}{2}a(v,v) - F(v)$, στον V_h,

$$J(u_h) = \min_{v \in V_h} J(v).$$

Απόδειξη. Ο $\big(V_h, a(\cdot,\cdot)\big)$ είναι, προφανώς, χώρος Hilbert, ως χώρος πεπερασμένης διάστασης. Επί πλέον $F|_{V_h} \in V_h'$ και το αποτέλεσμα έπεται από όσα προηγήθηκαν. □

Σημειώνουμε ότι για την απόδειξη ύπαρξης και μοναδικότητας της προσεγγιστικής λύσης u_h δεν χρειάζεται να ανατρέξει κανείς στο θεώρημα αναπαράστασης του Riesz. Πράγματι, αφού $\dim V_h < \infty$, το πρόβλημα (1.8) μπορεί να γραφεί ως γραμμικό σύστημα. Το γεγονός αυτό μας δίνει τη δυνατότητα να υπολογίσουμε πραγματικά την u_h, ενώ κάτι τέτοιο είναι γενικά ανέφικτο για τη λύση u του προβλήματος (1.7), βλ. την Άσκηση 1.4.

Από τις σχέσεις (1.7) και (1.8) έπεται αμέσως η ακόλουθη θεμελιώδης σχέση ορθογωνιότητας

(1.9)
$$a(u-u_h,\chi)=0 \quad \forall \chi \in V_h,$$

αφού ο V_h υπετέθη υπόχωρος του V. Ιδιαίτερα, η u_h είναι η *βέλτιστη* προσέγγιση της u από τον χώρο V_h, ως προς τη νόρμα ενέργειας που παράγεται από το εσωτερικό γινόμενο $a(\cdot,\cdot)$,

(1.10)
$$\|u-u_h\|_E = \min_{v\in V_h} \|u-v\|_E.$$

1.2 Η μη συμμετρική περίπτωση

Σε αυτό το εδάφιο θεωρούμε και πάλι το *μεταβολικό πρόβλημα*: Ζητείται $u \in V$ τέτοιο ώστε

(1.11)
$$a(u,v)=F(v) \quad \forall v \in V,$$

με όλες τις υποθέσεις για τους χώρους, το γραμμικό συναρτησιακό και τη διγραμμική μορφή, μόνο που τώρα *δεν* υποθέτουμε πλέον συμμετρία της διγραμμικής μορφής. Τη λύση στο πρόβλημά μας, τώρα που δεν μπορούμε να καταφύγουμε στο θεώρημα αναπαράστασης του Riesz, μας τη δίνει το Λήμμα των Lax–Milgram, το οποίο και διατυπώνουμε αμέσως. Σημειώνουμε ότι το Λήμμα των Lax–Milgram αποτελεί γενίκευση του θεωρήματος αναπαράστασης του Riesz.

Θεώρημα 1.2 (Λήμμα των Lax–Milgram.) *Έστω $\big(H,(\cdot,\cdot)\big)$ ένας χώρος Hilbert, a μια φραγμένη (με σταθερά C) και ελλειπτική (με σταθερά α) διγραμμική μορφή, και $F : H \to$*

$\mathbb{R}$ *ένα φραγμένο γραμμικό συναρτησιακό. Τότε υπάρχει ακριβώς ένα* $u \in H$, *τέτοιο ώστε*

$$\forall v \in H \quad a(u, v) = F(v),$$

και μάλιστα ισχύει $\|u\| \leq \|F\|/\alpha$.

Απόδειξη. Σταθεροποιούμε ένα $y \in H$. Τότε, το $\Phi : H \to \mathbb{R}$, $\Phi(x) = a(y, x)$, είναι προφανώς ένα γραμμικό και φραγμένο συναρτησιακό. Σύμφωνα με το θεώρημα αναπαράστασης του Riesz, υπάρχει συνεπώς ακριβώς ένα $\tilde{y} \in H$ τέτοιο ώστε

$$\forall x \in H \quad \Phi(x) = a(y, x) = (\tilde{y}, x).$$

Έστω $A : H \to H, y \mapsto Ay := \tilde{y}$. Ο A είναι τότε ένας γραμμικός τελεστής. Πράγματι, για $\lambda \in \mathbb{R}$ και $u, v \in H$, έχουμε

$$\begin{aligned}\forall x \in H \quad \big(A(\lambda u + v), x\big) &= a(\lambda u + v, x) = \lambda a(u, x) + a(v, x) \\ &= \lambda(Au, x) + (Av, x) = (\lambda Au + Av, x),\end{aligned}$$

συνεπώς $A(\lambda u + v) = \lambda Au + Av$.

Θα αποδείξουμε τώρα ότι

$$\forall x \in H \quad \|x\| \leq \frac{1}{\alpha}\|Ax\|. \tag{1.12}$$

Πράγματι, λόγω της ελλειπτικότητας της a έχουμε

$$\alpha\|x\|^2 \leq a(x, x) = (Ax, x) \leq \|Ax\|\,\|x\|,$$

από την οποία έπεται αμέσως η (1.12). Μια άμεση συνέπεια της (1.12) είναι το γεγονός ότι ο τελεστής A είναι ένα προς ένα. Επίσης, ο A είναι φραγμένος, αφού η σχέση

$$\forall x \in H \quad a(y, x) = (Ay, x)$$

δίνει, για $x := Ay$,

$$\|Ay\|^2 = a(y, Ay) \leq C\|Ay\|\,\|y\|,$$

οπότε

$$\forall y \in H \quad \|Ay\| \leq C\|y\|.$$

Θα αποδείξουμε τώρα ότι η εικόνα $A(H)$ του H μέσω του A είναι κλειστός υπόχωρος του H. Πράγματι, έστω $\tilde{x}_n := Ax_n$ και $\tilde{x}_n \to \tilde{x}$. Αρκεί να αποδείξουμε ότι $\tilde{x} \in A(H)$. Σύμφωνα με την (1.12) θα έχουμε τότε

$$\alpha\|x_n - x_m\| \le \|\tilde{x}_n - \tilde{x}_m\|,$$

συνεπώς η $(x_n)_{n\in\mathbb{N}}$ είναι ακολουθία Cauchy και, συνακόλουθα, συγκλίνουσα. Έστω $x_n \to x$. Λόγω της συνέχειας του A θα έχουμε τότε $\tilde{x}_n \to Ax$, συνεπώς $\tilde{x} = Ax \in A(H)$.

Θα αποδείξουμε τώρα ότι η εικόνα του H μέσω του A είναι ολόκληρος ο χώρος H, $A(H) = H$, δηλαδή ότι ο τελεστής A είναι επί, με απαγωγή σε άτοπο. Έστω ότι $A(H) \subsetneqq H$. Τότε, θα υπήρχε $z \in A(H)^\perp$, $z \neq 0$. Συνεπώς θα είχαμε $(z, v) = 0$, για κάθε $v \in A(H)$, οπότε

$$\forall x \in H \quad a(x, z) = (Ax, z) = 0.$$

Για $x = z$ η σχέση αυτή δίνει $0 = a(z, z) \ge \alpha\|z\|^2$, άτοπο. Επομένως ισχύει πράγματι $A(H) = H$.

Σύμφωνα με το θεώρημα αναπαράστασης του Riesz, υπάρχει ακριβώς ένα $x \in H$ τέτοιο ώστε

$$\forall v \in H \quad (x, v) = F(v).$$

Αλλά, αφού ο A είναι ένα προς ένα και επί, συμπεραίνουμε ότι υπάρχει ακριβώς ένα $u \in H$ τέτοιο ώστε $Au = x$. Ανακεφαλαιώνοντας, υπάρχει ακριβώς ένα $u \in H$ τέτοιο ώστε

$$\forall v \in H \quad a(u, v) = (Au, v) = (x, v) = F(v).$$

Επί πλέον, έχουμε

$$\alpha\|u\|^2 \le a(u, u) = F(u) \le \|F\|\,\|u\|,$$

συνεπώς

$$\|u\| \le \frac{1}{\alpha}\|F\|,$$

και η απόδειξη είναι πλήρης. □

Έχοντας στη διάθεσή μας τα απαιτούμενα μέσα, επιστρέφουμε τώρα στο μεταβολικό πρόβλημα (1.11). Αφού ο V είναι κλειστός υπόχωρος του χώρου Hilbert H, ο $\big(V, (\cdot,\cdot)\big)$ είναι χώρος Hilbert. Η ύπαρξη και μοναδικότητα της λύσης u του προβλήματος (1.11), καθώς και η εκτίμηση

$$\|u\| \le \frac{1}{\alpha}\|F\|, \tag{1.13}$$

έπονται αμέσως από το Θεώρημα 1.2.

Πρόταση 1.5 (Η μέθοδος του Galerkin.) *Έστω V_h ένας υπόχωρος του V με πεπερασμένη διάσταση. Τότε το πρόβλημα προσέγγισης: Ζητείται $u_h \in V_h$ τέτοιο ώστε*

$$a(u_h, \chi) = F(\chi) \quad \forall \chi \in V_h, \tag{1.14}$$

έχει ακριβώς μία λύση.

Απόδειξη. Ο $\big(V_h, (\cdot,\cdot)\big)$ είναι, προφανώς, χώρος Hilbert, ως χώρος πεπερασμένης διάστασης. Επί πλέον $F|_{V_h} \in V_h'$ και το αποτέλεσμα έπεται από το λήμμα των Lax–Milgram. □

Όσον αφορά τις προσεγγιστικές ιδιότητες της u_h, δεν μπορούμε πλέον να αναμένουμε αυτή να είναι βέλτιστη προσέγγιση της u από τον V_h, βλ. την (1.10), αφού η διγραμμική μορφή a δεν παράγει κάποια νόρμα. Παρά ταύτα μπορούμε να εκτιμήσουμε το σφάλμα $u - u_h$, στη νόρμα του V. Αυτό παρουσιάζεται στο επόμενο αποτέλεσμα.

Θεώρημα 1.3 (Λήμμα του Céa.) *Έστω ότι η διγραμμική μορφή a είναι συνεχής και ελλειπτική στον $\big(V, \|\cdot\|_V\big)$. Έστω V_h ένας υπόχωρος του V, πεπερασμένης διάστασης, και u, u_h οι λύσεις των προβλημάτων* (1.11) *και* (1.14)*, αντίστοιχα. Τότε για το σφάλμα $u - u_h$ ισχύει η εκτίμηση*

$$\|u - u_h\|_V \le \frac{C}{\alpha} \min_{v \in V_h} \|u - v\|_V, \tag{1.15}$$

με C και α τις σταθερές στη συνέχεια (1.1) *και στην ελλειπτικότητα* (1.2) *της διγραμμικής μορφής a στον χώρο V, με τη νόρμα $\|\cdot\|_V$ στη θέση της νόρμας $\|\cdot\|$.*

Απόδειξη. Αφού ο V_h είναι υπόχωρος του V, από τις (1.11) και (1.14) έπεται ότι

$$a(u - u_h, \chi) = 0 \quad \forall \chi \in V_h. \tag{1.16}$$

Συνεπώς, για $\chi \in V_h$, έχουμε, βάσει των (1.2) και (1.16),

$$\begin{aligned}
\alpha \|u - u_h\|_V^2 &\le a(u - u_h, u - u_h) \\
&= a(u - u_h, u - \chi) + a(u - u_h, \chi - u_h) \\
&= a(u - u_h, u - \chi).
\end{aligned}$$

Χρησιμοποιώντας τώρα και τη συνέχεια της a (δηλαδή την (1.1) με $\|\cdot\|_V$ στη θέση της $\|\cdot\|$), έχουμε

$$\alpha\|u-u_h\|_V^2 \le C\|u-u_h\|_V\,\|u-\chi\|_V,$$

οπότε

$$\alpha\|u-u_h\|_V \le C\|u-\chi\|_V \quad \forall \chi\in V_h. \tag{1.17}$$

Από αυτή τη σχέση έπεται αμέσως η ζητούμενη εκτίμηση (1.15). □

1.3 Τα λήμματα του Strang

Μέχρι τώρα, τόσο στη μέθοδο του Galerkin (1.14) όσο και στη μέθοδο του Ritz (1.8), υποθέσαμε ότι ο V_h είναι υπόχωρος του V και ότι η προσέγγιση u_h ορίζεται μέσω της διγραμμικής μορφής a και του γραμμικού συναρτησιακού F. Όταν ικανοποιούνται όλες αυτές οι συνθήκες, μιλάμε για *συμμορφικά* (conforming) πεπερασμένα στοιχεία. Στην πράξη, πολλές φορές, αναγκαζόμαστε να καταφύγουμε σε *μη συμμορφικά* (nonconforming) πεπερασμένα στοιχεία, για διάφορους λόγους. Στη συνέχεια θα ασχοληθούμε με δύο περιπτώσεις μη συμμορφικών πεπερασμένων στοιχείων (εδώ σε αφηρημένη μορφή).

Συγκεκριμένα, στην πρώτη περίπτωση θα υποθέσουμε ότι ο V_h εξακολουθεί να είναι υπόχωρος του V, αλλά το μη συμμορφικό οφείλεται στο γεγονός ότι η προσέγγιση u_h ορίζεται μέσω μιας *προσεγγιστικής* διγραμμικής μορφής a_h και ενός *προσεγγιστικού* γραμμικού συναρτησιακού F_h. Στις εφαρμογές, η a_h και το F_h ορίζονται μόνο για στοιχεία του V_h (δηλαδή γενικά δεν ορίζονται για στοιχεία του V), προκύπτουν δε από την a και το F, αντίστοιχα, προσεγγίζοντας τα ολοκληρώματα με τύπους αριθμητικής ολοκλήρωσης. Στη δεύτερη περίπτωση θα υποθέσουμε ότι ο V_h *δεν* είναι υπόχωρος του V, και ότι η προσέγγιση u_h ορίζεται μέσω μιας *προσεγγιστικής* διγραμμικής μορφής a_h και του γραμμικού συναρτησιακού F. Στις εφαρμογές, ο V_h δεν είναι υπόχωρος του V είτε γιατί τα στοιχεία του δεν έχουν την απαιτούμενη ομαλότητα, είτε γιατί δεν ικανοποιούν τις συνοριακές συνθήκες. Χρησιμοποιούμε επίσης μια προσεγγιστική διγραμμική μορφή a_h γιατί, γενικά, η a δεν ορίζεται για στοιχεία του V_h, πάλι λόγω έλλειψης της απαιτούμενης ομαλότητας. Αντίθετα, γενικά, το F ορίζεται και για στοιχεία με λιγότερη ομαλότητα· γιαυτό, λόγω του ότι στη δεύτερη αυτή περίπτωση επικεντρώνουμε την προσοχή μας στις δυσκολίες που προκύπτουν από το γεγονός ότι ο V_h δεν είναι υπόχωρος του V, υποθέτουμε ότι δεν χρειάζεται να το προσεγγίσουμε.

Αρχίζουμε με την πρώτη περίπτωση: Υποθέτουμε ότι ο V_h είναι υπόχωρος του V, πεπερασμένης διάστασης, και ότι η προσεγγιστική διγραμμική μορφή $a_h : V_h \times V_h \to \mathbb{R}$ είναι ομοιόμορφα ελλειπτική στον V_h, δηλαδή ότι υπάρχει μια θετική σταθερά $\tilde{\alpha}$, ανεξάρτητη του h, τέτοια ώστε

$$\forall v_h \in V_h \quad a_h(v_h, v_h) \geq \tilde{\alpha} \|v_h\|_V^2. \tag{1.18}$$

Επί πλέον, θεωρούμε ένα προσεγγιστικό γραμμικό συναρτησιακό $F_h : V_h \to \mathbb{R}$. Ορίζουμε τώρα την προσέγγιση $u_h \in V_h$ μέσω της

$$a_h(u_h, v_h) = F_h(v_h) \quad \forall v_h \in V_h. \tag{1.19}$$

Κατ' αρχάς, χρησιμοποιώντας την ελλειπτικότητα (1.18) της διγραμμικής μορφής a_h διαπιστώνει κανείς αμέσως ότι η u_h είναι καλώς ορισμένη, δηλαδή ότι το πρόβλημα (1.19) έχει μοναδική λύση. Στο ακόλουθο Θεώρημα, που αναφέρεται στη βιβλιογραφία ως *πρώτο λήμμα του Strang*, δίνουμε μια εκτίμηση του σφάλματος $u - u_h$.

Θεώρημα 1.4 (Το πρώτο λήμμα του Strang.) *Έστω ότι η διγραμμική μορφή a είναι συνεχής στον $\big(V, \|\cdot\|_V\big)$. Έστω V_h ένας υπόχωρος του V, πεπερασμένης διάστασης, και u, u_h οι λύσεις των προβλημάτων* (1.11) *και* (1.19)*, αντίστοιχα. Τότε για το σφάλμα $u - u_h$ ισχύει η εκτίμηση*

$$\begin{aligned} \|u - u_h\|_V \leq \tilde{C}\Big[\min_{v_h \in V_h} \Big(\|u - v_h\|_V + \max_{w_h \in V_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_V}\Big) \\ + \max_{w_h \in V_h} \frac{|F(w_h) - F_h(w_h)|}{\|w_h\|_V}\Big], \end{aligned} \tag{1.20}$$

με $\tilde{C}$ μια σταθερά, ανεξάρτητη του h.

Απόδειξη. Έστω $v_h \in V_h$. Τότε, σύμφωνα με την (1.18),

$$a_h(u_h - v_h, u_h - v_h) \geq \tilde{\alpha} \|u_h - v_h\|_V^2. \tag{1.21}$$

Εξ άλλου, $a(u, u_h - v_h) = F(u_h - v_h)$ και $a_h(u_h, u_h - v_h) = F_h(u_h - v_h)$, συνεπώς

$$\begin{aligned} a_h(u_h - v_h, u_h - v_h) &= a(u, u_h - v_h) - F(u_h - v_h) + F_h(u_h - v_h) - a_h(v_h, u_h - v_h) \\ &= a(u - v_h, u_h - v_h) + \big[a(v_h, u_h - v_h) - a_h(v_h, u_h - v_h)\big] \\ &\quad + \big[F_h(u_h - v_h) - F(u_h - v_h)\big]. \end{aligned}$$

Εκτιμώντας τον πρώτο όρο στο δεξιό μέλος βάσει της συνέχειας της a (δηλαδή της (1.1) με $\|\cdot\|_V$ στη θέση της $\|\cdot\|$) και συνδυάζοντας το αποτέλεσμα με την (1.21), λαμβάνουμε

$$\begin{aligned}\tilde{\alpha}\|u_h - v_h\|_V \le C\|u - v_h\|_V &+ \frac{a(v_h, u_h - v_h) - a_h(v_h, u_h - v_h)}{\|u_h - v_h\|_V} \\ &+ \frac{F_h(u_h - v_h) - F(u_h - v_h)}{\|u_h - v_h\|_V}.\end{aligned}$$

Επομένως, έχουμε

$$\begin{aligned}\tilde{\alpha}\|u_h - v_h\|_V \le C\|u - v_h\|_V &+ \max_{w_h \in V_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_V} \\ &+ \max_{w_h \in V_h} \frac{|F(w_h) - F_h(w_h)|}{\|w_h\|_V}.\end{aligned} \tag{1.22}$$

Συνδυάζοντας την (1.22) με την τριγωνική ανισότητα

$$\|u - u_h\|_V \le \|u - v_h\|_V + \|u_h - v_h\|_V$$

και παίρνοντας το ελάχιστο ως προς $v_h \in V_h$, οδηγούμαστε αμέσως στην επιθυμητή εκτίμηση (1.20). □

Προχωρούμε τώρα στη δεύτερη περίπτωση: Τώρα ο V_h έχει πεπερασμένη διάσταση, αλλά δεν είναι κατ' ανάγκη υπόχωρος του V. Υποθέτουμε ότι έχουμε μια νόρμα $\|\cdot\|_h$, η οποία ορίζεται στον χώρο $V + V_h$, δηλαδή για στοιχεία που είναι άθροισμα στοιχείων του V και του V_h, και στον V συμπίπτει με την $\|\cdot\|_V$, δηλαδή $\|v\|_h = \|v\|_V$, για κάθε $v \in V$. Υποθέτουμε επίσης ότι η προσεγγιστική διγραμμική μορφή $a_h : V_h \times V_h \to \mathbb{R}$ είναι ομοιόμορφα ελλειπτική στον $\big(V_h, \|\cdot\|_h\big)$, δηλαδή ότι υπάρχει μια θετική σταθερά $\tilde{\alpha}$, ανεξάρτητη του h, τέτοια ώστε

$$\forall v_h \in V_h \quad a_h(v_h, v_h) \ge \tilde{\alpha}\|v_h\|_h^2. \tag{1.23}$$

Επί πλέον, υποθέτουμε ότι ο περιορισμός της a_h στον V συμπίπτει με την a, και ότι η a_h είναι συνεχής στον $V + V_h$, δηλαδή ότι υπάρχει μια σταθερά C_1 τέτοια ώστε

$$\forall v, w \in V + V_h \quad |a_h(v, w)| \le C_1\|v\|_h \, \|w\|_h. \tag{1.24}$$

Ακόμα υποθέτουμε ότι το γραμμικό συναρτησιακό F ορίζεται και για στοιχεία του V_h. Ορίζουμε τώρα την προσέγγιση $u_h \in V_h$ μέσω της

$$a_h(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h. \tag{1.25}$$

Και αυτή τη φορά, χρησιμοποιώντας την ελλειπτικότητα (1.23) της διγραμμικής μορφής a_h διαπιστώνει κανείς αμέσως ότι η u_h είναι καλώς ορισμένη, δηλαδή ότι το πρόβλημα (1.25) έχει μοναδική λύση. Στο ακόλουθο Θεώρημα, που αναφέρεται στη βιβλιογραφία ως *δεύτερο λήμμα του Strang*, δίνουμε μια εκτίμηση του σφάλματος $u - u_h$.

Θεώρημα 1.5 (Το δεύτερο λήμμα του Strang.) *Έστω u, u_h οι λύσεις των προβλημάτων* (1.11) *και* (1.25)*, αντίστοιχα. Τότε για το σφάλμα $u - u_h$ ισχύει η εκτίμηση*

$$\|u - u_h\|_h \le \tilde{C}\Big[\min_{v_h \in V_h} \|u - v_h\|_h + \max_{w_h \in V_h} \frac{|a(u, w_h) - F(w_h)|}{\|w_h\|_h}\Big], \tag{1.26}$$

με $\tilde{C}$ μια σταθερά, ανεξάρτητη του h.

Απόδειξη. Έστω $v_h \in V_h$. Τότε, σύμφωνα με την (1.23), έχουμε

$$a_h(u_h - v_h, u_h - v_h) \ge \tilde{\alpha}\|u_h - v_h\|_h^2. \tag{1.27}$$

Εξ άλλου, $a_h(u_h, u_h - v_h) = F(u_h - v_h)$, συνεπώς

$$\begin{aligned} a_h(u_h - v_h, u_h - v_h) &= F(u_h - v_h) - a_h(v_h, u_h - v_h) \\ &= a_h(u - v_h, u_h - v_h) + \big[F(u_h - v_h) - a_h(u, u_h - v_h)\big]. \end{aligned}$$

Εκτιμώντας τον πρώτο όρο στο δεξιό μέλος βάσει της (1.24) και συνδυάζοντας το αποτέλεσμα με την (1.27), λαμβάνουμε

$$\tilde{\alpha}\|u_h - v_h\|_h \le C_1\|u - v_h\|_h + \frac{|F(u_h - v_h) - a_h(u, u_h - v_h)|}{\|u_h - v_h\|_h}.$$

Επομένως, έχουμε

$$\tilde{\alpha}\|u_h - v_h\|_h \le C_1\|u - v_h\|_h + \max_{w_h \in V_h} \frac{|F(w_h) - a_h(u, w_h)|}{\|w_h\|_h}. \tag{1.28}$$

Συνδυάζοντας την (1.28) με την τριγωνική ανισότητα

$$\|u - u_h\|_V \le \|u - v_h\|_V + \|u_h - v_h\|_V$$

και παίρνοντας το ελάχιστο ως προς $v_h \in V_h$, οδηγούμαστε αμέσως στην επιθυμητή εκτίμηση (1.26). □

Ασκήσεις

1.1 Έστω $\big(H,(\cdot,\cdot)\big)$ ένας πραγματικός χώρος με εσωτερικό γινόμενο και $\|\cdot\|$ η παραγόμενη από το εσωτερικό γινόμενο νόρμα. Έστω $v, w \in H$. Αποδείξτε ότι:

i. $\|v+w\|^2 = \|v\|^2 + \|w\|^2 + 2(v,w)$.

ii. Αν $(v,w)=0$, τότε $\|v+w\|^2 = \|v\|^2 + \|w\|^2$ (Πυθαγόρειο θεώρημα).

iii. $\|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2$ (ισότητα του παραλληλογράμμου).

1.2 Στην Πρόταση 1.1 υποθέτουμε τώρα ότι η διάσταση του υποχώρου $\tilde H$ είναι πεπερασμένη, $\dim \tilde H = n$. Έστω $\{x_1,\dots,x_n\}$ μια βάση του $\tilde H$. Αποδείξτε ότι η (1.5) γράφεται ισοδύναμα στη μορφή

$$(x_i, y) = (x, x_i), \quad i = 1,\dots,n.$$

Με $y = \alpha_1 x_1 + \cdots + \alpha_n x_n$, με άγνωστους συντελεστές $\alpha_1,\dots,\alpha_n$, αποδείξτε ότι αυτές οι σχέσεις γράφονται ως γραμμικό σύστημα κανονικών εξισώσεων

$$\sum_{j=1}^{n} (x_i, x_j)\alpha_j = (x, x_i), \quad i = 1,\dots,n.$$

Ο πίνακας αυτού του γραμμικού συστήματος, $G(x_1,\dots,x_n) := \big((x_i,x_j)\big)_{i,j=1,\dots,n}$, λέγεται *πίνακας του Gram*. Αποδείξτε ότι ο πίνακας του Gram είναι συμμετρικός και θετικά ορισμένος.

1.3 Έστω $A \in \mathbb{R}^{n,n}$ ένας συμμετρικός και θετικά ορισμένος πίνακας, και $b \in \mathbb{R}^n$. Αποδείξτε ότι η συνάρτηση $\varphi : \mathbb{R}^n \to \mathbb{R}$, $\varphi(x) := (Ax,x)/2 - (b,x)$, όπου $(\cdot,\cdot)$ είναι το Ευκλείδειο εσωτερικό γινόμενο στον $\mathbb{R}^n$, λαμβάνει το ελάχιστό της μόνο σε ένα σημείο, στη λύση του γραμμικού συστήματος $Ax = b$, χωρίς να χρησιμοποιήσετε την Πρόταση 1.3.

1.4 Θεωρούμε το μεταβολικό πρόβλημα (1.8) με $N_h := \dim V_h$ και $\{\chi_1,\dots,\chi_{N_h}\}$ μια βάση του V_h. Με $u_h = \alpha_1\chi_1 + \cdots + \alpha_{N_h}\chi_{N_h}$, με άγνωστους συντελεστές $\alpha_1,\dots,\alpha_{N_h}$, αποδείξτε ότι το πρόβλημα (1.8) γράφεται ισοδύναμα ως γραμμικό σύστημα N_h εξισώσεων με N_h αγνώστους,

$$\sum_{j=1}^{N_h} a(\chi_i,\chi_j)\alpha_j = F(\chi_i), \quad i = 1,\dots,N_h,$$

παράβαλε με την Άσκηση 1.2. Τι ιδιότητες έχει ο πίνακας αυτού του συστήματος;

1.5 Θεωρήστε την προηγούμενη Άσκηση, αυτή τη φορά για το μεταβολικό πρόβλημα (1.14) και, χρησιμοποιώντας μια βάση του V_h, γράψτε το πρόβλημα ως γραμμικό σύστημα. Ποια από τις ιδιότητες του αντίστοιχου πίνακα της προηγούμενης Άσκησης *δεν* έχει ο πίνακας στην προκειμένη περίπτωση;

2. Στοιχεία από τη θεωρία των χώρων του Sobolev σε μία διάσταση

Σε αυτό το βοηθητικό κεφάλαιο θα παρουσιάσουμε συνοπτικά βασικά αποτελέσματα για χώρους του Sobolev. Οι χώροι αυτοί είναι πολύ χρήσιμοι στα Εφαρμοσμένα Μαθηματικά, γενικά, και στη μελέτη προβλημάτων συνοριακών τιμών για διαφορικές εξισώσεις, ειδικότερα, αφού αποτελούν το φυσιολογικό πλαίσιο για τη μεταβολική διατύπωση τέτοιων προβλημάτων· φυσιολογικά εμφανίζονται επίσης στη διακριτοποίηση τέτοιων προβλημάτων με τη μέθοδο του Galerkin, και την ιδιαίτερα σημαντική ειδική εκδοχή της, τη μέθοδο των πεπερασμένων στοιχείων. Για να διευκολύνουμε τον αναγνώστη, θα ασχοληθούμε εδώ με τους χώρους του Sobolev σε μία διάσταση· με τους χώρους του Sobolev σε δύο διαστάσεις θα ασχοληθούμε στο τέταρτο κεφάλαιο, η γενίκευση σε περισσότερες διαστάσεις δεν παρουσιάζει ουσιαστικές δυσκολίες.

2.1 Ένα κίνητρο

Έστω $a, b \in C^1[0,1]$ και $c \in C[0,1]$ τρεις συναρτήσεις τέτοιες ώστε η συνάρτηση a να λαμβάνει μόνο θετικές τιμές και η συνάρτηση $c - 1/2b'$ να λαμβάνει μη αρνητικές τιμές,

$$a(x) \geq \alpha > 0, \;\; c(x) - \frac{1}{2}b'(x) \geq 0 \quad \forall x \in [0,1]. \tag{2.1}$$

Έστω επίσης $f \in C[0,1]$. Θεωρούμε τότε το εξής πρόβλημα δύο σημείων: Ζητείται μια συνάρτηση $u : [0,1] \to \mathbb{R}$, δύο φορές συνεχώς παραγωγίσιμη, τέτοια ώστε

$$\begin{cases} -(au')' + bu' + cu = f \quad \text{στο } (0,1), \\ u(0) = u(1) = 0. \end{cases} \tag{2.2}$$

Κάθε λύση u του προβλήματος (2.2), η οποία έχει τις ιδιότητες που προαναφέραμε, λέγεται *κλασική* ή *ισχυρή* λύση του (2.2). Τα επίθετα αυτά τα χρησιμοποιούμε, γιατί στη συνέχεια θα αναφερθούμε και σε άλλες "λύσεις", τις λεγόμενες *γενικευμένες* ή *ασθενείς* λύσεις του (2.2).

Προκειμένου να χρησιμοποιήσουμε τη θεωρία του πρώτου κεφαλαίου, για να αποδείξουμε ύπαρξη και μοναδικότητα της λύσης u του προβλήματος (2.2), θα θέλαμε να διατυπώσουμε αυτό το πρόβλημα σε μεταβολική μορφή. Προς τούτο, πολλαπλασιάζουμε κατ' αρχάς τη διαφορική εξίσωση επί μια συνάρτηση *δοκιμής* v, συνεχώς παραγωγίσιμη στο $[0,1]$, με μηδενικές τιμές στα σημεία 0 και 1, και ολοκληρώνουμε στο διάστημα $[0,1]$, οπότε παίρνουμε

$$-\int_0^1 (au')'v\,dx + \int_0^1 bu'v\,dx + \int_0^1 cuv\,dx = \int_0^1 fv\,dx.$$

Τώρα, ολοκληρώνοντας κατά μέρη, έχουμε

$$-\int_0^1 (au')'v\,dx = \big[au'v\big]_{x=0}^{x=1} + \int_0^1 au'v'\,dx.$$

Όμως, ο πρώτος όρος στο δεξιό μέλος αυτής της σχέσης ισούται με μηδέν, αφού $v(0) = v(1) = 0$ (γιαυτό ζητήσαμε άλλωστε να μηδενίζονται στα άκρα του διαστήματος οι συναρτήσεις δοκιμής), οπότε

$$-\int_0^1 (au')'v\,dx = \int_0^1 au'v'\,dx.$$

Επομένως, με $V := \{v \in C^1[0,1] : v(0) = v(1) = 0\}$, έχουμε αφ' ενός $u \in V$ και αφ' ετέρου

$$\int_0^1 au'v'\,dx + \int_0^1 bu'v\,dx + \int_0^1 cuv\,dx = \int_0^1 fv\,dx \quad \forall v \in V. \tag{2.3}$$

Για να γράψουμε το πρόβλημα στην επιθυμητή μορφή, συμβολίζουμε με $(\cdot,\cdot)$ το ακόλουθο εσωτερικό γινόμενο

$$(v,w) := \int_0^1 vw\,dx \quad \forall v,w \in C[0,1],$$

και με $\|\cdot\|$ την παραγόμενη νόρμα, και ορίζουμε τη διγραμμική μορφή $a : V \times V \to \mathbb{R}$,

$$a(v,w) := (av',w') + (bv',w) + (cv,w),$$

και το γραμμικό συναρτησιακό $F : V \to \mathbb{R}$, $F(v) := (f, v)$. Με αυτόν τον συμβολισμό, η (2.3) γράφεται στη μορφή: Ζητείται συνάρτηση $u \in V$ τέτοια ώστε

$$a(u, v) = F(v) \quad \forall v \in V. \tag{2.4}$$

Πριν προχωρήσουμε σημειώνουμε ότι στα προβλήματα (2.3) και (2.4) η λύση αρκεί να είναι *μία* φορά συνεχώς παραγωγίσιμη, πρόκειται, όπως λέμε, για μια *ασθενή* ή *γενικευμένη* λύση, σε αντίθεση με το πρόβλημα (2.2), όπου απαιτούμε η u να είναι δύο φορές συνεχώς παραγωγίσιμη.

Τώρα το πρόβλημα (2.4) είναι φαινομενικά της μορφής (1.11)· γράφουμε φαινομενικά γιατί ακόμα δεν έχουμε ελέγξει κατά πόσον ισχύουν οι υποθέσεις που θέσαμε στο πρόβλημα (1.11). Για να αποδείξουμε συνέχεια και ελλειπτικότητα της διγραμμικής μορφής φαίνεται φυσιολογικό να εφοδιάσουμε τον V με το εσωτερικό γινόμενο $(\cdot, \cdot)_1$,

$$(v, w)_1 := (v', w') + (v, w) \quad \forall v, w \in V.$$

Συμβολίζουμε με $\| \cdot \|_1$ την παραγόμενη από το εσωτερικό γινόμενο $(\cdot, \cdot)_1$ νόρμα. Η μονάδα χρησιμοποιείται στο εσωτερικό γινόμενο και στη νόρμα ως δείκτης, επειδή και στα δύο υπεισέρχεται η πρώτη παράγωγος. Ας ελέγξουμε τώρα μία προς μία τις συνθήκες. Αρχίζουμε με τη συνέχεια της διγραμμικής μορφής: Για $v, w \in V$ έχουμε

$$|a(v, w)| \le \max_{0 \le x \le 1} a(x) \int_0^1 |v'w'|\, dx + \max_{0 \le x \le 1} |b(x)| \int_0^1 |v'w|\, dx + \max_{0 \le x \le 1} |c(x)| \int_0^1 |vw|\, dx,$$

οπότε, βάσει της ανισότητας των Cauchy–Schwarz, με

$$C := 2 \max\big\{ \max_{0 \le x \le 1} a(x), \max_{0 \le x \le 1} |b(x)|, \max_{0 \le x \le 1} |c(x)| \big\},$$

έχουμε

$$|a(v, w)| \le \frac{C}{2} \big[\|v'\|\, \|w'\| + \|v'\|\, \|w\| + \|v\|\, \|w\| \big].$$

Επομένως, πάλι με βάση την ανισότητα των Cauchy–Schwarz, αυτή τη φορά όμως για αθροίσματα, έχουμε

$$|a(v, w)| \le \frac{C}{2} \big(\|v'\|^2 + \|v'\|^2 + \|v\|^2 \big)^{1/2} \big(\|w'\|^2 + \|w\|^2 + \|w\|^2 \big)^{1/2}$$

και εύκολα καταλήγουμε στην

$$|a(v, w)| \le C \|v\|_1\, \|w\|_1 \quad \forall v, w \in V. \tag{2.5}$$

Για την ελλειπτικότητα, παρατηρούμε κατ' αρχάς ότι, για $v \in V$,

$$\begin{aligned} a(v,v) &= (av',v') + (bv',v) + (cv,v) \\ &= (av',v') + \frac{1}{2}\big(b,(v^2)'\big) + (cv,v) \\ &= (av',v') - \frac{1}{2}(b',v^2) + (cv,v) \\ &= (av',v') + \big((c - \frac{1}{2}b')v,v\big) \ge (av',v'), \end{aligned}$$

επομένως

$$a(v,v) \ge \alpha \|v'\|^2 \quad \forall v \in V. \tag{2.6}$$

Για να ολοκληρώσουμε την απόδειξη της ελλειπτικότητας, απομένει να αντικαταστήσουμε την ποσότητα $\|v'\|^2$ στο δεξιό μέλος αυτής της σχέσης με ένα θετικό πολλαπλάσιο της $\|v\|_1^2$. Προς τούτο θα χρησιμοποιήσουμε την *ανισότητα των Poincaré–Friedrichs*,

$$\|v\| \le \|v'\| \quad \forall v \in V, \tag{2.7}$$

την απόδειξη της οποίας αναβάλλουμε για λίγο· σημειώνουμε επίσης ότι η σταθερά 1 στο δεξιό μέλος μπορεί να αντικατασταθεί με άλλη, αρκετά μικρότερη σταθερά, θέμα στο οποίο θα επανέλθουμε στο επόμενο κεφάλαιο. Βάσει της (2.7), η (2.6) δίνει

$$a(v,v) \ge \alpha \|v\|^2 \quad \forall v \in V,$$

οπότε, χρησιμοποιώντας και πάλι τη (2.6),

$$a(v,v) \ge \frac{\alpha}{2}\big(\|v\|^2 + \|v'\|^2\big) \quad \forall v \in V,$$

δηλαδή

$$a(v,v) \ge \frac{\alpha}{2}\|v\|_1^2 \quad \forall v \in V, \tag{2.8}$$

η a είναι συνεπώς όντως ελλειπτική. Για να ολοκληρώσουμε την απόδειξη, απομένει να αποδείξουμε την ανισότητα των Poincaré–Friedrichs (2.7). Για $v \in V$ έχουμε

$$v(x) = \int_0^x v'(s)\,ds,$$

συνεπώς

$$|v(x)|^2 \le \Big(\int_0^x 1^2\,ds\Big)^{1/2}\Big(\int_0^x \big(v'(s)\big)^2\,ds\Big)^{1/2},$$

δηλαδή

$$|v(x)|^2 \le \|v'\|^2 \quad \forall x \in [0,1] \;\; \forall v \in V. \tag{2.9}$$

Ολοκληρώνοντας αυτή τη σχέση στο διάστημα $[0,1]$ οδηγούμαστε στη (2.7).

Ας σημειωθεί ακόμη ότι από τη (2.9) προκύπτει αμέσως η *ανισότητα του Sobolev*

$$\max_{0\le x\le 1} |v(x)| \le \|v'\| \quad \forall v \in V. \tag{2.10}$$

Προχωρούμε τώρα στην απόδειξη του ότι το γραμμικό συναρτησιακό F είναι φραγμένο. Πράγματι, για $v \in V$, έχουμε, βάσει της ανισότητας των Cauchy–Schwarz,

$$|F(v)| = |(f,v)| \le \|f\|\,\|v\| \le \|f\|\,\|v\|_1,$$

συνεπώς

$$\|F\| \le \|f\|. \tag{2.11}$$

Απομένει τώρα να εξετάσουμε κατά πόσον ο χώρος $\big(V, (\cdot,\cdot)_1\big)$ είναι χώρος Hilbert. Δυστυχώς η απάντηση είναι αρνητική! Πράγματι, εύκολα διαπιστώνει κανείς ότι η ακολουθία $(\varphi_n)_{n\in\mathbb{N}}$,

$$\varphi_n(x) := \begin{cases} x, & 0 \le x \le \frac{1}{2}(1-\frac{1}{n^2}), \\ \frac{1}{2}\big[1-2n^2(x-\frac{1}{2})^2-\frac{1}{2n^2}\big], & \frac{1}{2}(1-\frac{1}{n^2}) < x < \frac{1}{2}(1+\frac{1}{n^2}), \\ 1-x, & \frac{1}{2}(1+\frac{1}{n^2}) \le x \le 1, \end{cases} \tag{2.12}$$

(σημειώστε ότι

$$\varphi_n'(x) = \begin{cases} 1, & 0 \le x \le \frac{1}{2}(1-\frac{1}{n^2}), \\ -\,2n^2(x-\frac{1}{2}), & \frac{1}{2}(1-\frac{1}{n^2}) < x < \frac{1}{2}(1+\frac{1}{n^2}), \\ -\,1, & \frac{1}{2}(1+\frac{1}{n^2}) \le x \le 1\,), \end{cases}$$

είναι ακολουθία Cauchy στον $\big(V, \|\cdot\|_1\big)$ και ότι στη νόρμα $\|\cdot\|_1$ συγκλίνει στη συνάρτηση φ,

$$\varphi(x) = \begin{cases} x, & 0 \le x \le \dfrac{1}{2}, \\ 1-x, & \dfrac{1}{2} \le x \le 1, \end{cases} \tag{2.13}$$

η οποία φυσικά δεν είναι στοιχείο του V, αφού δεν είναι παραγωγίσιμη στο $1/2$. [Το παράδειγμα αυτό μας βάζει στον πειρασμό να τροποποιήσουμε τον ορισμό του χώρου V, έτσι ώστε αυτός να αποτελείται από συνεχείς και κατά τμήματα συνεχώς παραγωγίσιμες συναρτήσεις. Δυστυχώς και αυτός ο χώρος *δεν* είναι πλήρης.]

Υπάρχουν πολλοί χώροι του Sobolev, συναρτήσεων ορισμένων στο διάστημα $[0.1]$. Ο απλούστερος έχει το εσωτερικό γινόμενο $(\cdot,\cdot)_1$, με μόνη διαφορά ότι αντί για το ολοκλήρωμα του Riemann χρησιμοποιούμε το ολοκλήρωμα του Lebesgue, και είναι πλήρης.

2.2 Προκαταρκτικά: Οι χώροι L^p

Υποθέτουμε ότι οι περισσότεροι αναγνώστες είναι εξοικειωμένοι με τους χώρους L^p. Περιοριζόμαστε ως εκ τούτου σε μια συνοπτική ανασκόπηση της θεωρίας τους, παραλείποντας όλες τις αποδείξεις. Ακολουθούμε το εξαιρετικό βιβλίο των Adams και Fournier, βλ. τη βιβλιογραφία, [1]. Όπως αναφέραμε στην προηγούμενη παράγραφο, στους χώρους του Sobolev δεν χρησιμοποιούμε το ολοκλήρωμα του Riemann, αλλά το ολοκλήρωμα του Lebesgue. Σε αυτή την ενότητα λοιπόν θα ορίσουμε το ολοκλήρωμα του Lebesgue και θα δούμε κάποιες ιδιότητες των χώρων L^p, που ορίζονται επίσης με βάση το ολοκλήρωμα του Lebesgue.

2.2.1 Το μέτρο του Lebesgue

Σε αυτό το εδάφιο θα ασχοληθούμε με το μέτρο του Lebesgue σε μία διάσταση, που μας παρέχει τη δυνατότητα να περιγράψουμε με ακρίβεια τι εννοούμε ως μήκος ορισμένων υποσυνόλων του $\mathbb{R}$.

Ορισμός 2.1 ($\sigma-$άλγεβρα.) Μια οικογένεια Σ υποσυνόλων του $\mathbb{R}$ καλείται *$\sigma-$άλγεβρα*, αν έχει τις ακόλουθες ιδιότητες:

i. Το $\mathbb{R}$ ανήκει στη Σ.

ii. Αν ένα σύνολο A ανήκει στη Σ, τότε και το συμπλήρωμά του $A^c := \mathbb{R} \setminus A$ ανήκει στη Σ.

iii. Αν $(A_n)_{n\in\mathbb{N}}$ είναι μια ακολουθία συνόλων που ανήκουν στη Σ, τότε και η ένωσή τους $\cup_{n=1}^{\infty} A_n$ ανήκει στη Σ.

Σχόλιο: Από τις *i.* και *ii.* έπεται αμέσως ότι το κενό σύνολο ανήκει στη Σ. Από τις *ii.* και *iii.* έπεται αμέσως ότι, αν $(A_n)_{n\in\mathbb{N}}$ είναι μια ακολουθία συνόλων που ανήκουν στη Σ, τότε και η τομή τους $\cap_{n=1}^{\infty} A_n$ ανήκει στη Σ,

$$\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} \left(\mathbb{R} \setminus A_n^c\right) = \mathbb{R} \setminus \bigcup_{n=1}^{\infty} A_n^c.$$

Επίσης, αν A και B είναι δύο σύνολα της Σ, τότε και η διαφορά $A \setminus B$ ανήκει στη Σ. $\square$

Ορισμός 2.2 (Θετικά μέτρα.) Έστω Σ μια $\sigma-$άλγεβρα στο $\mathbb{R}$. Μια απεικόνιση $\mu : \Sigma \to [0, \infty]$ λέγεται *θετικό μέτρο*, αν είναι *αριθμήσιμα προσθετική*, δηλαδή για κάθε ακολουθία $(A_n)_{n\in\mathbb{N}}$ συνόλων της Σ, που είναι ανά δύο ξένα μεταξύ τους, ισχύει

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n).$$

Σχόλιο: Από τον ορισμό έπεται αμέσως ότι, αν $A, B \in \Sigma$ και $A \subset B$, τότε $\mu(A) \leq \mu(B)$. Επίσης, αν $A_n \in \Sigma$, $n \in \mathbb{N}$, και $A_1 \subset A_2 \subset A_3 \subset \cdots$, τότε $\mu\big(\cup_{n=1}^{\infty} A_n\big) = \lim_{n\to\infty} \mu(A_n)$. $\square$

Θεώρημα 2.1 (Ύπαρξη του μέτρου του Lebesgue.) *Υπάρχει μια $\sigma-$άλγεβρα Σ υποσυνόλων του $\mathbb{R}$ και ένα θετικό μέτρο μ στη Σ με τις εξής ιδιότητες:*

i. Κάθε ανοικτό υποσύνολο του $\mathbb{R}$ ανήκει στη Σ.

ii. Αν $A \subset B$, $B \in \Sigma$ και $\mu(B) = 0$, τότε και το A ανήκει στη Σ και $\mu(A) = 0$.

iii. Αν $A = [a, b]$, τότε το A ανήκει στη Σ και $\mu(A) = b - a$,

iv. Το μ είναι αναλλοίωτο σε μετατοπίσεις, δηλαδή, αν $x \in \mathbb{R}$ και $A \in \Sigma$, τότε και $x + A \in \Sigma$, όπου $x + A := \{x + y : y \in A\}$, και $\mu(x + A) = \mu(A)$. $\square$

Σχόλιο: Από το $i.$ έπεται ότι και κάθε κλειστό υποσύνολο του $\mathbb{R}$ ανήκει στη Σ. □

Τα στοιχεία της $\sigma-$άλγεβρας Σ του Θεωρήματος 2.1 λέγονται *μετρήσιμα κατά Lebesgue* (και θα αναφέρονται στη συνέχεια απλώς ως μετρήσιμα) υποσύνολα του $\mathbb{R}$. Το μέτρο μ καλείται *μέτρο του Lebesgue*, το $\mu(A)$ θα αναφέρεται ως μέτρο ή μήκος του A. Τα στοιχεία της Σ με μέτρο μηδέν λέγονται *σύνολα μέτρου μηδέν*. Κάθε αριθμήσιμο υποσύνολο του $\mathbb{R}$ έχει μέτρο μηδέν.

Με το αξίωμα της επιλογής αποδεικνύεται ότι υπάρχουν μη μετρήσιμα υποσύνολα του $\mathbb{R}$.

Ορισμός 2.3 (Σχεδόν παντού.) Αν $B \subset A \subset \mathbb{R}$ και $\mu(B) = 0$, τότε κάθε συνθήκη που ικανοποιείται στο σύνολο $A \setminus B$ λέμε ότι ισχύει *σχεδόν παντού* (σ.π.) στο A.

Ορισμός 2.4 (Μετρήσιμες συναρτήσεις.) Έστω A ένα μετρήσιμο υποσύνολο του $\mathbb{R}$. Μια συνάρτηση $f : A \to [-\infty, \infty]$ λέγεται *μετρήσιμη*, αν τα σύνολα $\{x \in A : f(x) > a\}$ είναι μετρήσιμα, για κάθε $a \in \mathbb{R}$.

Θεώρημα 2.2 (Ιδιότητες μετρήσιμων συναρτήσεων.)

$i.$ *Αν μια συνάρτηση f είναι μετρήσιμη, τότε και η συνάρτηση $|f|$ είναι μετρήσιμη.*

$ii.$ *Αν f και g είναι δύο μετρήσιμες συναρτήσεις, τότε και οι $f+g$ και fg είναι μετρήσιμες.*

$iii.$ *Αν $(f_n)_{n\in\mathbb{N}}$ είναι μια ακολουθία μετρήσιμων συναρτήσεων, τότε και οι συναρτήσεις $\sup_{n\in\mathbb{N}} f_n$ (με $(\sup_{n\in\mathbb{N}} f_n)(x) := \sup_{n\in\mathbb{N}} f_n(x)$), $\inf_{n\in\mathbb{N}} f_n$, $\limsup_{n\in\mathbb{N}} f_n$ καθώς και $\liminf_{n\in\mathbb{N}} f_n$ είναι μετρήσιμες.*

$iv.$ *Κάθε συνεχής συνάρτηση, ορισμένη σε ένα μετρήσιμο σύνολο, είναι μετρήσιμη.*

$v.$ *Έστω $f : \mathbb{R} \to \mathbb{R}$ μια συνεχής συνάρτηση και g μια μετρήσιμη συνάρτηση. Τότε και η σύνθεση $f \circ g$ είναι μετρήσιμη.*

$vi.$ (Το θεώρημα του Lusin.) *Αν μια συνάρτηση f είναι μετρήσιμη και $f(x) = 0$ για $x \in A^c$, όπου A υποσύνολο του $\mathbb{R}$ τέτοιο ώστε $\mu(A) < \infty$, και $\varepsilon > 0$, τότε υπάρχει μια συνεχής συνάρτηση $g : \mathbb{R} \to \mathbb{R}$ με συμπαγή φορέα, τέτοια ώστε $\max_{x\in\mathbb{R}} g(x) \leq \sup_{x\in\mathbb{R}} f(x)$ και $\mu\Big(\{x \in \mathbb{R} : f(x) \neq g(x)\}\Big) < \varepsilon$.* □

Ορισμός 2.5 (Χαρακτηριστικές και απλές συναρτήσεις.) Έστω A ένα υποσύνολο του $\mathbb{R}$. Η συνάρτηση $\chi_A : \mathbb{R} \to \mathbb{R}$,

$$\chi_A(x) := \begin{cases} 1, & \text{αν } x \in A, \\ 0, & \text{αν } x \notin A, \end{cases}$$

λέγεται *χαρακτηριστική συνάρτηση του* A. Μια συνάρτηση $s : \mathbb{R} \to \mathbb{R}$ καλείται *απλή*, αν η εικόνα της είναι ένα πεπερασμένο σύνολο, δηλαδή αν υπάρχουν $a_1, \ldots, a_n \in \mathbb{R}$ (ανά δύο διαφορετικά μεταξύ τους), τέτοια ώστε $s(x) \in \{a_1, \ldots, a_n\}$, για κάθε $x \in \mathbb{R}$.

Σχόλιο: Έστω $s : \mathbb{R} \to \mathbb{R}$ μια απλή συνάρτηση, $s(x) \in \{a_1, \ldots, a_n\}$. Θέτουμε $A_i := \{x \in \mathbb{R} : s(x) = a_i\}$, $i = 1, \ldots, n$. Αμέσως διαπιστώνουμε ότι η s γράφεται στη μορφή

$$s = \sum_{i=1}^{n} a_i \chi_{A_i}.$$

Η s είναι μετρήσιμη, αν και μόνο αν τα σύνολα $A_1, \ldots, A_n$ είναι μετρήσιμα. □

Οι απλές συναρτήσεις έχουν πολύ καλές προσεγγιστικές ιδιότητες, όπως θα δούμε αμέσως τώρα.

Θεώρημα 2.3 (Προσέγγιση συναρτήσεων με απλές συναρτήσεις.) *Έστω A ένα υποσύνολο του $\mathbb{R}$ και $f : A \to \mathbb{R}$ μια συνάρτηση. Τότε ισχύουν:*

i. *Υπάρχει μια ακολουθία $(s_n)_{n\in\mathbb{N}}$ απλών συναρτήσεων, που συγκλίνει σημειακά (ή κατά σημείο) στην f, δηλαδή*

$$\lim_{n\to\infty} s_n(x) = f(x), \quad \text{για κάθε } x \in A.$$

ii. *Αν η f είναι φραγμένη, τότε η ακολουθία $(s_n)_{n\in\mathbb{N}}$ μπορεί να επιλεγεί έτσι ώστε η σύγκλιση να είναι ομοιόμορφη,*

$$\lim_{n\to\infty} \left(\sup_{x\in A} |f(x) - s_n(x)| \right) = 0.$$

iii. *Αν η f είναι μετρήσιμη, τότε η ακολουθία $(s_n)_{n\in\mathbb{N}}$ στο i. μπορεί να επιλεγεί έτσι ώστε κάθε s_n να είναι μετρήσιμη.*

iv. *Αν η f λαμβάνει μόνο μη αρνητικές τιμές, τότε η ακολουθία $(s_n)_{n\in\mathbb{N}}$ στο i. μπορεί να επιλεγεί έτσι ώστε, για κάθε $x \in \mathbb{R}$, η ακολουθία $\big(s_n(x)\big)_{n\in\mathbb{N}}$ να είναι αύξουσα.* □

2.2.2 Το ολοκλήρωμα του Lebesgue

Σε αυτό το εδάφιο θα ασχοληθούμε με το ολοκλήρωμα του Lebesgue. Αυτό το ολοκλήρωμα παίζει πολύ σημαντικό ρόλο στα σύγχρονα Μαθηματικά· ιδιαίτερα με αυτό ορίζονται οι χώροι L^p. Το ολοκλήρωμα του Riemann εξακολουθεί να χρησιμοποιείται για καθαρά παιδαγωγικούς λόγους.

Στη συνέχεια δίνουμε τον ορισμό του ολοκληρώματος του Lebesgue, σε τρία στάδια: αρχίζουμε με την περίπτωση απλών συναρτήσεων, συνεχίζουμε με την περίπτωση μη αρνητικών συναρτήσεων και καταλήγουμε στη γενική περίπτωση. Κάθε μία από αυτές τις περιπτώσεις χρησιμοποιείται για να μας οδηγήσει στον ορισμό στην περίπτωση που ακολουθεί.

Ορισμός 2.6 (Το ολοκλήρωμα του Lebesgue.) Έστω A ένα μετρήσιμο υποσύνολο του $\mathbb{R}$. *Το ολοκλήρωμα του Lebesgue* μιας απλής συνάρτησης s,

$$s = \sum_{i=1}^{n} a_i \chi_{A_i},$$

όπου $A_i \subset A$ μετρήσιμα σύνολα, ορίζεται ως εξής

$$\int_A s(x)\,dx := \sum_{i=1}^{n} a_i \mu(A_i).$$

Το ολοκλήρωμα του Lebesgue μιας μετρήσιμης, μη αρνητικής συνάρτησης $f : A \to \mathbb{R}$ ορίζεται ως εξής

$$\int_A f(x)\,dx := \sup \int_A s(x)\,dx,$$

όπου το supremum λαμβάνεται ως προς όλες τις μετρήσιμες, απλές συναρτήσεις s, που μηδενίζονται στο συμπλήρωμα του A, τέτοιες ώστε $0 \leq s(x) \leq f(x)$, για κάθε $x \in A$.

Το ολοκλήρωμα του Lebesgue μιας μετρήσιμης συνάρτησης $f : A \to \mathbb{R}$ ορίζεται ως εξής

$$\int_A f(x)\,dx := \int_A f^+(x)\,dx - \int_A f^-(x)\,dx,$$

υπό την προϋπόθεση ότι τουλάχιστον ένα από τα ολοκληρώματα στο δεξιό μέλος είναι πεπερασμένο, όπου $f^+(x) := \max\big(f(x), 0\big)$ και $f^-(x) := \min\big(f(x), 0\big)$.

Αν το ολοκλήρωμα του Lebesgue μιας συνάρτησης f στο A είναι πεπερασμένο, τότε λέμε ότι η f είναι *κατά Lebesgue ολοκληρώσιμη στο A*. Το σύνολο των κατά Lebesgue ολοκληρώσιμων συναρτήσεων στο A συμβολίζεται με $L^1(A)$.

Παρατήρηση 2.1 (Παράδειγμα κατά Lebesgue ολοκληρώσιμης συνάρτησης που δεν είναι κατά Riemann ολοκληρώσιμη.) Όπως είναι γνωστό, και πολύ εύκολο να το δει κανείς, η συνάρτηση f,

$$f(x) := \begin{cases} 1, & \text{αν } x \in [0,1] \text{ άρρητος}, \\ 0, & \text{αν } x \in [0,1] \text{ ρητός}, \end{cases}$$

δεν είναι ολοκληρώσιμη κατά Riemann. Όμως, το σύνολο $A_1 := \{x \in [0,1] : x \text{ ρητός}\}$ έχει μέτρο μηδέν, συνεπώς η f είναι ολοκληρώσιμη κατά Lebesgue, το ολοκλήρωμά της μάλιστα ισούται με τη μονάδα.

Θεώρημα 2.4 (Ιδιότητες ολοκληρώσιμων συναρτήσεων.) *Υποθέτουμε ότι όλες οι συναρτήσεις και τα σύνολα που εμφανίζονται στη συνέχεια είναι μετρήσιμες, και μετρήσιμα, αντίστοιχα.*

i. *Αν μια συνάρτηση f είναι φραγμένη σε ένα σύνολο A και $\mu(A) < \infty$, τότε η f είναι ολοκληρώσιμη στο A, $f \in L^1(A)$.*

ii. *Αν $\mu(A) < \infty$ και $a \le f(x) \le b$, για κάθε $x \in A$, τότε*

$$a\mu(A) \le \int_A f(x)\,dx \le b\mu(A).$$

iii. *Αν $f(x) \le g(x)$, για κάθε $x \in A$, και αν υπάρχουν και τα δύο ολοκληρώματα, τότε*

$$\int_A f(x)\,dx \le \int_A g(x)\,dx.$$

iv. *Αν $f, g \in L^1(A)$, τότε $f + g \in L^1(A)$ και*

$$\int_A (f+g)(x)\,dx = \int_A f(x)\,dx + \int_A g(x)\,dx.$$

v. *Αν $f \in L^1(A)$ και $c \in \mathbb{R}$, τότε $cf \in L^1(A)$ και*

$$\int_A (cf)(x)\,dx = c\int_A f(x)\,dx.$$

vi. *Αν $f \in L^1(A)$, τότε $|f| \in L^1(A)$ και*

$$\Big|\int_A f(x)\,dx\Big| \le \int_A |f(x)|\,dx.$$

vii. *Αν* $f \in L^1(A)$ *και* $B \subset A$, *τότε* $f \in L^1(B)$. *Αν επί πλέον* $f(x) \geq 0$, *για κάθε* $x \in A$, *τότε*

$$\int_B f(x)\,dx \leq \int_A f(x)\,dx.$$

viii. *Αν* $\mu(A) = 0$, *τότε*

$$\int_A f(x)\,dx = 0.$$

ix. *Αν* $f \in L^1(A)$ *και*

$$\int_B f(x)\,dx = 0,$$

για κάθε $B \subset A$, *τότε* $f(x) = 0$ *σ.π. στο* A. □

Παρατήρηση 2.2 (Ισότητα στοιχείων του $L^1(A)$.) Μια συνέπεια του $ix.$ του Θεωρήματος 2.4 και της προσθετικότητας του ολοκληρώματος είναι ότι σύνολα μέτρου μηδέν μπορούν να αγνοηθούν στην ολοκλήρωση. Δηλαδή, αν f και g είναι δύο μετρήσιμες συναρτήσεις σε ένα σύνολο A και $f(x) = g(x)$ σ.π. στο A, τότε

$$\int_A f(x)\,dx = \int_A g(x)\,dx.$$

Έτσι, δύο στοιχεία του $L^1(A)$ θεωρούμε ότι ταυτίζονται, αν είναι ίσα σχεδόν παντού. Επομένως, τα στοιχεία του $L^1(A)$ είναι κλάσεις ισοδυναμίας συναρτήσεων, δύο συναρτήσεις ανήκουν στην ίδια κλάση, αν είναι ίσες σχεδόν παντού. Παρά ταύτα τα στοιχεία του $L^1(A)$ τα αναφέρουμε συνήθως ως συναρτήσεις. Τονίζουμε πάντως ότι δεν έχει νόημα να μιλάμε για την τιμή σε κάποιο συγκεκριμένο σημείο ενός στοιχείου του $L^1(A)$.

2.2.3 Οι χώροι L^p

Έστω Ω ένα χωρίο στον $\mathbb{R}$, δηλαδή ένα *μη κενό, ανοικτό* υποσύνολο του $\mathbb{R}$. Έστω $p \geq 1$. Συμβολίζουμε με $L^p(\Omega)$ την κλάση των μετρήσιμων συναρτήσεων, που ορίζονται στο Ω και παίρνουν πραγματικές τιμές, με την ιδιότητα η $|u|^p$ να είναι ολοκληρώσιμη στο Ω, δηλαδή

$$\int_\Omega |u(x)|^p\,dx < \infty.$$

Στοιχεία του $L^p(\Omega)$, που είναι ίσα σ.π. στο Ω, θεωρούμε ότι ταυτίζονται. Χάριν ευκολίας θα αναφέρουμε τα στοιχεία του $L^p(\Omega)$ ως συναρτήσεις και θα γράφουμε $u \in L^p(\Omega)$, αν η u ικανοποιεί την ανισότητα που αναφέρθηκε προηγουμένως.

Πρόταση 2.1 (Ο $L^p(\Omega)$ είναι γραμμικός χώρος.) *Ο $L^p(\Omega)$ είναι γραμμικός χώρος και η απεικόνιση $\|\cdot\|_p : L^p(\Omega) \to \mathbb{R}$,*

$$\|u\|_p := \Big(\int_\Omega |u(x)|^p\,dx\Big)^{1/p},$$

είναι νόρμα.

Απόδειξη. Για $c \in \mathbb{R}$ και $u \in L^p(\Omega)$ έχουμε

$$\|cu\|_p = |c|\Big(\int_\Omega |u(x)|^p\,dx\Big)^{1/p} = |c|\,\|u\|_p,$$

ιδιαίτερα λοιπόν $cu \in L^p(\Omega)$.

Έστω τώρα $u, v \in L^p(\Omega)$. Θα αποδείξουμε ότι $u + v \in L^p(\Omega)$, και κατ' αυτόν τον τρόπο θα ολοκληρωθεί η απόδειξη ότι ο $L^p(\Omega)$ είναι γραμμικός χώρος. Αρχίζουμε με την πιο εύκολη περίπτωση, $p = 1$. Τότε έχουμε

$$\|u + v\|_1 = \int_\Omega |u(x) + v(x)|\,dx \le \int_\Omega \big(|u(x)| + |v(x)|\big)\,dx = \|u\|_1 + \|v\|_1,$$

ιδιαίτερα λοιπόν $u + v \in L^p(\Omega)$. Για $p > 1$ ισχύει επίσης η εκτίμηση

$$\|u + v\|_p \le \|u\|_p + \|v\|_p,$$

η οποία λέγεται *ανισότητα του Minkowski*, βλ. την Άσκηση 2.4. Η απόδειξη της ανισότητας του Minkowski βασίζεται στην *ανισότητα του Hölder*, σύμφωνα με την οποία ισχύει το εξής: αν $q := p/(p-1)$ είναι ο λεγόμενος *συζυγής εκθέτης* του p, δηλαδή το q είναι τέτοιο ώστε

$$\frac{1}{p} + \frac{1}{q} = 1,$$

$u \in L^p(\Omega)$ και $v \in L^q(\Omega)$, τότε $uv \in L^1(\Omega)$ και $\|uv\|_1 \le \|u\|_p\,\|v\|_q$, βλ. την Άσκηση 2.3.

Τέλος, ισχύει $\|u\|_p \ge 0$, και για τη μηδενική συνάρτηση u, $u = 0$, έχουμε προφανώς $\|u\|_p = 0$. Επί πλέον, για $u \in L^p(\Omega)$, αν $\|u\|_p = 0$, τότε $\int_\Omega |u(x)|^p\,dx = 0$, συνεπώς, σύμφωνα με το ix. του Θεωρήματος 2.4, $|u(x)|^p = 0$ σ.π. στο Ω, δηλαδή $u(x) = 0$ σ.π. στο Ω, οπότε $u = 0$.

Ανακεφαλαιώνοντας, αποδείξαμε ότι ο $L^p(\Omega)$ είναι γραμμικός χώρος και η $\|\cdot\|_p$ είναι νόρμα στον $L^p(\Omega)$. □

Με τη βοήθεια δύο θεμελιωδών αποτελεσμάτων της θεωρίας της ολοκλήρωσης κατά Lebesgue, του θεωρήματος της μονότονης σύγκλισης και του λήμματος του Fatou, μπορεί να αποδειχθεί ότι

Θεώρημα 2.5 (Πληρότητα των $L^p(\Omega)$.) *Οι χώροι* $L^p(\Omega)$, $1 \leq p < \infty$, *είναι πλήρεις.* □

Πόρισμα 2.1 (Ο $L^2(\Omega)$ είναι χώρος Hilbert.) *Ο χώρος* $L^2(\Omega)$ *είναι χώρος Hilbert με το εσωτερικό γινόμενο*

$$(u, v) := \int_\Omega uv\,dx. \qquad \square$$

Σχόλιο. (Ο χώρος $L^\infty(\Omega)$.) Λέμε ότι μια μετρήσιμη συνάρτηση u στο Ω είναι *ουσιαστικά φραγμένη* στο Ω, αν υπάρχει σταθερά K τέτοια ώστε

$$|u(x)| \leq K \ \text{ σ.π. στο } \ \Omega.$$

Το infimum των σταθερών K, για τις οποίες ισχύει αυτή η σχέση, λέγεται *ουσιαστικό supremum* της $|u|$ στο Ω και συμβολίζεται με

$$\operatorname{ess\,sup}_{x\in\Omega} |u(x)|.$$

Ο χώρος των ουσιαστικά φραγμένων συναρτήσεων στο Ω συμβολίζεται με $L^\infty(\Omega)$. Ο $L^\infty(\Omega)$ είναι γραμμικός χώρος και η $\|\cdot\|_\infty$,

$$\|u\|_\infty := \operatorname{ess\,sup}_{x\in\Omega} |u(x)|,$$

είναι νόρμα στον $L^\infty(\Omega)$. Μάλιστα ο $\big(L^\infty(\Omega), \|\cdot\|_\infty\big)$ είναι πλήρης χώρος.

Εύκολα πείθεται κανείς, ότι στην περίπτωση που μια συνάρτηση u είναι συνεχής στο Ω, το ουσιαστικό supremum της ταυτίζεται με το supremum της. □

Για $p \neq 2$ οι χώροι $L^p(\Omega)$ είναι, όπως αναφέραμε ήδη, πλήρεις, δηλαδή χώροι Banach, αλλά η νόρμα τους δεν παράγεται από εσωτερικό γινόμενο, δηλαδή αυτοί οι χώροι δεν είναι χώροι Hilbert.

Ορισμός 2.7 (Πυκνότητα και διαχωρισιμότητα.) Έστω $\big(X, \|\cdot\|\big)$ ένας χώρος με νόρμα. Λέμε ότι ένα υποσύνολο S του X είναι *πυκνό* στον X, αν κάθε $x \in X$ είναι όριο μιας ακολουθίας στοιχείων του X, ή με άλλα λόγια αν κάθε x μπορεί να προσεγγισθεί όσο καλά θέλουμε με στοιχεία του S,

$$\forall \varepsilon > 0 \ \exists y \in S \quad \|x - y\| < \varepsilon.$$

Λέμε ότι ο χώρος X είναι *διαχωρίσιμος*, αν έχει ένα αριθμήσιμο, πυκνό υποσύνολο.

Έστω Ω ένα χωρίο και u μια συνεχής συνάρτηση στο Ω. *Φορέας* της u καλείται η κλειστή θήκη του συνόλου $\{x \in \Omega : u(x) \neq 0\}$. Το σύνολο των συνεχών συναρτήσεων στο Ω με φορέα ένα φραγμένο σύνολο που περιέχεται στο Ω (προσοχή: στο Ω, όχι στο $\bar{\Omega}$) συμβολίζεται με $C_0(\Omega)$. Ο δείκτης 0 σε αυτόν τον συμβολισμό υποδηλώνει ότι οι συναρτήσεις αυτές μηδενίζονται σε περιοχές κοντά στο σύνορο του Ω.

Πρόταση 2.2 (Πυκνότητα του $C_0(\Omega)$ στους $L^p(\Omega)$.) *Για $1 \leq p < \infty$, ο χώρος $C_0(\Omega)$ είναι πυκνός στον $L^p(\Omega)$.* □

Σχόλιο. (Μη πυκνότητα του $C_0(\Omega)$ στον $L^\infty(\Omega)$.) Έστω $\Omega := (0,1)$ και $u : (0,1) \to \mathbb{R}$, $u(x) := 1$. Προφανώς $u \in L^\infty(\Omega)$. Βεβαιωθείτε ότι για κάθε $\varphi \in C_0(\Omega)$ ισχύει $\|u - \varphi\|_\infty \geq 1$, για να πεισθείτε ότι το αντίστοιχο της Πρότασης 2.2 για $p = \infty$ δεν ισχύει. □

Πρόταση 2.3 (Διαχωρισιμότητα των $L^p(\Omega)$.) *Για $1 \leq p < \infty$, ο χώρος $L^p(\Omega)$ είναι διαχωρίσιμος.* □

Σχόλιο. (Μη διαχωρισιμότητα του $L^\infty(\Omega)$.) Μπορεί να αποδειχθεί ότι ο χώρος $L^\infty(\Omega)$ *δεν* είναι διαχωρίσιμος. □

Αναφέρουμε ακόμα τον χώρο $L^1_{\text{loc}}(\Omega)$, ο οποίος αποτελείται από τις *τοπικά* (locally) ολοκληρώσιμες συναρτήσεις, δηλαδή, αν Ω_1 είναι ένα ανοικτό υποσύνολο του Ω τέτοιο ώστε το $\bar{\Omega}_1$ να είναι φραγμένο και να περιέχεται στο Ω, τότε οι περιορισμοί των στοιχείων του $L^1_{\text{loc}}(\Omega)$ στο Ω_1 περιέχονται στον $L^1(\Omega_1)$. Ο $L^1_{\text{loc}}(\Omega)$ είναι *γνήσιος* υπόχωρος του $L^1(\Omega)$. Φερ' ειπείν, η συνάρτηση $u : (0,1) \to \mathbb{R}$, $u(x) := 1/x$, ανήκει στον $L^1_{\text{loc}}(0,1)$, αφού, για κάθε $0 < a < b < 1$, το ολοκλήρωμα

$$\int_a^b |u(x)|\, dx$$

είναι πεπερασμένο, αλλά δεν ανήκει στον $L^1(0,1)$, αφού δεν είναι ολοκληρώσιμη στο διάστημα $(0,1)$. Σημειώστε ότι στον $L^1_{\text{loc}}(\Omega)$ δεν ορίζουμε κάποια νόρμα.

Ο *φορέας* $\operatorname{supp} u$ μιας συνάρτησης $u \in L^1(\Omega)$ ορίζεται ως εξής: Είναι το σύνολο των σημείων $x \in \bar{\Omega}$ με την ιδιότητα ότι για κάθε $\delta > 0$ υπάρχει μια συνάρτηση $\varphi \in C_0(x - \delta, x + \delta)$, τέτοια ώστε

$$\int_{x-\delta}^{x+\delta} u(y)\varphi(y)\, dy \neq 0.$$

Εύκολα πείθεται κανείς ότι ο παρών ορισμός γενικεύει εκείνον που έχουμε ήδη δώσει για τον φορέα συνεχών συναρτήσεων.

Εξομαλυντές. Θεωρούμε πρώτα τη συνάρτηση $\omega : \mathbb{R} \to \mathbb{R}$,

$$\omega(x) := \begin{cases} k \cdot \mathrm{e} \cdot \mathrm{e}^{-\frac{1}{1-x^2}}, & |x| < 1, \\ 0, & |x| \geq 1, \end{cases} \tag{2.14}$$

με τη θετική σταθερά k τέτοια ώστε το ολοκλήρωμα της ω στο $\mathbb{R}$ να ισούται με τη μονάδα. Προφανώς, $\operatorname{supp} \omega = [-1, 1]$. Για $x \in (-1, 1)$ η ω είναι άπειρες φορές συνεχώς παραγωγίσιμη και, για $j \in \mathbb{N}_0$,

$$\omega^{(j)}(x) = k \,\mathrm{e}\, \frac{P_j(x)}{(1-x^2)^{j+1}}\, \mathrm{e}^{-\frac{1}{1-x^2}},$$

με κατάλληλα πολυώνυμα P_j. Χρησιμοποιώντας το γεγονός ότι $s^j \mathrm{e}^{-s} \to 0,\ s \to \infty$, διαπιστώνουμε ότι $\omega^{(j)}(x) \to 0$, τόσο για $x \to 1-$ όσο και για $x \to 1+$, και οδηγούμεθα στο συμπέρασμα ότι η ω είναι άπειρες φορές συνεχώς παραγωγίσιμη. Σημειώνουμε ακόμα ότι

$$\omega(0) = k \ \text{ και } \ 0 \leq \omega(x) \leq k, \ \text{ για κάθε } \ x \in \mathbb{R}.$$

Για $\varepsilon > 0$, θεωρούμε τώρα τη συνάρτηση

$$\varphi_\varepsilon : \mathbb{R} \to \mathbb{R}, \quad \varphi_\varepsilon(x) := \frac{1}{\varepsilon} \omega\big(\frac{x}{\varepsilon}\big).$$

Αμέσως διαπιστώνουμε ότι $\varphi_\varepsilon \in C^\infty(\mathbb{R})$, $\operatorname{supp} \varphi_\varepsilon = [-\varepsilon, \varepsilon]$, $\varphi_\varepsilon(0) = k/\varepsilon$, $0 \leq \varphi_\varepsilon(x) \leq k/\varepsilon$, για κάθε $x \in \mathbb{R}$, και το ολοκλήρωμα της φ_ε στο $\mathbb{R}$ ισούται με τη μονάδα.

Οι συναρτήσεις φ_ε, όπως και κάθε άλλη συνάρτηση με τις τρεις θεμελιώδεις ιδιότητες α) να είναι μη αρνητικές και να έχουν συμπαγή φορέα, β) να είναι άπειρες φορές παραγωγίσιμες συναρτήσεις και γ) να έχουν ολοκλήρωμα ίσο με τη μονάδα, καλούνται *εξομαλυντές*. Αν για μια συνάρτηση $u : \mathbb{R} \to \mathbb{R}$ υπάρχει το ολοκλήρωμα

$$\int_{\mathbb{R}} u(y) \varphi_\varepsilon(x - y)\, dy,$$

για κάθε $x \in \mathbb{R}$, όπως συμβαίνει για $u \in L^1_{\text{loc}}(\mathbb{R})$, τότε η *συνέλιξη* $\varphi_\varepsilon \star u$,

$$(\varphi_\varepsilon \star u)(x) := \int_{\mathbb{R}} u(y) \varphi_\varepsilon(x - y)\, dy,$$

λέγεται *εξομάλυνση* ή *ομαλοποίηση* της u.

Σημειώνουμε ότι η συνέλιξη $u \star v$ ορίζεται γενικά για δύο συναρτήσεις $u, v : \mathbb{R} \to \mathbb{R}$,

$$(u \star v)(x) := \int_{\mathbb{R}} u(x-y)v(y)\,dy,$$

και ότι $u \star v = v \star u$. Οι μαθηματικοί αρέσκονται να αναφέρονται στη συνέλιξη ως τον τέλειο γάμο, αφού το 'παιδί' $u \star v$ κληρονομεί 'οποιαδήποτε' καλή ιδιότητα έχει ένας εκ των γονέων u και v.

Η εξομάλυνση δεν είναι απαραίτητο να γίνεται με τις συναρτήσεις φ_ε, αλλά μπορεί να γίνεται με οποιονδήποτε άλλον εξομαλυντή. Εμείς εδώ θα χρησιμοποιούμε τους εξομαλυντές φ_ε.

Στη συνέχεια αναφέρουμε ορισμένες ιδιότητες της εξομάλυνσης.

Θεώρημα 2.6 (Ιδιότητες εξομαλύνσεων.) *Έστω $\Omega \subset \mathbb{R}$ και $u : \Omega \to \mathbb{R}$ μια συνάρτηση. Επεκτείνουμε τη u σε όλο το $\mathbb{R}$ με $u(x) := 0$, για κάθε $x \notin \Omega$. Τότε ισχύουν:*

i. *Αν $u \in L^1_{\text{loc}}(\mathbb{R})$, τότε $\varphi_\varepsilon \star u \in C^\infty(\mathbb{R})$.*

ii. *Αν $u \in L^1_{\text{loc}}(\mathbb{R})$ και ο φορέας της u είναι κλειστός και φραγμένος και περιέχεται στο Ω, τότε $\varphi_\varepsilon \star u \in C^\infty(\Omega)$ και, για ε μικρότερο από την απόσταση μεταξύ του φορέα της u και του συνόρου του Ω, ο φορέας της συνέλιξης έχει τις ίδιες ιδιότητες.*

iii. *Αν $u \in L^p(\Omega)$ και $1 \le p < \infty$, τότε $\varphi_\varepsilon \star u \in L^p(\Omega)$. Επί πλέον*

$$\|\varphi_\varepsilon \star u\|_p \le \|u\|_p \quad \text{και} \quad \lim_{\varepsilon \to 0+} \|\varphi_\varepsilon \star u - u\|_p = 0.$$

iv. *Αν $u \in C(\Omega)$, τότε $\lim_{\varepsilon \to 0+}(\varphi_\varepsilon \star u)(x) = u(x)$ ομοιόμορφα σε κάθε κλειστό και φραγμένο υποσύνολο του Ω.*

v. *Αν $u \in C(\bar{\Omega})$, τότε $\lim_{\varepsilon \to 0+}(\varphi_\varepsilon \star u)(x) = u(x)$ ομοιόμορφα στο Ω.* □

Μια ενδιαφέρουσα απόρροια των $ii.$ και $iii.$ του Θεωρήματος 2.6 είναι το ακόλουθο Πόρισμα, παράβαλε με την Πρόταση 2.2, για τον χώρο $C_0^\infty(\Omega)$, που αποτελείται από τα στοιχεία του $C_0(\Omega)$ που είναι άπειρες φορές συνεχώς παραγωγίσιμα.

Πόρισμα 2.2 (Πυκνότητα του $C_0^\infty(\Omega)$ στους $L^p(\Omega)$.) *Για $1 \le p < \infty$, ο χώρος $C_0^\infty(\Omega)$ είναι πυκνός στον $L^p(\Omega)$.* □

2.3 Χώροι του Sobolev

Έχοντας πλέον εξοικειωθεί με τους χώρους L^p, προχωρούμε στο παρόν εδάφιο στο κύριο αντικείμενο αυτού του κεφαλαίου, την εισαγωγή των χώρων του Sobolev σε μία διάσταση. Ουσιαστικό ρόλο στους χώρους του Sobolev παίζουν οι λεγόμενες *γενικευμένες* παράγωγοι. Στο εδάφιο που ακολουθεί θα ασχοληθούμε λοιπόν με γενικευμένες παραγώγους και μετά θα είμαστε σε θέση να μιλήσουμε για χώρους του Sobolev.

2.3.1 Γενικευμένες παράγωγοι

Οι χώροι του Sobolev σε μία διάσταση μπορούν να ορισθούν σε οποιοδήποτε ανοικτό υποσύνολο του $\mathbb{R}$. Στις εφαρμογές, όμως, τα προβλήματα εμφανίζονται σε διαστήματα I, $I=(a,b)$ ή $I=(-\infty,b)$ ή $I=(a,\infty)$ ή $I=\mathbb{R}$, με $a,b\in\mathbb{R}$. Κατά συνέπεια θα περιοριστούμε εδώ στην περίπτωση τέτοιων χωρίων, και μόνο σε κάποια παραδείγματα θα παρεκκλίνουμε από αυτόν τον κανόνα.

Συμβολίζουμε με $C_0^\infty(I)$ το σύνολο των συναρτήσεων, που ορίζονται στο I, είναι άπειρες φορές παραγωγίσιμες και έχουν συμπαγή φορέα που περιέχεται στο I. Ο $C_0^\infty(I)$ λέγεται *χώρος δοκιμής*, ή *χώρος ελέγχου*, και τα στοιχεία του λέγονται *συναρτήσεις δοκιμής*. Συχνά ο χώρος $C_0^\infty(I)$ συμβολίζεται με $\mathscr{D}(I)$. Από το $ii.$ του Θεωρήματος 2.6 έπεται ιδιαίτερα ότι ο χώρος $C_0^\infty(I)$ περιέχει και μη τετριμμένα στοιχεία. Για αυτό το γεγονός μπορεί να πεισθεί κανείς και ως εξής: Για $x^\star\in\mathbb{R}$, θεωρούμε τη συνάρτηση $\psi_\varepsilon:\mathbb{R}\to\mathbb{R}$, $\psi_\varepsilon(x):=\omega\big((x-x^\star)/\varepsilon\big)$. Η ψ_ε έχει κατά τα λοιπά τις ιδιότητες του εξομαλυντή φ_ε, αλλά ο φορέας της είναι το διάστημα $[x^\star-\varepsilon,x^\star+\varepsilon]$. Τέλος, για $x^\star\in I$ και αρκετά μικρό ε, διαπιστώνουμε αμέσως ότι ο περιορισμός της συνάρτησης ψ_ε στο I ανήκει στον χώρο $C_0^\infty(I)$.

Έστω $u\in C^1(I)$. Τότε μπορούμε να ορίσουμε κατά τα γνωστά την παράγωγο της u σε κάθε σημείο του I. Αυτό δεν μπορεί να γίνει με συναρτήσεις, φερ' ειπείν, $u\in L^1_{\text{loc}}(I)$, αφού ούτε καν για τις τιμές τους σε κάθε σημείο $x\in I$ δεν μπορούμε να μιλήσουμε. Για να οδηγηθούμε σε μια γενίκευση της έννοιας της παραγώγου, ας δούμε κατ' αρχάς μια άλλη ιδιότητα της u', πάντα για $u\in C^1(I)$, στην οποία υπεισέρχονται μόνο ολοκληρώματα, και όχι τιμές σε σημεία.

Έστω λοιπόν $u\in C^1(I)$. Με ολοκλήρωση κατά μέρη, αμέσως διαπιστώνουμε ότι

$$\int_I u'\varphi\,dx=-\int_I u\varphi'\,dx\quad\forall\varphi\in C_0^\infty(I). \tag{2.15}$$

Είναι εύκολο να πεισθεί κανείς ότι η σχέση (2.15) χαρακτηρίζει τη u'. Πράγματι, έστω $v \in C(I)$ τέτοια ώστε

$$\int_I v\varphi\,dx = -\int_I u\varphi'\,dx \quad \forall \varphi \in C_0^\infty(I). \tag{2.16}$$

Από τις (2.15) και (2.16) έπεται ότι

$$\int_I (u'-v)\varphi\,dx = 0 \quad \forall \varphi \in C_0^\infty(I). \tag{2.17}$$

Θέλουμε να αποδείξουμε ότι $u' = v$. Υποθέτουμε λοιπόν ότι για κάποιο $x^\star \in I$ ισχύει $u'(x^\star) \neq v(x^\star)$, και μάλιστα χωρίς περιορισμό της γενικότητας ότι $u'(x^\star) > v(x^\star)$, και θα οδηγηθούμε σε άτοπο. Κατ' αρχάς, αφού οι u' και v είναι συνεχείς, υπάρχει $\delta > 0$ τέτοιο ώστε $u'(x) > v(x)$, για κάθε $x \in (x^\star - \delta, x^\star + \delta)$. Επιλέγοντας τη συνάρτηση δοκιμής φ έτσι ώστε να παίρνει μόνο μη αρνητικές τιμές και να έχει φορέα το διάστημα $[x^\star - \delta, x^\star + \delta]$, και αυτή η επιλογή είναι εφικτή, όπως έχουμε ήδη δει, διαπιστώνουμε ότι

$$\int_I (u'-v)\varphi\,dx = \int_{x^\star-\delta}^{x^\star+\delta} (u'-v)\varphi\,dx > 0,$$

άτοπο, αφού αντίκειται στη (2.17). Συνεπώς, η (2.15) χαρακτηρίζει όντως τη u'.

Η (2.15) αποτελεί το κίνητρο για τον ακόλουθο ορισμό.

Ορισμός 2.8 (Γενικευμένες παράγωγοι.) Έστω $u, v \in L^1_{\text{loc}}(I)$. Λέμε ότι η v είναι *γενικευμένη* (ή *ασθενής*) (πρώτη) παράγωγος της u, αν

$$\int_I u\varphi'\,dx = -\int_I v\varphi\,dx \quad \forall \varphi \in C_0^\infty(I). \tag{2.18}$$

Σχόλιο. (Μοναδικότητα γενικευμένων παραγώγων.) Φυσικά, είναι δυνατόν είτε να υπάρχει είτε να μην υπάρχει η ασθενής παράγωγος μιας συνάρτησης του $L^1_{\text{loc}}(I)$. Πάντως αποδεικνύεται μοναδικότητα της ασθενούς παραγώγου. Πράγματι, αν $v, \tilde{v} \in L^1_{\text{loc}}(I)$ ασθενείς παράγωγοι μιας συνάρτησης του $L^1_{\text{loc}}(I)$, τότε

$$\int_I (v-\tilde{v})\varphi\,dx = 0 \quad \forall \varphi \in C_0^\infty(I). \tag{2.19}$$

Θέτουμε $w := v - \tilde{v}$ και αρκεί να αποδείξουμε ότι η w μηδενίζεται σε κάθε υποδιάστημα $I_1 = (a_1, b_1)$ του I. Θεωρούμε προς τούτο ένα διάστημα $I_2 = (a_2, b_2)$ με

$a_2, b_2 \in I$, $a_2 < a_1$ και $b_2 > b_1$. Θέτουμε $\hat{w} := w \cdot \chi_{I_2}$, όπου χ_{I_2} η χαρακτηριστική συνάρτηση του I_2. Έστω $\varepsilon_1 := \min(a_1 - a_2, b_2 - b_1)$. Εύκολα διαπιστώνουμε ότι για $\varepsilon < \varepsilon_1$ και $x \in I_1$ ισχύει

$$\begin{aligned}(\varphi_\varepsilon \star \hat{w})(x) &= \int_{\mathbb{R}} \hat{w}(y)\varphi_\varepsilon(x-y)\,dy = \int_{x-\varepsilon}^{x+\varepsilon} w(y)\varphi_\varepsilon(x-y)\,dy \\ &= \int_I w(y)\varphi_\varepsilon(x-y)\,dy.\end{aligned}$$

Όμως, για τα x και ε που αναφέρθηκαν, ισχύει $\varphi_\varepsilon(x-\cdot) \in C_0^\infty(I)$, οπότε, σύμφωνα με τη (2.19), $(\varphi_\varepsilon \star w)(x) = 0$, δηλαδή

$$(\varphi_\varepsilon \star \hat{w})(x) = 0 \quad \forall x \in I_1. \tag{2.20}$$

Επομένως,

$$\int_{I_1} |w|\,dx = \int_{I_1} |\varphi_\varepsilon \star \hat{w} - \hat{w}|\,dx \le \int_{I_2} |\varphi_\varepsilon \star \hat{w} - \hat{w}|\,dx.$$

Αλλά, $\hat{w} \in L^1(I_2)$, οπότε, σύμφωνα με το $iii.$ του Θεωρήματος 2.6, το δεξιό μέλος αυτής της ανισότητας τείνει στο μηδέν καθώς το ε τείνει στο μηδέν. Συνεπώς, η w μηδενίζεται στο I_1. Σημειώνουμε ότι, αν κανείς υποθέσει ότι η ασθενής παράγωγος ανήκει σε κάποιον χώρο $L^p(I)$, με $1 < p < \infty$, τότε η απόδειξη μπορεί να απλοποιηθεί πολύ, βλ. την Άσκηση 2.10. □

Πόρισμα 2.3 (Μεταβολική σχέση που ισχύει μόνο για τη μηδενική συνάρτηση.) *Έστω* $v \in L^1_{\text{loc}}(I)$ *μια συνάρτηση τέτοια ώστε*

$$\int_I v\varphi\,dx = 0 \quad \forall \varphi \in C_0^\infty(I).$$

Τότε $v = 0$. □

Με παρόμοιο τρόπο ορίζεται και η παράγωγος $v \in L^1_{\text{loc}}(I)$, οποιασδήποτε τάξης $j \in \mathbb{N}$, μιας συνάρτησης $u \in L^1_{\text{loc}}(I)$ ως το μοναδικό στοιχείο, αν υπάρχει, τέτοιο ώστε

$$\int_I u\varphi^{(j)}\,dx = (-1)^j \int_I v\varphi\,dx \quad \forall \varphi \in C_0^\infty(I). \tag{2.21}$$

Σημειώνουμε ότι η ασθενής δεύτερη παράγωγος, φερ' ειπείν, δεν ορίζεται ως η πρώτη παράγωγος της πρώτης παραγώγου, αλλά κατ' ευθείαν. Αυτό είναι ιδιαίτερα σημαντικό στις περισσότερες διαστάσεις, όπου είναι παραδείγματος χάριν δυνατόν να υπάρχει η u_{xy} χωρίς να υπάρχει καμμία από τις u_x και u_y.

Τονίζουμε επίσης ότι στον ορισμό (2.21) οι συναρτήσεις δοκιμής θα μπορούσαν εξ ίσου καλά να ληφθούν ως συναρτήσεις j φορές συνεχώς παραγωγίσιμες με συμπαγή φορέα, δηλαδή ως στοιχεία του χώρου $C_0^j(I)$. Στη (2.18) ο αντίστοιχος χώρος δοκιμής θα ήταν φυσικά ο $C_0^1(I)$. Βλ. την Άσκηση 2.11.

Παραδείγματα 2.1 Δίνουμε τώρα τρία παραδείγματα ασθενών παραγώγων.

1. Έστω $I := (0,2)$ και $u : I \to \mathbb{R}$,

$$u(x) := \begin{cases} x, & \text{αν } 0 < x \le 1, \\ 1, & \text{αν } 1 < x < 2. \end{cases}$$

Τότε η ασθενής παράγωγος u' της u είναι

$$u'(x) = \begin{cases} 1, & \text{αν } 0 < x \le 1, \\ 0, & \text{αν } 1 < x < 2. \end{cases}$$

Πράγματι, για $\varphi \in C_0^\infty(I)$, έχουμε

$$\begin{aligned}\int_0^2 u\varphi'\,dx &= \int_0^1 x\varphi'(x)\,dx + \int_1^2 \varphi'(x)\,dx \\ &= \big[x\varphi(x)\big]_{x=0}^{x=1} - \int_0^1 \varphi(x)\,dx + \varphi(2) - \varphi(1) \\ &= \varphi(1) - \int_0^1 \varphi(x)\,dx - \varphi(1) \\ &= -\int_0^1 \varphi(x)\,dx = -\int_0^2 u'\varphi(x)\,dx.\end{aligned}$$

Σημειώστε ότι αν τροποποιήσουμε τη συνάρτηση u, αλλάζοντας τις τιμές της σε πεπερασμένου πλήθους σημεία (ή ακόμα και σε αριθμήσιμο πλήθος σημείων), αυτή δεν αλλάζει ως στοιχείο του L^1_{loc}, συνεπώς εξακολουθεί να έχει ασθενή παράγωγο, και μάλιστα την ίδια όπως και πριν την τροποποίηση.

Σημειώστε ακόμη ότι η ασθενής παράγωγος μιας *συνεχούς*, κατά τμήματα συνεχώς παραγωγίσιμης συνάρτησης σε ένα διάστημα I συμπίπτει με την κλασική παράγωγο, εκεί όπου η κλασική παράγωγος ορίζεται (τα εναπομείναντα σημεία έχουν μέτρο μηδέν και δεν μας ενοχλούν).

2. Έστω $I := (0,2)$ και $u : I \to \mathbb{R}$,

$$u(x) := \begin{cases} x, & \text{αν } 0 < x \le 1, \\ 2, & \text{αν } 1 < x < 2. \end{cases}$$

Αυτή η συνάρτηση *δεν* έχει ασθενή παράγωγο στον $L^1_{\text{loc}}(0,2)$. Για να το δούμε αυτό, υποθέτουμε ότι μια συνάρτηση $v \in L^1_{\text{loc}}(0,2)$ είναι ασθενής παράγωγος της u, δηλαδή ότι

$$\int_0^2 u\varphi'\,dx = -\int_0^2 v\varphi\,dx \quad \forall \varphi \in C_0^\infty(0,2), \tag{2.22}$$

και θα οδηγηθούμε σε άτοπο. Από τη (2.22) προκύπτει, με ολοκλήρωση κατά μέρη,

$$-\int_0^2 v\varphi\,dx = \int_0^2 u\varphi'\,dx = \int_0^1 x\varphi'(x)\,dx + 2\int_1^2 \varphi'(x)\,dx = -\int_0^1 \varphi(x)\,dx - \varphi(1),$$

δηλαδή

$$\int_0^1 v\varphi\,dx = \int_0^1 \varphi(x)\,dx + \varphi(1), \quad \forall \varphi \in C_0^\infty(0,2). \tag{2.23}$$

Με τη συνάρτηση ω της (2.14), θεωρούμε τώρα τις συναρτήσεις ω_m,

$$\omega_m(x) := \frac{1}{k}\omega\big(m(x-1)\big), \quad m \geq 2.$$

Αμέσως διαπιστώνουμε ότι $\omega_m(1) = 1$, $0 \leq \omega_m(x) \leq 1$, για κάθε $x \in (0,2)$, και $\operatorname{supp}\omega_m = [1-\frac{1}{m}, 1+\frac{1}{m}]$. Τώρα, σύμφωνα με τη (2.23), έχουμε, για $\varphi := \omega_m$,

$$\omega_m(1) = \int_0^2 v\omega_m\,dx - \int_0^1 \omega_m\,dx$$

ή

$$1 = \int_0^2 v\omega_m\,dx - \int_0^1 \omega_m\,dx.$$

Αλλά,

$$\lim_{m\to\infty}\int_0^2 v\omega_m\,dx = 0 \quad \text{και} \quad \lim_{m\to\infty}\int_0^1 \omega_m\,dx = 0,$$

οπότε η προηγούμενη σχέση δίνει $1 = 0$, άτοπο.

Σημειώστε ότι παρόμοια κατάσταση έχουμε όταν η συνάρτηση ορίζεται σε ένα διάστημα και δεν είναι σχεδόν παντού ίση με μια συνεχή συνάρτηση. Προσέξτε ότι αυτό είναι εντελώς διαφορετικό από το να είναι η συνάρτηση σχεδόν παντού συνεχής. Σε περισσότερες διαστάσεις υπάρχουν συναρτήσεις που δεν είναι σχεδόν παντού ίσες με μια συνεχή συνάρτηση, σε ένα ορθογώνιο, φερ' ειπείν, και παρά ταύτα έχουν ασθενείς παραγώγους πρώτης τάξεως.

3. Έστω $I := (0,2) \setminus \{1\}$ και $u : I \to \mathbb{R}$,

$$u(x) := \begin{cases} x, & \text{αν } 0 < x < 1, \\ 2, & \text{αν } 1 < x < 2. \end{cases}$$

Η συνάρτηση αυτή είναι άπειρες φορές συνεχώς παραγωγίσιμη στο I, ιδιαίτερα λοιπόν έχει ασθενείς παραγώγους κάθε τάξης. Σημειώστε ότι, αντίθετα με ότι συνέβαινε στο προηγούμενο παράδειγμα, εδώ όλες οι συναρτήσεις δοκιμής μηδενίζονται στο σημείο 1.

2.3.2 Χώροι του Sobolev

Έστω I ένα διάστημα, όπως στην αρχή της ενότητας 2.3.1. Έστω $p \in [1,\infty]$ και $k \in \mathbb{N}$.

Ορισμός 2.9 (Χώροι του Sobolev.) Ο *χώρος του Sobolev* $W^{k,p} := W^{k,p}(I)$ αποτελείται από τις συναρτήσεις $u \in L^p(I)$ που έχουν ασθενείς παραγώγους $u', u'', \dots, u^{(k)}$, τάξης έως και k, και όλες αυτές οι παράγωγοι ανήκουν στον $L^p(I)$. Στον $W^{k,p}$ ορίζουμε τη νόρμα $\|\cdot\|_{k,p}$,

$$\|u\|_{k,p} := \Big(\sum_{j=0}^{k} \|u^{(j)}\|_{L^p}^p\Big)^{1/p}, \quad \text{αν } p < \infty,$$

και, για $p = \infty$,

$$\|u\|_{k,\infty} := \sum_{j=0}^{k} \|u^{(j)}\|_{L^\infty}.$$

Ότι η $\|\cdot\|_{k,p}$ είναι όντως νόρμα στον $W^{k,p}$ αποδεικνύεται εύκολα, με την ανισότητα του Minkowski στον $\mathbb{R}^{k+1}$, αν χρησιμοποιήσουμε το γεγονός ότι η $\|\cdot\|_{L^p}$ (όπως γράφουμε τώρα αντί της $\|\cdot\|_p$ για σαφήνεια) είναι νόρμα στον $L^p(I)$, βλ. την Άσκηση 2.9.

Για $p = 2$ γράφουμε συνήθως H^k αντί για $W^{k,2}$. Ο λόγος είναι ότι, όπως θα δούμε στη συνέχεια, αυτοί οι χώροι είναι χώροι Hilbert. Το αντίστοιχο εσωτερικό γινόμενο $(\cdot,\cdot)_k$ δίνεται ως εξής

$$(u,v)_k := \sum_{j=0}^{k} (u^{(j)}, v^{(j)}), \quad \forall u, v \in H^k,$$

όπου $(\cdot,\cdot)$ το εσωτερικό γινόμενο στον $L^2(I)$,

$$(u,v) := \int_I uv\,dx.$$

Επίσης, θα συμβολίζουμε για απλότητα με $\|\cdot\|_k$, αντί για $\|\cdot\|_{k,2}$, τη νόρμα του H^k, στη συνέχεια· κίνδυνος να προκληθεί σύγχυση με τη νόρμα του L^p δεν θα υπάρξει, αφού τα συμφραζόμενα δεν θα αφήνουν τέτοια περιθώρια. Σημειώνουμε πάντως ότι ιστορικά ορίστηκαν οι χώροι $W^{k,p}$ και $H^{k,p}$ με διαφορετικούς τρόπους, απεδείχθη όμως εκ των υστέρων ότι αυτοί οι ορισμοί οδηγούσαν στους ίδιους χώρους.

Θεώρημα 2.7 (Πληρότητα των χώρων του Sobolev.) *Οι χώροι του Sobolev $W^{k,p} := W^{k,p}(I)$, $k \in \mathbb{N}$ και $p \in [1, \infty]$, είναι πλήρεις.*

Απόδειξη. Έστω $(u_m)_{m\in\mathbb{N}} \subset W^{k,p} := W^{k,p}(I)$ μια ακολουθία Cauchy, δηλαδή ακολουθία Cauchy ως προς τη νόρμα $\|\cdot\|_{k,p}$. Τότε, για κάθε $j \in \{0, \dots, k\}$, η ακολουθία $\big(u_m^{(j)}\big)_{m\in\mathbb{N}}$ είναι, προφανώς ακολουθία Cauchy στον $L^p(I)$. Λόγω της πληρότητας του $L^p(I)$ υπάρχουν επομένως συναρτήσεις $v_j \in L^p(I)$ τέτοιες ώστε

$$u_m^{(j)} \to v_j, \quad m \to \infty, \quad \text{στον} \quad L^p(I).$$

Ιδιαίτερα, λοιπόν,

$$u_m \to v_0 =: u, \quad m \to \infty, \quad \text{στον} \quad L^p(I).$$

Θα αποδείξουμε τώρα ότι $u \in W^{k,p}$ και ότι

$$u_m \to u, \quad m \to \infty, \quad \text{στον} \quad W^{k,p}(I),$$

και κατ' αυτόν τον τρόπο θα ολοκληρωθεί η απόδειξη της πληρότητας.

Έστω λοιπόν $\varphi \in C_0^\infty(I)$ μια συνάρτηση δοκιμής. Τότε, για $j \in \{1, \dots, k\}$,

$$\int_I u_m \varphi^{(j)}\, dx = (-1)^j \int_I u_m^{(j)} \varphi\, dx.$$

Αφήνοντας εδώ το m να τείνει στο άπειρο, το αριστερό και το δεξιό μέλος αυτής της σχέσης τείνουν, αντίστοιχα, στα

$$\int_I u\varphi^{(j)}\, dx \quad \text{και} \quad (-1)^j \int_I v_j \varphi\, dx,$$

βλ. και την Άσκηση 2.5, έχουμε δηλαδή ότι

$$\int_I u\varphi^{(j)}\, dx = (-1)^j \int_I v_j \varphi\, dx \quad \forall \varphi \in C_0^\infty(I).$$

Επομένως, υπάρχει η ασθενής παράγωγος $u^{(j)}$ της u, για $j \in \{1, \dots, k\}$, και μάλιστα $u^{(j)} = v_j$, $j = 1, \dots, k$. Χρησιμοποιώντας τώρα το γεγονός ότι $v_j \in L^p(I)$, διαπιστώνουμε αμέσως ότι $u^{(j)} \in L^p(I)$, $j = 0, \dots, k$, οπότε, φυσικά, $u \in W^{k,p}$. Επί πλέον, για $1 \le p < \infty$,

$$\|u - u_m\|_{k,p}^p = \sum_{j=0}^k \|u^{(j)} - u_m^{(j)}\|_{L^p}^p = \sum_{j=0}^k \|v_j - u_m^{(j)}\|_{L^p}^p \to 0, \quad m \to \infty,$$

και, παρόμοια, για $p = \infty$,

$$\|u - u_m\|_{k,\infty} = \sum_{j=0}^{k} \|u^{(j)} - u_m^{(j)}\|_{L^\infty} = \sum_{j=0}^{k} \|v_j - u_m^{(j)}\|_{L^\infty} \to 0, \quad m \to \infty,$$

δηλαδή $u_m \to u$, $m \to \infty$, στον $W^{k,p}(I)$, και η απόδειξη είναι πλήρης. □

Εύκολα πείθεται κανείς για τις ακόλουθες ιδιότητες ασθενών παραγώγων· η απόδειξη αυτών των ιδιοτήτων αφήνεται ως Άσκηση.

Πρόταση 2.4 (Στοιχειώδεις ιδιότητες ασθενών παραγώγων.) *Έστω $u, v \in W^{k,p}(I)$ και $j \in \{0, \dots, k\}$. Τότε ισχύουν*

i. $u^{(j)} \in W^{k-j,p}(I)$ *και* $\big(u^{(j)}\big)^{(i)} = \big(u^{(i)}\big)^{(j)} = u^{(j+i)}$, *για κάθε* $i \in \{0, \dots, k-j\}$.

ii. Για $\lambda, \mu \in \mathbb{R}$, *ισχύει* $\lambda u + \mu v \in W^{k,p}(I)$ *και* $(\lambda u + \mu v)^{(j)} = \lambda u^{(j)} + \mu v^{(j)}$, *για* $j \in \{0, \dots, k\}$.

iii. Αν J είναι ένα ανοικτό υποσύνολο (υποδιάστημα) του I, τότε ο περιορισμός της u στο J ανήκει στον $W^{k,p}(J)$.

iv. Αν $\zeta : I \to \mathbb{R}$ είναι μια συνάρτηση, k φορές παραγωγίσιμη στο I, και τέτοια ώστε οι παράγωγοι $\zeta^{(j)}$, $j = 0, \dots, k$, να είναι φραγμένες, τότε $\zeta u \in W^{k,p}(I)$ και οι ασθενείς παράγωγοι της ζu δίνονται από τον τύπο του Leibniz,

$$(\zeta u)^{(j)} = \sum_{i=0}^{j} \binom{j}{i} \zeta^{(i)}\, u^{(j-i)}, \quad j = 1, \dots, k.$$ □

Προχωρούμε τώρα κατ' αρχάς σε δύο προκαταρκτικά αποτελέσματα για να προετοιμάσουμε το έδαφος για να αποδείξουμε στη συνέχεια ένα βασικό θεώρημα.

Λήμμα 2.1 (Μηδενική ασθενής παράγωγος συνεπάγεται σταθερή συνάρτηση.) *Έστω $I = (a, b)$ ένα διάστημα και μια συνάρτηση $u \in L^1_{\text{loc}}(I)$ τέτοια ώστε*

$$\int_I u\varphi'\, dx = 0 \quad \forall \varphi \in C_0^\infty(I).$$

Τότε υπάρχει σταθερά C τέτοια ώστε $u(x) = C$ σ.π. στο I.

Απόδειξη. Έστω $\psi \in C_0^\infty(I)$ μια συνάρτηση τέτοια ώστε $\int_I \psi(x)\, dx = 1$. Θέτουμε

$$C := \int_I u(x)\psi(x)\, dx$$

και θα αποδείξουμε ότι $u(x) = C$ σ.π. στο I. Για $w \in C_0^\infty(I)$, ορίζουμε τις συναρτήσεις h και φ,

$$h := w - \int_I w(s)\,ds \cdot \psi, \quad \varphi(x) := \int_a^x h(s)\,ds.$$

Προφανώς, $h \in C_0^\infty(I)$ και, για κατάλληλα $a_1, b_1 \in (a,b)$, $\operatorname{supp} h \subset [a_1, b_1] \subset I$. Αμέσως διαπιστώνει κανείς ότι

$$\varphi(x) = 0 \quad \text{για} \;\; x \in (a, a_1) \cup (b_1, b).$$

Ιδιαίτερα, $\varphi \in C_0^\infty(I)$. Επίσης, $\varphi'(x) = w(x) - \int_I w(s)\,ds \cdot \psi(x)$. Για αυτή την επιλογή της συνάρτησης δοκιμής, η υπόθεσή μας δίνει

$$\begin{aligned} 0 = \int_I u\varphi'\,dx &= \int_I u(x)w(x)\,dx - \int_I w(s)\,ds \cdot \int_I u(x)\psi(x)\,dx \\ &= \int_I u(x)w(x)\,dx - \int_I Cw(s)\,ds = \int_I [u(x) - C]w(x)\,dx, \end{aligned}$$

συνεπώς

$$\int_I (u - C)w\,dx = 0 \quad \forall w \in C_0^\infty(I).$$

Από το Πόρισμα 2.3 έπεται ότι $u(x) = C$ σ.π. στο I. □

Για $a < b$ συμβολίζουμε με $\int_b^a u\,dx$, όπως και στην περίπτωση του ολοκληρώματος του Riemann, το $-\int_a^b u\,dx$.

Λήμμα 2.2 (Ασθενής παράγωγος ολοκληρώματος.) *Έστω $I = (a,b)$ ένα διάστημα, $v \in L^1_{\text{loc}}(I)$ και $y_0 \in I$. Τότε, όπως είναι γνωστό από τη Θεωρία Μέτρου, η συνάρτηση $u : I \to \mathbb{R}$,*

$$u(x) := \int_{y_0}^x v(s)\,ds,$$

είναι συνεχής στο I. Επί πλέον,

$$\int_I u\varphi'\,dx = -\int_I v\varphi\,dx \quad \forall \varphi \in C_0^\infty(I),$$

δηλαδή η v είναι ασθενής παράγωγος της u.

Απόδειξη. Έστω $\varphi \in C_0^\infty(I)$. Τότε

$$\begin{aligned} 0 = \int_I u\varphi'\,dx &= \int_I \Big(\int_{y_0}^x v(s)\,ds \Big) \varphi'(x)\,dx \\ &= -\int_a^{y_0} \Big(\int_x^{y_0} v(s)\,ds \Big) \varphi'(x)\,dx + \int_{y_0}^b \Big(\int_{y_0}^x v(s)\,ds \Big) \varphi'(x)\,dx. \end{aligned}$$

Τώρα, ο πρώτος όρος στο δεξιό μέλος αυτής της σχέσης γράφεται στη μορφή

$$-\int_a^{y_0}\Big(\int_x^{y_0} v(s)\,ds\Big)\varphi'(x)\,dx = -\int_a^{y_0}\Big(\int_x^{y_0} v(s)\varphi'(x)\,ds\Big)\,dx.$$

Σύμφωνα με ένα βασικό αποτέλεσμα της Θεωρίας Μέτρου, το θεώρημα του Fubini, το δεξιό μέλος είναι το διπλό ολοκλήρωμα στο ορθογώνιο τρίγωνο με κορυφές τα σημεία $(a,a),(a,y_0)$ και (y_0,y_0). Το ίδιο θεώρημα μας επιτρέπει να αλλάξουμε τη σειρά ολοκλήρωσης, οπότε έχουμε

$$\begin{aligned}-\int_a^{y_0}\Big(\int_x^{y_0} v(s)\,ds\Big)\varphi'(x)\,dx &= -\int_a^{y_0}\Big(\int_a^{s} v(s)\varphi'(x)\,dx\Big)\,ds\\ &= -\int_a^{y_0} v(s)\Big(\int_a^{s}\varphi'(x)\,dx\Big)\,ds = -\int_a^{y_0} v(s)\varphi(s)\,ds.\end{aligned}$$

Παρόμοια, για τον δεύτερο όρο έχουμε

$$\int_{y_0}^{b}\Big(\int_{y_0}^{x} v(s)\,ds\Big)\varphi'(x)\,dx = -\int_{y_0}^{b} v(s)\varphi(s)\,ds.$$

Από τα προηγούμενα έπεται αμέσως το ζητούμενο. □

Σχόλιο. (Το ολοκλήρωμα συνάρτησης του L^p ανήκει στον $W^{1,p}$.) Έστω $v \in L^1(I)$. Αν η συνάρτηση u που προκύπτει με ολοκλήρωση της v, όπως στο Λήμμα 2.2, ανήκει στον $L^p(I)$ (και αυτό συμβαίνει πάντα, αν το I είναι φραγμένο), τότε, σύμφωνα με το Λήμμα 2.2, ανήκει και στον $W^{1,p}$, $u \in W^{1,p}$. □

Μετά τα δύο βοηθητικά Λήμματα ερχόμαστε τώρα σε ένα βασικό αποτέλεσμα. Σύμφωνα με αυτό, στην κλάση κάθε συνάρτησης του $W^{1,p}(I)$, όπου I ένα διάστημα, υπάρχει μια συνεχής συνάρτηση, ένας συνεχής εκπρόσωπος της u. Με άλλα λόγια η u είναι σχεδόν παντού ίση με μια συνεχή συνάρτηση (ιδιότητα πολύ ισχυρότερη από το να είναι η u σχεδόν παντού συνεχής). Αυτό, μεταξύ άλλων, μας παρέχει τη δυνατότητα να μιλάμε για την τιμή της u σε ένα σημείο (κάτι μη εφικτό για στοιχεία του $L^p(I)$) και με αυτό εννοούμε την τιμή του συνεχούς εκπροσώπου της u.

Θεώρημα 2.8 (Συνέχεια στοιχείων του $W^{1,p}$.) *Έστω $I=(a,b)$ ένα διάστημα και μια συνάρτηση $u \in W^{1,p}(I)$. Τότε υπάρχει ένας συνεχής αντιπρόσωπος στην κλάση της u, δηλαδή μια συνάρτηση $\tilde u$, συνεχής στο $\bar I$, τέτοια ώστε $\tilde u = u$ σ.π. στο I. Μάλιστα ισχύει*

$$\tilde u(x) - \tilde u(y) = \int_y^x u'(s)\,ds \quad \forall x,y \in \bar I.$$

Απόδειξη. Έστω $y_0 \in I$. Προφανώς ισχύει $u' \in L^p(I)$, οπότε, ιδιαίτερα $u \in L^1_{\text{loc}}(I)$. Ορίζουμε τη συνάρτηση $\bar{u} : I \to \mathbb{R}$ ως

$$\bar{u}(x) := \int_{y_0}^{x} u'(s)\, ds.$$

Σύμφωνα με το Λήμμα 2.2 έχουμε $\bar{u} \in C(I)$ και

$$\int_I \bar{u}\varphi'\, dx = -\int_I u'\varphi\, dx \quad \forall \varphi \in C_0^\infty(I).$$

Χρησιμοποιώντας εδώ τον ορισμό της u' διαπιστώνουμε αμέσως ότι

$$\int_I (\bar{u} - u)\varphi'\, dx = 0 \quad \forall \varphi \in C_0^\infty(I).$$

Επομένως, σύμφωνα με το Λήμμα 2.1, υπάρχει σταθερά C τέτοια ώστε $\bar{u} - u = C$ σ.π. στο I. Ορίζουμε τώρα τη συνάρτηση $\tilde{u} : I \to \mathbb{R}$, $\tilde{u}(x) := \bar{u}(x) - C$. Από τα προηγούμενα έπεται αμέσως ότι $\tilde{u} \in C(I)$ και $\tilde{u} = u$ σ.π. στο I. Επί πλέον, για $x, y \in I$,

$$\tilde{u}(x) - \tilde{u}(y) = \bar{u}(x) - \bar{u}(y) = \int_{y_0}^{x} u'(s)\, ds - \int_{y_0}^{y} u'(s)\, ds = \int_{y}^{x} u'(s)\, ds. \qquad \square$$

Σχόλιο. (Ομαλότητα συνόρου συνόλων.) Αν το σύνολο I δεν είναι διάστημα, το συμπέρασμα του Θεωρήματος 2.9 δεν είναι αναγκαστικά σωστό. Φερ' ειπείν, για $I := (0,2) \setminus \{1\}$, η συνάρτηση $u : I \to \mathbb{R}$,

$$u(x) := \begin{cases} x, & \text{για } 0 < x < 1, \\ 2, & \text{για } 1 < x < 2, \end{cases}$$

βλ. και το τρίτο από τα Παραδείγματα 2.1, είναι άπειρες φορές παραγωγίσιμη στο I, συνεπώς είναι και στοιχείο των $W^{k,p}(I)$, για κάθε $k \in \mathbb{N}$ και $p \in [1, \infty]$, αλλά, προφανώς, δεν υπάρχει αντιπρόσωπος στην κλάση του u που να είναι συνεχής στο $\bar{I} = [0, 2]$. Τέτοια σύνολα δεν έχουν ιδιαίτερη σημασία στις εφαρμογές· σκοπός μας είναι απλώς να προϊδεάσουμε τον αναγνώστη ότι απαιτούνται κάποιες συνθήκες. Αυτά τα θέματα είναι πολύ σημαντικά στις πολλές διαστάσεις, όπου θα απαιτηθούν συνθήκες ομαλότητας των συνόρων των χωρίων. Η "μη ομαλότητα" στο παράδειγμά μας σχετίζεται με το γεγονός ότι μέρος του I βρίσκεται και από τις δύο πλευρές του σημείου 1, δηλαδή τόσο αριστερά του όσο και δεξιά του.

Προχωρούμε τώρα σε ένα θεώρημα επέκτασης συναρτήσεων που ανήκουν σε έναν χώρο Sobolev σε ένα διάστημα κατά τρόπον ώστε η επέκταση να ανήκει στον χώρο Sobolev με τους ίδιους δείκτες στον $\mathbb{R}$. Αυτό θα είναι χρήσιμο στη συνέχεια, γιατί συναρτήσεις σε διαστήματα θα κληρονομούν παρόμοιες ιδιότητες στοιχείων αντίστοιχων χώρων Sobolev σε όλη την πραγματική ευθεία.

Θεώρημα 2.9 (Τελεστής επέκτασης.) *Έστω $I = (a, b)$ ένα διάστημα και $p \in [1, \infty]$. Τότε υπάρχει ένας γραμμικός τελεστής επέκτασης $E : W^{1,p}(I) \to W^{1,p}(\mathbb{R})$ τέτοιος ώστε, για κάθε $u \in W^{1,p}(I)$,*

i. *$(Eu)(x) = u(x)$ σ.π. στο I, δηλαδή η συνάρτηση Eu είναι όντως επέκταση της u.*

ii. *$\|Eu\|_{L^p(\mathbb{R})} \le C\|u\|_{L^p(I)}$, με σταθερά C ανεξάρτητη της u, δηλαδή ο τελεστής E είναι φραγμένος ως προς αυτές τις νόρμες.*

iii. *$\|Eu\|_{W^{1,p}(\mathbb{R})} \le C\|u\|_{W^{1,p}(I)}$, με σταθερά C ανεξάρτητη της u, δηλαδή ο τελεστής $E : W^{1,p}(I) \to W^{1,p}(\mathbb{R})$ είναι φραγμένος.*

Απόδειξη. Θα δώσουμε την απόδειξη για $1 \le p < \infty$, η περίπτωση $p = \infty$ αντιμετωπίζεται με παρόμοιο τρόπο. Θα διακρίνουμε διάφορες περιπτώσεις για το διάστημα I, αρχίζοντας με το $I = (0, \infty)$. Θα δούμε ότι με *άρτια ανάκλαση* στο σημείο 0, δηλαδή με τον ορισμό

$$(Eu)(x) := u^\star(x) := \begin{cases} u(x), & \text{για } x \ge 0, \\ u(-x), & \text{για } x < 0, \end{cases}$$

οδηγούμαστε σε τελεστή επέκτασης με τις επιθυμητές ιδιότητες. Κατ' αρχάς, η $i.$ είναι προφανής. Σημειώστε εδώ ότι η τιμή $u(0)$ ορίζεται, σύμφωνα με το Θεώρημα 2.8. Επί πλέον, έχουμε

$$\|u^\star\|^p_{L^p(\mathbb{R})} = \int_{-\infty}^0 |u(-x)|^p\,dx + \int_0^\infty |u(x)|^p\,dx = 2\int_0^\infty |u(x)|^p\,dx = 2\|u\|^p_{L^p(I)},$$

και διαπιστώνουμε ότι ισχύει η $ii.$ με $C = 2^{1/p}$. Ορίζουμε τώρα τη συνάρτηση $v : \mathbb{R} \to \mathbb{R}$,

$$v(x) := \begin{cases} u'(x), & \text{για } x > 0, \\ -\,u'(-x), & \text{για } x < 0, \end{cases}$$

στο μηδέν δεν δίνουμε τιμή, ούτως ή άλλως μας ενδιαφέρει η v ως στοιχείο του $L^p(\mathbb{R})$. Προφανώς, $v \in L^p(\mathbb{R})$. Επί πλέον, σύμφωνα με το Θεώρημα 2.8,

$$u^\star(x) - u(0) = \int_0^x u'(s)\,ds = \int_0^x v(s)\,ds, \quad \text{για } x > 0,$$

και

$$u^\star(x) - u(0) = \int_0^{-x} u'(s)\,ds = \int_0^x [-u'(-s)]\,ds = \int_0^x v(s)\,ds, \quad \text{για } x < 0.$$

Συνεπώς,

$$u^\star(x) - u(0) = \int_0^x v(s)\,ds, \quad \text{για } x \in \mathbb{R}.$$

Από τα προηγούμενα και το Λήμμα 2.2 έπεται ότι $u^\star \in W^{1,p}(\mathbb{R})$ και $(u^\star)' = v$. Επομένως,

$$\|u^\star\|^p_{W^{1,p}(\mathbb{R})} = \|u^\star\|^p_{L^p(\mathbb{R})} + \|v\|^p_{L^p(\mathbb{R})} = 2\big(\|u^\star\|^p_{L^p(I)} + \|v\|^p_{L^p(I)}\big),$$

δηλαδή

$$\|u^\star\|_{W^{1,p}(\mathbb{R})} = 2^{1/p}\|u\|_{W^{1,p}(I)}. \tag{2.24}$$

Σύμφωνα με τα προηγούμενα, στην περίπτωση $I = (0,\infty)$ ο τελεστής επέκτασης E έχει όλες τις επιθυμητές ιδιότητες· μάλιστα, σε αυτή την περίπτωση, $C = 2^{1/p}$ στις $ii.$ και $iii.$ Η ανωτέρω απόδειξη γενικεύεται εύκολα για να καλύψει περιπτώσεις της μορφής $I = (a,\infty)$ ή $I = (-\infty, a)$ με $a \in \mathbb{R}$. Φερ' ειπείν, για $u \in W^{1,p}(0,\infty)$ ορίζουμε την επέκταση Eu με άρτια ανάκλαση ως προς το σημείο a,

$$(Eu)(x) := \begin{cases} u(x), & \text{για } x \ge a,\\ u(2a-x) = u\big(a + (a-x)\big), & \text{για } x < 0, \end{cases}$$

και προχωρούμε όπως προηγουμένως.

Θα ασχοληθούμε τώρα με την περίπτωση ενός φραγμένου διαστήματος. Πρώτα θα μελετήσουμε την περίπτωση $I = (0,1)$ και αργότερα θα δούμε ότι οι άλλες περιπτώσεις ανάγονται σε αυτήν με μια αλλαγή μεταβλητής. Θεωρούμε λοιπόν μια συνάρτηση $\eta \in C^1(\mathbb{R})$ τέτοια ώστε $0 \le \eta(x) \le 1$, για κάθε $x \in \mathbb{R}$, και

$$\eta(x) = \begin{cases} 1, & \text{αν } x < \dfrac{1}{4},\\ 0, & \text{αν } x > \dfrac{3}{4}. \end{cases}$$

Γράφουμε μια ομαλή συνάρτηση u, ορισμένη στο διάστημα $(0,1)$, στη μορφή $u = \eta u + (1-\eta)u$. Η ηu μηδενίζεται στο σημείο 1 και μπορεί εύκολα να επεκταθεί ως ομαλή συνάρτηση στο $[0,\infty)$, ως μηδενική συνάρτηση. Αντίστοιχα, η $(1-\eta)u$ μηδενίζεται στο σημείο 0 και μπορεί εύκολα να επεκταθεί ως ομαλή συνάρτηση στο

$(-\infty, 0]$, ως μηδενική συνάρτηση. Λεπτομέρειες δίνονται στη συνέχεια, εδώ θελήσαμε απλώς να παρουσιάσουμε το κίνητρο για την εισαγωγή της η.

Για οποιαδήποτε συνάρτηση $f : (0,1) \to \mathbb{R}$ συμβολίζουμε την επέκτασή της με μηδέν στο $[1,\infty)$ με $\tilde{f}$, $\tilde{f} : (0,\infty) \to \mathbb{R}$,

$$\tilde{f}(x) = \begin{cases} f(x), & \text{για } x \in (0,1), \\ 0, & \text{για } x \geq 1. \end{cases}$$

Τώρα, για $u \in W^{1,p}(I)$, ισχυριζόμαστε ότι αφ' ενός $\eta\tilde{u} \in W^{1,p}(0,\infty)$ και αφ' ετέρου ότι $(\eta\tilde{u}) = \eta'\tilde{u} + \eta\tilde{u'}$, όπου $\tilde{u'}$ η επέκταση με μηδέν στο $[1,\infty)$ της u'. Πράγματι, $\eta\tilde{u} \in L^p(0,\infty)$, αφού

$$\|\eta\tilde{u}\|^p_{L^p(0,\infty)} = \int_0^\infty \eta^p |\tilde{u}|^p\,dx \leq \int_0^{\frac{3}{4}} |u|^p\,dx \leq \|u\|^p_{L^p(0,1)},$$

δηλαδή

$$\|\eta\tilde{u}\|_{L^p(0,\infty)} \leq \|u\|_{L^p(0,1)}. \tag{2.25}$$

Επί πλέον, για $\varphi \in C_0^1(0,\infty)$ (βλ. την Άσκηση 2.11) έχουμε προφανώς $\eta\varphi \in C_0^1(0,1)$ και

$$\begin{aligned}\int_0^\infty \eta\tilde{u}\varphi'\,dx &= \int_0^1 \eta u\varphi'\,dx = \int_0^1 u\big((\eta\varphi)' - \eta'\varphi\big)\,dx \\ &= \int_0^1 u(\eta\varphi)'\,dx - \int_0^1 u\eta'\varphi\,dx \\ &= -\int_0^1 u'(\eta\varphi)\,dx - \int_0^1 u\eta'\varphi\,dx = -\int_0^1 (u'\eta + u\eta')\varphi\,dx \\ &= -\int_0^\infty g\varphi\,dx\end{aligned}$$

με

$$g(x) := \begin{cases} u'(x)\eta(x) + u(x)\eta'(x) & \text{για } x \in (0,1), \\ 0 & \text{για } x \geq 1, \end{cases}$$

δηλαδή $g = \eta'\tilde{u} + \eta\tilde{u'}$. Τώρα, $g \in L^p(0,\infty)$, αφού

$$\begin{aligned}\|g\|_{L^p(0,\infty)} &= \|u'\eta + u\eta'\|_{L^p(0,\infty)} = \|u'\eta + u\eta'\|_{L^p(0,1)} \\ &\leq \|u'\eta\|_{L^p(0,1)} + \|u\eta'\|_{L^p(0,1)} \\ &\leq \|u'\|_{L^p(0,1)} + \max_{0\leq x\leq 1} |\eta'(x)|\,\|u\|_{L^p(0,1)} \\ &\leq \max_{0\leq x\leq 1} |\eta'(x)|\big(\|u\|_{L^p(0,1)} + \|u'\|_{L^p(0,1)}\big).\end{aligned}$$

Η ανισότητα του Hölder στον $\mathbb{R}^2$ δίνει

$$\|u\|_{L^p(0,1)} + \|u'\|_{L^p(0,1)} \le 2^{\frac{p-1}{p}} \big(\|u\|^p_{L^p(0,1)} + \|u'\|^p_{L^p(0,1)}\big)^{1/p} = 2^{\frac{p-1}{p}} \|u\|_{W^{1,p}(0,1)},$$

βλ. την Άσκηση 2.3, οπότε συνολικά έχουμε

$$\|g\|_{L^p(0,\infty)} \le C_1 \|u\|_{W^{1,p}(0,1)} \tag{2.26}$$

με

$$C_1 := \max_{0\le x\le 1} |\eta'(x)|\, 2^{\frac{p-1}{p}} \|u\|_{W^{1,p}(0,1)}.$$

[Σημειώστε ότι, σύμφωνα με το θεώρημα της μέσης τιμής του διαφορικού λογισμού, έχουμε $\max_{0\le x\le 1} |\eta'(x)| \ge 2$. Εύκολα πάντως μπορεί κανείς να κατασκευάσει συναρτήσεις η με τις επιθυμητές ιδιότητες τέτοιες ώστε αυτό το μέγιστο να είναι οποιοσδήποτε αριθμός μεγαλύτερος του 2, φερ' ειπείν 2.5.] Επομένως, υπάρχει η ασθενής παράγωγος $(\eta\tilde u)'$ και $(\eta\tilde u)' = g \in L^p(0,\infty)$. Συνεπώς, $\eta\tilde u \in W^{1,p}(0,\infty)$. Σύμφωνα με τη (2.26) έχουμε

$$\|(\eta\tilde u)'\|_{L^p(0,\infty)} \le C_1 \|u\|_{W^{1,p}(0,1)}. \tag{2.27}$$

Οι (2.25) και (2.27) δίνουν

$$\|\eta\tilde u\|_{W^{1,p}(0,\infty)} \le C_2 \|u\|_{W^{1,p}(0,1)} \tag{2.28}$$

με

$$C_2 := (1 + C_1^p)^{1/p}.$$

Γράφουμε τώρα τη u στη μορφή

$$u = \eta u + (1-\eta) u.$$

Η ηu επεκτείνεται στο $(0,\infty)$ ως $\eta\tilde u$. Επεκτείνουμε τώρα την $\eta\tilde u$ στο $\mathbb{R}$ με άρτια ανάκλαση ως προς το 0 και οδηγούμαστε σε μια συνάρτηση $v_1 \in W^{1,p}(\mathbb{R})$. Σύμφωνα με τις (2.24) και (2.28) θα έχουμε

$$\|v_1\|_{W^{1,p}(\mathbb{R})} \le C_2\, 2^{1/p}\, \|u\|_{W^{1,p}(0,1)}. \tag{2.29}$$

Εντελώς ανάλογα επεκτείνεται η συνάρτηση $(1-\eta)u$ (σημειώστε ότι αυτή μηδενίζεται στο διάστημα $[0,1/4]$) στο $(-\infty,1)$ ως $(1-\eta)\hat u$ με

$$\hat u(x) := \begin{cases} u(x) & \text{για } x \in (0,1), \\ 0 & \text{για } -\infty < x \le 0. \end{cases}$$

Πάλι έχουμε $(1-\eta)\hat{u} \in W^{1,p}(-\infty, 1)$. Επεκτείνουμε τώρα την $(1-\eta)\hat{u}$ στο $\mathbb{R}$ με άρτια ανάκλαση ως προς το 1 και οδηγούμαστε σε μια συνάρτηση $v_2 \in W^{1,p}(\mathbb{R})$. Όπως στη (2.29) θα έχουμε

$$\|v_2\|_{W^{1,p}(\mathbb{R})} \le C_2\, 2^{1/p}\, \|u\|_{W^{1,p}(0,1)}. \tag{2.30}$$

Ορίζουμε τώρα τον τελεστή επέκτασης E, ως $Eu := v_1 + v_2$. Αμέσως διαπιστώνουμε ότι ο E έχει τις επιθυμητές ιδιότητες.

Τέλος, στην περίπτωση ενός φραγμένου διαστήματος $I = (a, b)$ η απόδειξη γίνεται εντελώς ανάλογα, αυτή τη φορά με τη βοηθητική συνάρτηση $\eta_{a,b}$,

$$\eta_{a,b}(x) := \eta\Big(\frac{x-a}{x-b}\Big). \qquad \square$$

Δίνουμε τώρα ένα αποτέλεσμα πυκνότητας των συναρτήσεων δοκιμής στον χώρο $W^{1,p}(\mathbb{R})$, $1 \le p < \infty$, το οποίο θα παίξει σημαντικό ρόλο στη συνέχεια.

Θεώρημα 2.10 (Πυκνότητα συναρτήσεων δοκιμής στον $W^{1,p}(\mathbb{R})$, $1 \le p < \infty$.) *Έστω $u \in W^{1,p}(\mathbb{R})$, $p \in [1, \infty)$. Τότε υπάρχει μια ακολουθία $(u_n)_{n\in\mathbb{N}} \subset C_0^\infty(\mathbb{R})$ συναρτήσεων δοκιμής που συγκλίνει στη u στον χώρο $W^{1,p}(\mathbb{R})$.*

Απόδειξη. Θεωρούμε μια συνάρτηση *αποκοπής* $\zeta \in C_0^\infty(\mathbb{R})$ τέτοια ώστε $0 \le \zeta(x) \le 1$, για κάθε $x \in \mathbb{R}$, και

$$\zeta(x) = \begin{cases} 1, & \text{αν } |x| \le 1, \\ 0, & \text{αν } |x| \ge 2. \end{cases}$$

Συναρτήσεις αποκοπής υπάρχουν όντως, βλ. την Άσκηση 2.14. Με τη βοήθεια της ζ, ορίζουμε και τις συναρτήσεις αποκοπής $\zeta_n \in C_0^\infty(\mathbb{R})$, $\zeta_n(x) := \zeta(x/n)$, $n \in \mathbb{N}$. Προφανώς ισχύει $0 \le \zeta_n(x) \le 1$, για κάθε $x \in \mathbb{R}$, και

$$\zeta_n(x) = 1 \ \text{ για } \ |x| \le n \quad \text{και} \quad \zeta_n(x) = 0 \ \text{ για } \ |x| \ge 2n.$$

Επί πλέον, με $C := \max_{x\in\mathbb{R}} |\zeta'(x)|$, έχουμε

$$\max_{x\in\mathbb{R}} |\zeta_n'(x)| = \frac{C}{n}\,. \tag{2.31}$$

Κατασκευάζουμε τώρα τις συναρτήσεις u_n σε δύο στάδια: πρώτα εξομαλύνουμε τη u με συνέλιξη με εξομαλυντές ω_n, $\omega_n(x) := n\omega(nx)$, βλ. τη (2.14), δηλαδή πρώτα σχηματίζουμε τη συνάρτηση $\omega_n \star u$, και εν συνεχεία αποκόπτουμε πολλαπλασιάζοντας με συναρτήσεις αποκοπής ζ_n. Θέτουμε δηλαδή

$$u_n := \zeta_n(\omega_n \star u).$$

Αφού $\omega_n \star u \in C^\infty(\mathbb{R})$ και $\zeta_n \in C_0^\infty(\mathbb{R})$, ισχύει πράγματι ότι $u_n \in C_0^\infty(\mathbb{R})$. Τώρα έχουμε

$$u - u_n = (u - \zeta_n u) + \zeta_n(u - \omega_n \star u).$$

Επομένως

$$\|u - u_n\|_{L^p(\mathbb{R})} \le \|u - \zeta_n u\|_{L^p(\mathbb{R})} + \|\zeta_n(u - \omega_n \star u)\|_{L^p(\mathbb{R})},$$

οπότε, προφανώς,

$$\|u - u_n\|_{L^p(\mathbb{R})} \le \|u - \zeta_n u\|_{L^p(\mathbb{R})} + \|u - \omega_n \star u\|_{L^p(\mathbb{R})}. \tag{2.32}$$

Λόγω του ότι $|(\zeta_n u)(x)| \le |u(x)|$ και $(\zeta_n u)(x) \to u(x)$, $n \to \infty$, σ.π. στο $\mathbb{R}$, σύμφωνα με ένα θεμελιώδες θεώρημα της Θεωρίας Μέτρου, το θεώρημα κυριαρχημένης σύγκλισης του Lebesgue, έχουμε

$$\|u - \zeta_n u\|_{L^p(\mathbb{R})} \to 0, \quad n \to \infty.$$

Χρησιμοποιώντας και το $iii.$ του Θεωρήματος 2.6, από τα προηγούμενα συμπεραίνουμε ότι

$$\|u - u_n\|_{L^p(\mathbb{R})} \to 0, \quad n \to \infty. \tag{2.33}$$

Επί πλέον, εύκολα διαπιστώνει κανείς ότι $(\omega_n \star u)' = \omega_n \star u'$, βλ. και τη σχετική υπόδειξη στην Άσκηση 2.11, οπότε

$$u' - u_n' = (u' - \zeta_n u') + \zeta_n(u' - \omega_n \star u') + \zeta_n'(\omega_n \star u).$$

Οι δύο πρώτοι όροι στο δεξιό μέλος αυτής της σχέσης μπορούν να εκτιμηθούν με τον ίδιο ακριβώς τρόπο όπως έγινε και προηγουμένως με u στη θέση της u'. Όσον αφορά τον τρίτο όρο, σύμφωνα με τη (2.31) και το $iii.$ του Θεωρήματος 2.6, έχουμε

$$\|\zeta_n'(\omega_n \star u)\|_{L^p(\mathbb{R})} \le \frac{C}{n}\|\omega_n \star u\|_{L^p(\mathbb{R})} \le \frac{C}{n}\|u\|_{L^p(\mathbb{R})} \to 0, \quad n \to \infty.$$

Επομένως έχουμε

$$\|u' - u_n'\|_{L^p(\mathbb{R})} \to 0, \quad n \to \infty. \tag{2.34}$$

Από τις (2.33) και (2.34) έπεται αμέσως το ζητούμενο. $\square$

Πόρισμα 2.4 (Πυκνότητα ομαλών συναρτήσεων στον $W^{1,p}(I)$, $1 \le p < \infty$.) *Έστω I ένα διάστημα και $u \in W^{1,p}(I)$, $p \in [1,\infty)$. Τότε υπάρχει μια ακολουθία $(u_n)_{n\in\mathbb{N}} \subset C_0^\infty(\mathbb{R})$ συναρτήσεων δοκιμής τέτοια ώστε η ακολουθία $(v_n)_{n\in\mathbb{N}} \subset C^\infty(I)$ των περιορισμών των u_n στο I να συγκλίνει στη u στον χώρο $W^{1,p}(I)$.*

Απόδειξη. Επεκτείνουμε τη u στο $\mathbb{R}$ και σχηματίζουμε τη συνάρτηση $Eu \in W^{1,p}(\mathbb{R})$ σύμφωνα με το Θεώρημα 2.9. Σύμφωνα με το Θεώρημα 2.9 υπάρχει μια ακολουθία $(u_n)_{n\in\mathbb{N}} \subset C_0^\infty(\mathbb{R})$ που συγκλίνει στην Eu στον χώρο $W^{1,p}(\mathbb{R})$. Συμβολίζουμε με v_n τους περιορισμούς των u_n στο I. Τότε, προφανώς, θα έχουμε, σύμφωνα με το $iii.$ του Θεωρήματος 2.7,

$$\|u - v_n\|_{W^{1,p}(I)} \le \|Eu - u_n\|_{W^{1,p}(\mathbb{R})} \to 0, \quad n \to \infty. \qquad \square$$

Σχόλιο. (Πυκνότητα του $C^\infty(\bar{I})$ στον $W^{1,p}(I)$, $1 \le p < \infty$, για φραγμένο I.) Έστω ότι το διάστημα I είναι φραγμένο. Σύμφωνα με το Πόρισμα 2.4, αν $u \in W^{1,p}(I)$, $p \in [1,\infty)$, τότε υπάρχει μια ακολουθία $(v_n)_{n\in\mathbb{N}} \subset C_0^\infty(\bar{I})$, οι περιορισμοί των $(u_n)_{n\in\mathbb{N}} \subset C_0^\infty(\mathbb{R})$ στο $\bar{I}$, που συγκλίνει στη v στον χώρο $W^{1,p}(I)$. Τονίζουμε ότι ο χώρος $C_0^\infty(I)$ είναι πυκνός στον $W^{1,p}(I)$, $p \in [1,\infty)$, μόνο στην περίπτωση $I = \mathbb{R}$. $\square$

Προχωρούμε τώρα στο θεώρημα εμφύτευσης του Sobolev. Η απόδειξη απλουστεύεται σημαντικά στην περίπτωση του χώρου H^1. Για το λόγο αυτόν, αλλά και επειδή αυτός ο χώρος μας ενδιαφέρει κυρίως, περιοριζόμαστε σε αυτή την περίπτωση.

Θεώρημα 2.11 (Θεώρημα εμφύτευσης του Sobolev.) *Έστω I ένα διάστημα. Τότε κάθε στοιχείο του $H^1(I)$ ανήκει στον $L^\infty(I)$ και μάλιστα ισχύει*

$$\|u\|_{L^\infty(I)} \le C\|u\|_{H^1(I)} \tag{2.35}$$

με μια σταθερά C, ανεξάρτητη του u. (Μάλιστα η C εξαρτάται μόνο από το, ενδεχομένως άπειρο, μήκος του I.)

Απόδειξη. Αρχίζουμε με την περίπτωση $I = \mathbb{R}$. Υποθέτουμε πρώτα ότι $v \in C_0^\infty(\mathbb{R})$. Τότε, για κάθε $x \in \mathbb{R}$,

$$\begin{aligned} |v(x)|^2 &= \int_{-\infty}^{x} (v^2)'\,ds = 2\int_{-\infty}^{x} vv'\,ds \\ &\le 2\|v\|_{L^2(\mathbb{R})}\,\|v'\|_{L^2(\mathbb{R})} \\ &\le \|v\|^2_{L^2(\mathbb{R})} + \|v'\|^2_{L^2(\mathbb{R})} = \|v\|^2_{H^1(\mathbb{R})}, \end{aligned}$$

συνεπώς

$$\|u\|_{L^\infty(\mathbb{R})} \le \|u\|_{H^1(\mathbb{R})} \quad \forall u \in C_0^\infty(\mathbb{R}). \tag{2.36}$$

Έστω τώρα $u \in H^1(\mathbb{R})$. Σύμφωνα με το Θεώρημα 2.10, υπάρχει μια ακολουθία $(u_n)_{n\in\mathbb{N}} \subset C_0^\infty(\mathbb{R})$ που συγκλίνει στη u στον χώρο $H^1(\mathbb{R})$. Ιδιαίτερα, η $(u_n)_{n\in\mathbb{N}}$ είναι ακολουθία Cauchy στον $H^1(\mathbb{R})$. Από τη (2.36) προκύπτει τότε αμέσως, ότι η $(u_n)_{n\in\mathbb{N}}$ είναι ακολουθία Cauchy στον $L^\infty(\mathbb{R})$. Λόγω της πληρότητας του $L^\infty(\mathbb{R})$, η $(u_n)_{n\in\mathbb{N}}$ θα συγκλίνει σε ένα στοιχείο $\tilde{u} \in L^\infty(\mathbb{R})$ στον $L^\infty(\mathbb{R})$. Θέλουμε τώρα να αποδείξουμε ότι $\tilde{u}(x) = u(x)$ σ.π. στο $\mathbb{R}$. Προφανώς, η $(u_n)_{n\in\mathbb{N}}$ συγκλίνει στη u και στον $L^2(\mathbb{R})$, αφού συγκλίνει στη u στον $H^1(\mathbb{R})$. Είναι τότε γνωστό από τη Θεωρία Μέτρου, ότι υπάρχει μια συγκλίνουσα υπακολουθία της $(u_n)_{n\in\mathbb{N}}$, την οποία χωρίς περιορισμό της γενικότητας και για απλοποίηση του συμβολισμού θεωρούμε ως την αρχική μας ακολουθία, τέτοια ώστε $u_n(x) \to u(x)$, $n \to \infty$, σ.π. στο $\mathbb{R}$. Συνεπώς, όντως ισχύει $\tilde{u}(x) = u(x)$ σ.π. στο $\mathbb{R}$. Αφήνοντας στην ανισότητα

$$\|u_n\|_{L^\infty(\mathbb{R})} \le \|u_n\|_{H^1(\mathbb{R})},$$

που ισχύει σύμφωνα με τη (2.36), το n να τείνει στο άπειρο, οδηγούμαστε στο συμπέρασμα

$$\|u\|_{L^\infty(\mathbb{R})} \le \|u\|_{H^1(\mathbb{R})}, \quad \forall u \in H^1(\mathbb{R}), \tag{2.37}$$

και έτσι ολοκληρώνεται η απόδειξη για την περίπτωση $I = \mathbb{R}$. Στη γενική περίπτωση, επεκτείνουμε κατ' αρχάς τη $u \in H^1(I)$ σε μια συνάρτηση $Eu \in H^1(\mathbb{R})$, σύμφωνα με το Θεώρημα 2.9. Σύμφωνα με τα προηγούμενα έχουμε $Eu \in L^\infty(\mathbb{R})$, οπότε βέβαια $u \in L^\infty(I)$. Επίσης, σύμφωνα με το Θεώρημα 2.9, προφανώς έχουμε

$$\|u\|_{L^\infty(I)} \le \|Eu\|_{L^\infty(\mathbb{R})}.$$

Επί πλέον, σύμφωνα με τη (2.37), $\|Eu\|_{L^\infty(\mathbb{R})} \le \|Eu\|_{H^1(\mathbb{R})}$. Τέλος, σύμφωνα με το $iii.$ του Θεωρήματος 2.7, $\|Eu\|_{H^1(\mathbb{R})} \le C\|u\|_{H^1(I)}$. Συνδυάζοντας τις τρεις τελευταίες ανισότητες οδηγούμαστε στην επιθυμητή εκτίμηση (2.35). □

Παρατήρηση 2.3 (Συμπεριφορά στοιχείων του $H^1(\mathbb{R})$ για μεγάλα x.) Από το Θεώρημα 2.10 και τη (2.31) έπεται ότι, για $u \in H^1(\mathbb{R})$, ισχύει

$$\lim_{|x|\to\infty} u(x) = 0.$$

Πράγματι, αν $(u_n)_{n\in\mathbb{N}} \subset C_0^\infty(\mathbb{R})$ είναι μια ακολουθία που συγκλίνει στη u στον χώρο $H^1(\mathbb{R})$, τότε, σύμφωνα με τη (2.35), θα ισχύει $\lim_{n\to\infty} \|u-u_n\|_{L^\infty(\mathbb{R})} = 0$. Έστω τώρα $\varepsilon > 0$. Τότε υπάρχει $N \in \mathbb{N}$ τέτοιο ώστε $\|u-u\|_{L^\infty(\mathbb{R})} < \varepsilon$. Για $|x|$ αρκετά μεγάλο ισχύει $u_N(x) = 0$, συνεπώς $|u(x)| < \varepsilon$. □

Παρατήρηση 2.4 (Συμπαγής εμφύτευση του $H^1(I)$ στον $C(\bar{I})$, για φραγμένο διάστημα.) Το Θεώρημα 2.11 διατυπώνεται συνήθως ως εξής: Ο $H^1(I)$ *εμφυτεύεται* στον $L^\infty(I)$, $H^1(I) \subset L^\infty(I)$, και αυτή η *εμφύτευση* είναι συνεχής. Στην περίπτωση που το διάστημα I είναι φραγμένο, έχουμε μάλιστα $H^1(I) \subset C^1(\bar{I})$, βλ. το Θεώρημα 2.8. Επί πλέον, στην τελευταία αυτή περίπτωση, η εμφύτευση είναι *συμπαγής*, δηλαδή κάθε φραγμένη ακολουθία στον $H^1(I)$ έχει συγκλίνουσα υπακολουθία στον $C^1(\bar{I})$. Πράγματι, για $u \in H^1(I)$ και $x, y \in I$, σύμφωνα με το Θεώρημα 2.8,

$$|u(x)-u(y)| = |\int_y^x u'(s)\,ds| \le \|u'\|_{L^2(I)}\,|x-y|^{1/2} \le \|u'\|_{H^1(I)}\,|x-y|^{1/2},$$

και το αποτέλεσμα έπεται αμέσως από το θεώρημα των Arzela–Ascoli. □

Τα δύο αποτελέσματα που ακολουθούν βασίζονται στο Θεώρημα 2.11 και στην Πρόταση 2.4. Και τα δύο αυτά αποτελέσματα θα χρησιμοποιηθούν στο τρίτο κεφάλαιο.

Πρόταση 2.5 (Το γινόμενο στοιχείων του $H^1(I)$ ανήκει στον ίδιο χώρο.) *Έστω I ένα διάστημα και $u, v \in H^1(I)$. Τότε $uv \in H^1(I)$ και $(uv)' = u'v + uv'$. Επί πλέον, μπορούμε να ολοκληρώσουμε κατά μέρη,*

$$\int_y^x u'v\,ds = u(x)v(x) - u(y)v(y) - \int_y^x uv'\,ds,$$

για κάθε $x, y \in \bar{I}$.

Απόδειξη. Αφού $u \in H^1(I)$, σύμφωνα με το Θεώρημα 2.11, θα έχουμε και $u \in L^\infty(I)$. Συνδυάζοντας αυτή την ιδιότητα με το γεγονός ότι $v \in L^2(I)$, συμπεραίνουμε αμέσως ότι $uv \in L^2(I)$. Σύμφωνα με την Πρόταση 2.4 υπάρχουν ακολουθίες $(\hat{u}_n)_{n\in\mathbb{N}}, (\hat{v}_n)_{n\in\mathbb{N}} \subset C_0^\infty(\mathbb{R})$ τέτοιες ώστε οι περιορισμοί τους u_n και v_n, αντίστοιχα, στο I, να συγκλίνουν στις u και v, αντίστοιχα, στον $H^1(I)$. Τότε, σύμφωνα με το Θεώρημα 2.11, οι ακολουθίες $(u_n)_{n\in\mathbb{N}}$ και $(v_n)_{n\in\mathbb{N}}$ θα συγκλίνουν στις u και v, αντίστοιχα, και στον $L^\infty(I)$.

Τώρα έχουμε

$$\begin{aligned}\|uv-u_nv_n\|_{L^2(I)} &\le \|uv-u_nv\|_{L^2(I)}+\|u_nv-u_nv_n\|_{L^2(I)}\\ &\le \|v\|_{L^\infty(I)}\,\|u-u_n\|_{L^2(I)}+\|u_n\|_{L^\infty(I)}\,\|v-v_n\|_{L^2(I)}\to 0,\ \ n\to\infty,\end{aligned}$$

δηλαδή η ακολουθία $(u_nv_n)_{n\in\mathbb{N}}$ συγκλίνει στη uv στον $L^2(I)$.

Επί πλέον,

$$\begin{aligned}\|u'v-u_n'v_n\|_{L^2(I)} &\le \|u'v-u'v_n\|_{L^2(I)}+\|u'v_n-u_n'v_n\|_{L^2(I)}\\ &\le \|u'\|_{L^2(I)}\,\|v-v_n\|_{L^\infty(I)}+\|v_n\|_{L^\infty(I)}\,\|u'-u_n'\|_{L^2(I)}\to 0,\ \ n\to\infty,\end{aligned}$$

δηλαδή $u_n'v_n \to u'v$, $n\to\infty$, στον $L^2(I)$. Αντίστοιχα, $u_nv_n' \to uv'$, $n\to\infty$, στον $L^2(I)$. Ανακεφαλαιώνοντας τα προηγούμενα, έχουμε

$$u_nv_n\to uv,\ \ n\to\infty,\quad \text{στον }\ L^2(I)$$

και

$$(u_nv_n)'\to u'v+uv',\ \ n\to\infty,\quad \text{στον }\ L^2(I).$$

Όπως στην απόδειξη της πληρότητας των χώρων του Sobolev στο Θεώρημα 2.7 (μάλιστα στην προκειμένη περίπτωση απλουστεύεται, απλώς χρησιμοποιούμε τη συνέχεια του εσωτερικού γινομένου), εύκολα διαπιστώνουμε ότι η uv έχει ασθενή παράγωγο την $u'v+uv'$, η οποία ανήκει στον $L^2(I)$, συνεπώς $uv\in H^1(I)$, και

$$u_nv_n\to uv,\ \ n\to\infty,\quad \text{στον }\ H^1(I).$$

Τέλος, ολοκληρώνοντας τη σχέση $(uv)'=u'v+uv'$ λαμβάνουμε

$$(uv)(x)-(uv)(y)=\int_y^x u'v\,ds+\int_y^x uv'\,ds,$$

δηλαδή μπορούμε όντως να ολοκληρώσουμε κατά μέρη. □

Σχόλιο. (Γινόμενο στοιχείων του $L^2(I)$.) Είναι γνωστό ότι αποτέλεσμα αντίστοιχο της Πρότασης 2.5 *δεν* ισχύει στους χώρους L^p. Παραδείγματος χάριν, η συνάρτηση $u:(0,1)\to\mathbb{R}$, $u(x):=x^{-1/4}$, ανήκει στον $L^2(0,1)$, αφού το αντίστροφο της τετραγωνικής ρίζας είναι ολοκληρώσιμο στο $(0,1)$, αλλά η u^2 δεν ανήκει στον $L^2(0,1)$, αφού το αντίστροφο της ταυτότητας δεν είναι ολοκληρώσιμο στο $(0,1)$. □

Πρόταση 2.6 (Σύνθεση συνάρτησης του $H^1(I)$ και ομαλής συνάρτησης.) *Έστω I ένα διάστημα, $u \in H^1(I)$ και $G : \mathbb{R} \to \mathbb{R}$ μια συνεχώς παραγωγίσιμη συνάρτηση. Τότε και η σύνθεση $G \circ u$ ανήκει στον $H^1(I)$ και $(G \circ u)' = (G' \circ u) \cdot u'$.*

Απόδειξη. Αφού $u \in H^1(I)$, σύμφωνα με το Θεώρημα 2.11, θα έχουμε και $u \in L^\infty(I)$. Θέτουμε $M := \|u\|_{L^\infty(I)}$. Εξ άλλου, σύμφωνα με το θεώρημα της μέσης τιμής του διαφορικού λογισμού, $G(s) - G(0) = G'(\vartheta)s$, δηλαδή $G(s) = G'(\vartheta)s$, για κάποιο ϑ μεταξύ 0 και s. Συνεπώς, για $\delta > 0$, υπάρχει σταθερά $C = C(G, M, \delta)$ τέτοια ώστε

$$\forall s \in [-M-\delta, M+\delta] \quad |G'(s)| \le C \ \text{ και } \ |G(s)| \le C|s|. \tag{2.38}$$

Επομένως, θα έχουμε $|G\big(u(x)\big)| \le C|u(x)|$ σ.π. στο I. Από αυτή την εκτίμηση έπεται αμέσως ότι $G \circ u \in L^2(I)$, αφού $u \in L^2(I)$. Επίσης, $|G'\big(u(x)\big)| \le C$ σ.π. στο I, οπότε $|G'\big(u(x)\big)|\,|u'(x)| \le C|u'(x)|$ σ.π. στο I, και συνεπώς $G'(u) \cdot u' \in L^2(I)$, αφού $u' \in L^2(I)$. Απομένει να αποδείξουμε ότι

$$\int_I G(u)\varphi'\,dx = -\int_I G'(u) \cdot u'\varphi\,dx \quad \forall \varphi \in C_0^\infty(I).$$

Σύμφωνα με το Πόρισμα 2.4 υπάρχει μια ακολουθία $(\hat{u}_n)_{n\in\mathbb{N}} \subset C_0^\infty(\mathbb{R})$ τέτοια ώστε η ακολουθία των περιορισμών u_n των $\hat{u}_n$ στο I να συγκλίνει στη u στον $H^1(I)$. Επί πλέον, σύμφωνα με το Θεώρημα 2.11, η ακολουθία $(u_n)_{n\in\mathbb{N}}$ θα συγκλίνει στη u και στον $L^\infty(I)$. Χρησιμοποιώντας τη συνέχεια της G συμπεραίνουμε ότι $G(u_n) \to G(u)$, $n \to \infty$, στον $L^\infty(I)$. Τώρα, $\|u_n\|_{L^\infty(I)} \to \|u\|_{L^\infty(I)}$, $n \to \infty$, συνεπώς, για αρκετά μεγάλο n, $\|u_n\|_{L^\infty(I)} \le M + \delta$. Επομένως, για αρκετά μεγάλο n, η (2.38) δίνει $|G\big(u_n(x)\big)| \le C|u_n(x)|$ σ.π. στο I. Από αυτή την εκτίμηση έπεται αμέσως ότι $G \circ u_n \in L^2(I)$ και $\|G(u_n)\|_{L^2(I)} \le C\|u_n\|_{L^2(I)}$. Σύμφωνα με τα ανωτέρω, ένα θεμελιώδες αποτέλεσμα της Θεωρίας Μέτρου, το θεώρημα κυριαρχημένης σύγκλισης του Lebesgue, μας επιτρέπει να συμπεράνουμε ότι $G(u_n) \to G(u)$, $n \to \infty$, στον $L^2(I)$. Επομένως,

$$\forall \varphi \in C_0^\infty(I) \quad \int_I G(u_n)\varphi'\,dx \to \int_I G(u)\varphi'\,dx, \ \ n \to \infty. \tag{2.39}$$

Τώρα,

$$\begin{aligned} \|G'(u)u' - G'(u_n)u_n'\|_{L^2(I)} &\le \|G'(u)u' - G'(u_n)u'\|_{L^2(I)} + \|G'(u_n)u' - G'(u_n)u_n'\|_{L^2(I)} \\ &\le \|u'\|_{L^2(I)}\,\|G'(u) - G'(u_n)\|_{L^\infty(I)} + \|G'(u_n)\|_{L^\infty(I)}\,\|u' - u_n'\|_{L^2(I)}. \end{aligned}$$

Όσον αφορά τους όρους στο δεξιό μέλος αυτής της εκτίμησης σημειώνουμε ότι:

- $\|u' - u_n'\|_{L^2(I)} \to 0$, $n \to \infty$, αφού $u_n \to u$, $n \to \infty$, στον $H^1(I)$.
- Για αρκετά μεγάλο n, $\|u_n\|_{L^\infty(I)} \leq M + \delta/2$, συνεπώς, σύμφωνα με τη (2.38), $\|G'(u_n)\|_{L^\infty(I)} \leq C$.
- Αφού $u_n \to u$, $n \to \infty$, στον $L^\infty(I)$, η συνέχεια της G' δίνει $G'(u_n) \to G'(u)$, $n \to \infty$, στον $L^\infty(I)$.

Επομένως οδηγούμαστε στο συμπέρασμα ότι $G'(u_n)u_n' \to G'(u)u'$, $n \to \infty$, στον $L^2(I)$, οπότε θα έχουμε και

$$\forall \varphi \in C_0^\infty(I) \quad \int_I G'(u_n)u_n'\varphi\, dx \to \int_I G'(u)u'\varphi\, dx, \ \ n \to \infty. \tag{2.40}$$

Όμως, αφού $\varphi \in C_0^\infty(I)$ και $u_n \in C_0^\infty(\mathbb{R})$, θα έχουμε

$$\int_I G(u_n)\varphi'\, dx = -\int_I G'(u_n)u_n'\varphi\, dx, \ \ n \in \mathbb{N}, \quad \forall \varphi \in C_0^\infty(I).$$

Βάσει των (2.39) και (2.40), αφήνοντας εδώ το n να τείνει στο άπειρο, διαπιστώνουμε ότι

$$\int_I G(u)\varphi'\, dx = -\int_I G'(u)u'\varphi\, dx \quad \forall \varphi \in C_0^\infty(I). \qquad \square$$

Οι χώροι του Sobolev $\boldsymbol{H^k}$. Όπως έχουμε ήδη δει, οι χώροι Sobolev H^k είναι χώροι Hilbert. Μάλιστα, για $k \geq 1$, ισχύει

$$H^k(I) = \{u \in H^{k-1} : u' \in H^{k-1}\},$$

με u', φυσικά, την ασθενή πρώτη παράγωγο της u. Έστω τώρα I ένα φραγμένο διάστημα. Εύκολα διαπιστώνουμε ότι κάθε συνάρτηση $u \in C^{k-1}(\bar{I})$, με κλασική παράγωγο $u^{(k)}$ τμηματικά συνεχή στο $\bar{I}$, έχει ασθενή παράγωγο τάξης k, η οποία συμπίπτει (σ.π.) με την $u^{(k)}$, και η u ανήκει στον $H^k(I)$.

Μια προφανής τροποποίηση του Θεωρήματος 2.8 οδηγεί στο συμπέρασμα ότι, για $u \in H^k(I)$, υπάρχει $\tilde{u} \in C^{k-1}(\bar{I})$ τέτοια ώστε $\tilde{u} = u$ σ.π. στο I και

$$\tilde{u}^{(i)}(x) - \tilde{u}^{(i)}(y) = \int_y^x u^{(i+1)}(s)\, ds, \quad \forall x, y \in I,$$

για $i \in \{0, \dots, k-1\}$. Επίσης, για $u \in H^k(I)$, υπάρχει μια ακολουθία $(\hat{u}_n)_{n\in\mathbb{N}} \subset C_0^\infty(\mathbb{R})$ τέτοια ώστε οι περιορισμοί u_n στο I των $\hat{u}_n$ να συγκλίνουν στη u στον $H^k(I)$. Επί πλέον, $H^k(I) \subset C^{k-1}(I)$ και

$$\sum_{i=0}^{k-1} \|u^{(i)}\|_{L^\infty(I)} \leq C\|u\|_{H^k(I)} \quad \forall u \in H^k(I).$$

Ακόμα, για $u, v \in H^k(I)$ ισχύει $uv \in H^k(I)$ και

$$(uv)^{(k)} = \sum_{i=0}^{k} \binom{k}{i} u^{(i)} v^{(k-i)},$$

δηλαδή η ασθενής παράγωγος του γινομένου δίνεται από τον τύπο του Leibniz. Τα αποτελέσματα που αναφέρθηκαν ανωτέρω προκύπτουν εύκολα με τεχνικές παρόμοιες με αυτές που χρησιμοποιήθηκαν για την απόδειξη αντίστοιχων αποτελεσμάτων στον $H^1(I)$.

Τέλος, αναφέρουμε χωρίς απόδειξη ένα *αποτέλεσμα παρεμβολής*. Για $i \in \{1, \dots, k-1\}$ και $\varepsilon > 0$, υπάρχει μια σταθερά C_ε, η οποία εξαρτάται από το ε και από το, ενδεχομένως άπειρο, μήκος του διαστήματος I, τέτοια ώστε

$$\|u^{(i)}\|_{L^2(I)} \le C_\varepsilon \|u\|_{L^2(I)} + \varepsilon \|u^{(k)}\|_{L^2(I)} \quad \forall u \in H^k(I). \tag{2.41}$$

Από τη (2.41) έπεται αμέσως ότι υπάρχουν θετικές σταθερές c_1 και c_2, τέτοιες ώστε

$$c_1 \|u\|_k \le C_\varepsilon \|u\|_{L^2(I)} + \varepsilon \|u^{(k)}\|_{L^2(I)} \le c_2 \|u\|_k \quad \forall u \in H^k(I), \tag{2.42}$$

δηλαδή ότι η ποσότητα $\|u\|_{L^2(I)} + \|u^{(k)}\|_{L^2(I)}$ ορίζει στον $H^k(I)$ μια νόρμα ισοδύναμη με τη νόρμα $\|\cdot\|_k$.

Ο χώρος H_0^1. Ο χώρος H_0^1 είναι πλήρης υπόχωρος του H^1 και θα παίξει πολύ σημαντικό ρόλο στο επόμενο κεφάλαιο.

Ορισμός 2.10 (Ο χώρος H_0^1.) Έστω I ένα διάστημα. Ο χώρος $H_0^1 := H_0^1(I)$ αποτελείται από τα στοιχεία του $H^1(I)$ που είναι όρια στη νόρμα $\|\cdot\|_1$ ακολουθιών συναρτήσεων δοκιμής, δηλαδή στοιχείων του $C_0^\infty(I)$.

Εφοδιάζουμε τον H_0^1 με τη νόρμα $\|\cdot\|_1$ του H^1. Από τον ορισμό προκύπτει αμέσως ότι ο $\big(H_0^1, \|\cdot\|_1\big)$ είναι πλήρης. Επί πλέον, αφού η νόρμα $\|\cdot\|_1$ παράγεται από εσωτερικό γινόμενο, ο H_0^1 είναι χώρος Hilbert.

Παρατήρηση 2.5 (Σχέση μεταξύ H^1 και H_0^1.)

i. Σύμφωνα με το Θεώρημα 2.10, στην περίπτωση $I = \mathbb{R}$, ο χώρος δοκιμής $C_0^\infty(\mathbb{R})$ είναι πυκνός στον $H^1(\mathbb{R})$. Επομένως, $H_0^1(\mathbb{R}) = H^1(\mathbb{R})$.

ii. Για $I \neq \mathbb{R}$, ο $H_0^1(I)$ είναι γνήσιος υπόχωρος του $H^1(I)$. Στην περίπτωση που το I είναι φραγμένο, φερ' ειπείν, η συνάρτηση $u \in C^\infty(\bar{I})$, $u(x) := c_1 \mathrm{e}^x + c_2 \mathrm{e}^{-x}$, με σταθερές c_1 και c_2, όχι και τις δύο μηδέν, ανήκει στον $H^1(I)$, και λόγω του ότι $u'' = u$ έχουμε

$$(u, \varphi)_1 = \int_I u\varphi \, dx + \int_I u'\varphi' \, dx = \int_I (u - u'')\varphi \, dx = 0, \quad \forall \varphi \in C_0^\infty(I).$$

Συνεπώς θα ισχύει και

$$(u, v)_1 = 0, \quad \forall v \in H_0^1(I).$$

Ιδιαίτερα, η βέλτιστη προσέγγιση της u από τον H_0^1 στη νόρμα $\|\cdot\|_1$ είναι η μηδενική συνάρτηση, δηλαδή $u \notin H_0^1$.

iii. Κάθε $u \in H^1(I)$ με φορέα κλειστό και φραγμένο υποσύνολο του I είναι επίσης στοιχείο του $H_0^1(I)$. Στη συνέχεια θα δούμε μάλιστα ένα πολύ ισχυρότερο αποτέλεσμα. □

Τέλος, αναφέρουμε, χωρίς απόδειξη, έναν πολύ χρήσιμο χαρακτηρισμό του H_0^1.

Θεώρημα 2.12 (Χαρακτηρισμός του H_0^1.) *Έστω I ένα διάστημα. Μια συνάρτηση $u \in H^1(I)$ ανήκει στον $H_0^1(I)$, αν και μόνο αν η u μηδενίζεται στα άκρα του I.* □

Σχόλιο. (Ρόλος του H_0^1 στο πρόβλημα δύο σημείων.) Για $u \in H^1(I)$ έχει νόημα να λέμε ότι η u μηδενίζεται στο σύνορο του I, βλ. το Θεώρημα 2.8. Ο χαρακτηρισμός που δίνεται στο Θεώρημα 2.12 καθιστά τον χώρο $H_0^1(I)$ πολύ χρήσιμο για το πρόβλημα δύο σημείων (2.2), αφού τα στοιχεία του ικανοποιούν αυτόματα τις ομογενείς συνοριακές συνθήκες. Στο θέμα θα επανέλθουμε στο επόμενο κεφάλαιο. □

Στην αρχή αυτού του κεφαλαίου, βλ. τη (2.7), είδαμε την ανισότητα των Poincaré–Friedrichs για συνεχώς παραγωγίσιμες συναρτήσεις που μηδενίζονται σε ένα σημείο ενός φραγμένου διαστήματος. Από το Θεώρημα 2.12 έπεται ότι η εν λόγω ανισότητα ισχύει και στον $H_0^1(I)$, για φραγμένο διάστημα I.

Πόρισμα 2.5 (Ανισότητα των Poincaré–Friedrichs.) *Έστω $I = (a, b)$ ένα φραγμένο διάστημα. Τότε*

$$\|u\|_1 \leq \big[1 + (b-a)^2\big]^{1/2} \|u'\|_{L^2(I)} \quad \forall u \in H_0^1(I), \tag{2.43}$$

δηλαδή η $|\cdot|_1$, $|u|_1 := \|u'\|_{L^2(I)}$, που είναι ημινόρμα στον $H^1(I)$, είναι νόρμα στον $H_0^1(I)$, ισοδύναμη με τη νόρμα $\|\cdot\|_1$.

Απόδειξη. Σύμφωνα με το Θεώρημα 2.12 ισχύει $u(a) = 0$. Επομένως, σύμφωνα με το Θεώρημα 2.8,

$$|u(x)| = |\int_a^x u'(s)\, ds| \leq \int_a^b |u'(s)|\, ds \leq (b-a)^{1/2} \|u'\|_{L^2(I)},$$

οπότε

$$\|u\|^2_{L^2(I)} \leq (b-a)^2 \|u'\|^2_{L^2(I)}. \tag{2.44}$$

Άρα έχουμε

$$\|u\|_1^2 = \|u\|^2_{L^2(I)} + \|u'\|^2_{L^2(I)} \leq \big(1 + (b-a)^2\big) \|u'\|^2_{L^2(I)}$$

και η ζητούμενη ανισότητα προκύπτει αμέσως. □

Για $k \geq 2$, οι χώροι $H_0^k(I)$ αποτελούνται από τα στοιχεία του $H^k(I)$ που είναι όρια ακολουθιών συναρτήσεων δοκιμής. Μπορεί να αποδειχθεί ότι

$$H_0^k(I) = \{u \in H^k(I) : \ u = u' = \cdots = u^{(k-1)} = 0 \ \text{ στο } \ \partial I\}.$$

Στο επόμενο κεφάλαιο θα χρησιμοποιήσουμε τον χώρο $H^2(I) \cap H_0^1(I)$. Σημειώνουμε ότι οι τιμές των πρώτων παραγώγων των στοιχείων αυτού του χώρου δεν προκαθορίζονται, κατ' αντιδιαστολή προς τις αντίστοιχες στοιχείων του $H_0^2(I)$ που μηδενίζονται.

Ασκήσεις

2.1 Με τους συμβολισμούς που χρησιμοποιήσαμε στην ανισότητα των Poincaré–Friedrichs (2.7), αποδείξτε ότι *δεν* υπάρχει σταθερά c τέτοια ώστε

$$\|v'\| \leq c\|v\| \quad \forall v \in V,$$

δηλαδή ότι η νόρμα $\|v'\|$ είναι *γνησίως* ισχυρότερη της νόρμας $\|v\|$.
[*Υπόδειξη*: Θεωρήστε την ακολουθία συναρτήσεων $(\varphi_n)_{n\in\mathbb{N}} \subset V$, $\varphi_n(x) := x^n(1-x)$, και βεβαιωθείτε ότι

$$\|\varphi_n\|^2 = \frac{1}{(n+1)(2n+1)(2n+3)}, \ \ \|\varphi_n'\|^2 = \frac{n}{(2n-1)(2n+1)},$$

οπότε

$$\frac{\|\varphi_n'\|^2}{\|\varphi_n\|^2} = \frac{n(n+1)(2n+3)}{(2n-1)} \to \infty, \ n \to \infty.]$$

2.2 (Η ανισότητα του Young.) Έστω $1 < p, q < \infty$ τέτοια ώστε $1/p + 1/q = 1$ και a, b θετικοί αριθμοί. Αποδείξτε την ανισότητα του Young

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Επί πλέον, για $\varepsilon > 0$, αποδείξτε και την ακόλουθη παραλλαγή της ανισότητας του Young

$$ab \leq \varepsilon a^p + C(\varepsilon) b^q \quad \text{με} \ C(\varepsilon) := \frac{1}{q}\,(\varepsilon p)^{-q/p}$$

[*Υπόδειξη*: Η συνάρτηση $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi(x) := \mathrm{e}^x$, είναι ως γνωστόν κυρτή, δηλαδή, για $x, y \in \mathbb{R}$ και $\lambda \in [0,1]$, ισχύει $\varphi(\lambda x + (1-\lambda)y) \leq \lambda\varphi(x) + (1-\lambda)\varphi(y)$. Ιδιαίτερα, για $\lambda := 1/p, x := \log a^p$ και $y := \log b^q$, έχουμε

$$\mathrm{e}^{\frac{1}{p}\log a^p + \frac{1}{q}\log b^q} \leq \frac{1}{p}\mathrm{e}^{\log a^p} + \frac{1}{q}\mathrm{e}^{\log b^q}.$$

Γράψτε σε κατάλληλη μορφή τα δύο μέλη αυτής της ανισότητας.]

2.3 (Η ανισότητα του Hölder.) Έστω $1 \leq p, q \leq \infty$ τέτοια ώστε $1/p + 1/q = 1$. Αν $u \in L^p(\Omega)$ και $v \in L^q(\Omega)$, αποδείξτε ότι $uv \in L^1(\Omega)$ και ότι ισχύει η ανισότητα του Hölder για ολοκληρώματα

$$\int_\Omega |uv|\,dx \leq \|u\|_{L^p}\,\|v\|_{L^q}.$$

[*Υπόδειξη*: Για $1 < p, q < \infty$ έχουμε, σύμφωνα με την ανισότητα του Young,

$$\int_\Omega |uv|\,dx \leq \frac{1}{p}\int_\Omega |u|^p\,dx + \frac{1}{q}\int_\Omega |u|^q\,dx.$$

Αν $\|u\|_{L^p} = \|v\|_{L^q} = 1$, η ανωτέρω σχέση οδηγεί αμέσως στο ζητούμενο. Βεβαιωθείτε ότι η υπόθεση $\|u\|_{L^p} = \|v\|_{L^q} = 1$ δεν περιορίζει τη γενικότητα.]

Για $x \in \mathbb{R}^n$ θέτουμε

$$\|x\|_p := \Big(\sum_{i=1}^n |x_i|^p\Big)^{1/p}, \quad \text{για} \ p < \infty, \quad \|x\|_\infty := \max_{1\leq i\leq n} |x_i|.$$

Αποδείξτε την ανισότητα του Hölder για αθροίσματα

$$\sum_{i=1}^n |x_i y_i| \leq \|x\|_p\,\|y\|_q.$$

Χρησιμοποιήστε την τελευταία ανισότητα για να πεισθείτε, για $1 < p < \infty$, ότι

$$\Big(\sum_{i=1}^n |x_i|\Big)^p \leq n^{p-1}\sum_{i=1}^n |x_i|^p.$$

2.4 (Η ανισότητα του Minkowski.) Έστω $1 \leq p \leq \infty$ και $u, v \in L^p(\Omega)$. Αποδείξτε την ανισότητα του Minkowski για ολοκληρώματα

$$\|u+v\|_{L^p} \leq \|u\|_{L^p} + \|v\|_{L^p}.$$

[*Υπόδειξη*: Για $p = 1$ καθώς και για $p = \infty$ η ανισότητα είναι τετριμμένη. Για $1 < p < \infty$ έχουμε, κατ' αρχάς,

$$\|u+v\|_{L^p}^p = \int_\Omega |u+v|\,|u+v|^{p-1}\,dx \leq \int_\Omega |u|\,|u+v|^{p-1}\,dx + \int_\Omega |v|\,|u+v|^{p-1}\,dx.$$

Επί πλέον, με την ανισότητα του Hölder, με q τέτοιο ώστε $1/p + 1/q = 1$, έχουμε, παραδείγματος χάριν,

$$\begin{aligned}\int_\Omega |u|\,|u+v|^{p-1}\,dx &\leq \Big(\int_\Omega |u|^p\,dx\Big)^{1/p}\Big(\int_\Omega |u+v|^{(p-1)q}\,dx\Big)^{1/q}\\ &= \|u\|_{L^p}\Big(\int_\Omega |u+v|^p\,dx\Big)^{\frac{p-1}{p}} = \|u\|_{L^p}\,\|u+v\|_{L^p}^{p-1}.\end{aligned}$$

Επομένως,

$$\|u+v\|_{L^p}^p \leq \|u+v\|_{L^p}^{p-1}\big(\|u\|_{L^p} + \|v\|_{L^p}\big).]$$

Επίσης, με τους συμβολισμούς της προηγούμενης Άσκησης, αποδείξτε την ανισότητα του Minkowski για αθροίσματα

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p, \quad x, y \in \mathbb{R}^n.$$

2.5 Έστω $A \subset \mathbb{R}$ ένα μετρήσιμο σύνολο και $1 \leq p \leq \infty$. Έστω q ο συζυγής εκθέτης του p, δηλαδή τέτοιος ώστε $1/p + 1/q = 1$.

i. Αν $u \in L^p(A)$, και $(v_n)_{n\in\mathbb{N}}$ μια συγκλίνουσα ακολουθία του $L^q(A)$, $v_n \to v$, $n \to \infty$, αποδείξτε ότι

$$\int_A u v_n\,dx \to \int_A uv\,dx, \quad n \to \infty.$$

ii. Αν $(u_n)_{n\in\mathbb{N}}$ μια συγκλίνουσα ακολουθία του $L^p(A)$, $u_n \to u$, $n \to \infty$, και $(v_n)_{n\in\mathbb{N}}$ μια συγκλίνουσα ακολουθία του $L^q(A)$, $v_n \to v$, $n \to \infty$, αποδείξτε ότι

$$\int_A u_n v_n\,dx \to \int_A uv\,dx, \quad n \to \infty.$$

2.6 Έστω $A \subset \mathbb{R}$ ένα μετρήσιμο σύνολο και $1 < p < \infty$. Έστω q ο συζυγής εκθέτης του p, δηλαδή τέτοιος ώστε $1/p + 1/q = 1$. Έστω $u \in L^p(A)$. Αποδείξτε ότι η συνάρτηση v,

$$v(x) := \begin{cases} |u(x)|^{p-2}u(x), & \text{αν } u(x) \neq 0,\\ 0, & \text{αν } u(x) = 0,\end{cases}$$

ανήκει στον $L^q(A)$.

2.7 Έστω $I = (a, b)$ ένα φραγμένο διάστημα και $1 \le \tilde{p} \le p < \infty$. Αν $u \in L^p(I)$, αποδείξτε ότι $u \in L^{\tilde{p}}(I)$ και

$$\|u\|_{\tilde{p}} \le (b-a)^{\frac{1}{\tilde{p}}-\frac{1}{p}} \|u\|_p.$$

[*Υπόδειξη*: Χρησιμοποιήστε την ανισότητα του Hölder.]

2.8 Έστω $I = (a, b)$ ένα φραγμένο διάστημα και $u \in L^\infty(I)$. Βεβαιωθείτε ότι $u \in L^p(I)$, για κάθε $p \in [1, \infty)$, και αποδείξτε ότι

$$\lim_{p\to\infty} \|u\|_p = \|u\|_\infty.$$

[*Υπόδειξη*: Κατ' αρχάς, προφανώς, ισχύει $\|u\|_p \le (b-a)^{1/p}\|u\|_\infty$, οπότε

$$\limsup_{p\to\infty} \|u\|_p \le \|u\|_\infty.$$

Επίσης, για κάθε $\varepsilon > 0$, υπάρχει $A \subset I$ με $\mu(A) > 0$ τέτοιο ώστε

$$|u(x)| \ge \|u\|_\infty - \varepsilon \quad \forall x \in A.$$

Επομένως

$$\|u\|_p^p \ge \int_A |u(x)|^p \, dx \ge \mu(A)\big(\|u\|_\infty - \varepsilon\big)^p,$$

οπότε

$$\liminf_{p\to\infty} \|u\|_p \ge \|u\|_\infty.]$$

2.9 Το μόνο ενδιαφέρον σημείο στην απόδειξη ότι η $\|\cdot\|_{k,p}$ είναι όντως νόρμα στον $W^{k,p} := W^{k,p}(I)$, βλ. τον Ορισμό 2.9, είναι η τριγωνική ανισότητα. Αποδείξτε ότι, για $u, v \in W^{k,p}$, ισχύει

$$\|u+v\|_{k,p} \le \|u\|_{k,p} + \|v\|_{k,p}.$$

[*Υπόδειξη*: Για $p = 1$ και $p = \infty$ η απόδειξη είναι τετριμμένη. Για $1 < p < \infty$, έχουμε, κατ' αρχάς, προφανώς,

$$\|u+v\|_{k,p} = \Big(\sum_{j=0}^{k} \|u^{(j)} + v^{(j)}\|_{L^p}^p\Big)^{1/p} \le \Big(\sum_{j=0}^{k} \big(\|u^{(j)}\|_{L^p} + \|v^{(j)}\|_{L^p}\big)^p\Big)^{1/p}.$$

Με $x, y \in \mathbb{R}^{k+1}$, $x_i := \|u^{(i+1)}\|_{L^p}$ και $y_i := \|v^{(i+1)}\|_{L^p}$, εκτιμώντας το δεξιό μέλος της ανωτέρω ανισότητας $\|x+y\|_p$ με $\|x\|_p+\|y\|_p$, δηλαδή με τη βοήθεια της ανισότητας του Minkowski στον $\mathbb{R}^{k+1}$, παίρνουμε

$$\|u+v\|_{k,p} \le \Big(\sum_{j=0}^{k} \|u^{(j)}\|_{L^p}^p\Big)^{1/p} + \Big(\sum_{j=0}^{k} \|v^{(j)}\|_{L^p}^p\Big)^{1/p},$$

που είναι η ζητούμενη ανισότητα.]

2.10 Έστω $1 < p < \infty$ και $v \in L^p(I)$ τέτοιο ώστε

$$\int_I v\varphi\,dx = 0 \quad \forall \varphi \in C_0^\infty(I).$$

Αποδείξτε ότι $v = 0$.
[*Υπόδειξη*: Έστω $w \in L^q(I)$, όπου q ο συζυγής εκθέτης του p. Λόγω της πυκνότητας του $C_0^\infty(I)$ στον $L^q(I)$, υπάρχει μια ακολουθία $(\varphi_n)_{n\in\mathbb{N}} \subset C_0^\infty(I)$, τέτοια ώστε $\varphi_n \to w$, $n \to \infty$, στον $L^q(I)$. Επομένως ισχύει

$$\int_I vw\,dx = 0 \quad \forall w \in L^q(I),$$

βλ. την Άσκηση 2.5. Επιλέγοντας εδώ το w ως

$$w(x) := \begin{cases} |v(x)|^{p-2}v(x), & \text{αν } v(x) \neq 0, \\ 0, & \text{αν } v(x) = 0, \end{cases}$$

βλ. την Άσκηση 2.6, έχουμε αφ' ενός $w \in L^q(I)$ και αφ' ετέρου

$$\int_I |v|^p\,dx = 0,$$

οπότε $\|v\|_p = 0$, δηλαδή $v = 0$. Σημειώστε ότι για $p = 2$ η απόδειξη απλουστεύεται.]

2.11 Έστω ένα διάστημα και $u, v \in L^1_{\text{loc}}$ δύο συναρτήσεις τέτοιες ώστε

$$(\star) \qquad \int_I u\varphi'\,dx = -\int_I v\varphi\,dx \quad \forall \varphi \in C_0^\infty(I),$$

δηλαδή η v είναι η ασθενής πρώτη παράγωγος της u. Αποδείξτε ότι ισχύει και

$$\int_I u\varphi'\,dx = -\int_I v\varphi\,dx \quad \forall \varphi \in C_0^1(I),$$

δηλαδή ότι στον ορισμό της ασθενούς παραγώγου ο χώρος δοκιμής $C_0^\infty(I)$ μπορεί να αντικατασταθεί από τον $C_0^1(I)$, και, φυσικά, ακριβώς αντίστοιχα στη σχέση (2.21) από τον χώρο $C_0^j(I)$.
[*Υπόδειξη*: Έστω $\varphi \in C_0^1(I)$. Τότε, για αρκετά μικρό ε, $\varphi_\varepsilon \star \varphi \in C_0^\infty(I)$. Συνεπώς, σύμφωνα με την $(\star)$,

$$\int_I u(\varphi_\varepsilon \star \varphi')\,dx = -\int_I v(\varphi_\varepsilon \star \varphi)\,dx.$$

Βεβαιωθείτε κατ' αρχάς ότι $(\varphi_\varepsilon \star \varphi)' = \varphi_\varepsilon \star \varphi'$. Χρησιμοποιήστε το γεγονός ότι υπάρχει ένα κλειστό και φραγμένο υποδιάστημα J του I, τέτοιο ώστε, για ε αρκετά μικρό, $\operatorname{supp} \varphi_\varepsilon \star \varphi \subset J$, και το ότι $u \in L^1(J)$, $\varphi_\varepsilon \star \varphi \to \varphi$ και $\varphi_\varepsilon \star \varphi' \to \varphi'$ στον $L^\infty(J)$ (αφού, σύμφωνα με το v. του Θεωρήματος 2.6, η σύγκλιση είναι ομοιόμορφη), καθώς και το i. της Άσκησης 2.5 με $p = 1$ και $q = \infty$, για να οδηγηθείτε στο συμπέρασμα ότι

$$\int_I u\varphi'\,dx = -\int_I v\varphi\,dx.]$$

2.12 Αποδείξτε την Πρόταση 2.4.

2.13 Έστω $I := (0,2) \setminus \{1\}$ και $u : I \to \mathbb{R}$,

$$u(x) := \begin{cases} 1, & \text{αν } 0 < x < 1, \\ 2, & \text{αν } 1 < x < 2. \end{cases}$$

Αποδείξτε ότι

$$\int_I u\varphi'\,dx = 0 \quad \forall \varphi \in C_0^\infty(I).$$

Έρχεται αυτό το αποτέλεσμα σε αντίφαση προς το Λήμμα 2.1;

2.14 Έστω χ η χαρακτηριστική συνάρτηση του διαστήματος $[-3/2, 3/2]$, και ω_2 ο εξομαλυντής $\omega_2(x) := 2\omega(2x)$, βλ. τη (2.14). Αποδείξτε ότι η συνέλιξή τους $\zeta, \zeta := \omega_2 \star \chi$, είναι συνάρτηση αποκοπής, δηλαδή $\zeta \in C_0^\infty(\mathbb{R})$, $0 \le \zeta(x) \le 1$, για κάθε $x \in \mathbb{R}$, $\zeta(x) = 1$, αν $|x| \le 1$, και $\zeta(x) = 0$, αν $|x| \ge 2$.

2.15 Έστω $I = (a,b)$ ένα διάστημα και $u \in H^2(I) \cap H_0^1(I)$. Αποδείξτε ότι

$$\|u'\|_{L^2(I)} \le \|u\|_{L^2(I)}^{1/2}\,\|u''\|_{L^2(I)}^{1/2}.$$

[*Υπόδειξη*:

$$\int_a^b (u')^2\,dx = -\int_a^b uu''\,dx.]$$

2.16 Έστω I ένα διάστημα. Τροποποιώντας κατάλληλα την απόδειξη του Θεωρήματος 2.11 βελτιώστε τη (2.35) στη μορφή: για κάθε $\delta > 0$, υπάρχει σταθερά C_δ τέτοια ώστε για κάθε $v \in H^1(I)$ να ισχύει

$$\|v\|_{L^\infty(I)} \le C_\delta \|v\|_{L^2(I)} + \delta \|v'\|_{L^2(I)}.$$

2.17 Έστω $I = (a,b)$ ένα διάστημα και $u \in H^2(I)$. Αποδείξτε ότι, για κάθε $\varepsilon > 0$, υπάρχει C_ε, ανεξάρτητο της u, τέτοιο ώστε

$$\|u'\|_{L^2(I)} \le C_\varepsilon \|u\|_{L^2(I)} + \varepsilon \|u''\|_{L^2(I)},$$

βλ. τη (2.41).

[*Υπόδειξη*:

$$\begin{aligned}\int_a^b (u')^2\,dx &= u(b)u'(b) - u(a)u'(a) - \int_a^b uu''\,dx\\ &\le 2\|u\|_{L^\infty(I)}\,\|u'\|_{L^\infty(I)} - \int_a^b uu''\,dx.\end{aligned}$$

Χρησιμοποιήστε την εκτίμηση της Άσκησης 2.16.]

3. 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 τρεις: Οι μέθοδοι των πεπερασμένων διαφορών, οι μέθοδοι των πεπερασμένων στοιχείων και οι μέθοδοι των πεπερασμένων χωρίων. Η μέθοδος των πεπερασμένων διαφορών αναπτύχθηκε πάρα πολύ στις δεκαετίες του 1950 και του 1960, η μέθοδος των πεπερασμένων στοιχείων κυρίως μετά το 1970· η μέθοδος των πεπερασμένων χωρίων είναι πιο πρόσφατη και προέκυψε από μια προσπάθεια εκσυγχρονισμού και γενίκευσης της μεθόδου των πεπερασμένων διαφορών, κυρίως σε μη Καρτεσιανές συντεταγμένες. Η μέθοδος των πεπερασμένων στοιχείων εφαρμόζεται σε περισσότερα προβλήματα, η ανάλυσή της είναι πιο συστηματοποιημένη 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 και των πεπερασμένων χωρίων είναι αποτελεσματικότερες για κάποιες κατηγορίες προβλημάτων, όπως είναι τα υπερβολικά προβλήματα. Υπάρχουν πάντως πολλές ομοιότητες μεταξύ των τριών αυτών κατηγοριών μεθόδων και έτσι όταν γνωρίζει κανείς και τις τρεις κατανοεί καλύτερα το θέμα. Σε αυτό το κεφάλαιο θα ασχοληθούμε με μεθόδους πεπερασμένων στοιχείων για ένα απλό πρόβλημα, το πρόβλημα δύο σημείων, με στόχο να εξοικειωθεί ο αναγνώστης με τις βασικές τεχνικές και έννοιες.

Στην πρώτη ενότητα θα παρουσιάσουμε συνοπτικά μερικές ιδιότητες της λύσης του προβλήματος δύο σημείων σε τέτοια μορφή ώστε να οδηγηθούμε εύκολα στις ενότητες που ακολουθούν σε ορισμένα διακριτά ανάλογά τους. Στη δεύτερη ενότητα θα γνωρίσουμε τη μέθοδο πεπερασμένων στοιχείων και θα εκτιμήσουμε το σφάλμα της, θα γνωρίσουμε μάλιστα τόσο εκ των προτέρων όσο και εκ των υστέρων εκτιμήσεις του σφάλματος.

3.1 Η θεωρία του προβλήματος

Θεωρούμε το εξής πρόβλημα δύο σημείων για μια συνήθη διαφορική εξίσωση (Σ.Δ.Ε.) δεύτερης τάξης: Ζητείται μια συνάρτηση u, $u \in C^2[0,1]$, τέτοια ώστε

$$\begin{cases} -(au')' + bu' + cu = f \quad \text{στο } [0,1], \\ u(0) = u(1) = 0, \end{cases} \tag{3.1}$$

όπου a και b συνεχώς παραγωγίσιμες συναρτήσεις στο διάστημα $[0,1]$, $a, b \in C^1[0,1]$, και c, f συνεχείς συναρτήσεις στο διάστημα $[0,1]$, $c, f \in C[0,1]$. Σε όλο το κεφάλαιο θα υποθέτουμε ότι η συνάρτηση a λαμβάνει μόνο θετικές τιμές, $a(x) \geq \alpha > 0$. Συχνά θα υποθέτουμε επίσης ότι η συνάρτηση $c - b'/2$ λαμβάνει μη αρνητικές τιμές,

$$a(x) \geq \alpha > 0, \quad c(x) - \frac{1}{2}b'(x) \geq 0, \quad \forall x \in [0,1]. \tag{3.2}$$

Είναι γνωστό ότι υπό τις ανωτέρω συνθήκες το πρόβλημα (3.1) έχει ακριβώς μία *κλασική* λύση, δηλαδή μια συνάρτηση που είναι δύο φορές συνεχώς παραγωγίσιμη στο διάστημα $[0,1]$, μηδενίζεται στα άκρα 0 και 1, και ικανοποιεί τη διαφορική εξίσωση σε κάθε σημείο του διαστήματος $[0,1]$.

Παρατήρηση 3.1 (Κίνητρο για τη συνθήκη μεταξύ b και c.) Χωρίς κάποιον περιορισμό στο πρόσημο των τιμών της συνάρτησης $c - b'/2$, το πρόβλημα (3.1) είναι δυνατόν να έχει πολλές λύσεις, φερ' ειπείν το πρόβλημα

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

έχει τις λύσεις $u(x) = \alpha \sin(\pi x)$ με $\alpha \in \mathbb{R}$. Ακόμη, όταν η συνάρτηση $c - b'/2$ λαμβάνει και αρνητικές τιμές, μπορεί να μην υπάρχει λύση, όπως, επί παραδείγματι, στην περίπτωση του προβλήματος

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

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 $\alpha, \beta \in \mathbb{R}$. Αν και αυτό το γεγονός είναι γνωστό από τη Θεωρία των Διαφορικών Εξισώσεων, επειδή 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, δηλαδή το πρόβλημά μας δεν επιδέχεται λύση. Επιστρέφουμε τώρα στη γενική λύση της Δ.Ε. Έστω w μια άλλη λύση. Τότε για την $v := u - w$ έχουμε

$$-v'' - \pi^2 v = 0 \tag{3.3}$$

και με κατάλληλη επιλογή των α και β μπορούμε να επιτύχουμε

$$v(0) = v'(0) = 0.$$

Πολλαπλασιάζοντας την (3.3) με v' τη γράφουμε στη μορφή

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

και συμπεραίνουμε ότι

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

Σύμφωνα με την $v(0) = v'(0) = 0$ 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 (3.1). Ορισμένες από αυτές θα μας οδηγήσουν στην απόδειξη διακριτών αναλόγων για την προσεγγιστική λύση.

Συμβολίζουμε, ως συνήθως, με $(\cdot,\cdot)$ το εσωτερικό γινόμενο του $L^2(0,1)$ και με $\|\cdot\|$ την παραγόμενη από αυτό νόρμα, δηλαδή

$$(v,w) := \int_0^1 v(x)w(x)\,dx, \quad \|v\| := (v,v)^{1/2}, \quad v,w \in L^2(0,1).$$

Λαμβάνοντας το εσωτερικό γινόμενο και των δύο μελών της εξίσωσης του προβλήματος (3.1) με μια συνάρτηση $v \in H_0^1 := H_0^1(0,1)$, έχουμε

$$-\big((au')',v\big) + (bu',v) + (cu,v) = (f,v),$$

και, συνεπώς, με ολοκλήρωση κατά μέρη,

$$(au',v') + (bu',v) + (cu,v) = (f,v) \quad \forall v \in H_0^1. \tag{3.4}$$

Η (3.4) αποτελεί το κίνητρο για την *ασθενή* ή *μεταβολική* διατύπωση του προβλήματος (3.1): Ζητείται $u \in H_0^1$ τέτοια ώστε

$$(au',v') + (bu',v) + (cu,v) = (f,v) \quad \forall v \in H_0^1. \tag{3.5}$$

Σύμφωνα με την (3.4), κάθε κλασική λύση του προβλήματος (3.1) ικανοποιεί και την (3.5), αποτελεί δηλαδή και ασθενή λύση του προβλήματος.

Η ασθενής διατύπωση (3.5) μας οδηγεί στην εισαγωγή της διγραμμικής μορφής $a : H_0^1 \times H_0^1 \to \mathbb{R}$, $a(v,w) := (av',w') + (bv',w) + (cv,w)$. Το πρόβλημα (3.5) διατυπώνεται τώρα ως

$$a(u,v) = (f,v) \quad \forall v \in H_0^1. \tag{3.6}$$

Ιδιότητες της διγραμμικής μορφής: Κατ' αρχάς θα αποδείξουμε ότι η διγραμμική μορφή a είναι συνεχής στον $\big(H_0^1, \|\cdot\|_1\big)$, χωρίς μάλιστα να χρησιμοποιήσουμε την υπόθεση για το πρόσημο των τιμών της συνάρτησης $c - b'/2$. Για $v,w \in H_0^1$, έχουμε

$$|a(v,w)| \le \max_{0\le x\le 1} a(x)\,\|v'\|\,\|w'\| + \max_{0\le x\le 1} |b(x)|\,\|v'\|\,\|w\| + \max_{0\le x\le 1} |c(x)|\,\|v\|\,\|w\|,$$

οπότε, με

$$C := 2\max\{\max_{0\le x\le 1} a(x), \max_{0\le x\le 1} |b(x)|, \max_{0\le x\le 1} |c(x)|\},$$

έχουμε,

$$|a(v,w)| \le \frac{C}{2}\big(\|v'\|\,\|w'\| + \|v'\|\,\|w\| + \|v\|\,\|w\|\big).$$

Εκτιμώντας το δεξιό μέλος με την ανισότητα των Cauchy–Schwarz για αθροίσματα, παίρνουμε

$$|a(v,w)| \le \frac{C}{2}\big(\|v'\|^2 + \|v'\|^2 + \|v\|^2\big)^{1/2}\big(\|w'\|^2| + \|w\|^2 + \|w\|^2\big)^{1/2}$$

και εύκολα διαπιστώνουμε ότι

$$|a(v,w)| \le C\|v\|_1\,\|w\|_1 \quad \forall v,w \in H_0^1. \tag{3.7}$$

Θα αποδείξουμε τώρα ελλειπτικότητα της διγραμμικής μορφής a στον $\big(H_0^1, \|\cdot\|_1\big)$. Εδώ θα παίξει καθοριστικό ρόλο η συνθήκη (3.2). Για $v \in H_0^1$, έχουμε $a(v,v) = (av',v') + \big(b,(v^2)'\big)/2 + (cv,v)$, οπότε ολοκληρώνοντας στον δεύτερο όρο κατά μέρη, λαμβάνουμε

$$a(v,v) = (av',v') + \big((c - \frac{1}{2}b')v, v\big).$$

Χρησιμοποιώντας εδώ την (3.2), παίρνουμε

$$a(v,v) \ge \alpha\|v'\|^2 \quad \forall v \in H_0^1. \tag{3.8}$$

και σε συνδυασμό με την ανισότητα των Poincaré–Friedrichs (2.43) καταλήγουμε στην ελλειπτικότητα της διγραμμικής μορφής,

$$a(v,v) \ge \frac{\alpha}{2}\|v\|_1^2 \quad \forall v \in H_0^1. \tag{3.9}$$

Ορίζουμε επί πλέον στον H_0^1 το γραμμικό συναρτησιακό F, $F(v) := (f,v)$. Η ανισότητα των Cauchy–Schwarz δίνει αμέσως $|F(v)| \le \|f\|\,\|v\|$, οπότε, ιδιαίτερα, $|F(v)| \le \|f\|\,\|v\|_1$. Το F είναι συνεπώς φραγμένο και $\|F\| \le \|f\|$. Σύμφωνα με το λήμμα των Lax–Milgram, το πρόβλημα (3.6) έχει μία ακριβώς λύση. Αποδείξαμε λοιπόν ύπαρξη και μοναδικότητα της ασθενούς λύσης του προβλήματος δύο σημείων (3.1). Σημειώνουμε ακόμα ότι η ύπαρξη και η μοναδικότητα της ασθενούς λύσης μπορούν να αποδειχθούν και υπό ασθενέστερες συνθήκες στα δεδομένα, παραδείγματος χάριν αρκεί να έχουμε $f \in L^2(0,1)$. Στη συνέχει θα αποδείξουμε ότι η ασθενής λύση μας έχει περισσότερη ομαλότητα, συγκεκριμένα ότι ανήκει στον $H^2(0,1)$.

Ελλειπτική ομαλότητα: Θέτοντας $v := u$ στην (3.6) και χρησιμοποιώντας την (3.8), παίρνουμε $\alpha\|u'\|^2 \le (f,u)$, οπότε $\alpha\|u'\|^2 \le \|f\|\,\|u\|$. Εκτιμώντας τη $\|u\|$ με την $\|u'\|$ στο δεξιό μέλος με την ανισότητα των Poincaré–Friedrichs παίρνουμε

$$\alpha\|u'\| \le \|f\|. \tag{3.10}$$

Πάλι με την ανισότητα των Poincaré–Friedrichs, από την (3.10) λαμβάνουμε

$$\alpha\|u\| \le \|f\|. \tag{3.11}$$

Θα αποδείξουμε τώρα ότι $u \in H^2$. Κατ' αρχάς, η (3.5) δίνει

$$\int_0^1 au'\varphi'\,dx = -\int_0^1 (f - bu' - cu)\varphi\,dx \quad \forall \varphi \in C_0^\infty(0,1).$$

Λαμβάνοντας εδώ υπ' όψιν το γεγονός ότι $f - bu' - cu \in L^2(0,1)$, συμπεραίνουμε ότι η au' έχει ασθενή παράγωγο στον $L^2(0,1)$, οπότε, αφού προφανώς και η ίδια ανήκει στον $L^2(0,1)$, θα ανήκει και στον H^1 και μάλιστα

$$(au')' = f - bu' - cu.$$

Τώρα $u' = (au')/a$, και αφού η a είναι ομαλή και θετική, θα έχουμε $u' \in H^1$. Επομένως, όντως ισχύει $u \in H^2$. Συνολικά δηλαδή έχουμε $u \in H^2 \cap H_0^1$. Επί πλέον, $(au')' = au'' + a'u$, συνεπώς, σύμφωνα με τα προηγούμενα,

$$au'' = f - bu' - cu - a'u',$$

οπότε

$$\alpha\|u''\| \le \|f\| + \max_{0\le x\le 1} |a'(x) + b(x)|\,\|u'\| + \max_{0\le x\le 1} |c(x)|\,\|u\|.$$

Χρησιμοποιώντας εδώ τις (3.10) και (3.11), λαμβάνουμε

$$\alpha\|u''\| \le c_1\|f\|,$$

με μια σταθερά c_1 ανεξάρτητη της f. Τέλος, συνδυάζοντας το τελευταίο αποτέλεσμα με τις (3.10) και (3.11), οδηγούμαστε στην επιθυμητή εκτίμηση ελλειπτικής ομαλότητας (στην απλούστερη μορφή της)

$$\|u\|_2 \le c_2\|f\|, \tag{3.12}$$

με μια σταθερά c_2 ανεξάρτητη της f.

Το συζυγές πρόβλημα: Το πρόβλημα δύο σημείων

$$\begin{cases} -(au')' - bu' + (c-b')u = f \quad \text{στο } [0,1], \\ u(0) = u(1) = 0, \end{cases} \tag{3.13}$$

λέγεται *συζυγές* του προβλήματος (3.1). Είναι γνωστό ότι η ύπαρξη και μοναδικότητα της λύσης του (3.1) είναι ισοδύναμη με την ύπαρξη και μοναδικότητα λύσης του συζυγούς προβλήματος (3.13). Επίσης, χρησιμοποιώντας το γεγονός ότι

$$-\int_0^1 (bv' + b'v)w\,dx = \int_0^1 bvw'\,dx \quad \forall v, w \in H_0^1,$$

διαπιστώνουμε αμέσως ότι η ασθενής διατύπωση του προβλήματος (3.13) είναι:

$$a(v,u) = (f,v) \quad \forall v \in H_0^1. \tag{3.14}$$

Τονίζουμε τη διαφορά μεταξύ των (3.6) και (3.14): Η θέση της λύσης είναι διαφορετική στη διγραμμική μορφή. Αν υποθέσουμε ότι ικανοποιείται η συνθήκη (3.2), εύκολα διαπιστώνουμε ότι το πρόβλημα (3.14) έχει ακριβώς μία ασθενή λύση, και ότι ισχύει η εκτίμηση ελλειπτικής ομαλότητας

$$\|u\|_2 \le \tilde{c}_2 \|f\|, \tag{3.15}$$

με μια σταθερά $\tilde{c}_2$ ανεξάρτητη της f. Στην περίπτωση $b = 0$, το πρόβλημα (3.1) συμπίπτει με το συζυγές του, γιαυτό και λέγεται *αυτοσυζυγές*. Σημειώστε ότι στην περίπτωση $b = 0$ η διγραμμική μορφή a είναι συμμετρική.

Παρατήρηση 3.2 (Η ασθενής λύση είναι και κλασική λύση, για ομαλά δεδομένα.) Υποθέτουμε τώρα ότι τα δεδομένα είναι ομαλά, ιδιαίτερα ότι η f είναι συνεχής στο $[0,1]$. Αφού $u \in H^2$, θα έχουμε $u' \in C[0,1]$. Επομένως $u \in C^1[0,1]$. Από τη σχέση $au'' = f - bu' - cu - a'u'$ διαπιστώνουμε τότε, αφού το δεξιό μέλος είναι συνεχής συνάρτηση, ότι $u \in C^2[0,1]$. Η τελευταία σχέση ισχύει επομένως σε κάθε σημείο, γεγονός που σημαίνει ότι η u ικανοποιεί τη διαφορική εξίσωση. Η u ικανοποιεί κατά τετριμμένο τρόπο και τις συνοριακές συνθήκες, άρα είναι κλασική λύση. Μια διαδικασία σαν αυτή που ακολουθήθηκε εδώ αναφέρεται ως *απόδειξη ομαλότητας* της λύσης.

Παρατήρηση 3.3 (Άλλες συνοριακές συνθήκες.)
(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 (3.1) 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

$$\begin{cases} -(au')' + bu' + cu = f \quad \text{στο } [0,1], \\ u(0) = \gamma, \\ u(1) = \delta, \end{cases} \tag{3.16}$$

με $\gamma, \delta \in \mathbb{R}$. Το πρόβλημα αυτό ανάγεται εύκολα σε ένα πρόβλημα της μορφής (3.1). Πράγματι, κατ' αρχάς η συνάρτηση $w, w(x) := \gamma(1-x) + \delta x$, ικανοποιεί τις συνοριακές συνθήκες του προβλήματος (3.16), $w(0) = \gamma$ και $w(1) = \delta$. Για τη συνάρτηση $v,\ v := u - w$, έχουμε τώρα προφανώς $v(0) = v(1) = 0$, και επί πλέον

$$-(av')' + bv' + cv = -(au')' + bu' + cu + (aw')' - bw' - cw = f + (\gamma - \delta)(a' + b) - cw.$$

Συνεπώς, με $g := f + (\gamma - \delta)(a' + b) - cw$, η v λύνει το πρόβλημα

$$\begin{cases} -(av')' + bv' + cv = g \quad \text{στο } [0,1], \\ v(0) = v(1) = 0, \end{cases} \tag{3.17}$$

που είναι της ίδιας μορφής με το (3.1). Επίσης εύκολα διαπιστώνει κανείς ότι αν η v λύνει το πρόβλημα (3.17), τότε η $u := v + w$ λύνει το πρόβλημα (3.16). Μπορούμε λοιπόν χωρίς περιορισμό της γενικότητας να ασχοληθούμε με την περίπτωση ομογενών συνοριακών συνθηκών.

(ii) Τροποποιούμε τώρα τις συνοριακές συνθήκες στο πρόβλημα (3.1) και θεωρούμε το εξής πρόβλημα δύο σημείων με ομογενείς συνοριακές *συνθήκες Neumann*

$$\begin{cases} -(au')' + cu = f \quad \text{στο } [0,1], \\ u'(0) = u'(1) = 0, \end{cases} \tag{3.18}$$

όπου $a \in C^1[0,1]$, $c, f \in C[0,1]$, και η c λαμβάνει μη αρνητικές τιμές, $c(x) \geq 0$, για κάθε $x \in [0,1]$.

Η συνθήκη $c \neq 0$ είναι αναγκαία στην προκειμένη περίπτωση για να έχουμε μοναδικότητα της λύσης, αφού για $c = 0$ 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,

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

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 3.5.

Γενικότερα μπορούμε να θεωρήσουμε το πρόβλημα

$$\begin{cases} -(au')' + cu = f & \text{στο } [0,1], \\ -u'(0) + \sigma_1 u(0) = 0, \\ u'(1) + \sigma_2 u(1) = 0, \end{cases} \tag{3.19}$$

με $c(x) \geq 0$, για κάθε $x \in [0,1]$, και $\sigma_1, \sigma_2 \geq 0$. Στην περίπτωση $c = 0$ υποθέτουμε επί πλέον ότι $\sigma_1 + \sigma_2 > 0$, ούτως ώστε, πέραν των παραγώγων της, να εμφανίζεται και η u σε κάποιο σημείο του προβλήματος. Και σε αυτή την περίπτωση έχουμε ακριβώς μία λύση, για τη μοναδικότητα βλ. την Άσκηση 3.6.

(iii) 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 (3.19) 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

$$\begin{cases} -(au')' + cu = f & \text{στο } [0,1], \\ u(0) = u'(1) = 0, \end{cases}$$

με a, c και f όπως στο πρόβλημα (3.1).

(iv) Ας υποθέσουμε ότι οι συναρτήσεις a, b, c και f μπορούν να επεκταθούν περιοδικά στο $\mathbb{R}$ ως συνεχώς παραγωγίσιμες οι πρώτες δύο και συνεχείς συναρτήσεις οι άλλες, με περίοδο ένα, ότι ισχύει δηλαδή $c(1) = c(0)$, $f(1) = f(0)$, $a(1) = a(0)$, $b(1) = b(0)$ και $a'(1) = a'(0)$, $b'(1) = b'(0)$. Αν θέλουμε οι επεκτάσεις αυτές να είναι περισσότερο ομαλές, πρέπει να απαιτήσουμε ισότητα αντιστοίχων παραγώγων στα άκρα. Σε αυτή την περίπτωση μπορούμε να θεωρήσουμε το ακόλουθο *περιοδικό* πρόβλημα: Ζητείται μια περιοδική συνάρτηση u, $u \in C^2(\mathbb{R})$, με περίοδο 1, τέτοια ώστε

$$-(au')' + bu' + cu = f \quad \text{στο } \mathbb{R}. \tag{3.20}$$

Αν η συνάρτηση $c - b'/2$ λαμβάνει μη αρνητικές τιμές, $c(x) - b'(x)/2 \geq 0$, για κάθε $x \in \mathbb{R}$, και δεν είναι η μηδενική συνάρτηση, τότε το πρόβλημα αυτό έχει μία ακριβώς λύση. □

3.2 Μέθοδοι πεπερασμένων στοιχείων

Σε αυτή την ενότητα θα μελετήσουμε μεθόδους πεπερασμένων στοιχείων για το πρόβλημα δύο σημείων. Η ενότητα αποτελείται από τέσσερεις παραγράφους, οι δύο πρώτες εκ των οποίων αναφέρονται σε εκ των προτέρων εκτιμήσεις ενώ η τρίτη σε εκ των υστέρων εκτιμήσεις· στην τέταρτη δίνονται κάποιες προσεγγιστικές ιδιότητες χώρων που αποτελούνται από συνεχείς συναρτήσεις, οι οποίες είναι τμηματικά πολυώνυμα μέχρι ενός συγκεκριμένου βαθμού και παίζουν σημαντικό ρόλο στις τρεις πρώτες ενότητες. Συγκεκριμένα, στην πρώτη παράγραφο θα ασχοληθούμε με το "ορισμένο" πρόβλημα, θα υποθέσουμε δηλαδή ότι η συνάρτηση $c - b'/2$ λαμβάνει μη αρνητικές τιμές, ενώ στη δεύτερη με το "μη ορισμένο" επιτρέποντας και αρνητικές τιμές της $c - b'/2$, 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.

3.2.1 Το ορισμένο πρόβλημα

Θεωρούμε το πρόβλημα δύο σημείων (3.1). Η μέθοδος των πεπερασμένων στοιχείων βασίζεται στη μεταβολική διατύπωση (3.6) του προβλήματος (3.1).

Έστω $r \geq 2$ ένας φυσικός αριθμός και $(S_h^r)_{0<h\leq 1} \subset H_0^1 := H_0^1(0,1)$ μια οικογένεια

υποχώρων του H_0^1, πεπερασμένης διάστασης, με την εξής *προσεγγιστική ιδιότητα*

$$\begin{aligned}&\exists c > 0 \quad \forall v \in H^s \cap H_0^1 \quad \exists \chi \in S_h^r \\ &\|v - \chi\| + h\|v' - \chi'\| \le ch^s\|v\|_s, \quad s = 2,\ s = r\end{aligned} \tag{3.21}$$

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

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

Παράδειγμα 3.1 (Χώροι με την προσεγγιστική ιδιότητα (3.21).) Έστω $S_h^r \subset H_0^1$ ο χώρος των splines ως προς τον ομοιόμορφο διαμερισμό του $[0,1]$ με βήμα h, οι οποίες σε κάθε υποδιάστημα αυτού του διαμερισμού είναι πολυώνυμα βαθμού το πολύ $r-1$, και στο $[0,1]$ είναι $r-2$ φορές συνεχώς παραγωγίσιμες συναρτήσεις. Είναι γνωστό ότι σε αυτή την περίπτωση ισχύει η (3.21). Με το θέμα θα ασχοληθούμε αναλυτικά στην τελευταία ενότητα αυτού του κεφαλαίου. Εδώ θα δώσουμε την απόδειξη για την περίπτωση $r = 2$. Έστω λοιπόν $v \in H^2 \cap H_0^1$ και $v_h \in S_h^2$ τέτοια ώστε $v_h(x_i) = v(x_i)$, $i = 0,\dots,J+1$, όπου $x_i = ih$, $i = 0,\dots,J+1, (J+1)h = 1$. Τότε, θέτοντας $w := v - v_h$ έχουμε, σύμφωνα με τη (2.44),

$$\int_{x_i}^{x_{i+1}} [w(x)]^2 dx \le h^2 \int_{x_i}^{x_{i+1}} [w'(x)]^2 dx, \quad i = 0,\dots,J. \tag{3.22}$$

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

$$\|w\| \le h\|w'\|. \tag{3.23}$$

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

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

δηλαδή

$$\int_{x_i}^{x_{i+1}} |w'(x)|^2 dx \le h^2 \int_{x_i}^{x_{i+1}} |v''(s)|^2 ds. \tag{3.24}$$

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

$$\|w'\| \le h\|v''\|. \tag{3.25}$$

Από τις (3.23) και (3.25) έπεται η (3.21) για $s = r = 2$. □

Σε αυτή την παράγραφο υποθέτουμε ότι η συνάρτηση $c - b'/2$ λαμβάνει μόνο μη αρνητικές τιμές. Θεωρούμε τώρα ένα σταθερό h. Προσεγγίζουμε τη λύση u του προβλήματος (3.1) με ένα στοιχείο $u_h \in S_h^r$ τέτοιο ώστε

$$a(u_h, \chi) = (f, \chi) \quad \forall \chi \in S_h^r, \tag{3.26}$$

βλ. την (3.6). Ύπαρξή και μοναδικότητα της προσεγγιστικής λύσης έπονται αμέσως από το λήμμα των Lax–Milgram, αφού ο S_h^r είναι πλήρης χώρος, ως χώρος πεπερασμένης διάστασης. Στην προκειμένη περίπτωση υπάρχει και ευκολότερη απόδειξη, η οποία μας παρέχει και τη δυνατότητα υπολογισμού της u_h : Έστω $N_h := \dim S_h^r$, και $\{\varphi_1, \dots, \varphi_{N_h}\}$ μια βάση του S_h^r. Η (3.26) γράφεται ισοδύναμα στη μορφή

$$a(u_h, \varphi_i) = (f, \varphi_i), \quad i = 1, \dots, N_h. \tag{3.27}$$

Παριστάνοντας τη u_h στη μορφή $u_h = \alpha_1\varphi_1 + \dots + \alpha_{N_h}\varphi_{N_h}$ διαπιστώνουμε εύκολα ότι το (3.27) είναι ένα γραμμικό σύστημα και ο αντίστοιχος πίνακας είναι 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 (3.26) (και χρησιμοποιώντας τη (2.43)) λαμβάνουμε αμέσως $u_h = 0$.

Το σύστημα (3.27) είναι της μορφής

$$A_h \alpha_h = f_h, \tag{3.28}$$

όπου $\alpha_h = (\alpha_1, \dots, \alpha_{N_h})^T, (f_h)_i = (f, \varphi_i), i = 1, \dots, N_h$, και $(A_h)_{ij} = a(\varphi_i, \varphi_j)$, $i, j = 1, \dots, N_h$. 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 (3.28) 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 3.2 (Βάσεις με υπολογιστικά πλεονεκτήματα.) Έστω $h := \frac{1}{J+1}, x_i := ih, i = 0, \dots, J+1$, και S_h^2 ο χώρος των συνεχών συναρτήσεων οι οποίες μηδενίζονται στα σημεία 0 και 1 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

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

διαπιστώνουμε αμέσως ότι ο πίνακας A_h είναι τριδιαγώνιος. $\square$

Μια προσεγγιστική ιδιότητα της u_h δίνεται στο ακόλουθο θεώρημα.

Θεώρημα 3.1 (Εκτιμήσεις του σφάλματος.) *Έστω $r \ge 2$, και $u \in H^r \cap H_0^1$ η λύση του προβλήματος* (3.1)*. Έστω $S_h^r \subset V$ τέτοιος ώστε να ισχύει η* (3.21) *και $u_h \in S_h^r$ η λύση του* (3.26)*. Τότε υπάρχει μια σταθερά C, ανεξάρτητη των u και h, τέτοια ώστε*

$$\|u' - u_h'\| \le Ch^{r-1}\|u\|_r, \tag{3.29}$$

και

$$\|u - u_h\| \le Ch^r\|u\|_r. \tag{3.30}$$

Απόδειξη. . Οι (3.6) και (3.26) γράφονται τώρα στη μορφή

$$a(u, v) = (f, v) \quad \forall v \in H_0^1 \tag{3.31}$$

$$a(u_h, \chi) = (f, \chi) \quad \forall \chi \in S_h^r. \tag{3.32}$$

Ιδιαίτερα, η (3.31) ισχύει για $v \in S_h^r$, δηλαδή

$$a(u, \chi) = (f, \chi) \quad \forall \chi \in S_h^r. \tag{3.33}$$

Αφαιρώντας κατά μέλη τις (3.32) και (3.33) λαμβάνουμε

$$a(u - u_h, \chi) = 0 \quad \forall \chi \in S_h^r. \tag{3.34}$$

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 (3.7) και (3.9),

$$\|u - u_h\|_1 \le 2\frac{C}{\alpha} \min_{\chi \in S_h^r} \|u - \chi\|_1. \tag{3.35}$$

Χρησιμοποιώντας εδώ την προσεγγιστική ιδιότητα (3.21) οδηγούμαστε αμέσως στην (3.29).

Για να αποδείξουμε την (3.30) θα χρησιμοποιήσουμε το λεγόμενο *τέχνασμα του Nitsche*. Θεωρούμε κατ' αρχάς το πρόβλημα

$$\begin{cases} -(aw')' - bw' + (c - b')w = u - u_h & \text{στο } [0,1], \\ w(0) = w(1) = 0, \end{cases} \tag{3.36}$$

βλ. το πρόβλημα (3.13). Σύμφωνα με την (3.15) ισχύει τότε

$$\|w\|_2 \le \tilde{c}\|u - u_h\|. \tag{3.37}$$

Επί πλέον, σύμφωνα με την (3.14), έχουμε

$$a(v, w) = (u - u_h, v) \quad \forall v \in H_0^1.$$

Συνεπώς, για $v := u - u_h$,

$$\|u - u_h\|^2 = a(u - u_h, w). \tag{3.38}$$

Επομένως, σύμφωνα με την (3.34), για $\chi \in S_h^r$, έχουμε

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

οπότε, λόγω της (3.7),

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

Σύμφωνα τώρα με τις (3.21) για $s = 2$ και (3.37), αυτή η σχέση δίνει

$$\|u - u_h\| \le Ch\|u - u_h\|_1, \tag{3.39}$$

η οποία συνδυαζόμενη με την (3.29) δίνει την (3.30). □

Παρατήρηση 3.4 (Μέθοδος του Ritz.) Η μεταβολική διατύπωση (3.32) (ή ισοδύναμα η (3.26)) 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 στην περίπτωση του αυτοσυζυγούς προβλήματος, δηλαδή όταν $b = 0,$ μπορεί να διατυπωθεί και ως πρόβλημα ελαχιστοποίησης, και τότε αναφέρεται ως *μέθοδος του Ritz.* Συγκεκριμένα, υποθέτοντας ότι $b = 0,$ θα αποδείξουμε ότι η λύση u_h του προβλήματος (3.32) 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 του Galerkin είναι γενικότερη, αφού εφαρμόζεται και σε περιπτώσεις όπου η αντίστοιχη διγραμμική μορφή δεν είναι συμμετρική, όπως, φερ' ειπείν, στην περίπτωση του προβλήματος (3.1) με $b \neq 0.$

3.2.2 Ένα μη ορισμένο πρόβλημα

Στη μέχρι τώρα μελέτη αριθμητικών μεθόδων για το πρόβλημα δύο σημείων (3.1) περιοριστήκαμε στην περίπτωση όπου η συνάρτηση $c - b'/2$ λαμβάνει μόνο μη αρνητικές τιμές. Ένα φυσιολογικό ερώτημα είναι πώς συμπεριφέρονται οι αριθμητικές μέθοδοι στην περίπτωση που το πρόβλημα (3.1) έχει μεν ακριβώς μία λύση, αρκετά ομαλή, αλλά η $c - b'/2$ επιτρέπεται να παίρνει και *αρνητικές* τιμές. Η μη αρνητικότητα της $c - b'/2$ χρησιμοποιήθηκε στην προηγούμενη παράγραφο για να αποδείξουμε την ελλειπτικότητα της διγραμμικής μορφής $a,$ βλ. την (3.9) (και για να αποδείξουμε βέβαια ότι ο πίνακας του γραμμικού συστήματος (3.30) είναι θετικά ορισμένος, αλλά αυτό το γεγονός σχετίζεται με την ελλειπτικότητα).

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 (3.30) και (3.29). Υπάρχει ανάλυση και μεθόδων πεπερασμένων διαφορών για μη ορισμένα προβλήματα, παρά το ότι σε 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 (3.1) και υποθέτουμε ότι αυτό έχει μία ακριβώς λύση, η οποία υποτίθεται αρκετά ομαλή. Επιτρέπουμε στη $c - b'/2$ να παίρνει και αρνητικές τιμές. Θεωρούμε μια οικογένεια $(S_h^r)_{0<h\leq 1}$ υποχώρων πεπερασμένης διάστασης του H_0^1 με την προσεγγιστική ιδιότητα (3.21). Έστω $u_h \in S_h^r$ μια προσεγγιστική λύση πεπερασμένων στοιχείων τέτοια ώστε

$$a(u_h, \chi) = (f, \chi) \qquad \forall \chi \in S_h^r, \tag{3.40}$$

βλ. την (3.26). Εύκολα μπορεί κανείς να δώσει παραδείγματα στα οποία δεν εξασφαλίζεται ύπαρξη και μοναδικότητα της u_h, δηλαδή τέτοια ώστε ο πίνακας A_h του γραμμικού συστήματος (3.28) να μην είναι αντιστρέψιμος, όταν η συνάρτηση $c-b'/2$ 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 (3.29) και (3.30).

Στη συνέχεια θα χρησιμοποιήσουμε, χωρίς να την αποδείξουμε, την ελλειπτική ομαλότητα και για την περίπτωση που η συνάρτηση $c - b'/2$ παίρνει και αρνητικές τιμές. Είναι γνωστό ότι, αν το πρόβλημα (3.1) έχει μία ακριβώς ομαλή λύση για μια 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 (3.9). 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 μια εκτίμηση της μορφής $\|u\| \leq c_1\|f\|$, τότε η απόδειξη της εκτίμησης της ελλειπτικής ομαλότητας ολοκληρώνεται εύκολα με τη μέθοδο της ενέργειας, την οποία χρησιμοποιήσαμε επανειλημμένα εδώ. Επίσης, όταν το πρόβλημα (3.1) έχει μία ακριβώς λύση, το ίδιο συμβαίνει και για το συζυγές πρόβλημα (3.13), ισχύει μάλιστα και για αυτό η εκτίμηση της ελλειπτικής ομαλότητας.

Θεωρούμε τη διγραμμική μορφή a που ορίστηκε λίγο πριν την (3.6). Όπως είχαμε ήδη τονίσει, η a είναι συνεχής και στην περίπτωση που μελετάμε τώρα, δηλαδή ισχύει η (3.7), δεν είναι όμως κατ' ανάγκη ελλειπτική, δεν ισχύει δηλαδή αναγκαστικά η (3.9), αλλά αντ' αυτής έχουμε την *ανισότητα του Gårding*

$$a(v,v) \geq \frac{\alpha}{2}\|v\|_1^2 - \beta\|v\|^2 \qquad \forall v \in H_0^1, \tag{3.41}$$

με τη σταθερά α της (3.2) και $\beta := -\min_{0\le x\le 1}[c(x) - b'(x)/2]$, βλ. την απόδειξη της (3.9). Ακριβώς επειδή η a δεν είναι ελλειπτική στην παρούσα περίπτωση, αναφέρεται το πρόβλημα ως *μη ορισμένο*.

Θεώρημα 3.2 (Εκτιμήσεις του σφάλματος.) *Έστω ότι το πρόβλημα* (3.1) *έχει μία ακριβώς ομαλή λύση u, $u \in H^r \cap H_0^1$. Έστω $r \geq 2$, και $S_h^r \subset H_0^1$ χώροι πεπερασμένης διάστασης τέτοιοι ώστε να ισχύει η* (3.21). *Τότε υπάρχουν δύο θετικές σταθερές h_0 και C, τέτοιες ώστε για κάθε $h \leq h_0$ το πρόβλημα* (3.40) *έχει μία ακριβώς λύση u_h, για την οποία μάλιστα ισχύουν οι εκτιμήσεις*

$$\|u' - u_h'\| \leq Ch^{r-1}\|u\|_r, \tag{3.42}$$

και

$$\|u - u_h\| \leq Ch^r\|u\|_r. \tag{3.43}$$

Απόδειξη. Θα αποδείξουμε πρώτα τις εκτιμήσεις (3.42) και (3.43) για κάθε προσεγγιστική λύση u_h της (3.40) 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 3.1, εδώ όμως φυσικά η συνάρτηση $c-b'/2$ επιτρέπεται να παίρνει και αρνητικές τιμές. Αφαιρώντας κατά μέλη τις (3.33) και (3.40) λαμβάνουμε

$$a(u-u_h,\chi) = 0 \quad \forall \chi \in S_h^r, \tag{3.44}$$

βλ. την (3.34). Σύμφωνα με τις (3.41) και (3.44) έχουμε, για $\chi \in S_h^r$,

$$\begin{aligned}\frac{\alpha}{2}\|u-u_h\|_1^2 &\leq a(u-u_h,u-u_h) + \beta\|u-u_h\|^2\\ &= a(u-u_h,u-\chi) + \beta\|u-u_h\|^2,\end{aligned}$$

οπότε, χρησιμοποιώντας την (3.7),

$$\frac{\alpha}{2}\|u-u_h\|_1^2 \le C\|u-u_h\|_1\,\|u-\chi\|_1 + \beta\|u-u_h\|^2 \quad \forall \chi \in S_h^r.$$

Επομένως, χρησιμοποιώντας την προσεγγιστική ιδιότητα (3.21), έχουμε

$$\frac{\alpha}{2}\|u-u_h\|_1^2 \le Ch^{r-1}\|u-u_h\|_1\,\|u\|_r + \beta\|u-u_h\|^2. \tag{3.45}$$

Για να εκτιμήσουμε την L^2-νόρμα θα χρησιμοποιήσουμε πάλι το τέχνασμα του Nitsche. Θεωρούμε το πρόβλημα (3.36). Σύμφωνα με όσα αναφέραμε, η w ορίζεται κατά μοναδικό τρόπο και ισχύει

$$\|w\|_2 \le \tilde{c}\|u-u_h\|. \tag{3.46}$$

Επί πλέον, εύκολα διαπιστώνουμε ότι

$$a(v,w) = (u-u_h, v) \quad \forall v \in H_0^1.$$

Συνεπώς, για $v := u-u_h$,

$$\|u-u_h\|^2 = a(u-u_h, w). \tag{3.47}$$

Επομένως, σύμφωνα με την (3.44), για $\chi \in S_h^r$, έχουμε

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

οπότε, λόγω της (3.7),

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

Σύμφωνα τώρα με τις (3.21) και (3.46), αυτή η σχέση δίνει

$$\|u-u_h\| \le Ch\|u-u_h\|_1. \tag{3.48}$$

Από τις (3.45) και (3.48) λαμβάνουμε

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

και επομένως, για αρκετά μικρό h,

$$\|u-u_h\|_1 \le Ch^{r-1}\|u\|_r. \tag{3.49}$$

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 (3.40) μπορεί να γραφεί ως γραμμικό σύστημα με τετραγωνικό πίνακα, βλ. την (3.28). Για $f = 0$ οδηγούμαστε στο αντίστοιχο ομογενές σύστημα. Τότε όμως $u = 0$, και, σύμφωνα με την (3.49), $u_h = 0$, δηλαδή το αντίστοιχο ομογενές σύστημα έχει μόνο την τετριμμένη λύση. Συνεπώς, η u_h είναι καλώς ορισμένη, για αρκετά μικρό h. Επί πλέον, η (3.42) έπεται αμέσως από την (3.49), και η (3.43) από τις (3.48) και (3.49). □

3.2.3 Εκ των υστέρων εκτιμήσεις του σφάλματος

Εκτιμήσεις του σφάλματος της μορφής (3.29) και (3.30), καθώς και όλες οι άλλες που έχουμε δει μέχρι τώρα, λέγονται εκτιμήσεις *εκ των προτέρων* για τον 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 (3.29), 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 στην πρώτη ενότητα. Πράγματι κατ' αυτόν τον τρόπο οδηγείται κανείς σε μια ποσότητα που θα μπορούσε να υπολογισθεί, αυτή όμως δεν είναι πρακτικά χρήσιμη αφού είναι γενικά υπερβολικά απαισιόδοξη, το φράγμα που προκύπτει είναι πολύ μεγαλύτερο από το πραγματικό σφάλμα. Μια αιτία γι' αυτό αποτελεί το γεγονός ότι στο φράγμα στην (3.29) χρησιμοποιείται μόνο το μέγιστο εύρος h, αλλά όχι περισσότερες λεπτομέρειες για τον 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 (3.29) εκ των υστέρων εκτίμηση είναι της μορφής

$$\|u' - u_h'\| \le \eta(u_h), \tag{3.50}$$

στην οποία το *συναρτησιακό εκτίμησης του σφάλματος* (error estimator), η, εξαρτάται από τα δεδομένα του προβλήματος, όπως οι συναρτήσεις a, b, c 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 $c - b'/2$ παίρνει μόνο μη αρνητικές τιμές.

Θεωρούμε έναν διαμερισμό $0 = x_0 < x_1 < \cdots < x_J < x_{J+1} = 1$ του $[0,1]$ 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 0 και 1.

Έστω $u_h \in S_h^r$ η λύση πεπερασμένων στοιχείων για το πρόβλημα (3.1), βλ. την (3.32). Η u_h ικανοποιεί τις συνοριακές συνθήκες του προβλήματος (3.1). Ένα μέγεθος για το πόσο η προσεγγιστική λύση αποτυγχάνει να είναι ακριβής λύση του (3.1) αποτελεί το λεγόμενο *υπόλοιπο* r_h,

$$r_h(x) := -(au_h')'(x) + b(x)u_h'(x) + c(x)u_h(x) - f(x). \tag{3.51}$$

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 3.3 (Εκ των υστέρων εκτίμηση στην H^1-νόρμα.) *Έστω $u_h \in S_h^r$ η λύση του* (3.32) *και r_h το υπόλοιπό της. Έστω $u \in H^2 \cap H_0^1$ η λύση του* (3.1)*. Τότε, για το σφάλμα $u - u_h$ ισχύει η εκ των υστέρων εκτίμηση*

$$\|u' - u_h'\| \le \frac{1}{2\alpha}\Big(\sum_{i=0}^{J} h_i^2 \|r_h\|^2_{L^2(x_i,x_{i+1})}\Big)^{1/2}, \tag{3.52}$$

όπου α η σταθερά στην (3.8) *και $\|\cdot\|_{L^2(x_i,x_{i+1})}$ 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

$$\alpha \|e_h'\|^2 \le a(e_h, e_h). \tag{3.53}$$

Για να οδηγηθούμε στην (3.52) αρκεί να εκτιμήσουμε κατάλληλα το δεξιό μέλος της (3.53). Υποθέτουμε κατ' αρχάς, για να γίνουν ευκολότερα οι υπολογισμοί, ότι έχουμε μια 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 $a =$ σταθερά και $b = c = 0$, βλ. την Άσκηση 3.23, 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 $[0,1]$, αφού η v δεν είναι ομαλή στους κόμβους, γράφουμε το $a(e_h, v)$ ως

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

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

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

και οδηγούμαστε εύκολα στη σχέση

$$\begin{aligned} a(e_h, v) &= -\big((a e_h')', v\big) + (b e_h', v) + (c e_h, v) \\ &= -\big((a(u' - u_h'))', v\big) + (b(u - u_h)', v) + (c(u - u_h), v), \end{aligned}$$

δηλαδή

$$a(e_h, v) = \big(-(a u')' + b u' + c u, v\big) - \big(-(a u_h')' + b u_h' + q u_h, v\big).$$

Τώρα $-(au')' + bu' + cu = f$ και χρησιμοποιώντας και τον ορισμό του r_h οδηγούμαστε στην επιθυμητή παράσταση

$$a(e_h, v) = -(r_h, v). \tag{3.54}$$

Τώρα, λόγω της (3.34) 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

$$\|e_h' - \chi'\|_{L^2(x_i,x_{i+1})} \le \|e_h'\|_{L^2(x_i,x_{i+1})}. \tag{3.55}$$

Εξ άλλου, χρησιμοποιώντας την ανισότητα των Poincaré–Friedrichs, βλ. την Άσκηση 3.2, έχουμε

$$\|e_h - \chi\|_{L^2(x_i,x_{i+1})} \le \frac{1}{2}h_i\|e_h' - \chi'\|_{L^2(x_i,x_{i+1})}. \tag{3.56}$$

Συνδυάζοντας τις (3.55) και (3.56) οδηγούμαστε σε ένα χρήσιμο για τη συνέχεια αποτέλεσμα,

$$\|e_h - \chi\|_{L^2(x_i,x_{i+1})} \le \frac{1}{2}h_i\|e_h'\|_{L^2(x_i,x_{i+1})}. \tag{3.57}$$

Σύμφωνα με τα προηγούμενα έχουμε τώρα, χρησιμοποιώντας πρώτα την ανισότητα των Cauchy–Schwarz για ολοκληρώματα, μετά την (3.57) 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

$$a(e_h, e_h) \le \frac{1}{2}\Big(\sum_{i=0}^{J} h_i^2 \|r_h\|^2_{L^2(x_i,x_{i+1})}\Big)^{1/2} \|e_h'\|. \tag{3.58}$$

Από τις (3.53) και (3.58) έπεται αμέσως η εκτίμηση (3.52). □

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 (3.36). Κατ' αρχάς, σύμφωνα με την ανισότητα των Poincaré–Friedrichs, βλ. την Άσκηση 3.2, έχουμε

$$\|w\| \le \frac{1}{2}\|w'\|. \tag{3.59}$$

Παίρνοντας το εσωτερικό γινόμενο στη διαφορική εξίσωση του (3.36) με την w διαπιστώνουμε αμέσως ότι

$$(aw', w') + \big((c - \frac{b'}{2})w, w\big) = (u - u_h, w),$$

συνεπώς $\alpha\|w'\|^2 \le \|u - u_h\|\,\|w\|$, οπότε η (3.59) δίνει

$$\|w\| \le \frac{1}{4\alpha}\|u - u_h\|. \tag{3.60}$$

Από την $\alpha\|w'\|^2 \le \|u - u_h\|\,\|w\|$ και την (3.59) παίρνουμε επίσης

$$\|w'\| \le \frac{1}{2\alpha}\|u - u_h\|. \tag{3.61}$$

Χρησιμοποιώντας τώρα τη διαφορική εξίσωση του (3.36) παίρνουμε

$$-aw'' = (a' + b)w' + (b' - c)w + u - u_h,$$

οπότε

$$\alpha\|w''\| \leq \|a' + b\|_\infty \, \|w'\| + \|b' - c\|_\infty \, \|w\| + \|u - u_h\|,$$

όπου $\|\cdot\|_\infty$ είναι η νόρμα μεγίστου στο διάστημα $[0,1]$. Συνδυάζοντας αυτή την εκτίμηση με τις (3.60) και (3.61) λαμβάνουμε τη σημαντική για τη συνέχεια εκτίμηση

$$\|w''\| \leq \gamma\|u - u_h\| \tag{3.62}$$

με

$$\gamma := \frac{1}{\alpha}\Big[\frac{1}{2\alpha}\|a' + b\|_\infty + \frac{1}{4\alpha}\|b - c\|_\infty + 1\Big]. \tag{3.63}$$

Θεώρημα 3.4 (Εκ των υστέρων εκτίμηση στην L^2-νόρμα.) *Έστω $u_h \in S_h^r$ η λύση του* (3.32) *και r_h το υπόλοιπό της. Έστω $u \in H^2 \cap H_0^1$ η λύση του* (3.1)*. Τότε, για το σφάλμα $u - u_h$ ισχύει η εκ των υστέρων εκτίμηση*

$$\|u - u_h\| \leq \frac{1}{2}\gamma\Big(\sum_{i=0}^{J} h_i^4\|r_h\|^2_{L^2(x_i,x_{i+1})}\Big)^{1/2}, \tag{3.64}$$

με γ της σταθερά της (3.63).

Απόδειξη. Συμβολίζουμε πάλι με e_h το σφάλμα, $e_h := u - u_h$. Σύμφωνα με την (3.38) έχουμε

$$\|e_h\|^2 = a(e_h, w). \tag{3.65}$$

Έστω τώρα χ η τμηματικά γραμμική παρεμβάλλουσα 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 (3.65) γράφεται στη μορφή

$$\|e_h\|^2 = a(e_h, w - \chi). \tag{3.66}$$

Τώρα

$$\|w - \chi\|_{L^2(x_i,x_{i+1})} \leq \frac{1}{2}h_i\|w' - \chi'\|_{L^2(x_i,x_{i+1})}, \tag{3.67}$$

βλ. την (3.56). 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

$$\|w' - \chi'\|_{L^2(x_i,x_{i+1})} \le h_i \|w''\|_{L^2(x_i,x_{i+1})}. \tag{3.68}$$

Οι (3.67) και (3.68) δίνουν

$$\|w - \chi\|_{L^2(x_i,x_{i+1})} \le \frac{1}{2} h_i^2 \|w''\|_{L^2(x_i,x_{i+1})}. \tag{3.69}$$

Προχωρώντας όπως στην απόδειξη της (3.58) και χρησιμοποιώντας τις (3.66), (3.54) και (3.69), 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

$$\|e_h\|^2 \le \frac{1}{2} \Big(\sum_{i=0}^J h_i^4 \|r_h\|^2_{L^2(x_i,x_{i+1})} \Big)^{1/2} \|w''\|. \tag{3.70}$$

Από τις (3.62) και (3.70) έπεται αμέσως η εκτίμηση (3.64). □

Παρατήρηση 3.5 (Εναλλακτική μορφή του υπολοίπου· άλλη εκ των υστέρων εκτίμηση.)

i. Χρησιμοποιώντας τη διαφορική εξίσωση του (3.1) στην (3.51), μπορούμε να γράψουμε το υπόλοιπο στη μορφή

$$r_h = \big(a(u-u_h)'\big)' - b(u-u_h)' - c(u-u_h). \tag{3.71}$$

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

ii. Στην απόδειξη του Θεωρήματος 3.3 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 3.16, λαμβάνουμε

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

οπότε, με τη βοήθεια και της (3.55), έχουμε

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

Συνδυάζοντας αυτό το αποτέλεσμα με την (3.53) οδηγούμαστε στην εξής εκ των υστέρων εκτίμηση, βλ. την (3.52),

$$\|u' - u_h'\| \le \frac{1}{\alpha}\|r_h\|_{-1}. \tag{3.72}$$

Παρατήρηση 3.6 (Αυτόματη επιλογή διαμερισμών.) 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 $[0, 1]$. 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 (3.52), εύκολα διαπιστώνουμε ότι αυτό εξασφαλίζεται από τις συνθήκες

$$h_i \|r_h\|^2_{L^2(x_i,x_{i+1})} \le 4\alpha^2\varepsilon^2 \tag{3.73}$$

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, βλ. την αντίστοιχη συζήτηση στην παράγραφο 3.5 του [2], 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 $\varepsilon_1 := 4\alpha^2\varepsilon^2$. 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

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

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 στον 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 στον 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 στον 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 (3.73) για όλα τα υποδιαστήματα με μήκος μεγαλύτερο του δ, επαναλαμβάνεται η όλη διαδικασία για τον νέο διαμερισμό. □

3.2.4 Χώροι με την προσεγγιστική ιδιότητα (3.21)

Σε αυτή την ενότητα θα δώσουμε παραδείγματα υποχώρων του $H_0^1(0,1)$, οι οποίοι έχουν την προσεγγιστική ιδιότητα (3.21). Ένα θεμελιώδες εργαλείο σε αποδείξεις προσεγγιστικών ιδιοτήτων αποτελεί το λήμμα των Bramble–Hilbert. [Οι Bramble και Hilbert είναι σύγχρονοι μαθηματικοί· ιδιαίτερα, ο Hilbert δεν είναι το ίδιο πρόσωπο με τον Hilbert, κατά τον οποίο αποκαλούνται οι αντίστοιχοι χώροι.]

Πρόταση 3.1 (Το λήμμα των Bramble–Hilbert.) *Έστω $r \geq 2$ ένας φυσικός αριθμός. Έστω $F : H^r(0,1) \to [0,\infty)$ ένα συναρτησιακό τέτοιο ώστε*

$$\begin{aligned} &F(v+w) \leq F(v) + F(w) && \forall v, w \in H^r,\\ &F(v) \leq C\|v\|_r && \forall v \in H^r,\\ &F(p) = 0 && \forall p \in \mathbb{P}_{r-1}, \end{aligned}$$

με μια σταθερά C ανεξάρτητη της v. Τότε υπάρχει μια θετική σταθερά c, τέτοια ώστε

$$F(v) \leq c|v|_r \quad \forall v \in H^r, \tag{3.74}$$

όπου $|\cdot|_r$ *η ημινόρμα* $|v|_r := \|v^{(r)}\| := \|v^{(r)}\|_{L^2(0,1)}$.

Απόδειξη. Έστω $v \in H^r$ και $p_{r-1} \in \mathbb{P}_{r-1}$ το πολυώνυμο Taylor της v περί ένα σημείο του διαστήματος $[0,1]$, το σημείο 0, ας πούμε. Σύμφωνα με τις υποθέσεις για το συναρτησιακό F έχουμε τότε

$$F(v) \le F(v - p_{r-1}) + F(p_{r-1}) = F(v - p_{r-1}) \le C\|v - p_{r-1}\|_r.$$

Για να ολοκληρωθεί η απόδειξη, αρκεί, προφανώς, να αποδείξουμε ότι υπάρχει μια σταθερά c, ανεξάρτητη της v, τέτοια ώστε

$$\|v - p_{r-1}\|_r \le c|v|_r. \tag{3.75}$$

Κατ' αρχάς έχουμε

$$v(x) = p_{r-1}(x) + \frac{1}{(r-1)!}\int_0^x (s-x)^{r-1} v^{(r)}(s)\,ds \quad \forall x \in [0,1]. \tag{3.76}$$

Επομένως, σύμφωνα με την ανισότητα των Cauchy–Schwarz,

$$\begin{aligned}|v(x) - p_{r-1}(x)|^2 &\le \frac{1}{[(r-1)!]^2}\int_0^1 s^{2(r-1)}\,ds \int_0^1 |v^{(r)}(s)|^2\,ds\\ &= \frac{1}{[(r-1)!]^2}\,\frac{1}{2r-1}\int_0^1 |v^{(r)}(s)|^2\,ds \quad \forall x \in [0,1],\end{aligned}$$

οπότε

$$\|v - p_{r-1}\|^2 \le \frac{1}{[(r-1)!]^2}\,\frac{1}{2r-1}\,|v|_r^2. \tag{3.77}$$

Γενικότερα, για $i \in \{0, \dots, r-1\}$, έχουμε

$$\|v^{(i)} - p_{r-1}^{(i)}\|^2 \le \frac{1}{[(r-1-i)!]^2}\,\frac{1}{2(r-i)-1}\,|v|_r^2. \tag{3.78}$$

Στην (3.78) μπορούμε να καταλήξουμε με δύο, ουσιαστικά ισοδύναμους, τρόπους: είτε παραγωγίζοντας την (3.76) και προχωρώντας όπως προηγουμένως, είτε αναλογιζόμενοι ότι το $p_{r-1}^{(i)} \in \mathbb{P}_{r-1-i}$ είναι το πολυώνυμο Taylor της $v^{(i)}$, ως προς το σημείο 0, βαθμού το πολύ $r-1-i$, οπότε μπορούμε να πάμε απ' ευθείας στην (3.78), που είναι η (3.77) για κατάλληλες συναρτήσεις και τιμές. Τώρα, από την (3.78) έπεται αμέσως ότι

$$\|v - p_{r-1}\|_r^2 \le \Big[1 + \sum_{i=0}^{r-1}\Big(\frac{1}{[(r-1-i)!]^2}\,\frac{1}{2(r-i)-1}\Big)\Big]\,|v|_r^2. \tag{3.79}$$

Από την (3.79) οδηγούμαστε αμέσως στην (3.75) και η απόδειξη της Πρότασης είναι πλήρης. □

Σχόλιο. (Παραδείγματα συναρτησιακών στο λήμμα των Bramble–Hilbert.) Στην Πρόταση 3.1 το συναρτησιακό F θα μπορούσε, φερ' ειπείν, να είναι:

i. Η απόλυτη τιμή του σφάλματος ενός τύπου αριθμητικής ολοκλήρωσης, αν ο τύπος αυτός ολοκληρώνει ακριβώς πολυώνυμα βαθμού μέχρι και $r-1$. Λεπτομέρειες, στην περίπτωση του τύπου του τραπεζίου, δίνονται στην Άσκηση 3.29.

ii. Μια νόρμα του σφάλματος παρεμβολής με πολυώνυμα βαθμού το πολύ $r-1$.

Σχόλιο. (Το λήμμα των Deny–Lions.) Η ανισότητα (3.75) αναφέρεται συχνά στη βιβλιογραφία ως λήμμα των Deny–Lions.

Πρόταση 3.2 (Επιχείρημα ομοιογένειας.) *Έστω $a, b \in \mathbb{R}, a < b$, $r \geq s \geq 2$ φυσικοί αριθμοί και $x_1, \dots, x_r \in [a, b]$ ανά δύο διαφορετικά μεταξύ τους σημεία. Έστω $v \in H^s(a, b)$ και $p \in \mathbb{P}_{r-1}$ το πολυώνυμο παρεμβολής της v στα σημεία $x_1, \dots, x_r$, $p(x_i) = v(x_i)$, $i = 1, \dots, r$. Τότε ισχύει*

$$\|v - p\|_{L^2(a,b)} \leq C(b-a)^s \|v^{(s)}\|_{L^2(a,b)} \tag{3.80}$$

και

$$\|v' - p'\|_{L^2(a,b)} \leq C(b-a)^{s-1} \|v^{(s)}\|_{L^2(a,b)}, \tag{3.81}$$

με μια σταθερά C, ανεξάρτητη των a, b και v.

Απόδειξη. Θεωρούμε την απεικόνιση $[0, 1] \to [a, b], x \mapsto a+(b-a)x$. Έστω $\hat{x}_1, \dots, \hat{x}_r \in [0, 1]$ τα σημεία που απεικονίζονται στα σημεία $x_1, \dots, x_r$, αντίστοιχα, μέσω της ανωτέρω συνάρτησης. Για κάθε $u : [a, b] \to \mathbb{R}$ συμβολίζουμε με $\hat{u} : [0, 1] \to \mathbb{R}$ τη συνάρτηση $\hat{u}(x) := u\big(a + (b-a)x\big)$. Προφανώς, αν $u \in H^s(a, b)$, τότε και $\hat{u} \in H^s(0, 1)$, καθώς επίσης, αν $p \in \mathbb{P}_{r-1}$, τότε και $\hat{p} \in \mathbb{P}_{r-1}$. Επί πλέον, από τις υποθέσεις μας έπεται αμέσως ότι

$$\hat{p}(\hat{x}_i) = \hat{v}(\hat{x}_i), \quad i = 1, \dots, r,$$

δηλαδή το $\hat{p} \in \mathbb{P}_{r-1}$ είναι το πολυώνυμο παρεμβολής της $\hat{v}$ στα σημεία $\hat{x}_1, \dots, \hat{x}_r$. Ορίζουμε τώρα το συναρτησιακό $F : H^s(0, 1) \to [0, \infty)$, $F(\hat{v}) := \|\hat{v} - \hat{p}\|_{L^2(0,1)}$. Είναι προφανώς σαφές ότι το F ορίζεται για κάθε $\hat{v} \in H^s(0,1)$, αφού κάθε τέτοια συνάρτηση αντιστοιχεί σε κάποια $v \in H^s(a,b)$. Ας δούμε τώρα ότι το F ικανοποιεί τις υποθέσεις της Πρότασης 3.1, με s στη θέση του r. Η πρώτη ιδιότητα έπεται αμέσως από την τριγωνική ανισότητα. Η τρίτη είναι επίσης προφανής, αφού $s \le r$ και, για κάθε $\hat{v} \in \mathbb{P}_{r-1}$, το πολυώνυμο παρεμβολής συμπίπτει με τη $\hat{v}$. Απομένει η δεύτερη ιδιότητα. Χρησιμοποιούμε την παράσταση του πολυωνύμου παρεμβολής σε μορφή Lagrange,

$$\hat{p} = \sum_{i=1}^{r} \hat{v}(\hat{x}_i) L_i,$$

όπου $L_1, \dots, L_r \in \mathbb{P}_{r-1}$ τα πολυώνυμα Lagrange ως προς τα σημεία $\hat{x}_1, \dots, \hat{x}_r$, δηλαδή τέτοια ώστε $L_i(\hat{x}_j) = \delta_{ij}$, $i, j = 1, \dots, r$. Χρησιμοποιώντας την ανισότητα του Sobolev (2.35) (η σταθερά για το διάστημα $[0,1]$ είναι ένα, βλ. και τη (2.10)), έχουμε

$$\|\hat{p}\|_{L^2(0,1)} \le \sum_{i=1}^{r} |\hat{v}(\hat{x}_i)| \, \|L_i\|_{L^2(0,1)} \le \Big(\sum_{i=1}^{r} \|L_i\|_{L^2(0,1)} \Big) \|\hat{v}\|_{H^1(0,1)},$$

συνεπώς

$$F(\hat{v}) \le \|\hat{v}\|_{L^2(0,1)} + \|\hat{p}\|_{L^2(0,1)} \le \Big(1 + \sum_{i=1}^{r} \|L_i\|_{L^2(0,1)} \Big) \|\hat{v}\|_{H^s(0,1)},$$

δηλαδή ικανοποιείται όντως και η δεύτερη ιδιότητα στην Πρόταση 3.1. Σύμφωνα με το λήμμα των Bramble–Hilbert έχουμε επομένως

$$F(\hat{v}) \le c \|\hat{v}^{(s)}\|_{L^2(0,1)}. \tag{3.82}$$

Τώρα, με την αλλαγή μεταβλητής $y = a + (b-a)x$,

$$\begin{aligned} \|v - p\|_{L^2(a,b)}^2 &= \int_a^b \big(v(y) - p(y) \big)^2 \, dy \\ &= (b-a) \int_0^1 \big(v(a + (b-a)x) - p(a + (b-a)x) \big)^2 \, dx \\ &= (b-a) \int_0^1 \big(\hat{v}(x) - \hat{p}(x) \big)^2 \, dx, \end{aligned}$$

συνεπώς

$$\|v - p\|_{L^2(a,b)}^2 = (b-a) \|\hat{v} - \hat{p}\|_{L^2(0,1)}^2. \tag{3.83}$$

Όμως, σύμφωνα με την (3.82), πάλι με την αλλαγή μεταβλητής $y = a + (b - a)x$, έχουμε

$$\|\hat{v} - \hat{p}\|^2_{L^2(0,1)} \leq c^2 \|\hat{v}^{(s)}\|^2_{L^2(0,1)} = c^2 \int_0^1 \big(\hat{v}^{(s)}(x)\big)^2\, dx$$
$$= c^2 \frac{1}{b-a} \int_a^b \big[(b-a)^s v^{(s)}(y)\big]^2\, dy,$$

δηλαδή

$$\|\hat{v} - \hat{p}\|^2_{L^2(0,1)} \leq c^2 \frac{1}{b-a} (b-a)^{2s} \|v^{(s)}\|^2_{L^2(a,b)}. \tag{3.84}$$

Από τις (3.83) και (3.84) έπεται αμέσως η (3.80). Η (3.81) αποδεικνύεται εντελώς ανάλογα. □

Σχόλιο. (Η σταθερά στις εκτιμήσεις (3.80) και (3.81).) Όπως προκύπτει από την απόδειξη της Πρότασης 3.2, η σταθερά στις εκτιμήσεις (3.80) και (3.81) εξαρτάται από το r και τα σημεία $\hat{x}_1, \dots, \hat{x}_r$.

Παρατήρηση 3.7 (Παρεμβολή τύπου Hermite.) Στην Πρόταση 3.2 ασχοληθήκαμε με παρεμβολή τύπου Lagrange. Οι εκτιμήσεις (3.80) και (3.81) ισχύουν και για την περίπτωση παρεμβολής τύπου Hermite, όπως εύκολα διαπιστώνει κανείς. Θα αναφερθούμε ιδιαίτερα στις ακόλουθες δύο περιπτώσεις:

1η *Περίπτωση*: Έστω $r \geq 4$ και $a = x_1 < \dots < x_{r-2} = b$. Για $2 \leq s \leq r$ και $v \in H^s(a,b)$, θεωρούμε το πολυώνυμο παρεμβολής $p \in \mathbb{P}_{r-1}$ τέτοιο ώστε

$$p(x_i) = v(x_i),\ i = 1, \dots, r-2, \quad \text{και} \quad p'(x_i) = v'(x_i),\ i = 1, r-2. \tag{3.85}$$

Μια κατάλληλη τροποποίηση της απόδειξης της Πρότασης 3.2 οδηγεί στις εκτιμήσεις (3.80) και (3.81) και στην προκειμένη περίπτωση. Η σημαντικότερη αλλαγή στην απόδειξη συνίσταται στην κατάλληλη επιλογή των πολυωνύμων $L_0, \dots, L_{r-1} \in \mathbb{P}_{r-1}$ κατά τρόπον ώστε να κάθε ένα από αυτά να ικανοποιεί την αντίστοιχη της (3.85) στο διάστημα $[0,1]$ με μία μόνο τιμή στο δεξιό μέλος ίση με τη μονάδα, διαφορετική κάθε φορά, φυσικά, και όλες τις άλλες ίσες με το μηδέν. Συγκεκριμένα, για $i = 0, \dots, r-1$, απαιτούμε να ικανοποιούνται οι συνθήκες

$$L_i(x_j) = \delta_{ij},\ j = 1, \dots, r-2, \quad \text{και} \quad L_i'(x_i) = \delta_{i0},\ L_i'(x_{r-2}) = \delta_{i,r-1}. \tag{3.86}$$

Τότε, με συμβολισμό αντίστοιχο εκείνου στην απόδειξη της Πρότασης 3.2, το πολυώνυμο παρεμβολής $\hat{p}$ της συνάρτησης $\hat{v} \in H^s(0,1)$ γράφεται στη μορφή

$$\hat{p} = \sum_{i=1}^{r-2} \hat{v}(\hat{x}_i)L_i + \hat{v}'(\hat{x}_1)L_0 + \hat{v}'(\hat{x}_{r-2})L_{r-1}.$$

Προφανώς ισχύει

$$\|\hat{p}\|_{L^2(0,1)} \le \sum_{i=1}^{r} |\hat{v}(\hat{x}_i)|\, \|L_i\|_{L^2(0,1)} + |\hat{v}'(\hat{x}_1)|\, \|L_0\|_{L^2(0,1)} + |\hat{v}'(\hat{x}_{r-2})|\, \|L_{r-1}\|_{L^2(0,1)},$$

οπότε, σύμφωνα με την ανισότητα του Sobolev,

$$\|\hat{p}\|_{L^2(0,1)} \le \Big(\sum_{i=0}^{r-1} \|L_i\|_{L^2(0,1)}\Big)\|\hat{v}\|_{H^2(0,1)}.$$

Η απόδειξη προχωράει μετά ακριβώς όπως στην περίπτωση της Πρότασης 3.2.

2η *Περίπτωση*: Έστω m το ακέραιο μέρος του $r/2$, $m = [r/2]$. Για $m+1 \le s \le r$ και $v \in H^s(a,b)$, θεωρούμε το πολυώνυμο παρεμβολής $p \in \mathbb{P}_{r-1}$ τέτοιο ώστε

$$p^{(i)}(a) = v^{(i)}(a), \quad p^{(i)}(b) = v^{(i)}(b), \quad i = 0, \dots, m. \tag{3.87}$$

Στην περίπτωση που το r είναι περιττός αριθμός, χρειαζόμαστε μία επί πλέον συνθήκη, παραδείγματος χάριν την

$$p\Big(\frac{a+b}{2}\Big) = v\Big(\frac{a+b}{2}\Big). \tag{3.88}$$

Μια κατάλληλη τροποποίηση της απόδειξης της Πρότασης 3.2 οδηγεί στις εκτιμήσεις (3.80) και (3.81) και στην προκειμένη περίπτωση. Η σημαντικότερη αλλαγή στην απόδειξη συνίσταται στην κατάλληλη επιλογή των πολυωνύμων L_i, τα οποία επιλέγονται τώρα ως εξής: Κάθε ένα ικανοποιεί τις (3.87) και (3.88) με μία μόνο τιμή στο δεξιό μέλος ίση με τη μονάδα, διαφορετική κάθε φορά, φυσικά, και όλες τις άλλες ίσες με το μηδέν. □

Μετά από την προηγηθείσα προεργασία σε αυτή την ενότητα, είμαστε τώρα σε θέση να δώσουμε παραδείγματα χώρων με την προσεγγιστική ιδιότητα (3.21).

Θεώρημα 3.5 (Χώροι με την προσεγγιστική ιδιότητα (3.21).) *Έστω* $r \geq 2$ *ένας φυσικός αριθμός,* $m = 0$ *ή, στην περίπτωση* $r \geq 4$, $m = 1$, $J \in \mathbb{N}$, $0 = x_0 < x_1 < \cdots < x_{J+1} = 1$ *ένας διαμερισμός του διαστήματος* $[0,1]$ *και*

$$S_h^r := \{v \in C^m[0,1] : v(0) = v(1) = 0 \ \textit{και} \ v|_{[x_i,x_{i+1}]} \in \mathbb{P}_{r-1}, \ i = 0, \dots, J\}.$$

Τότε η οικογένεια των χώρων S_h^r *έχει την προσεγγιστική ιδιότητα* (3.21)*, με*

$$h := \max_{0 \leq i \leq J} (x_{i+1} - x_i).$$

Απόδειξη. Ας εξετάσουμε πρώτα την περίπτωση $m = 0$, δηλαδή την περίπτωση που τα στοιχεία του S_h^r είναι απλώς συνεχή στους εσωτερικούς κόμβους $x_1, \dots, x_J$ του διαμερισμού. Έστω $0 = \hat{\tau}_1 < \hat{\tau}_2 < \cdots < \hat{\tau}_r = 1$. Με $h_i := x_{i+1} - x_i$, $i = 0, \dots, J$, θέτουμε τότε

$$x_{ij} := x_i + h_i \hat{\tau}_j, \quad j = 1, \dots, r, \ i = 0, \dots, J.$$

Έστω $v \in H^2(0,1) \cap H_0^1(0,1)$. Για $i \in \{0, \dots, J\}$, θεωρούμε το πολυώνυμο $p_i : [x_i, x_{i+1}] \to \mathbb{R}$, βαθμού το πολύ $r-1$, που παρεμβάλλεται στη v στα σημεία $x_{i1}, \dots, x_{ir} \in [x_i, x_{i+1}]$, $p_i(x_{ij}) = v(x_{ij})$, $j = 1, \dots, r$. Θεωρούμε τώρα τη συνάρτηση

$$\chi : [0,1] \to \mathbb{R}, \ \ \chi(x) := p_i(x), \ \ \text{για} \ \ x \in [x_i, x_{i+1}], \ \ i = 0, \dots, J.$$

Σημειώστε ότι η χ είναι καλώς ορισμένη, αφού $x_{i-1,r} = x_{i0}$, οπότε οι συνθήκες παρεμβολής δίνουν $p_{i-1}(x_i) = p_i(x_i)$. Είναι τώρα προφανές ότι η χ είναι συνεχής, τμηματικά πολυώνυμο βαθμού το πολύ $r-1$, και μηδενίζεται στα σημεία 0 και 1, αφού $v(0) = v(1) = 0$. Επομένως, η χ είναι όντως στοιχείο του S_h^r. Τώρα, σύμφωνα με την (3.80), για $s = 2$ ή $s = r$, η δεύτερη περίπτωση, φυσικά, υπό την προϋπόθεση ότι $v \in H^r(0,1) \cap H_0^1(0,1)$,

$$\|v - p_i\|^2_{L^2(x_i,x_{i+1})} \leq C^2 h_i^{2s} \|v^{(s)}\|^2_{L^2(x_i,x_{i+1})},$$

οπότε

$$\|v - p_i\|^2_{L^2(x_i,x_{i+1})} \leq C^2 h^{2s} \|v^{(s)}\|^2_{L^2(x_i,x_{i+1})}.$$

Αθροίζοντας αυτές τις σχέσεις, λαμβάνουμε

$$\|v - \chi\|^2_{L^2(0,1)} \leq C^2 h^{2s} \|v^{(s)}\|^2_{L^2(0,1)}$$

ή

$$\|v - \chi\|_{L^2(0,1)} \leq C h^s \|v^{(s)}\|_{L^2(0,1)}. \tag{3.89}$$

Επίσης, σύμφωνα με την (3.81), για $s = 2$ ή $s = r$,

$$\|v' - p_i'\|^2_{L^2(x_i,x_{i+1})} \le C^2 h_i^{2(s-1)} \|v^{(s)}\|^2_{L^2(x_i,x_{i+1})},$$

και, όπως προηγουμένως,

$$\|v' - \chi'\|_{L^2(0,1)} \le C h^{s-1} \|v^{(s)}\|_{L^2(0,1)}. \tag{3.90}$$

Από τις (3.89) και (3.90) έπεται αμέσως ότι

$$\|v - \chi\|_{L^2(0,1)} + h\|v' - \chi'\|_{L^2(0,1)} \le 2Ch^s \|v^{(s)}\|_{L^2(0,1)}, \quad s = 2, r\,, \tag{3.91}$$

δηλαδή ικανοποιείται πράγματι η (3.21).

Προχωρούμε τώρα στην περίπτωση $m = 1$, οπότε τα στοιχεία του S_h^r είναι συνεχώς παραγωγίσιμες συναρτήσεις στους εσωτερικούς κόμβους του διαμερισμού. Έστω τώρα $r \ge 4$ και $0 = \hat{\tau}_1 < \cdots < \hat{\tau}_{r-2} = 1$. Με $h_i := x_{i+1} - x_i$, $i = 0, \dots, J$, θέτουμε

$$x_{ij} := x_i + h_i \hat{\tau}_j, \quad j = 1, \dots, r-2, \ i = 0, \dots, J.$$

Έστω $v \in H^2(0,1) \cap H_0^1(0,1)$. Για $i \in \{0, \dots, J\}$, θεωρούμε το πολυώνυμο $p_i : [x_i, x_{i+1}] \to \mathbb{R}$, βαθμού το πολύ $r-1$, που παρεμβάλλεται στη v ως εξής

$$p_i(x_{ij}) = v(x_{ij}), \ j = 1, \dots, r-2, \quad \text{και} \quad p'(x_{ij}) = v'(x_{ij}), \ j = 1, r-2,$$

βλ. την (3.85). Θεωρούμε τώρα τη συνάρτηση

$$\chi : [0,1] \to \mathbb{R}, \ \ \chi(x) := p_i(x), \quad \text{για} \ \ x \in [x_i, x_{i+1}], \ \ i = 0, \dots, J.$$

Είναι τώρα προφανές ότι η χ είναι συνεχώς παραγωγίσιμη, τμηματικά πολυώνυμο βαθμού το πολύ $r-1$, και μηδενίζεται στα σημεία 0 και 1, επομένως, η χ είναι στοιχείο του S_h^r. Χρησιμοποιώντας τώρα τις (3.80) και (3.81), βλ. την Παρατήρηση 3.7, $1^{\underline{\eta}}$ Περίπτωση, και προχωρώντας ακριβώς όπως και στην περίπτωση $m = 0$, που προηγήθηκε, οδηγούμαστε στο επιθυμητό αποτέλεσμα. □

Σχόλιο. (Πιο ομαλές splines.) Τα στοιχεία του S_h^r λέγονται (πολυωνυμικές) splines. Στο Θεώρημα 3.5 υποθέσαμε ότι $m = 0$ ή $m = 1$. Ο περιορισμός αυτός χρησιμοποιήθηκε στην απόδειξη σε δύο σημεία: αφ' ενός μας επέτρεψε να ορίσουμε την παρεμβάλλουσα συνάρτηση τοπικά, δηλαδή σε κάθε υποδιάστημα χρησιμοποιήσαμε μόνο τιμές της παρεμβαλλόμενης συνάρτησης και παραγώγων της στο ίδιο υποδιάστημα, και αφ' ετέρου αυτή η παρεμβάλλουσα ορίζεται και για συναρτήσεις στον

$H^2(0,1)$, αφού ορίστηκε μόνο συναρτήσει τιμών της παρεμβαλλόμενης συνάρτησης και της πρώτης παραγώγου της. Πάντως η συνθήκη $m \leq r-2$ αρκεί για να έχει ο αντίστοιχος χώρος την προσεγγιστική ιδιότητα (3.21). Σημειώνουμε κατ' αρχάς ότι η συνθήκη αυτή είναι η γενικότερη δυνατή, αφού για $m \geq r-1$ έχουμε $S_h^r = \mathbb{P}_{r-1}$, δηλαδή οι splines εκφυλίζονται σε πολυώνυμα. Στην περίπτωση $m = r-2$ αναφερόμαστε συνήθως με τους όρους *ομαλές splines* ή *splines με απλούς κόμβους*. Προφανώς, για ένα συγκεκριμένο r, οι ομαλές splines περιέχονται και στους χώρους S_h^r για κάθε $m \in \{0,\dots,r-3\}$. Αρκεί συνεπώς να αποδειχθεί ότι οι ομαλές splines ικανοποιούν την προσεγγιστική ιδιότητα (3.21). Αυτό γίνεται συνήθως στη βιβλιογραφία χρησιμοποιώντας μια οιονεί παρεμβάλλουσα spline, που ορίζεται συναρτήσει καταλλήλων συναρτήσεων βάσης, των λεγόμενων $B-$splines, και έχει νόημα και για στοιχεία του $H^2(0,1)$. Δεν θα δώσουμε εδώ τις λεπτομέρειες, αφού κάτι τέτοιο θα απαιτούσε λεπτομερή παρουσίαση των $B-$splines και των ιδιοτήτων τους· αντ' αυτού παραπέμπουμε τον ενδιαφερόμενο αναγνώστη στη βιβλιογραφία, φερ' ειπείν στην §6.4, και ιδιαίτερα το Corollary 6.26, στο βιβλίο [40] του L. Schumaker και στο Θεώρημα 7.4 στο βιβλίο των DeVore και Lorentz [19].

Στη συνέχεια θα αρκεσθούμε σε δύο αποτελέσματα: Το πρώτο αφορά την προσεγγιστική ιδιότητα (3.21) *μόνο* για $s = r$, δηλαδή όχι για $s = 2$, στην περίπτωση που το m είναι το πολύ όσο το ακέραιο μέρος του $r/2$, και το δεύτερο αφορά την περίπτωση των κυβικών splines, δηλαδή των ομαλών splines για $r = 4$.

Πρόταση 3.3 (Πιο ομαλές splines με την προσεγγιστική ιδιότητα (3.21) για $s = r$.) *Έστω $r \geq 4$ ένας φυσικός αριθμός και $m \in \{0,\dots,[r/2]\}$, με $[r/2]$ το ακέραιο μέρος του $r/2$. Έστω $J \in \mathbb{N}$, $0 = x_0 < x_1 < \cdots < x_{J+1} = 1$ ένας διαμερισμός του διαστήματος $[0,1]$ και*

$$S_h^r := \{v \in C^m[0,1] : v(0) = v(1) = 0 \ \textit{και} \ v|_{[x_i,x_{i+1}]} \in \mathbb{P}_{r-1},\ i = 0,\dots,J\}.$$

Έστω

$$h := \max_{0\leq i\leq J}(x_{i+1} - x_i).$$

Τότε υπάρχει μια θετική σταθερά c, ανεξάρτητη του h, τέτοια ώστε, για κάθε $v \in H^r(0,1) \cap H_0^1(0,1)$, υπάρχει $\chi \in S_h^r$ με την προσεγγιστική ιδιότητα

$$\|v-\chi\|_{L^2(0,1)} + h\|v'-\chi'\|_{L^2(0,1)} \leq ch^r\|v^{(r)}\|_{L^2(0,1)}. \tag{3.92}$$

Απόδειξη. Έστω $v \in H^2(0,1) \cap H_0^1(0,1)$. Για $i \in \{0,\dots,J\}$, θεωρούμε το πολυώνυμο $p_i : [x_i,x_{i+1}] \to \mathbb{R}$, βαθμού το πολύ $r-1$, που παρεμβάλλεται στη v ως εξής

$$p_i^{(j)}(x_\ell) = v^{(j)}(x_\ell),\ j = 1,\dots,[\frac{r}{2}],\ \ \ell = i, i+1,$$

και επί πλέον, στην περίπτωση που το r είναι περιττός αριθμός,

$$p_i\big(\frac{1}{2}(x_i + x_{i+1})\big) = v\big(\frac{1}{2}(x_i + x_{i+1})\big),$$

βλ. τις (3.87) και (3.88). (Σημειώστε ότι το p_i *δεν* ορίζεται για $v \in H^2(0,1) \cap H_0^1(0,1)$.) Θεωρούμε τώρα τη συνάρτηση

$$\chi : [0,1] \to \mathbb{R}, \ \ \chi(x) := p_i(x), \quad \text{για} \ \ x \in [x_i, x_{i+1}], \ \ i = 0, \dots, J.$$

Είναι προφανές ότι η χ είναι $[r/2]$ φορές συνεχώς παραγωγίσιμη, τμηματικά πολυώνυμο βαθμού το πολύ $r-1$, και μηδενίζεται στα σημεία 0 και 1, επομένως, η χ είναι στοιχείο των S_h^r. Χρησιμοποιώντας τώρα τις (3.80) και (3.81), για $s = r$, βλ. την Παρατήρηση 3.7, 2η Περίπτωση, και προχωρώντας ακριβώς όπως και στην απόδειξη του Θεωρήματος 3.5, οδηγούμαστε στην (3.92). □

Τέλος, θα αποδείξουμε ότι μια οικογένεια ομαλών splines, που χρησιμοποιούνται πολύ στην πράξη, οι κυβικές splines, έχει την προσεγγιστική ιδιότητα (3.21).

Πρόταση 3.4 (Κυβικές splines.) *Έστω $J \in \mathbb{N}$, $0 = x_0 < x_1 < \dots < x_{J+1} = 1$ ένας διαμερισμός του διαστήματος $[0,1]$ και S_h^4 ο χώρος των ομαλών κυβικών splines ως προς αυτόν τον διαμερισμό,*

$$S_h^4 := \{v \in C^2[0,1] : v(0) = v(1) = 0 \ \textit{και} \ v|_{[x_i,x_{i+1}]} \in \mathbb{P}_3, \ i = 0, \dots, J\}.$$

Έστω

$$h := \max_{0 \le i \le J} (x_{i+1} - x_i).$$

Τότε υπάρχει μια θετική σταθερά c, ανεξάρτητη του h, τέτοια ώστε, για κάθε $v \in H^s(0,1) \cap H_0^1(0,1)$, με $s = 2$ ή $s = 4$, υπάρχει $\chi \in S_h^4$ με την προσεγγιστική ιδιότητα

$$\|v - \chi\|_{L^2(0,1)} + h\|v' - \chi'\|_{L^2(0,1)} \le ch^s \|v^{(s)}\|_{L^2(0,1)}, \tag{3.93}$$

δηλαδή αυτή η οικογένεια χώρων έχει την προσεγγιστική ιδιότητα (3.21).

Απόδειξη. Έστω $v \in H^2(0,1) \cap H_0^1(0,1)$. Θεωρούμε την κυβική spline χ που παρεμβάλλεται στη v ως ακολούθως

$$\chi(x_i) = v(x_i), \ i = 0, \dots, J+1, \quad \text{και} \quad \chi'(x_i) = v'(x_i), \ i = 0, J+1. \tag{3.94}$$

Η ύπαρξη και η μοναδικότητα της χ μπορούν να αποδειχθούν με κατάλληλη τροποποίηση της απόδειξης του Θεωρήματος 4.7 του [2], το οποίο αναφέρεται στο ίδιο πρόβλημα παρεμβολής στην περίπτωση ομοιόμορφου διαμερισμού. Συμβολίζουμε τώρα με e το σφάλμα της παρεμβολής, $e := v - \chi$. Αρκεί τώρα να αποδείξουμε ότι, για $s = 2$ και $s = 4$,

$$\|e\|_{L^2(0,1)} \le \frac{1}{2} h^s \|v^{(s)}\|_{L^2(0,1)} \tag{3.95}$$

και

$$\|e'\|_{L^2(0,1)} \le h^{s-1} \|v^{(s)}\|_{L^2(0,1)}. \tag{3.96}$$

Αφού η συνάρτηση e μηδενίζεται στους κόμβους, χρησιμοποιώντας την ανισότητα των Poincaré–Friedrichs, βλ. την Άσκηση 3.2, έχουμε

$$\|e\|_{L^2(x_i,x_{i+1})} \le \frac{1}{2} h_i \|e'\|_{L^2(x_i,x_{i+1})}, \quad i = 0, \dots, J,$$

με $h_i := x_{i+1} - x_i$, και οδηγούμαστε εύκολα στην εκτίμηση

$$\|e\|_{L^2(0,1)} \le \frac{1}{2} h \|e'\|_{L^2(0,1)}. \tag{3.97}$$

Η (3.95) έπεται αμέσως από τις (3.96) και (3.97). Απομένει συνεπώς να αποδείξουμε την (3.96). Τώρα, σύμφωνα με το θεώρημα του Rolle, υπάρχει $\xi_i \in (x_i, x_{i+1})$, τέτοιο ώστε $e'(\xi_i) = 0$, $i = 0, \dots, J$. Με $\xi_{-1} := 0$ και $\xi_{J+1} := 1$ έχουμε επί πλέον, προφανώς, $e'(\xi_{-1}) = e'(\xi_{J+1}) = 0$. Επομένως, πάλι με την ανισότητα των Poincaré–Friedrichs,

$$\|e'\|_{L^2(\xi_i,\xi_{i+1})} \le \frac{1}{2} (\xi_{i+1} - \xi_i) \|e''\|_{L^2(\xi_i,\xi_{i+1})}, \quad i = -1, \dots, J,$$

οπότε

$$\|e'\|_{L^2(\xi_i,\xi_{i+1})} \le h \|e''\|_{L^2(\xi_i,\xi_{i+1})}, \quad i = -1, \dots, J. \tag{3.98}$$

Υψώνοντας αυτή τη σχέση στο τετράγωνο, αθροίζοντας και εξάγοντας τετραγωνικές ρίζες, λαμβάνουμε

$$\|e'\|_{L^2(0,1)} \le h \|e''\|_{L^2(0,1)}. \tag{3.99}$$

Τώρα, αφού, προφανώς, $\chi \in H^3(0,1)$, χρησιμοποιώντας το γεγονός ότι $e'(0) = e'(1) = 0$, έχουμε

$$\int_0^1 e''(x)\chi''(x)\,dx = -\int_0^1 e'(x)\chi'''(x)\,dx = -\sum_{i=0}^{J} \int_{x_i}^{x_{i+1}} e'(x)\chi'''(x)\,dx = 0,$$

αφού η χ''' είναι σταθερή στο (x_i, x_{i+1}) και

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

Επομένως, έχουμε

$$\|e''\|^2_{L^2(0,1)} = \int_0^1 e''(x)v''(x)\,dx \le \|e''\|_{L^2(0,1)}\,\|v''\|_{L^2(0,1)},$$

οπότε

$$\|e''\|_{L^2(0,1)} \le \|v''\|_{L^2(0,1)}. \tag{3.100}$$

Συνδυάζοντας την (3.100) με την (3.99) διαπιστώνουμε αμέσως ότι ισχύει η (3.96) για $s = 2$.

Για $s = 4$, σημειώνουμε ακόμα ότι

$$\|e''\|^2_{L^2(0,1)} = \int_0^1 e''(x)v''(x)\,dx = -\int_0^1 e'(x)v'''(x)\,dx = \int_0^1 e(x)v^{(4)}(x)\,dx,$$

οπότε

$$\|e''\|^2_{L^2(0,1)} \le \|e\|_{L^2(0,1)}\,\|v^{(4)}\|_{L^2(0,1)}.$$

Συνδυάζοντας αυτή τη σχέση με τις (3.99) και (3.97) λαμβάνουμε

$$\|e''\|_{L^2(0,1)} \le \frac{1}{2}h^2\|v^{(4)}\|_{L^2(0,1)}. \tag{3.101}$$

Οι (3.99) και (3.101) μας οδηγούν στο συμπέρασμα ότι η (3.96) ισχύει και για $s = 4$.
$\square$

Ασκήσεις

3.1 Αποδείξτε την ανισότητα των Poincaré–Friedrichs στην ακόλουθη μορφή

$$\|v\| \le \frac{b-a}{\sqrt{2}}\|v'\| \quad \forall v \in H_0^1(a,b).$$

Συγκρίνετε με την ανισότητα (2.7) για την περίπτωση του διαστήματος $(0,1)$.
[*Υπόδειξη*: Λόγω του ότι $v(a) = 0$, έχουμε

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

δηλαδή

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

Ολοκληρώστε στο διάστημα $[a,b]$.]

3.2 Αποδείξτε την ανισότητα των Poincaré–Friedrichs στην ακόλουθη μορφή

$$\|v\| \le \frac{b-a}{2}\|v'\| \quad \forall v \in H_0^1(a,b).$$

[*Υπόδειξη*: Χρησιμοποιήστε τις σχέσεις

$$\begin{aligned} v(x) &= \int_a^x v'(s)\,ds \qquad \text{για } x \in [a, \frac{a+b}{2}],\\ v(x) &= -\int_x^b v'(s)\,ds \quad \text{για } x \in [\frac{a+b}{2}, b].\,] \end{aligned}$$

3.3 Βελτιώστε τη σταθερά στην ανισότητα των Poincaré–Friedrichs στην Άσκηση 3.1, εκμεταλλευόμενοι και το γεγονός ότι $v(b)=0$, αποδεικνύοντας ότι

$$\|v\| \le \frac{b-a}{2\sqrt{2}}\|v'\| \quad \forall v \in H_0^1(a,b).$$

[*Υπόδειξη*: Παρατηρήστε ότι

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

και αντίστοιχα για το διάστημα $[\frac{a+b}{2}, b]$, και συνδυάστε αυτό το γεγονός με την υπόδειξη στην Άσκηση 3.2.]

3.4 Χρησιμοποιήστε τις ιδέες που παρατέθηκαν στις προηγούμενες Ασκήσεις για να αποδείξετε την ανισότητα του Sobolev στην ακόλουθη μορφή

$$\max_{a\le x\le b} |v(x)| \le \sqrt{\frac{b-a}{2}}\|v'\| \quad \forall v \in H_0^1(a,b).$$

Συγκρίνετε με την ανισότητα (2.10) στην περίπτωση του διαστήματος $(0,1)$.

3.5 Πολλαπλασιάστε τη Δ.Ε. του προβλήματος (3.18) επί u, ολοκληρώστε στο $[0,1]$ και ολοκληρώστε κατά μέρη. Συμπεράνετε ότι το πρόβλημα, στην περίπτωση $c(x) \ge 0$, για κάθε $x \in [0,1]$, και $c \ne 0$, έχει μοναδική λύση. Υποθέτοντας τώρα επί πλέον ότι η c είναι γνήσια θετική στο $[0,1]$, αποδείξτε μια ανισότητα της μορφής (3.12).

3.6 Αποδείξτε ότι το πρόβλημα (3.19) έχει το πολύ μία ομαλή λύση.

3.7 Θεωρούμε το ακόλουθο πρόβλημα δύο σημείων

$$(\star) \qquad \begin{cases} -u'' + bu' + cu = f \quad \text{στο } [0,1],\\ u(0) = u(1) = 0, \end{cases}$$

όπου $b, c, f \in C[0,1]$, και $c(x) \geq 0$, για κάθε $x \in [0,1]$. Αποδείξτε ότι πολλαπλασιάζοντας τη Δ.Ε. στο $(\star)$ με $\mathrm{e}^{-\int_0^x b(s)ds}$ μπορούμε να αναγάγουμε το πρόβλημα αυτό σε ένα πρόβλημα της μορφής (3.1) με $b = 0$.

3.8 Θεωρούμε το πρόβλημα δύο σημείων

$$(\star) \qquad \begin{cases} -u'' + bu' + cu = 0 \quad \text{στο } [0,1], \\ u(0) = \gamma, \\ u(1) = \delta, \end{cases}$$

με $b, c \in C[0,1]$ και $c(x) > 0$, για κάθε $x \in [0,1]$. Αποδείξτε την εξής *αρχή μεγίστου*

$$\max_{0 \leq x \leq 1} |u(x)| \leq \max(|\gamma|, |\delta|).$$

[*Υπόδειξη*: Αν η u δεν είναι σταθερά, αποδείξτε ότι δεν μπορεί να έχει θετικό τοπικό μέγιστο ή αρνητικό τοπικό ελάχιστο στο $(0,1)$.]

3.9 Αποδείξτε για το πρόβλημα (3.20) μια ανισότητα της μορφής (3.12) και οδηγηθείτε στο συμπέρασμα ότι αυτό έχει το πολύ μία ομαλή λύση.
[*Υπόδειξη*: Αποδείξτε πρώτα ότι για $v \in H^1(0,1)$ ισχύει

$$\|v\|^2 \leq C\{\min_{0 \leq x \leq 1} |v(x)|^2 + \|v'\|^2\}$$

με μια σταθερά C ανεξάρτητη της v.]

3.10 Αποδείξτε για το πρόβλημα (3.18) ότι ισχύει

$$\max_{0 \leq x \leq 1} |u(x)| \leq C \max_{0 \leq x \leq 1} |f(x)|$$

με μια σταθερά C εξαρτώμενη μόνο από τα a και c.
[*Υπόδειξη*: Έστω $v \in H^1(0,1)$. Αποδείξτε τότε ότι

$$\max_{0 \leq x \leq 1} |v(x)| \leq |v(y)| + \|v'\| \quad \forall y \in [0,1].$$

Επί πλέον έχουμε προφανώς

$$(cv, v) \geq \min_{0 \leq x \leq 1} |v(x)|^2 \int_0^1 c(x)dx.]$$

3.11 Αποδείξτε για το πρόβλημα (3.18) μια ανισότητα της μορφής (3.12), γενικεύοντας έτσι το αποτέλεσμα της Άσκησης 3.5 και για την περίπτωση στην οποία το c παίρνει και την τιμή μηδέν, $\min_{0 \leq x \leq 1} c(x) = 0$.

3.12 Πώς μπορεί το πρόβλημα

$$\begin{cases} -u'' + cu = f \qquad \text{στο } [0,1], \\ u'(0) = \gamma, \\ u'(1) = \delta, \end{cases}$$

να αναχθεί σε ένα αντίστοιχο πρόβλημα με ομογενείς συνοριακές συνθήκες Neumann;
[*Υπόδειξη*: Τι ιδιότητες έχει η συνάρτηση $v(x) := \gamma x + \frac{\delta-\gamma}{2}x^2$, $x \in [0,1]$;]

3.13 Έστω $\rho > 0$ και $x^\star \in (0,1)$. Θεωρούμε το εξής πρόβλημα δύο σημείων με *διεπιφάνεια*: Ζητείται μια συνεχής συνάρτηση $u : [0,1] \to \mathbb{R}$, ομαλή στα υποδιαστήματα $[0,x^\star]$ και $[x^\star,1]$, και τέτοια ώστε

$$\begin{cases} -u'' + cu = f & \text{στο } [0,x^\star) \cup (x^\star,1], \\ u'(x^\star-) = \rho u'(x^\star+), \\ u(0) = u(1) = 0, \end{cases}$$

όπου c, f συνεχείς συναρτήσεις στα $[0,x^\star)$, $(x^\star,1]$, οι οποίες μπορούν να επεκταθούν συνεχώς στα $[0,x^\star]$ και $[x^\star,1]$, και $c(x) \ge 0$, για κάθε $x \in [0,1]$. Αποδείξτε ότι το πρόβλημα αυτό έχει το πολύ μία ομαλή λύση.
[*Υπόδειξη*. Χρησιμοποιήστε το εσωτερικό γινόμενο $(\cdot,\cdot)$,

$$(v,w) := \int_0^{x^\star} v(x)w(x)dx + \rho\int_{x^\star}^1 v(x)w(x)dx.\,]$$

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

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

έχει το πολύ μία ομαλή λύση.

3.15 Έστω $f : [0,1] \times \mathbb{R} \to \mathbb{R}$ μια συνεχώς παραγωγίσιμη συνάρτηση τέτοια ώστε $-\gamma \le \partial_2 f(x,s) \le 0$, για κάθε $x \in [0,1], s \in \mathbb{R}$, με μια θετική σταθερά γ. Αποδείξτε ότι το *μη γραμμικό* πρόβλημα δύο σημείων

$$\begin{cases} -(au')' = f(x,u) & \text{στο } [0,1], \\ u(0) = u(1) = 0, \end{cases}$$

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

3.16 Εισάγουμε στον $L^2(0,1)$ την "αρνητική" (δηλαδή αρνητικής τάξης) νόρμα $\|\cdot\|_{-1}$ ως

$$\|f\|_{-1} := \Big(\int_0^1 \big[\int_0^x f(s)ds\big]^2 dx\Big)^{1/2}.$$

Για $f \in L^2(0,1)$ και $v \in H_0^1(0,1)$, αποδείξτε την εξής ανισότητα "τύπου Cauchy–Schwarz"

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

Αποδείξτε για τη λύση u του προβλήματος (3.1) μια εκτίμηση της μορφής

$$\|u\| + \|u'\| \leq C\|f\|_{-1}$$

με μια σταθερά C εξαρτώμενη μόνο από τις συναρτήσεις a και $c - b'/2$.

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

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

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

3.18 Θεωρούμε το πρόβλημα (3.18). Θεωρούμε μια οικογένεια $(S_h^r)_{0<h\leq 1}$ γραμμικών χώρων πεπερασμένης διάστασης συνεχών και τμηματικά συνεχώς παραγωγίσιμων συναρτήσεων, με προσεγγιστική ιδιότητα ανάλογη της (3.21), αλλά για $v \in H^s(a,b)$ (χωρίς να απαιτείται δηλαδή $v(0) = v(1) = 0$). Ορίστε την προσεγγιστική λύση $u_h \in S_h^r$ για αυτό το πρόβλημα και αποδείξτε εκτιμήσεις αντίστοιχες των (3.29) και (3.30) γι' αυτή την περίπτωση.

3.19 Τι χώρους πεπερασμένων στοιχείων θα επιλέγατε για το πρόβλημα δύο σημείων

$$\begin{cases} -u'' + cu = f & \text{στο } [0,1], \\ u(0) = u'(1) = 0; \end{cases}$$

3.20 Έστω ότι $b = 0$ και ότι η συνάρτηση c λαμβάνει μόνο μη αρνητικές τιμές. Αποδείξτε ότι η λύση u του προβλήματος (3.1) είναι το μόνο στοιχείο του χώρου $H_0^1(0,1)$ στο οποίο το συναρτησιακό J, $J(v) := a(v,v) - 2(f,v)$, λαμβάνει την ελάχιστη τιμή του στον $H_0^1(0,1)$.

3.21 Θεωρούμε το πρόβλημα δύο σημείων με διεπιφάνεια της Άσκησης 3.13. Δώστε μια μέθοδο πεπερασμένων στοιχείων για αυτό το πρόβλημα, αντίστοιχη της (3.26), και αποδείξτε εκτιμήσεις αντίστοιχες των (3.29) και (3.30).

3.22 Θεωρούμε το πρόβλημα δύο σημείων

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

Έστω S_h^r ένας υπόχωρος του $H_0^1(0,1)$ αποτελούμενος από συνεχείς συναρτήσεις οι οποίες σε καθένα των υποδιαστημάτων $[x_i, x_{i+1}]$, $i = 0, \ldots, J$, ενός διαμερισμού $0 = x_0 < x_1 < \cdots < x_{J+1} = 1$ του $[0,1]$ είναι πολυώνυμα βαθμού το πολύ r. Αποδείξτε ότι

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

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

$$\chi(x) = \begin{cases} \dfrac{x}{x_i}, & 1 \le x \le x_i, \\ \dfrac{x-1}{x_i - 1}, & x_i < x \le 1, \end{cases}$$

είναι στοιχείο του S_h^r.]

3.23 Θεωρούμε το πρόβλημα (3.1), με $a = 1$ και $b = 0$, και την προσεγγιστική λύση πεπερασμένων στοιχείων $u_h \in S_h^2$, με S_h^2 όπως στο Παράδειγμα 3.1. Αποδείξτε ότι

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

με τις συναρτήσεις φ_i όπως στο Παράδειγμα 3.2. Προσεγγίζοντας τα ολοκληρώματα με τον σύνθετο τύπο του τραπεζίου, δηλαδή το $\int_{x_{i-1}}^{x_{i+1}} g(x)dx$ με $h[g(x_{i-1})/2 + g(x_i) + g(x_{i+1})/2]$, οδηγηθείτε στο σχήμα πεπερασμένων διαφορών

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

3.24 Θεωρούμε το πρόβλημα δύο σημείων (3.1), με $a = 1$ και $b = 0$. Έστω S_h^2 ένας υπόχωρος του $H_0^1(0,1)$ αποτελούμενος από συνεχείς συναρτήσεις οι οποίες σε καθένα των υποδιαστημάτων $[x_i, x_{i+1}]$, $i = 0, \dots, J$, ενός διαμερισμού $0 = x_0 < x_1 < \cdots < x_{J+1} = 1$ του $[0,1]$ είναι πολυώνυμα βαθμού το πολύ ένα. Χρησιμοποιήστε τις εκ των υστέρων εκτιμήσεις (3.52) και (3.64) για να οδηγηθείτε στις εκ των προτέρων εκτιμήσεις (3.29) και (3.30), με $r = 2$. Αυτή η διαδικασία επιβεβαιώνει ότι οι εκ των υστέρων εκτιμήσεις, τουλάχιστον στην εν λόγω περίπτωση, είναι βέλτιστης τάξεως.

[*Υπόδειξη*: Από την (3.52), 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 (3.71) τώρα γράφεται στη μορφή $r_h = u'' + c(u - u_h)$ και αν χρησιμοποιήσουμε το γεγονός ότι $\|u_h\| \le \frac{1}{2}\|f\|$ οδηγούμεθα εύκολα στο ζητούμενο.]

3.25 Θεωρούμε το πρόβλημα δύο σημείων (3.1). Αντί της δεύτερης υπόθεσης στην (3.2), υποθέτουμε ότι $c(x) > 0$ και

$$\max_x |b(x)|^2 \le 4 \min_x a(x)\ \min_x c(x).$$

Αποδείξτε ότι η διγραμμική μορφή a, που ορίστηκε λίγο πριν την (3.6), είναι ελλειπτική.

3.26 Αποδείξτε την εκτίμηση (3.93) για $s = 3$.

3.27 (Αντίστροφη ανισότητα.) Έστω $0 = x_0 < x_1 < \cdots < x_{J+1} = 1$ ένας διαμερισμός του διαστήματος $[0,1]$, $h_i := x_i - x_{i-1}$, $i = 1, \dots, J+1$, και

$$S_h^r := \{\chi \in H_0^1(0,1) : \chi|_{[x_{i-1},x_i]} \in \mathbb{P}_{r-1}, i = 1, \dots, J+1\}.$$

Θεωρούμε μια οικογένεια τέτοιων διαμερισμών, η οποία υποθέτουμε ότι είναι *ημιομοιόμορφη*, δηλαδή ότι

$$\frac{\max_i(x_{i+1} - x_i)}{\min_i(x_{i+1} - x_i)} \le c,$$

με την ίδια σταθερά c για όλους τους διαμερισμούς. Αποδείξτε ότι ισχύει η ακόλουθη *αντίστροφη ανισότητα*: Υπάρχει μια σταθερά C, ανεξάρτητη του διαμερισμού, τέτοια ώστε

$$(\star) \qquad \|\chi'\| \le Ch^{-1}\|\chi\| \quad \forall \chi \in S_h^r$$

με $h := \max_i h_i$. Η ανισότητα αυτή λέγεται αντίστροφη, γιατί εκτιμούμε μια ισχυρή νόρμα με μια ασθενέστερη. Κάτι τέτοιο δεν είναι εφικτό στον $H_0^1(0,1)$. Σε χώρους πεπερασμένης διάστασης όλες οι νόρμες είναι ισοδύναμες, φυσικά. Το ουσιαστικό σε αυτή την εκτίμηση είναι η ειδική μορφή του συντελεστή στο δεξιό μέλος.

α) Έστω $J \in \mathbb{N}, h := \frac{1}{J+1}$ και $x_i := ih, i = 0, \dots, J+1$, ένας ομοιόμορφος διαμερισμός του διαστήματος $[0,1]$. Θεωρούμε τη συνάρτηση $v : [0,1] \to \mathbb{R}$,

$$v(x) := \begin{cases} \dfrac{x - x_0}{h}, & x_0 \le x \le x_1, \\ \dfrac{x_2 - x}{h}, & x_1 < x \le x_2, \\ 0, & \text{διαφορετικά}. \end{cases}$$

Αποδείξτε ότι $\|v'\| = \sqrt{3}\, h^{-1}\|v\|$, και οδηγηθείτε στο συμπέρασμα ότι ο εκθέτης του h στο δεξιό μέλος της αντίστροφης ανισότητας $(\star)$ είναι ο καλύτερος δυνατός.

β) Έστω $v \in S_h^r$ και $i \in \{0, \dots, J\}$. 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,

$$\|v'\|^2_{L^2} \le C(\min_i h_i)^{-2} \|v\|^2_{L^2}.$$

Λαμβάνοντας υπ' όψιν την ημιομορφία του διαμερισμού, οδηγηθείτε στην αντίστροφη ανισότητα $(\star)$.

3.28 Θεωρούμε το πρόβλημα δύο σημείων

$$(\star) \qquad \begin{cases} -(au')' + bu' + cu = f \quad \text{στο } [0,1], \\ u'(0) = u'(1) = 0, \end{cases}$$

Υποθέτουμε ότι $a(x), c(x) > 0$, για κάθε $x \in [0,1]$, και

$$\max_x |b(x)|^2 < 4 \min_x a(x) \ \min_x c(x).$$

Αποδείξτε ότι το εν λόγω πρόβλημα έχει το πολύ μία λύση.
[*Υπόδειξη*: Με $\alpha := \min_x a(x), \gamma := \min_x c(x)$ και $\beta := \max_x |b(x)|$, έχουμε

$$\begin{aligned} \int_0^1 \big[-(au')' + bu' + cu\big] u \, dx &= \int_0^1 \big[a(u')^2 + bu'u + cu^2\big] \, dx \\ &\ge \alpha \|u'\|^2 - \beta \|u'\| \, \|u\| + \gamma \|u\|^2. \big] \end{aligned}$$

3.29 Έστω $a = x_0 < x_1 < \cdots < x_J = b$ ένας διαμερισμός ενός διαστήματος $[a,b]$. Προσεγγίζουμε το ολοκλήρωμα μιας συνάρτησης $v \in H^2(a,b)$ στο (a,b) με τον σύνθετο τύπο του τραπεζίου Q,

$$Q(v) := \int_a^b v(x) \, dx - \frac{1}{2} \sum_{i=1}^J h_i [v(x_{i-1}) + v(x_i)],$$

με $h_i := x_i - x_{i-1}, i = 1, \dots, J$. Αποδείξτε για το σφάλμα ότι

$$\Big| \int_a^b v(x) \, dx - Q(v) \Big| \le ch^2 \|v''\|_{L^2(a,b)},$$

με $h := \max_{1 \le i \le J} h_i$ και μια σταθερά c ανεξάρτητη της v και του διαμερισμού.

[*Υπόδειξη*: Έστω R το σφάλμα του απλού τύπου του τραπεζίου στο διάστημα $[0,1]$,

$$R(u) := \Big| \int_0^1 u(x)\,dx - \frac{1}{2}[u(0)+u(1)]\Big|.$$

Χρησιμοποιήστε το λήμμα των Bramble–Hilbert (Πρόταση 3.1 με $r=2$) για να βεβαιωθείτε ότι

$$R(u) \le c\|u''\|_{L^2(0,1)}.$$

Θέτουμε τώρα

$$R_i(v) := \Big| \int_{x_{i-1}}^{x_i} v(x)\,dx - \frac{1}{2}h_i[v(x_{i-1})+v(x_i)]\Big|.$$

Με την αλλαγή μεταβλητής $x = x_{i-1} + h_i s$, αποδείξτε ότι

$$R_i(v) \le c h_i^{5/2}\|v''\|_{L^2(x_{i-1},x_i)}, \quad i=1,\dots,J.$$

Τέλος, χρησιμοποιώντας την ανισότητα των Cauchy–Schwarz για αθροίσματα, έχουμε

$$\begin{aligned}\sum_{i=1}^J R_i(v) &\le ch^2 \sum_{i=1}^J h_i^{1/2}\|v''\|_{L^2(x_{i-1},x_i)} \le ch^2 \Big(\sum_{i=1}^J h_i\Big)^{1/2}\Big(\sum_{i=1}^J |v''\|^2_{L^2(x_{i-1},x_i)}\Big)^{1/2}\\ &= c\sqrt{b-a}\,h^2\|v''\|_{L^2(a,b)}.\end{aligned}$$]

4. Ελλειπτικές εξισώσεις

Σε αυτό το κεφάλαιο θα μελετήσουμε μεθόδους πεπερασμένων στοιχείων για ελλειπτικές εξισώσεις σε δύο διαστάσεις. Στην πρώτη ενότητα θα παρουσιάσουμε εν συντομία κάποια βασικά αποτελέσματα για τους χώρους L^p και για τους χώρους του Sobolev σε πολλές διαστάσεις. Στη δεύτερη ενότητα θα μελετήσουμε μεθόδους πεπερασμένων στοιχείων· το μεγαλύτερο μέρος της παρουσίασης αφορά χώρους πεπερασμένων στοιχείων που αποτελούνται από συνεχείς, τμηματικά ομοπαραλληλικές (δηλαδή πολυώνυμα βαθμού το πολύ ένα) συναρτήσεις.

4.1 Χώροι του Sobolev σε πολλές διαστάσεις

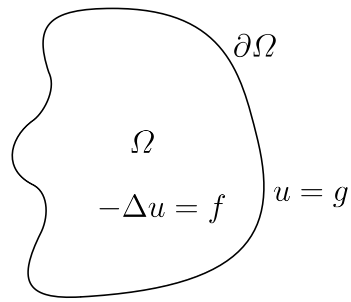
Σε αυτή την ενότητα θα παρουσιάσουμε βασικά αποτελέσματα για χώρους του Sobolev σε πολλές διαστάσεις. Πολλά από όσα αναφέραμε στο δεύτερο κεφάλαιο, κυρίως για τους χώρους L^p, αλλά ακόμα και για χώρους του Sobolev σε μία διάσταση, γενικεύονται αμέσως και σε πολλές διαστάσεις. Ως εκ τούτου η παρουσίαση εδώ θα είναι ακόμα συνοπτικότερη· υπάρχουν πάντως βασικά αποτελέσματα που εξαρτώνται κατά ουσιαστικό τρόπο από τη διάσταση, και σε κάποια από αυτά θα εστιάσουμε την προσοχή μας.

Τα προβλήματα συνοριακών τιμών για διαφορικές εξισώσεις διατυπώνονται συνήθως σε ανοικτά, συνεκτικά, μη κενά υποσύνολα του $\mathbb{R}^d$. Τέτοια υποσύνολα λέγονται *χωρία*. Έστω λοιπόν $\Omega \subset \mathbb{R}^d$ ένα φραγμένο χωρίο.

Για ομαλές συναρτήσεις $u : \Omega \to \mathbb{R}$ συμβολίζουμε με Δ τον τελεστή του Laplace (τη Λαπλασιανή),

$$\Delta u := u_{x_1x_1} + \cdots + u_{x_dx_d}.$$

Μια εξίσωση της μορφής $\Delta u = f$, με δεδομένη τη συνάρτηση f και ζητούμενη τη συνάρτηση u, λέγεται *εξίσωση του Poisson*, και στην ειδική περίπτωση $f = 0$ *εξίσωση του Laplace*.



Σχήμα 4.1: Το πρόβλημα για την εξίσωση του Poisson με συνοριακές συνθήκες Dirichlet.

Έστω $\Omega \subset \mathbb{R}^d$ ένα φραγμένο χωρίο, με ομαλό σύνορο $\partial\Omega$. Έστω $f \in C(\bar{\Omega})$ και $g \in C(\partial\Omega)$. Θεωρούμε το πρόβλημα συνοριακών τιμών για την εξίσωση του Poisson, με συνοριακές συνθήκες Dirichlet, στο Ω: Ζητείται συνάρτηση $u \in C^2(\Omega) \cap C(\bar{\Omega})$ τέτοια ώστε

$$
\begin{cases} -\Delta u = f & \text{στο } \Omega, \\ u = g & \text{στο } \partial\Omega. \end{cases} \tag{4.1}
$$

Για να απλοποιήσουμε κάπως το πρόβλημα, θεωρούμε μια συνάρτηση $v \in C^2(\bar{\Omega})$ τέτοια ώστε $v = g \in \partial\Omega$. Αυτό είναι ένα πρόβλημα παρεμβολής και δεν σχετίζεται με διαφορικές εξισώσεις. Τότε η συνάρτηση $w, w := u - v$, ικανοποιεί την εξίσωση $-\Delta w = -\Delta u + \Delta v = f + \Delta v =: \tilde{f}$, συνεπώς αποτελεί λύση του αντίστοιχου προβλήματος συνοριακών τιμών με *ομογενείς* συνοριακές συνθήκες Dirichlet,

$$
\begin{cases} -\Delta w = \tilde{f} & \text{στο } \Omega, \\ w = 0 & \text{στο } \partial\Omega. \end{cases}
$$

Χωρίς περιορισμό της γενικότητας θεωρούμε λοιπόν, αντί του (4.1), το πρόβλημα

$$
\begin{cases} -\Delta u = f & \text{στο } \Omega, \\ u = 0 & \text{στο } \partial\Omega. \end{cases} \tag{4.2}
$$

Μια λύση του προβλήματος (4.2), στον χώρο που αναφέρθηκε προηγουμένως, λέγεται *κλασική* λύση του προβλήματος.

Έστω $\varphi : \Omega \to \mathbb{R}$ μια αρκετά ομαλή συνάρτηση. Τότε θα έχουμε $-\Delta u\,\varphi = f\,\varphi$, οπότε

$$
-\int_\Omega \Delta u\,\varphi\,dx = \int_\Omega f\,\varphi\,dx.
$$

Τώρα, σύμφωνα με τον τύπο του Green,

$$-\int_\Omega \Delta u \,\varphi \,dx = -\int_{\partial\Omega} \varphi \nabla u \cdot \nu dS + \int_\Omega \nabla u \cdot \nabla \varphi \,dx,$$

με $\cdot$ το Ευκλείδειο εσωτερικό γινόμενο στον $\mathbb{R}^d$ και ν το εξωτερικό μοναδιαίο κάθετο διάνυσμα στο $\partial\Omega$. Υποθέτουμε τώρα ότι η συνάρτηση φ μηδενίζεται στο σύνορο, όπως μηδενίζεται και η λύση u, και το ολοκλήρωμα στο σύνορο εξαφανίζεται, οπότε συνολικά λαμβάνουμε

$$\int_\Omega \nabla u \cdot \nabla \varphi \,dx = \int_\Omega f \,\varphi \,dx, \tag{4.3}$$

για όλες τις ομαλές συναρτήσεις φ, οι οποίες μηδενίζονται στο σύνορο $\partial\Omega$.

Έστω $\big(H, (\cdot,\cdot)\big)$ ένας χώρος συναρτήσεων, με στοιχεία που μηδενίζονται στο σύνορο $\partial\Omega$, με εσωτερικό γινόμενο

$$(v, w) := \int_\Omega \nabla v \cdot \nabla w \,dx. \tag{4.4}$$

Φυσικά ο χώρος H πρέπει να είναι τέτοιος ώστε το εσωτερικό γινόμενο (v, w) να ορίζεται για κάθε $v, w \in H$. Είναι πολύ εύκολο να διαπιστώσουμε ότι αυτό συμβαίνει, αν και μόνο αν, για κάθε $v \in H$, η ποσότητα $\|v\|$,

$$\|v\| := \Big[\int_\Omega |\nabla v|^2 \,dx \Big]^{1/2}, \tag{4.5}$$

με $|\cdot|$ την Ευκλείδεια νόρμα στον $\mathbb{R}^d$, είναι πεπερασμένη.

Υποθέτουμε ότι το γραμμικό συναρτησιακό $F : H \to \mathbb{R}$,

$$F(\varphi) := \int_\Omega f \varphi \,dx,$$

είναι φραγμένο, κάτι που είναι εύκολο να επιτευχθεί. Διατυπώνουμε τότε το πρόβλημα (4.2) στη μορφή: Ζητείται $u \in H$ τέτοια ώστε

$$(u, \varphi) = F(\varphi) \qquad \forall \varphi \in H. \tag{4.6}$$

Η διατύπωση αυτή του προβλήματος λέγεται *ασθενής* ή *γενικευμένη* (γιατί σε αυτή χρησιμοποιούνται μόνο οι πρώτες παράγωγοι κατ' αντιδιαστολή προς το αρχικό πρόβλημα, όπου εμφανίζονται και δεύτερες παράγωγοι) ή *μεταβολική* (γιατί μελετάμε τις ιδιότητες της u καθώς η φ μεταβάλλεται).

Αν ο χώρος $\big(H, (\cdot,\cdot)\big)$ είναι πλήρης, δηλαδή χώρος Hilbert, η ύπαρξη και μοναδικότητα, καθώς και μια εκτίμηση στη νόρμα $\|\cdot\|$, της λύσης $u \in H$ του προβλήματος (4.6) έπονται αμέσως από το θεώρημα αναπαράστασης του Riesz. Οι χώροι των συνεχών, και αρκετά ομαλών, συναρτήσεων *δεν* είναι δυστυχώς χώροι Hilbert ως προς το εν λόγω εσωτερικό γινόμενο. Ο ορισμός κατάλληλων τέτοιων χώρων οδήγησε στην εισαγωγή χώρων του Sobolev.

Με τη μεταβολική διατύπωση επιτυγχάνονται συγχρόνως δύο στόχοι: Αφ' ενός ορίζονται λύσεις με φυσική σημασία και σε περιπτώσεις που οι κλασικές λύσεις δεν έχουν έννοια, και αφ' ετέρου μας παρέχεται η δυνατότητα να χρησιμοποιήσουμε μέσα της Συναρτησιακής Ανάλυσης για να μελετήσουμε θέματα ύπαρξης, μοναδικότητας και συνεχούς εξάρτησης της λύσης από τα δεδομένα.

Οι ασθενείς λύσεις ορίζονται για μια κατάλληλη κλάση συναρτήσεων f. Στην περίπτωση που η f είναι ομαλή, και το σύνορο $\partial\Omega$ αρκετά "ομαλό", τότε η ασθενής λύση δεν ανήκει απλώς στον H, αλλά έχει καλύτερες ιδιότητες ομαλότητας. Τέτοιες μελέτες αναφέρονται ως "ομαλότητα" της λύσης. Ο κύκλος κλείνει, αποδεικνύοντας ότι, υπό κατάλληλες συνθήκες, η ασθενής λύση είναι και ομαλή, και έτσι οδηγούμαστε σε μια απάντηση στο αρχικό μας ερώτημα για το πρόβλημα (4.2).

Για να αποκλείσουμε την πιθανότητα να αφήσουμε τον αναγνώστη με την εντύπωση ότι με τη μεταβολική διατύπωση οι διαφορικές εξισώσεις αντιμετωπίζονται πλήρως με μέσα της Συναρτησιακής Ανάλυσης, τονίζουμε ότι η Συναρτησιακή Ανάλυση παίζει μεν χρήσιμο βοηθητικό ρόλο, δεν επαρκεί όμως για να απαντηθούν καθοριστικής σημασίας ερωτήματα για διαφορικές εξισώσεις· τέτοια θέματα απαιτούν συνήθως λεπτομερέστερη κάθε φορά μελέτη με ειδικά μέσα Ανάλυσης.

4.1.1 Προκαταρκτικά: Οι χώροι L^p

Οι αναγνώστες είναι ήδη εξοικειωμένοι με τους χώρους L^p σε μία διάσταση. Περιοριζόμαστε ως εκ τούτου σε μια συνοπτική ανασκόπηση της θεωρίας των αντίστοιχων χώρων σε πολλές διαστάσεις, παραλείποντας όλες τις αποδείξεις. Ακολουθούμε το εξαιρετικό βιβλίο των Adams και Fournier, βλ. τη βιβλιογραφία, [1], όπως πράξαμε και στην περίπτωση της μίας διάστασης. Όπως γνωρίζουμε, στους χώρους του Sobolev δεν χρησιμοποιούμε το ολοκλήρωμα του Riemann, αλλά το ολοκλήρωμα του Lebesgue. Σε αυτή την ενότητα λοιπόν θα ορίσουμε το ολοκλήρωμα του Lebesgue και θα δούμε κάποιες ιδιότητες των χώρων L^p, χώρων που ορίζονται επίσης με βάση το ολοκλήρωμα του Lebesgue.

Το μέτρο του Lebesgue. Σε αυτό το εδάφιο θα ασχοληθούμε με το μέτρο του Lebesgue σε πολλές διαστάσεις, που μας παρέχει τη δυνατότητα να περιγράψουμε με ακρίβεια τι εννοούμε ως όγκο (εμβαδόν, στις δύο διαστάσεις) ορισμένων υποσυνόλων του $\mathbb{R}^d$.

Οι ορισμοί της $\sigma-$άλγεβρας και του θετικού μέτρου γενικεύονται κατά προφανή τρόπο στις πολλές διαστάσεις. Προχωρούμε λοιπόν κατ' ευθείαν στο μέτρο του Lebesgue.

Θεώρημα 4.1 (Ύπαρξη του μέτρου του Lebesgue.) *Υπάρχει μια $\sigma-$άλγεβρα Σ υποσυνόλων του $\mathbb{R}^d$ και ένα θετικό μέτρο μ στη Σ με τις εξής ιδιότητες:*

$i.$ *Κάθε ανοικτό υποσύνολο του $\mathbb{R}^d$ ανήκει στη Σ.*

$ii.$ *Αν $A \subset B$, $B \in \Sigma$ και $\mu(B) = 0$, τότε και το A ανήκει στη Σ και $\mu(A) = 0$.*

$iii.$ *Αν $A = [a_1, b_1]\times\cdots\times[a_d, b_d]$, τότε το A ανήκει στη Σ και $\mu(A) = (b_1-a_1)\cdots(b_d-a_d)$.*

$iv.$ *Το μ είναι αναλλοίωτο σε μετατοπίσεις, δηλαδή, αν $x \in \mathbb{R}^d$ και $A \in \Sigma$, τότε και $x + A \in \Sigma$, όπου $x + A := \{x + y : y \in A\}$, και $\mu(x + A) = \mu(A)$.* □

Σχόλιο: Από το $i.$ έπεται ότι και κάθε κλειστό υποσύνολο του $\mathbb{R}^d$ ανήκει στη Σ. □

Τα στοιχεία της $\sigma-$άλγεβρας Σ του Θεωρήματος 4.1 λέγονται *μετρήσιμα κατά Lebesgue* (και θα αναφέρονται στη συνέχεια απλώς ως μετρήσιμα) υποσύνολα του $\mathbb{R}^d$. Το μέτρο μ καλείται *μέτρο του Lebesgue*, το $\mu(A)$ θα αναφέρεται ως μέτρο ή όγκος (εμβαδόν στις δύο διαστάσεις) του A. Τα στοιχεία της Σ με μέτρο μηδέν λέγονται *σύνολα μέτρου μηδέν*. Κάθε υποσύνολο του $\mathbb{R}^{d-1}$ έχει μέτρο μηδέν στον $\mathbb{R}^d$.

Η έννοια *σχεδόν παντού* (σ.π.) το A γενικεύεται κατά προφανή τρόπο και για $A \subset \mathbb{R}^d$, βλ. τον Ορισμός 2.3.

Ορισμός 4.1 (Μετρήσιμες συναρτήσεις.) Έστω A ένα μετρήσιμο υποσύνολο του $\mathbb{R}^d$. Μια συνάρτηση $f : A \to [-\infty, \infty]$ λέγεται *μετρήσιμη*, αν τα σύνολα $\{x \in A : f(x) > a\}$ είναι μετρήσιμα, για κάθε $a \in \mathbb{R}$.

Θεώρημα 4.2 (Ιδιότητες μετρήσιμων συναρτήσεων.)

$i.$ *Αν μια συνάρτηση f είναι μετρήσιμη, τότε και η συνάρτηση $|f|$ είναι μετρήσιμη.*

$ii.$ *Αν f και g είναι δύο μετρήσιμες συναρτήσεις, τότε και οι $f+g$ και fg είναι μετρήσιμες.*

iii. *Αν* $(f_n)_{n\in\mathbb{N}}$ *είναι μια ακολουθία μετρήσιμων συναρτήσεων, τότε και οι συναρτήσεις* $\sup_{n\in\mathbb{N}} f_n$ *(με* $(\sup_{n\in\mathbb{N}} f_n)(x) := \sup_{n\in\mathbb{N}} f_n(x)$ *),* $\inf_{n\in\mathbb{N}} f_n$*,* $\limsup_{n\in\mathbb{N}} f_n$ *καθώς και* $\liminf_{n\in\mathbb{N}} f_n$ *είναι μετρήσιμες.*

iv. *Κάθε συνεχής συνάρτηση, ορισμένη σε ένα μετρήσιμο σύνολο, είναι μετρήσιμη.*

v. *Έστω* $f:\mathbb{R}\to\mathbb{R}$ *μια συνεχής συνάρτηση και* $g:\mathbb{R}^d\to\mathbb{R}$ *μια μετρήσιμη συνάρτηση. Τότε και η σύνθεση* $f\circ g$ *είναι μετρήσιμη.*

vi. (*Το θεώρημα του Lusin.*) *Αν μία συνάρτηση* f *είναι μετρήσιμη και* $f(x)=0$ *για* $x\in A^c$*, όπου* A *υποσύνολο του* $\mathbb{R}^d$ *τέτοιο ώστε* $\mu(A)<\infty$*, και* $\varepsilon>0$*, τότε υπάρχει μια συνεχής συνάρτηση* $g:\mathbb{R}^d\to\mathbb{R}$ *με συμπαγή φορέα, τέτοια ώστε* $\max_{x\in\mathbb{R}^d} g(x)\le\sup_{x\in\mathbb{R}^d} f(x)$ *και* $\mu\Big(\{x\in\mathbb{R}^d : f(x)\neq g(x)\}\Big)<\varepsilon$*.* □

Ορισμός 4.2 (Χαρακτηριστικές και απλές συναρτήσεις.) Έστω A ένα υποσύνολο του $\mathbb{R}^d$. Η συνάρτηση $\chi_A:\mathbb{R}^d\to\mathbb{R}$,

$$\chi_A(x) := \begin{cases} 1, & \text{αν } x\in A,\\ 0, & \text{αν } x\notin A,\end{cases}$$

λέγεται *χαρακτηριστική συνάρτηση του* A. Μια συνάρτηση $s:\mathbb{R}^d\to\mathbb{R}$ καλείται *απλή*, αν η εικόνα της είναι ένα πεπερασμένο σύνολο, δηλαδή αν υπάρχουν $a_1,\dots,a_n\in\mathbb{R}$ (ανά δύο διαφορετικά μεταξύ τους), τέτοια ώστε $s(x)\in\{a_1,\dots,a_n\}$, για κάθε $x\in\mathbb{R}^d$.

Σχόλιο: Έστω $s:\mathbb{R}^d\to\mathbb{R}$ μια απλή συνάρτηση, $s(x)\in\{a_1,\dots,a_n\}$. Θέτουμε $A_i := \{x\in\mathbb{R}^d : s(x)=a_i\}$, $i=1,\dots,n$. Αμέσως διαπιστώνουμε ότι η s γράφεται στη μορφή

$$s=\sum_{i=1}^n a_i\chi_{A_i}.$$

Η s είναι μετρήσιμη, αν και μόνο αν τα σύνολα $A_1,\dots,A_n$ είναι μετρήσιμα. □

Οι απλές συναρτήσεις έχουν πολύ καλές προσεγγιστικές ιδιότητες, όπως θα δούμε αμέσως τώρα.

Θεώρημα 4.3 (Προσέγγιση συναρτήσεων με απλές συναρτήσεις.) *Έστω* A *ένα υποσύνολο του* $\mathbb{R}^d$ *και* $f:A\to\mathbb{R}$ *μια συνάρτηση. Τότε ισχύουν:*

i. *Υπάρχει μια ακολουθία* $(s_n)_{n\in\mathbb{N}}$ *απλών συναρτήσεων, που συγκλίνει σημειακά (ή κατά σημείο) στην* f*, δηλαδή*

$$\lim_{n\to\infty} s_n(x)=f(x),\quad \text{για κάθε } x\in A.$$

ii. *Αν η f είναι φραγμένη, τότε η ακολουθία $(s_n)_{n\in\mathbb{N}}$ μπορεί να επιλεγεί έτσι ώστε η σύγκλιση να είναι ομοιόμορφη,*

$$\lim_{n\to\infty}\Big(\sup_{x\in A}|f(x)-s_n(x)|\Big)=0.$$

iii. *Αν η f είναι μετρήσιμη, τότε η ακολουθία $(s_n)_{n\in\mathbb{N}}$ στο i. μπορεί να επιλεγεί έτσι ώστε κάθε s_n να είναι μετρήσιμη.*

iv. *Αν η f λαμβάνει μόνο μη αρνητικές τιμές, τότε η ακολουθία $(s_n)_{n\in\mathbb{N}}$ στο i. μπορεί να επιλεγεί έτσι ώστε, για κάθε $x\in\mathbb{R}$, η ακολουθία $\big(s_n(x)\big)_{n\in\mathbb{N}}$ να είναι αύξουσα.* □

Το ολοκλήρωμα του Lebesgue.

Σε αυτό το εδάφιο θα ασχοληθούμε με το ολοκλήρωμα του Lebesgue σε πολλές διαστάσεις. Στη συνέχεια δίνουμε τον ορισμό του ολοκληρώματος του Lebesgue, σε τρία στάδια, όπως ακριβώς κάναμε και στην περίπτωση μίας διάστασης: αρχίζουμε με την περίπτωση απλών συναρτήσεων, συνεχίζουμε με την περίπτωση μη αρνητικών συναρτήσεων και καταλήγουμε στη γενική περίπτωση. Κάθε μία από αυτές τις περιπτώσεις χρησιμοποιείται για να μας οδηγήσει στον ορισμό στην περίπτωση που ακολουθεί.

Ορισμός 4.3 (Το ολοκλήρωμα του Lebesgue.) Έστω A ένα μετρήσιμο υποσύνολο του $\mathbb{R}^d$. *Το ολοκλήρωμα του Lebesgue* μιας απλής συνάρτησης s,

$$s=\sum_{i=1}^n a_i\chi_{A_i},$$

όπου $A_i\subset A$ μετρήσιμα σύνολα, ορίζεται ως εξής

$$\int_A s(x)\,dx:=\sum_{i=1}^n a_i\mu(A_i).$$

Το ολοκλήρωμα του Lebesgue μιας μετρήσιμης, μη αρνητικής συνάρτησης $f:A\to\mathbb{R}$ ορίζεται ως εξής

$$\int_A f(x)\,dx:=\sup\int_A s(x)\,dx,$$

όπου το supremum λαμβάνεται ως προς όλες τις μετρήσιμες, απλές συναρτήσεις s, που μηδενίζονται στο συμπλήρωμα του A, τέτοιες ώστε $0\le s(x)\le f(x)$, για κάθε

$x \in A$.

Το ολοκλήρωμα του Lebesgue μιας μετρήσιμης συνάρτησης $f : A \to \mathbb{R}$ ορίζεται ως εξής
$$\int_A f(x)\,dx := \int_A f^+(x)\,dx - \int_A f^-(x)\,dx,$$
υπό την προϋπόθεση ότι τουλάχιστον ένα από τα ολοκληρώματα στο δεξιό μέλος είναι πεπερασμένο, όπου $f^+(x) := \max\big(f(x), 0\big)$ και $f^-(x) := \min\big(f(x), 0\big)$.
Αν το ολοκλήρωμα του Lebesgue μιας συνάρτησης f στο A είναι πεπερασμένο, τότε λέμε ότι η f είναι *κατά Lebesgue ολοκληρώσιμη στο A*. Το σύνολο των κατά Lebesgue ολοκληρώσιμων συναρτήσεων στο A συμβολίζεται με $L^1(A)$.

Θεώρημα 4.4 (Ιδιότητες ολοκληρώσιμων συναρτήσεων.) *Υποθέτουμε ότι όλες οι συναρτήσεις και τα σύνολα που εμφανίζονται στη συνέχεια είναι μετρήσιμες, και μετρήσιμα, αντίστοιχα.*

i. *Αν μια συνάρτηση f είναι φραγμένη σε ένα σύνολο A και $\mu(A) < \infty$, τότε η f είναι ολοκληρώσιμη στο A, $f \in L^1(A)$.*

ii. *Αν $\mu(A) < \infty$ και $a \le f(x) \le b$, για κάθε $x \in A$, τότε*
$$a\mu(A) \le \int_A f(x)\,dx \le b\mu(A).$$

iii. *Αν $f(x) \le g(x)$, για κάθε $x \in A$, και αν υπάρχουν και τα δύο ολοκληρώματα, τότε*
$$\int_A f(x)\,dx \le \int_A g(x)\,dx.$$

iv. *Αν $f, g \in L^1(A)$, τότε $f + g \in L^1(A)$ και*
$$\int_A (f+g)(x)\,dx = \int_A f(x)\,dx + \int_A g(x)\,dx.$$

v. *Αν $f \in L^1(A)$ και $c \in \mathbb{R}$, τότε $cf \in L^1(A)$ και*
$$\int_A (cf)(x)\,dx = c\int_A f(x)\,dx.$$

vi. *Αν $f \in L^1(A)$, τότε $|f| \in L^1(A)$ και*
$$\Big|\int_A f(x)\,dx\Big| \le \int_A |f(x)|\,dx.$$

vii. *Αν $f \in L^1(A)$ και $B \subset A$, τότε $f \in L^1(B)$. Αν επί πλέον $f(x) \geq 0$, για κάθε $x \in A$, τότε*

$$\int_B f(x)\,dx \leq \int_A f(x)\,dx.$$

viii. *Αν $\mu(A) = 0$, τότε*

$$\int_A f(x)\,dx = 0.$$

ix. *Αν $f \in L^1(A)$ και*

$$\int_B f(x)\,dx = 0,$$

για κάθε $B \subset A$, τότε $f(x) = 0$ σ.π. στο A. □

Οι χώροι L^p. Έστω Ω ένα χωρίο στον $\mathbb{R}^d$, δηλαδή ένα *μη κενό*, *ανοικτό* υποσύνολο του $\mathbb{R}^d$. Έστω $p \geq 1$. Συμβολίζουμε με $L^p(\Omega)$ την κλάση των μετρήσιμων συναρτήσεων, που ορίζονται στο Ω και παίρνουν πραγματικές τιμές, με την ιδιότητα η $|u|^p$ να είναι ολοκληρώσιμη στο Ω, δηλαδή

$$\int_\Omega |u(x)|^p\,dx < \infty.$$

Στοιχεία του $L^p(\Omega)$, που είναι ίσα σ.π. στο Ω, θεωρούμε ότι ταυτίζονται. Χάριν ευκολίας θα αναφέρουμε τα στοιχεία του $L^p(\Omega)$ ως συναρτήσεις και θα γράφουμε $u \in L^p(\Omega)$, αν η u ικανοποιεί την ανισότητα που αναφέρθηκε προηγουμένως.

Όπως και στη μία διάσταση, βλ. την Πρόταση 2.1, έχουμε:

Πρόταση 4.1 (Ο $L^p(\Omega)$ είναι γραμμικός χώρος.) *Ο $L^p(\Omega)$ είναι γραμμικός χώρος και η απεικόνιση $\|\cdot\|_p : L^p(\Omega) \to \mathbb{R}$,*

$$\|u\|_p := \Big(\int_\Omega |u(x)|^p\,dx\Big)^{1/p},$$

είναι νόρμα. □

Θεώρημα 4.5 (Πληρότητα των $L^p(\Omega)$.) *Οι χώροι $L^p(\Omega)$, $1 \leq p < \infty$, είναι πλήρεις.* □

Πόρισμα 4.1 (Ο $L^2(\Omega)$ είναι χώρος Hilbert.) *Ο χώρος $L^2(\Omega)$ είναι χώρος Hilbert με το εσωτερικό γινόμενο*

$$(u, v) := \int_\Omega uv\,dx.$$ □

Σχόλιο. (Ο χώρος $L^\infty(\Omega)$.) Λέμε ότι μια μετρήσιμη συνάρτηση u στο Ω είναι *ουσιαστικά φραγμένη* στο Ω, αν υπάρχει σταθερά K τέτοια ώστε

$$|u(x)| \leq K \ \text{ σ.π. στο } \Omega.$$

Το infimum των σταθερών K, για τις οποίες ισχύει αυτή η σχέση, λέγεται *ουσιαστικό supremum* της $|u|$ στο Ω και συμβολίζεται με

$$\operatorname{ess\,sup}_{x\in\Omega} |u(x)|.$$

Ο χώρος των ουσιαστικά φραγμένων συναρτήσεων στο Ω συμβολίζεται με $L^\infty(\Omega)$. Ο $L^\infty(\Omega)$ είναι γραμμικός χώρος και η $\|\cdot\|_\infty$,

$$\|u\|_\infty := \operatorname{ess\,sup}_{x\in\Omega} |u(x)|,$$

είναι νόρμα στον $L^\infty(\Omega)$. Μάλιστα ο $\big(L^\infty(\Omega), \|\cdot\|_\infty\big)$ είναι πλήρης χώρος.

Εύκολα πείθεται κανείς, ότι στην περίπτωση που μια συνάρτηση u είναι συνεχής στο Ω, το ουσιαστικό supremum της ταυτίζεται με το supremum της. □

Για $p \neq 2$ οι χώροι $L^p(\Omega)$ είναι, όπως αναφέραμε ήδη, πλήρεις, δηλαδή χώροι Banach, αλλά η νόρμα τους δεν παράγεται από εσωτερικό γινόμενο, δηλαδή αυτοί οι χώροι δεν είναι χώροι Hilbert.

Πρόταση 4.2 (Πυκνότητα του $C_0(\Omega)$ στους $L^p(\Omega)$.) *Για $1 \leq p < \infty$, ο χώρος $C_0(\Omega)$ είναι πυκνός στον $L^p(\Omega)$.* □

Πρόταση 4.3 (Διαχωρισιμότητα των $L^p(\Omega)$.) *Για $1 \leq p < \infty$, ο χώρος $L^p(\Omega)$ είναι διαχωρίσιμος.* □

Σχόλιο. (Μη διαχωρισιμότητα του $L^\infty(\Omega)$.) Μπορεί να αποδειχθεί ότι ο χώρος $L^\infty(\Omega)$ *δεν* είναι διαχωρίσιμος. □

Αναφέρουμε ακόμα τον χώρο $L^1_{\text{loc}}(\Omega)$, ο οποίος αποτελείται από τις *τοπικά* (locally) ολοκληρώσιμες συναρτήσεις, δηλαδή, αν Ω_1 είναι ένα ανοικτό υποσύνολο του Ω τέτοιο ώστε το $\bar{\Omega}_1$ να είναι φραγμένο και να περιέχεται στο Ω, τότε οι περιορισμοί των στοιχείων του $L^1_{\text{loc}}(\Omega)$ στο Ω_1 περιέχονται στον $L^1(\Omega_1)$. Ο $L^1_{\text{loc}}(\Omega)$ είναι *γνήσιος* υπόχωρος του $L^1(\Omega)$, όπως και στη μία διάσταση. Σημειώστε ότι στον $L^1_{\text{loc}}(\Omega)$ δεν ορίζουμε κάποια νόρμα.

Ο *φορέας* $\operatorname{supp} u$ μιας συνάρτησης $u \in L^1(\Omega)$ ορίζεται ως εξής: Είναι το σύνολο των σημείων $x \in \bar{\Omega}$ με την ιδιότητα ότι για κάθε $\delta > 0$ υπάρχει μια συνάρτηση $\varphi \in C_0\big(B(x;\delta)\big)$, τέτοια ώστε

$$\int_{B(x;\delta)} u(y)\varphi(y)\,dy \neq 0,$$

όπου $B(x;\delta)$ η σφαίρα (κύκλος στις δύο διαστάσεις) με κέντρο x και ακτίνα δ, $B(x;\delta) := \{y \in \mathbb{R}^d : |x-y| < \delta\}$, με $|\cdot|$ την Ευκλείδεια νόρμα στον $\mathbb{R}^d$. Εύκολα πείθεται κανείς ότι ο παρών ορισμός γενικεύει εκείνον που έχουμε ήδη δώσει για τον φορέα συνεχών συναρτήσεων.

Εξομαλυντές. Θεωρούμε πρώτα τη συνάρτηση $\omega : \mathbb{R}^d \to \mathbb{R}$,

$$\omega(x) := \begin{cases} k \cdot \mathrm{e} \cdot \mathrm{e}^{-\frac{1}{1-|x|^2}}, & x \in B(0;1), \\ 0, & x \notin B(0;1), \end{cases} \tag{4.7}$$

με τη θετική σταθερά k τέτοια ώστε το ολοκλήρωμα της ω στον $\mathbb{R}^d$ να ισούται με τη μονάδα. Προφανώς, $\operatorname{supp}\omega = \overline{B(0;1)} = \{y \in \mathbb{R}^d : |y| \leq 1\}$. Η ω είναι άπειρες φορές συνεχώς παραγωγίσιμη, βλ. την αντίστοιχη συζήτηση στη μία διάσταση. Σημειώνουμε ακόμα ότι

$$\omega(0) = k \ \text{ και } \ 0 \leq \omega(x) \leq k, \ \text{ για κάθε } \ x \in \mathbb{R}^d.$$

Για $\varepsilon > 0$, θεωρούμε τώρα τη συνάρτηση

$$\varphi_\varepsilon : \mathbb{R}^d \to \mathbb{R}, \quad \varphi_\varepsilon(x) := \frac{1}{\varepsilon^d}\omega(\frac{x}{\varepsilon}).$$

Αμέσως διαπιστώνουμε ότι $\varphi_\varepsilon \in C^\infty(\mathbb{R}^d)$, $\operatorname{supp}\varphi_\varepsilon = \overline{B(0;\varepsilon)}$, $\varphi_\varepsilon(0) = k/\varepsilon^d$, $0 \leq \varphi_\varepsilon(x) \leq k/\varepsilon^d$, για κάθε $x \in \mathbb{R}^d$, και το ολοκλήρωμα της φ_ε στον $\mathbb{R}^d$ ισούται με τη μονάδα.

Οι συναρτήσεις φ_ε, όπως και κάθε άλλη συνάρτηση με τις τρεις θεμελιώδεις ιδιότητες α) να είναι μη αρνητικές και να έχουν συμπαγή φορέα, β) να είναι άπειρες φορές παραγωγίσιμες συναρτήσεις και γ) να έχουν ολοκλήρωμα ίσο με τη μονάδα, καλούνται *εξομαλυντές*. Αν για μια συνάρτηση $u : \mathbb{R}^d \to \mathbb{R}$ υπάρχει το ολοκλήρωμα

$$\int_{\mathbb{R}^d} u(y)\varphi_\varepsilon(x-y)\,dy,$$

για κάθε $x \in \mathbb{R}^d$, όπως συμβαίνει για $u \in L^1_{\rm loc}(\mathbb{R}^d)$, τότε η *συνέλιξη* $\varphi_\varepsilon \star u$,

$$(\varphi_\varepsilon \star u)(x) := \int_{\mathbb{R}^d} u(y)\varphi_\varepsilon(x-y)\,dy,$$

λέγεται *εξομάλυνση* ή *ομαλοποίηση* της u.

Σημειώνουμε ότι η συνέλιξη $u \star v$ ορίζεται γενικά για δύο συναρτήσεις $u, v : \mathbb{R}^d \to \mathbb{R}$,

$$(u \star v)(x) := \int_{\mathbb{R}^d} u(x-y)v(y)\,dy,$$

και ότι $u \star v = v \star u$. Οι μαθηματικοί αρέσκονται να αναφέρονται στη συνέλιξη ως τον τέλειο γάμο, αφού το 'παιδί' $u \star v$ κληρονομεί 'οποιαδήποτε' καλή ιδιότητα έχει ένας εκ των γονέων u και v.

Η εξομάλυνση δεν είναι απαραίτητο να γίνεται με τις συναρτήσεις φ_ε, αλλά μπορεί να γίνεται με οποιονδήποτε άλλον εξομαλυντή. Εμείς εδώ θα χρησιμοποιούμε τους εξομαλυντές φ_ε.

Στη συνέχεια αναφέρουμε ορισμένες ιδιότητες της εξομάλυνσης.

Θεώρημα 4.6 (Ιδιότητες εξομαλύνσεων.) *Έστω $\Omega \subset \mathbb{R}^d$ και $u : \Omega \to \mathbb{R}$ μια συνάρτηση. Επεκτείνουμε τη u σε όλον τον $\mathbb{R}^d$ με $u(x) := 0$, για κάθε $x \notin \Omega$. Τότε ισχύουν:*

i. Αν $u \in L^1_{\rm loc}(\mathbb{R}^d)$, τότε $\varphi_\varepsilon \star u \in C^\infty(\mathbb{R}^d)$.

ii. Αν $u \in L^1_{\rm loc}(\mathbb{R}^d)$ και ο φορέας της u είναι κλειστός και φραγμένος και περιέχεται στο Ω, τότε $\varphi_\varepsilon \star u \in C^\infty(\Omega)$ και, για ε μικρότερο από την απόσταση μεταξύ του φορέα της u και του συνόρου του Ω, ο φορέας της συνέλιξης έχει τις ίδιες ιδιότητες.

iii. Αν $u \in L^p(\Omega)$ και $1 \le p < \infty$, τότε $\varphi_\varepsilon \star u \in L^p(\Omega)$. Επί πλέον

$$\|\varphi_\varepsilon \star u\|_p \le \|u\|_p \quad \textit{και} \quad \lim_{\varepsilon\to 0+} \|\varphi_\varepsilon \star u - u\|_p = 0.$$

iv. Αν $u \in C(\Omega)$, τότε $\lim_{\varepsilon\to0+}(\varphi_\varepsilon \star u)(x) = u(x)$ ομοιόμορφα σε κάθε κλειστό και φραγμένο υποσύνολο του Ω.

v. Αν $u \in C(\bar{\Omega})$, τότε $\lim_{\varepsilon\to0+}(\varphi_\varepsilon \star u)(x) = u(x)$ ομοιόμορφα στο Ω. □

Μια ενδιαφέρουσα απόρροια των $ii.$ και $iii.$ του Θεωρήματος 4.6 είναι το ακόλουθο Πόρισμα, παράβαλε με την Πρόταση 4.2, για τον χώρο $C_0^\infty(\Omega)$, που αποτελείται από τα στοιχεία του $C_0(\Omega)$ που είναι άπειρες φορές συνεχώς παραγωγίσιμα.

Πόρισμα 4.2 (Πυκνότητα του $C_0^\infty(\Omega)$ στους $L^p(\Omega)$.) *Για $1 \le p < \infty$, ο χώρος $C_0^\infty(\Omega)$ είναι πυκνός στον $L^p(\Omega)$.* □

4.1.2 Χώροι του Sobolev

Έχοντας πλέον εξοικειωθεί με τους χώρους L^p, προχωρούμε στο παρόν εδάφιο στο κύριο αντικείμενο αυτής της παραγράφου, την εισαγωγή των χώρων του Sobolev σε πολλές διαστάσεις. Ουσιαστικό ρόλο στους χώρους του Sobolev παίζουν οι λεγόμενες *γενικευμένες* παράγωγοι. Στο εδάφιο που ακολουθεί θα ασχοληθούμε λοιπόν με γενικευμένες παραγώγους και μετά θα είμαστε σε θέση να μιλήσουμε για χώρους του Sobolev.

Γενικευμένες παράγωγοι. Οι χώροι του Sobolev σε πολλές διαστάσεις μπορούν να ορισθούν σε οποιοδήποτε ανοικτό υποσύνολο Ω του $\mathbb{R}^d$.

Συμβολίζουμε με $C_0^\infty(\Omega)$ το σύνολο των συναρτήσεων, που ορίζονται στο Ω, είναι άπειρες φορές παραγωγίσιμες και έχουν συμπαγή φορέα που περιέχεται στο Ω. Ο $C_0^\infty(\Omega)$ λέγεται *χώρος δοκιμής*, ή *χώρος ελέγχου*, και τα στοιχεία του λέγονται *συναρτήσεις δοκιμής*. Συχνά ο χώρος $C_0^\infty(\Omega)$ συμβολίζεται με $\mathscr{D}(\Omega)$. Από το $ii.$ του Θεωρήματος 4.6 έπεται ιδιαίτερα ότι ο χώρος $C_0^\infty(\Omega)$ περιέχει και μη τετριμμένα στοιχεία. Για αυτό το γεγονός μπορεί να πεισθεί κανείς και ως εξής: Για $x^\star \in \mathbb{R}^d$, θεωρούμε τη συνάρτηση $\psi_\varepsilon : \mathbb{R}^d \to \mathbb{R}$, $\psi_\varepsilon(x) := \omega\big((x - x^\star)/\varepsilon\big)$. Η ψ_ε έχει κατά τα λοιπά τις ιδιότητες του εξομαλυντή φ_ε, αλλά ο φορέας της είναι η κλειστή σφαίρα $\overline{B(x^\star;\varepsilon)}$. Τέλος, για $x^\star \in \Omega$ και αρκετά μικρό ε, διαπιστώνουμε αμέσως ότι ο περιορισμός της συνάρτησης ψ_ε στο Ω ανήκει στον χώρο $C_0^\infty(\Omega)$.

Έστω $u \in C^m(\Omega)$. Τότε μπορούμε να ορίσουμε κατά τα γνωστά τις μερικές παραγώγους τάξεως το πολύ m της u σε κάθε σημείο του Ω. Αυτό δεν είναι εφικτό με συναρτήσεις, φερ' ειπείν, $u \in L^1_{\text{loc}}(\Omega)$, αφού ούτε καν για τις τιμές τους σε κάθε σημείο $x \in \Omega$ δεν μπορούμε να μιλήσουμε. Όπως και στη μία διάσταση, οδηγούμαστε εύκολα στον ακόλουθο ορισμό.

Ορισμός 4.4 (Γενικευμένες παράγωγοι.) Έστω $u, v \in L^1_{\text{loc}}(\Omega)$ και $\alpha \in \mathbb{N}_0^d$ ένας πολυδείκτης. Λέμε ότι η v είναι *γενικευμένη* (ή *ασθενής*) μερική παράγωγος $D^\alpha u$ της u, αν

$$\int_\Omega u D^\alpha \varphi \, dx = (-1)^{|\alpha|} \int_\Omega v \varphi \, dx \quad \forall \varphi \in C_0^\infty(\Omega), \tag{4.8}$$

όπου $|\alpha| := \alpha_1 + \cdots + \alpha_d$ η *τάξη* του πολυδείκτη α.

Σχόλιο. (Μοναδικότητα γενικευμένων μερικών παραγώγων.) Φυσικά, είναι δυνατόν είτε να υπάρχει είτε να μην υπάρχει η ασθενής παράγωγος μιας συνάρτησης του $L^1_{\text{loc}}(\Omega)$.

Πάντως, όπως και στην περίπτωση της μίας διάστασης, αποδεικνύεται μοναδικότητα της ασθενούς παραγώγου. Επίσης, εύκολα διαπιστώνουμε ότι, αν μια συνάρτηση $u : \Omega \to \mathbb{R}$ είναι m φορές συνεχώς παραγωγίσιμη, τότε υπάρχουν οι γενικευμένες παράγωγοί της $D^\alpha u$, για κάθε πολυδείκτη $\alpha \in \mathbb{N}_0^d$, τάξεως το πολύ m, και μάλιστα αυτές συμπίπτουν με τις αντίστοιχες κλασικές παραγώγους της u.

Σημειώνουμε ακόμα ότι η ασθενής παράγωγος για κάποιον πολυδείκτη α, φερ' ειπείν, ορίζεται απ' ευθείας για αυτόν τον πολυδείκτη, και είναι δυνατόν αυτή να υπάρχει παρ' ότι δεν υπάρχουν οι ασθενείς μερικές παράγωγοι για πολυδείκτες χαμηλότερης τάξης (σε αντιδιαστολή με ότι συμβαίνει στην περίπτωση των κλασικών μερικών παραγώγων)! Πράγματι, αν, με $\Omega := \{(x, y) \in \mathbb{R}^2 : \ -1 < x, y < 1\}$, θεωρήσουμε τις συναρτήσεις $u, v : \Omega \to \mathbb{R}$, $u(x, y) := \operatorname{sgn} x$, $v(x, y) := \operatorname{sgn} x - \operatorname{sgn} y$, εύκολα διαπιστώνουμε ότι η γενικευμένη παράγωγος $D^{(1,0)}u$ δεν υπάρχει, ενώ η $D^{(0,1)}u$ υπάρχει και είναι μάλιστα η μηδενική συνάρτηση, καθώς και ότι οι $D^{(1,0)}v$ και $D^{(0,1)}v$ δεν υπάρχουν, ενώ η $D^{(1,1)}v$ υπάρχει και είναι μάλιστα η μηδενική συνάρτηση, βλ. την Άσκηση 4.1.

Τονίζουμε επίσης ότι στον Ορισμό 4.4 οι συναρτήσεις δοκιμής θα μπορούσαν εξ ίσου καλά να ληφθούν ως συναρτήσεις $|\alpha|$ φορές συνεχώς παραγωγίσιμες με συμπαγή φορέα, δηλαδή ως στοιχεία του χώρου $C_0^{|\alpha|}(\Omega)$.

Χώροι του Sobolev. Έστω Ω ένα ανοικτό υποσύνολο του $\mathbb{R}^d$. Έστω $p \in [1, \infty]$ και $k \in \mathbb{N}$.

Ορισμός 4.5 (Χώροι του Sobolev.) *Ο χώρος του Sobolev $W^{k,p} := W^{k,p}(\Omega)$ αποτελείται από τις συναρτήσεις $u \in L^p(\Omega)$ που έχουν ασθενείς μερικές παραγώγους $D^\alpha u$, $\alpha \in \mathbb{N}_0^d$, τάξης έως και k, $|\alpha| \le k$, και όλες αυτές οι μερικές παράγωγοι ανήκουν στον $L^p(\Omega)$. Στον $W^{k,p}$ ορίζουμε τη νόρμα $\|\cdot\|_{k,p}$,*

$$\|u\|_{k,p} := \Big(\sum_{|\alpha| \le k} \|D^\alpha u\|_{L^p}^p \Big)^{1/p}, \quad \text{αν } p < \infty,$$

και, για $p = \infty$,

$$\|u\|_{k,\infty} := \sum_{|\alpha| \le k} \|D^\alpha u\|_{L^\infty}.$$

Ότι η $\|\cdot\|_{k,p}$ είναι όντως νόρμα στον $W^{k,p}$ αποδεικνύεται εύκολα, με την ανισότητα του Minkowski στον $\mathbb{R}^m$, όπου $m := \binom{d+k}{k}$ είναι το πλήθος των πολυδεικτών $\alpha \in \mathbb{N}_0^d$ τάξης το πολύ k, βλ. την Άσκηση 4.2, αν χρησιμοποιήσουμε το γεγονός ότι η $\|\cdot\|_{L^p}$ (όπως γράφουμε τώρα αντί της $\|\cdot\|_p$ για σαφήνεια) είναι νόρμα στον $L^p(\Omega)$, βλ. την Άσκηση 2.9.

Για $p = 2$ γράφουμε συνήθως H^k αντί για $W^{k,2}$. Ο λόγος είναι ότι, όπως θα δούμε στη συνέχεια, αυτοί οι χώροι είναι χώροι Hilbert. Το αντίστοιχο εσωτερικό γινόμενο $(\cdot,\cdot)_k$ δίνεται ως εξής

$$(u,v)_k := \sum_{|\alpha|\le k} (D^a u, D^\alpha v), \quad \forall u,v \in H^k,$$

όπου $(\cdot,\cdot)$ το εσωτερικό γινόμενο στον $L^2(\Omega)$,

$$(u,v) := \int_\Omega uv\,dx.$$

Επίσης, θα συμβολίζουμε για απλότητα με $\|\cdot\|_k$, αντί για $\|\cdot\|_{k,2}$, τη νόρμα του H^k, στη συνέχεια· κίνδυνος να προκληθεί σύγχυση με τη νόρμα του L^p δεν θα υπάρξει, αφού τα συμφραζόμενα δεν θα αφήνουν τέτοια περιθώρια. Σημειώνουμε πάντως ότι ιστορικά ορίστηκαν οι χώροι $W^{k,p}$ και $H^{k,p}$ με διαφορετικούς τρόπους, απεδείχθη όμως εκ των υστέρων ότι αυτοί οι ορισμοί οδηγούσαν στους ίδιους χώρους.

Η πληρότητα των χώρων του Sobolev σε πολλές διαστάσεις αποδεικνύεται με επιχειρήματα εντελώς ανάλογα εκείνων που χρησιμοποιήσαμε στη μία διάσταση. Παραθέτουμε ως εκ τούτου απλώς το αποτέλεσμα.

Θεώρημα 4.7 (Πληρότητα των χώρων του Sobolev.) *Έστω Ω ένα ανοικτό υποσύνολο του $\mathbb{R}^d$. Τότε, οι χώροι του Sobolev $W^{k,p} := W^{k,p}(\Omega)$, $k \in \mathbb{N}$ και $p \in [1,\infty]$, είναι πλήρεις.* □

Παράδειγμα 4.1 (Αναμενόμενη ομαλότητα στοιχείων του $W^{k,p}$.) Έστω $\Omega := B(0;1)$ η μοναδιαία σφαίρα στον $\mathbb{R}^d$ και $\lambda > 0$. Θεωρούμε τη συνάρτηση u, $u(x) := |x|^{-\lambda}$, $x \in \Omega$, $x \neq 0$. (Στο μηδέν θα μπορούσαμε να δώσουμε στη u οποιαδήποτε τιμή, αυτό όμως δεν θα έπαιζε ρόλο στη συζήτηση που θα ακολουθήσει.) Η u δεν είναι φραγμένη στο Ω και έχει ουσιαστική ασυνέχεια στο μηδέν, δηλαδή ασυνέχεια που δεν μπορεί να αρθεί δίνοντας στη u μια κατάλληλη τιμή (ή τροποποιώντας τις τιμές της u σε πεπερασμένου πλήθους σημεία). Θέλουμε να δούμε για ποιες τιμές των λ και p ανήκει η u στους χώρους $W^{1,p}(\Omega)$.

Ας αρχίσουμε με τους χώρους $L^p(\Omega)$. Για $\mu > 0$, χρησιμοποιώντας πολικές συντεταγμένες, έχουμε

$$\begin{aligned}\int_\Omega \frac{1}{|x|^\mu}\,dx &= \int_0^1 \Big(\int_{\partial B(0;r)} \frac{1}{r^\mu}\,dS\Big)\,dr\\ &= \omega_d \int_0^1 \frac{1}{r^\mu}\,r^{d-1}\,dr = \omega_d \int_0^1 \frac{1}{r^{\mu-d+1}}\,dr,\end{aligned}$$

όπου ω_d το εμβαδόν της μοναδιαίας σφαίρας $\partial B(0;1)$ στον $\mathbb{R}^d$. Το τελευταίο ολοκλήρωμα υπάρχει, βέβαια, αν και μόνο αν $\mu - d + 1 < 1$, δηλαδή $\mu < d$. Επομένως ισχύει

$$u \in L^p(\Omega) \iff \lambda < \frac{d}{p}. \tag{4.9}$$

Για $x \neq 0$ η u είναι (άπειρες φορές) συνεχώς παραγωγίσιμη, και οι πρώτες μερικές παράγωγοί της u_{x_i} δίνονται ως

$$u_{x_i}(x) = -\frac{\lambda x_i}{|x|^{\lambda+2}}, \quad i = 1, \dots, d,$$

όπως διαπιστώνει κανείς αμέσως. Επομένως, η Ευκλείδεια νόρμα του ανάδελτα $\nabla u = (u_{x_1}, \dots, u_{x_d})$ της u δίνεται ως

$$|\nabla u(x)| = \frac{|\lambda|}{|x|^{\lambda+1}}.$$

Σύμφωνα με την (4.9) έχουμε, συνεπώς,

$$|\nabla u(\cdot)| \in L^p(\Omega) \iff \lambda < \frac{d-p}{p}. \tag{4.10}$$

Υποθέτουμε λοιπόν στη συνέχεια ότι $\lambda < d-1$ και θα αποδείξουμε ότι οι u_{x_i} είναι τότε όντως οι ασθενείς μερικές παράγωγοι της u. Έστω λοιπόν $\varphi \in C_0^\infty(\Omega)$ μια συνάρτηση δοκιμής. Τότε, για $0 < \varepsilon < 1$, έχουμε (αφού ολοκληρώνουμε εκεί όπου η u είναι ομαλή)

$$\int_{\Omega \setminus B(0;\varepsilon)} u\varphi_{x_i}\, dx = -\int_{\Omega \setminus B(0;\varepsilon)} u_{x_i}\varphi\, dx + \int_{\partial B(0;\varepsilon)} u\varphi n_i\, dS, \tag{4.11}$$

για $i = 1, \dots, d$, όπου $n = (n_1, \dots, n_d)$ το μοναδιαίο κανονικό διάνυσμα στην $\partial B(0;\varepsilon)$ με κατεύθυνση προς το εσωτερικό της $B(0;\varepsilon)$, σύμφωνα με το θεώρημα των Gauss–Green. (Γιατί σε αυτή τη σχέση δεν εμφανίζεται το ολοκλήρωμα στο σύνορο της μοναδιαίας σφαίρας;) Τώρα,

$$\Big| \int_{\partial B(0;\varepsilon)} u\varphi n_i\, dS \Big| \le \|\varphi\|_{L^\infty(\Omega)} \int_{\partial B(0;\varepsilon)} \varepsilon^{-\lambda}\, dS \le C\varepsilon^{d-1-\lambda} \to 0, \text{ για } \varepsilon \to 0.$$

Επομένως, αφήνοντας στην (4.11) το ε να τείνει στο μηδέν, λαμβάνουμε

$$\int_\Omega u\varphi_{x_i}\, dx = -\int_\Omega u_{x_i}\varphi\, dx \quad \forall \varphi \in C_0^\infty(\Omega), \tag{4.12}$$

αν $0 \le \lambda < d-1$. Τα ανωτέρω μας οδηγούν αμέσως στο συμπέρασμα ότι

$$u \in W^{1,p}(\Omega) \iff \lambda < \frac{d-p}{p}. \tag{4.13}$$

Σημειώστε τη διαφορά στην προκειμένη περίπτωση, όπου, για $d \ge 2$, τα στοιχεία του $W^{1,p}(\Omega)$, για ένα χωρίο Ω με πολύ ομαλό σύνορο, μπορεί να είναι ασυνεχή, κατ' αντιδιαστολή με την περίπτωση ενός διαστήματος!

Για να μην αφήσουμε τον αναγνώστη με την εντύπωση ότι ουσιαστική ασυνέχεια μπορεί να εμφανισθεί μόνο σε ένα σημείο (ή σε πεπερασμένου πλήθους σημεία), γενικεύοντας λίγο τα προηγούμενα, θεωρούμε ένα αριθμήσιμο σύνολο $\{r_1, r_2, \dots\}$, πυκνό στη μοναδιαία σφαίρα $\Omega = B(0;1)$, και τη συνάρτηση

$$u(x) := \sum_{\ell=1}^{\infty} \frac{1}{2^\ell} |x - r_\ell|^{-\lambda}, \quad x \in \Omega,\, x \neq 0.$$

Όπως προηγουμένως, διαπιστώνουμε εύκολα ότι $u \in W^{1,p}(\Omega)$, αν $\lambda < (d-p)/p$. Σημειώστε ότι, για $0 < \lambda < (d-p)/p$, η u ανήκει στον $W^{1,p}(\Omega)$ και δεν είναι φραγμένη σε κανένα ανοικτό υποσύνολο του Ω! $\square$

Όπως γνωρίζουμε ήδη, για να συμπεράνουμε ότι τα στοιχεία χώρων του Sobolev σε ένα χωρίο έχουν κάποια ομαλότητα, είναι φερ' ειπείν συνεχή, απαιτείται ομαλότητα του συνόρου $\partial\Omega$ του χωρίου, βλ. το σχόλιο που προηγείται του Θεωρήματος 2.11. Υπάρχουν πολλών ειδών ομαλότητες του $\partial\Omega$, π.χ. αυτό να ικανοποιεί τη συνθήκη του ευθυγράμμου τμήματος, ή τη συνθήκη του κώνου, ή τη συνθήκη του Lipschitz, ή να είναι της κλάσης $\mathcal{C}^m$, $m \in \mathbb{N}$, βλ. [1] σελίδες 81–84. Για ευκολία, εμείς εδώ θα περιοριστούμε στην περίπτωση που το Ω είναι φραγμένο και θα δώσουμε τους ορισμούς μόνο για τις δύο τελευταίες συνθήκες.

Ορισμός 4.6 (Ομαλότητα συνόρου χωρίων.) Έστω $\Omega \subset \mathbb{R}^d$ ένα φραγμένο χωρίο. Υποθέτουμε ότι για κάθε $x^\star \in \partial\Omega$ υπάρχει $r > 0$ και συνάρτηση $\gamma : \mathbb{R}^{d-1} \to \mathbb{R}$, τέτοια ώστε (με κατάλληλη κάθε φορά επιλογή του συστήματος συντεταγμένων) να ισχύει

$$\Omega \cap B(x^\star, r) = \{x \in B(x^\star, r) : x_d > \gamma(x_1, \dots, x_{d-1})\}.$$

Αν όλες οι συναρτήσεις γ μπορούν να επιλεγούν ώστε να ικανοποιούν τη συνθήκη του Lipschitz, λέμε ότι το σύνορο $\partial\Omega$ *ικανοποιεί τη συνθήκη του Lipschitz*. Αν $m \in \mathbb{N}$ και όλες οι συναρτήσεις γ μπορούν να επιλεγούν ώστε να είναι m φορές συνεχώς παραγωγίσιμες, λέμε ότι το σύνορο $\partial\Omega$ *ανήκει στην κλάση* $\mathcal{C}^m$, ή απλώς ότι το σύνορο είναι $\mathcal{C}^m$.

Θεώρημα 4.8 (Ομαλότητα στοιχείων του $W^{k,p}$.) *Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^d$ τέτοιο ώστε το σύνορό του $\partial\Omega$ να ικανοποιεί τη συνθήκη του Lipschitz. Έστω $j, k \in \mathbb{N}$ και $1 \le p < \infty$.*

i. *Αν είτε $kp > d$, είτε $k = d$ και $p = 1$, τότε κάθε συνάρτηση $u \in W^{j+k,p}(\Omega)$ ανήκει και στον χώρο $C^j(\bar{\Omega})$. Μάλιστα, στην περίπτωση $kp > d > (k-1)p$, οι μερικές παράγωγοι τάξης j της u ικανοποιούν τη συνθήκη του Hölder με εκθέτη οποιονδήποτε αριθμό $\lambda \in (0, k - \frac{d}{p}]$, δηλαδή ανήκουν στον χώρο $C^{j,\lambda}(\bar{\Omega})$. Στην περίπτωση $d = (k-1)p$, έχουμε $u \in C^{j,\lambda}(\bar{\Omega})$, για οποιοδήποτε $\lambda \in (0,1)$. Τέλος, για $k - 1 = d$ και $p = 1$, ισχύει $u \in C^{j,\lambda}(\bar{\Omega})$, για οποιοδήποτε $\lambda \in (0,1]$.*

ii. *Αν είτε $kp > d$, είτε $k = d$ και $p = 1$, τότε κάθε συνάρτηση $u \in W^{j+k,p}(\Omega)$ ανήκει και στον χώρο $W^{j,q}(\Omega)$, για κάθε $q \in [p, \infty]$.*

iii. *Αν $kp = d$, τότε κάθε συνάρτηση $u \in W^{j+k,p}(\Omega)$ ανήκει και στον χώρο $W^{j,q}(\Omega)$, για κάθε $q \in [p, \infty)$.*

iv. *Αν είτε $kp < d$ και, είτε $d - kp < d$, είτε $p = 1$ και $d - k \leq d$, τότε κάθε συνάρτηση $u \in W^{j+k,p}(\Omega)$ ανήκει και στον χώρο $W^{j,q}(\Omega)$, για κάθε $q \in [p, dp/(d-kp)]$.* □

Στη συνέχεια, περιοριζόμαστε στην περίπτωση του χώρου $H^1(\Omega)$, επειδή αυτός κυρίως μας ενδιαφέρει για τα προβλήματα συνοριακών τιμών με τα οποία θα ασχοληθούμε.

Από τη σχέση (4.13) έπεται εύκολα ότι το γινόμενο δύο στοιχείων του $H^1(\Omega)$ δεν είναι, στις πολλές διαστάσεις, γενικά στοιχείο του $H^1(\Omega)$, σε αντιδιαστολή με τη μία διάσταση, βλ. την Πρόταση 2.5. Αν όμως υποθέσουμε, επί πλέον, ότι δύο στοιχεία του $H^1(\Omega)$ είναι και στοιχεία του $L^\infty(\Omega)$, τότε και το γινόμενό τους είναι στοιχείο του $H^1(\Omega)$ και, προφανώς και του $L^\infty(\Omega)$.

Πρόταση 4.4 (Γινόμενο στοιχείων του $H^1(\Omega)$.) *Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^d$ και $u, v \in H^1(\Omega) \cap L^\infty(\Omega)$. Τότε $uv \in H^1(\Omega) \cap L^\infty(\Omega)$ και οι ασθενείς παράγωγοι $(uv)_{x_i}$ δίνονται ως εξής*

$$(uv)_{x_i} = u_{x_i} v + u v_{x_i}, \quad i = 1, \dots, d. \qquad \square$$

Θεώρημα 4.9 (Πυκνότητα ομαλών συναρτήσεων στον $H^1(\Omega)$.) *Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^d$ με σύνορο $\partial\Omega$ της κλάσης $\mathcal{C}^1$. Τότε, για κάθε $u \in H^1(\Omega)$, υπάρχει μια ακολουθία $(u_n)_{n\in\mathbb{N}} \subset C_0^\infty(\mathbb{R}^d)$ τέτοια ώστε οι περιορισμοί των u_n στο Ω να συγκλίνουν στη u στον $H^1(\Omega)$, καθώς το n τείνει στο άπειρο. Με άλλα λόγια, οι περιορισμοί στο Ω των στοιχείων του $C_0^\infty(\mathbb{R}^d)$ σχηματίζουν έναν πυκνό υπόχωρο του $H^1(\Omega)$.* □

Στη μελέτη προβλημάτων συνοριακών τιμών θα μας είναι χρήσιμες γνώσεις για τη συμπεριφορά στοιχείων του $H^1(\Omega)$ στο σύνορο $\partial\Omega$ του Ω. Υποθέτουμε ότι το σύνορο $\partial\Omega$ του Ω είναι της κλάσης $\mathcal{C}^1$, παρά το ότι τα αποτελέσματα που ακολουθούν ισχύουν και υπό ασθενέστερες συνθήκες. Τότε είναι δυνατόν να ορίσουμε το μέτρο Lebesgue στο $\partial\Omega$, συνεπώς και τον χώρο $L^2(\partial\Omega)$, αποτελούμενον από συναρτήσεις ορισμένες στο $\partial\Omega$ τέτοιες ώστε η νόρμα $\|\cdot\|_{L^2(\partial\Omega)}$,

$$\|v\|_{L^2(\partial\Omega)} := \Big(\int_{\partial\Omega} |v(y)|^2\,dy\Big)^{1/2},$$

να είναι πεπερασμένη.

Αποδεικνύεται ότι η νόρμα του $L^2(\partial\Omega)$ του περιορισμού μιας ομαλής συνάρτησης v σε ένα φραγμένο χωρίο Ω με ομαλό σύνορο, φράσσεται από την H^1 νόρμα της v στο Ω, επί μια σταθερά, ανεξάρτητη της v :

Λήμμα 4.1 (Εκτίμηση της νόρμας του $L^2(\partial\Omega)$ από τη νόρμα του $H^1(\Omega)$.) *Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^d$ με σύνορο $\partial\Omega$ της κλάσης $\mathcal{C}^1$. Τότε, υπάρχει μια σταθερά C τέτοια ώστε για κάθε συνάρτηση $v \in C^1(\bar{\Omega})$ να ισχύει*

$$\|v\|_{L^2(\partial\Omega)} \le C\|v\|_{H^1(\Omega)}. \tag{4.14}$$

□

Το Λήμμα 4.1, σε συνδυασμό με το Θεώρημα 4.9, μας επιτρέπει να ορίσουμε συνοριακές συνθήκες για στοιχεία του $H^1(\Omega)$, για φραγμένα χωρία του $\mathbb{R}^d$ με σύνορο $\partial\Omega$ της κλάσης $\mathcal{C}^1$. Όντως, έστω $u \in H^1(\Omega)$. Σύμφωνα με το Θεώρημα 4.7, υπάρχει μια ακολουθία $(u_n)_{n\in\mathbb{N}} \subset C^\infty(\bar{\Omega})$ τέτοια ώστε $u_n \to u, n \to \infty$, στον $H^1(\Omega)$. Σύμφωνα με την (4.14), οι περιορισμοί των u_n στο $\partial\Omega$ σχηματίζουν ακολουθία Cauchy στον $L^2(\partial\Omega)$, συνεπώς αυτή συγκλίνει σε μια συνάρτηση v στον $L^2(\partial\Omega)$. Εύκολα διαπιστώνουμε ότι το όριο v είναι ανεξάρτητο της ακολουθίας $(u_n)_{n\in\mathbb{N}}$. Η συνάρτηση v αποτελεί τις “συνοριακές τιμές” στο $\partial\Omega$ της u, ή το *ίχνος* της u στο $\partial\Omega$. Μερικές φορές λέμε ότι η $v = u|_{\partial\Omega}$ αποτελεί τις συνοριακές τιμές της u με την έννοια του ίχνους. Ανακεφαλαιώνοντας έχουμε:

Θεώρημα 4.10 (Θεώρημα του ίχνους.) *Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^d$ με σύνορο $\partial\Omega$ της κλάσης $\mathcal{C}^1$. Τότε, κάθε $u \in H^1(\Omega)$ έχει συνοριακές τιμές στο $\partial\Omega$, τις οποίες συμβολίζουμε επίσης με u. Επί πλέον, υπάρχει μια σταθερά C τέτοια ώστε για κάθε συνάρτηση $u \in H^1(\Omega)$ να ισχύει*

$$\|u\|_{L^2(\partial\Omega)} \le C\|u\|_{H^1(\Omega)}. \tag{4.15}$$

□

Ο τύπος του Green. Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^d$ με σύνορο $\partial\Omega$ της κλάσης $\mathcal{C}^1$. Αποδεικνύεται τότε ότι ισχύει ο *τύπος του Green* (που αναφέρεται επίσης συχνά ως *θεώρημα του Gauss*)

$$\forall u, v \in H^1(\Omega) \quad \int_\Omega u_{x_i} v\,dx = -\int_\Omega u v_{x_i}\,dx + \int_{\partial\Omega} u v n_i\,dy, \quad i = 1, \dots, d, \tag{4.16}$$

με n_i την $i-$στή συνιστώσα του μοναδιαίου κάθετου εξωτερικού διανύσματος n στο $\partial\Omega$. Από την (4.16) προκύπτει αμέσως ότι

$$\forall u \in H^2(\Omega) \ \forall v \in H^1(\Omega) \quad \int_\Omega v\Delta u\,dx = -\int_\Omega \nabla u \cdot \nabla v\,dx + \int_{\partial\Omega} v\frac{\partial u}{\partial n}\,dy \tag{4.17}$$

με $\frac{\partial u}{\partial n}$ την κατά κατεύθυνση παράγωγο της u στην κατεύθυνση του μοναδιαίου κάθετου εξωτερικού διανύσματος n στο $\partial\Omega$, βλ. την Άσκηση 4.3. Σημειώστε ότι οι ασθενείς μερικές παράγωγοι u_{x_i} ανήκουν στον $H^1(\Omega)$, αφού $u \in H^2(\Omega)$, συνεπώς ορίζεται και το ίχνος τους στο σύνορο $\partial\Omega$, οπότε $\frac{\partial u}{\partial n} \in L^2(\partial\Omega)$.

Ο χώρος H_0^1. Ο χώρος H_0^1 είναι πλήρης υπόχωρος του H^1 και θα παίξει πολύ σημαντικό ρόλο στην επόμενη ενότητα, όταν θα μελετήσουμε προβλήματα συνοριακών τιμών και τη διακριτοποίησή τους με τη μέθοδο των πεπερασμένων στοιχείων.

Ορισμός 4.7 (Ο χώρος H_0^1.) Έστω Ω ένα χωρίο του $\mathbb{R}^d$. Ο χώρος $H_0^1 := H_0^1(\Omega)$ αποτελείται από τα στοιχεία του $H^1(\Omega)$ που είναι όρια στη νόρμα $\|\cdot\|_1$ ακολουθιών συναρτήσεων δοκιμής, δηλαδή στοιχείων του $C_0^\infty(\Omega)$.

Εφοδιάζουμε τον H_0^1 με τη νόρμα $\|\cdot\|_1$ του H^1. Από τον ορισμό προκύπτει αμέσως ότι ο $\big(H_0^1, \|\cdot\|_1\big)$ είναι πλήρης. Επί πλέον, αφού η νόρμα $\|\cdot\|_1$ παράγεται από εσωτερικό γινόμενο, ο H_0^1 είναι χώρος Hilbert.

Παρατήρηση 4.1 (Σχέση μεταξύ H^1 και H_0^1.)

i. Όπως και στη μία διάσταση, βλ. την Παρατήρηση 2.5, μπορεί να αποδειχθεί ότι στην περίπτωση που το Ω είναι όλος ο χώρος, $\Omega = \mathbb{R}^d$, ο χώρος $H_0^1(\Omega)$ συμπίπτει με τον $H^1(\Omega)$. Στην περίπτωση $\Omega \neq \mathbb{R}^d$, γενικά, ο $H_0^1(\Omega)$ είναι γνήσιος υπόχωρος του $H^1(\Omega)$.

ii. Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^d$ με σύνορο $\partial\Omega$ της κλάσης $\mathcal{C}^1$. Τότε είναι γνωστό ότι ένα στοιχείο u του $H^1(\Omega)$ ανήκει στον $H_0^1(\Omega)$, αν και μόνο αν το ίχνος της u στο $\partial\Omega$ είναι μηδέν, δηλαδή

$$H_0^1(\Omega) = \{v \in H^1(\Omega) : \ v|_{\partial\Omega} = 0\}. \tag{4.18}$$

Πρόταση 4.5 (Η ανισότητα των Poincaré–Friedrichs.) *Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^d$. Τότε υπάρχει μια σταθερά C, που εξαρτάται μόνο από το Ω, τέτοια ώστε*

$$\|u\| \le C\Big(\sum_{i=1}^{d}\|u_{x_i}\|^2\Big)^{1/2} \quad \forall u \in H_0^1(\Omega). \tag{4.19}$$

Ειδικότερα, η $|\cdot|_1$,

$$|u|_1 := \Big(\sum_{i=1}^{d}\|u_{x_i}\|^2\Big)^{1/2},$$

είναι νόρμα στον $H_0^1(\Omega)$, ισοδύναμη με τη νόρμα $\|\cdot\|_1$. Προφανώς, η $|\cdot|_1$ παράγεται από το εσωτερικό γινόμενο

$$\sum_{i=1}^{d}\int_{\Omega} u_{x_i}v_{x_i}\,dx. \qquad \square$$

Αν το σύνορο του Ω είναι αρκετά ομαλό, αν φερ' ειπείν ανήκει στην κλάση $\mathcal{C}^1$, τότε είναι γνωστό ότι η νόρμα $\|\cdot\|_k$ του $H^k(\Omega)$, $k\in\mathbb{N}$, είναι ισοδύναμη με τη νόρμα

$$\Big(\|u\|^2 + \sum_{|\alpha|=k}\|D^\alpha u\|^2\Big)^{1/2}.$$

Μάλιστα, μπορεί να αποδειχθεί το ακόλουθο *αποτέλεσμα παρεμβολής*: Για κάθε πολυδείκτη $\alpha \in \mathbb{N}_0^d$, τάξης το πολύ k, και κάθε $\varepsilon > 0$, υπάρχει σταθερά $C = C(\alpha,\varepsilon,\Omega)$, τέτοια ώστε

$$\|D^\alpha u\| \le \varepsilon \sum_{|\beta|=k}\|D^\beta u\| + C\|u\| \quad \forall u \in H^k(\Omega).$$

Κατ' αναλογίαν προς τον $H_0^1(\Omega)$, ορίζονται και οι χώροι $H_0^k(\Omega)$ για $k\in\mathbb{N}$. Ο $H_0^k(\Omega)$ αποτελείται από τα στοιχεία του $H^k(\Omega)$ που είναι όρια , στη νόρμα του $H^k(\Omega)$, ακολουθιών συναρτήσεων δοκιμής, δηλαδή στοιχείων του $C_0^\infty(\Omega)$. Στην περίπτωση που το σύνορο $\partial\Omega$ είναι αρκετά ομαλό, φερ' ειπείν της κλάσης $\mathcal{C}^k$, μπορεί να αποδειχθεί ότι μια συνάρτηση $u \in H^k(\Omega)$ ανήκει στον $H_0^k(\Omega)$, αν και μόνο αν $(D^\alpha u)|_{\partial\Omega} = 0$ για κάθε πολυδείκτη $\alpha \in \mathbb{N}_0^d$ τάξης το πολύ $k-1$. Τέλος, ας σημειωθεί η διαφορά μεταξύ των χώρων $H^2 \cap H_0^1$ και H_0^2,

$$\begin{aligned} &H^2(\Omega)\cap H_0^1(\Omega) = \{u \in H^2(\Omega):\ \ u|_{\partial\Omega} = 0\}\\ &H_0^2(\Omega) = \{u \in H^2(\Omega):\ \ u|_{\partial\Omega} = 0,\ \ u_{x_i}|_{\partial\Omega} = 0,\ \ i = 1,\dots,d\}.\end{aligned}$$

4.2 Πεπερασμένα στοιχεία για ελλειπτικές εξισώσεις

Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^d$ με σύνορο $\partial\Omega$ της κλάσης $\mathcal{C}^1$. Θεωρούμε το ακόλουθο πρόβλημα συνοριακών συνθηκών: Ζητείται μια συνάρτηση $u \in C^2(\bar{\Omega})$ τέτοια ώστε

$$\begin{cases} -\Delta u + u = f & \text{στο } \Omega, \\ u = 0 & \text{στο } \partial\Omega, \end{cases} \tag{4.20}$$

όπου $f : \bar{\Omega} \to \mathbb{R}$ μια δεδομένη συνάρτηση. Στο πρόβλημα αυτό έχουμε μία ομογενή συνοριακή συνθήκη Dirichlet. Κάθε λύση u του προβλήματος (4.20) με την ομαλότητα που αναφέρθηκε προηγουμένως, λέγεται *κλασική* λύση του.

Πολλαπλασιάζοντας τη διαφορική εξίσωση στο πρόβλημα (4.20) επί μια ομαλή συνάρτηση v ολοκληρώνοντας στο Ω και εφαρμόζοντας στον πρώτο όρο τον τύπο του Green, καταλήγουμε στη σχέση $(u, v)_1 = (f, v)$. Αυτή η μεταβολική διατύπωση απετελεί το κίνητρο για τον ακόλουθο ορισμό: Μια συνάρτηση $u \in H_0^1(\Omega)$ τέτοια ώστε

$$(u, v)_1 = (f, v) \quad \forall v \in H_0^1(\Omega) \tag{4.21}$$

λέγεται *ασθενής* λύση του προβλήματος (4.20). Υπενθυμίζουμε ότι τα εσωτερικά γινόμενα $(\cdot,\cdot)_1$ και $(\cdot,\cdot)$ του H^1 και του L^2, αντίστοιχα, ορίζονται ως

$$(v, w)_1 = (v, w) + \sum_{i=1}^{d} (v_{x_i}, w_{x_i}), \quad (v, w) = \int_\Omega vw \, dx.$$

(Συνήθως γράφουμε $(\nabla v, \nabla w)$ αντί του αθροίσματος από $i = 1$ έως d στον ανωτέρω τύπο.) Για $f \in L^2(\Omega)$, από την ανισότητα των Cauchy–Schwarz έπεται αμέσως ότι $|(f, v)| \le \|\varphi\| \, \|v\|_1$. Η ύπαρξη και η μοναδικότητα της ασθενούς λύσης του (4.20) καθώς και η εκτίμηση

$$\|u\|_1 \le \|f\| \tag{4.22}$$

έπονται αμέσως από το θεώρημα αναπαράστασης του Riesz. Λόγω της συμμετρίας της διγραμμικής μορφής, η λύση u του προβλήματος (4.21) ελαχιστοποιεί στον $H_0^1(\Omega)$ ένα συγκεκριμένο συναρτησιακό, βλ. την Άσκηση 4.4· αυτό το γεγονός στη συγκεκριμένη περίπτωση αναφέρεται ως *αρχή του Dirichlet*.

Είναι γνωστό ότι, αν $f \in H^m(\Omega)$ και το σύνορο $\partial\Omega$ του Ω είναι της κλάσης $\mathcal{C}^m$, τότε η ασθενής λύση u του προβλήματος (4.20) είναι στοιχείο του $H^{m+2}(\Omega)$ και ισχύει η εκτίμηση

$$\|u\|_{m+2} \le C\|f\|_m, \tag{4.23}$$

με μια σταθερά C, ανεξάρτητη της f. Η ανισότητα (4.23) αναφέρεται ως *ελλειπτική ομαλότητα* και η απόδειξή της είναι μακροσκελής. Αποτελέσματα αυτής της μορφής αναφέρονται γενικά ως αποτελέσματα *ομαλότητας*.

Είναι εύκολο να δούμε ότι κάθε κλασική λύση του (4.20) είναι και ασθενής λύση του ιδίου προβλήματος. Επίσης, εύκολα πείθεται κανείς ότι αν η ασθενής λύση είναι αρκετά ομαλή ώστε να ανήκει στον χώρο $C^2(\bar{\Omega})$, τότε αυτή αποτελεί και κλασική λύση, βλ. την Άσκηση 4.5. Το γεγονός ότι η ασθενής λύση του (4.20) ανήκει στον $C(\bar{\Omega})$, αν $f \in H^m$ για $m > d/2$ και το σύνορο του Ω είναι αρκετά ομαλό, έπεται από την (4.23) και το Θεώρημα 4.7. Υπάρχει πάντως και καταλληλότερη για την περίσταση θεωρία, οι καλούμενες *εκτιμήσεις Schauder*· σύμφωνα με αυτή τη θεωρία, αν το σύνορο $\partial\Omega$ του Ω ανήκει στην κλάση των χώρων του Hölder $\mathcal{C}^{2,\alpha}$, με $0 < \alpha < 1$, και $f \in C^{0,\alpha}(\bar{\Omega})$, τότε υπάρχει λύση $u \in C^{2,\alpha}(\bar{\Omega})$ του (4.20). Επί πλέον, αν το $\partial\Omega$ ανήκει στην κλάση $\mathcal{C}^{m+1}$, με $m \in \mathbb{N}$, και $f \in C^{m,\alpha}(\bar{\Omega})$, τότε υπάρχει λύση $u \in C^{m+2,\alpha}(\bar{\Omega})$ του (4.20). Οι χώροι του Hölder ορίζονται ως

$$\begin{aligned}
&C^{0,\alpha}(\bar{\Omega}) := \{u \in C(\bar{\Omega}) : \sup_{x,y\in\Omega,\, x\neq y} \frac{|u(x)-u(y)|}{|x-y|^\alpha} < \infty\},\\
&C^{m,\alpha}(\bar{\Omega}) := \{u \in C^m(\bar{\Omega}) : \ D^\beta u \in C^{0,\alpha}(\bar{\Omega}) \text{ για κάθε πολυδείκτη } \beta \in \mathbb{N}_0^d \text{ τάξης } m\}.
\end{aligned}$$

Με τη βοήθεια αυτών των χώρων ορίζονται οι κλάσεις $\mathcal{C}^{m,\alpha}$ κατά προφανή τρόπο.

Η προηγηθείσα συζήτηση γενικεύεται εύκολα στην περίπτωση γενικότερων αυτοσυζυγών γραμμικών ελλειπτικών τελεστών με μεταβλητούς συντελεστές. Έστω $a_{ij} \in C^1(\bar{\Omega})$, $i,j = 1,\dots,d$, συναρτήσεις τέτοιες ώστε $a_{ij} = a_{ji}$, $i,j = 1,\dots,d$, και ο συμμετρικός πίνακας $A(x)$ με στοιχεία a_{ij}, $i,j = 1,\dots,d$, να είναι *ομοιόμορφα* θετικά ορισμένος, δηλαδή να υπάρχει μια θετική σταθερά α τέτοια ώστε

$$\forall x \in \bar{\Omega} \ \ \forall \xi \in \mathbb{R}^d \quad \sum_{i,j=1}^d a_{ij}(x)\xi_i\xi_j \ge \alpha \sum_{i=1}^d \xi_i^2. \tag{4.24}$$

Έστω ακόμα $a_0 \in C(\bar{\Omega})$ μια μη αρνητική συνάρτηση, $a_0(x) \ge 0$, για κάθε $x \in \bar{\Omega}$.

Θεωρούμε τότε το ακόλουθο πρόβλημα συνοριακών συνθηκών: Ζητείται μια συνάρτηση $u \in C^2(\bar{\Omega})$ τέτοια ώστε

$$\begin{cases} -\displaystyle\sum_{i,j=1}^{d} \big(a_{ij}u_{x_i}\big)_{x_j} + a_0 u = f & \text{στο } \Omega, \\ u = 0 & \text{στο } \partial\Omega, \end{cases} \tag{4.25}$$

όπου $f : \bar{\Omega} \to \mathbb{R}$ μια δεδομένη συνάρτηση. Κάθε λύση u του προβλήματος (4.25), με την ομαλότητα που αναφέρθηκε προηγουμένως, λέγεται *κλασική* λύση του.

Η συνθήκη (4.24) καθιστά τη διαφορική εξίσωση στο πρόβλημα (4.25) *ελλειπτική*.

Μια συνάρτηση $u \in H_0^1(\Omega)$ τέτοια ώστε

$$a(u, v) = (f, v) \quad \forall v \in H_0^1(\Omega), \tag{4.26}$$

με τη διγραμμική μορφή $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$,

$$a(v, w) = \int_\Omega \Big(\sum_{i,j=1}^{d} a_{ij} v_{x_i} w_{x_j} + a_0 v w \Big) dx \quad \forall v, w \in H_0^1(\Omega), \tag{4.27}$$

λέγεται *ασθενής* λύση του προβλήματος (4.25).

Υποθέτουμε πάλι ότι Ω είναι ένα φραγμένο χωρίο του $\mathbb{R}^d$ με σύνορο $\partial\Omega$ της κλάσης $\mathcal{C}^1$ και θεωρούμε το ακόλουθο πρόβλημα συνοριακών συνθηκών, με ομογενείς συνοριακές *συνθήκες Neumann* αυτή τη φορά: Ζητείται μια συνάρτηση $u \in C^2(\bar{\Omega})$ τέτοια ώστε

$$\begin{cases} -\Delta u + u = f & \text{στο } \Omega, \\ \dfrac{\partial u}{\partial n} = 0 & \text{στο } \partial\Omega, \end{cases} \tag{4.28}$$

με $f : \bar{\Omega} \to \mathbb{R}$ μια δεδομένη συνάρτηση. Ως συνήθως συμβολίσαμε με $\frac{\partial u}{\partial n}$ το εσωτερικό γινόμενο του ανάδελτα της u με το μοναδιαίο κάθετο εξωτερικό διάνυσμα n στο $\partial\Omega$, $\nabla u \cdot n$, δηλαδή την κατά κατεύθυνση παράγωγο της u στην κατεύθυνση του κάθετου εξωτερικού διανύσματος. Κάθε λύση u του προβλήματος (4.28) με την ομαλότητα που αναφέρθηκε προηγουμένως, λέγεται *κλασική* λύση του. Μια συνάρτηση $u \in H^1(\Omega)$ τέτοια ώστε

$$(u, v)_1 = (f, v) \quad \forall v \in H^1(\Omega) \tag{4.29}$$

λέγεται *ασθενής* λύση του προβλήματος (4.28). Για $f \in L^2(\Omega)$, από την ανισότητα των Cauchy–Schwarz έπεται αμέσως ότι $|(f,v)| \le \|\varphi\| \, \|v\|_1$. Η ύπαρξη και η μοναδικότητα της ασθενούς λύσης του (4.28) καθώς και η εκτίμηση

$$\|u\|_1 \le \|f\| \tag{4.30}$$

έπονται αμέσως από το θεώρημα αναπαράστασης του Riesz. Ακριβώς τα ίδια αποτελέσματα, όπως στην περίπτωση συνοριακών συνθηκών Dirichlet, σχετικά με την ελλειπτική ομαλότητα ισχύουν και στην προκειμένη περίπτωση. Επίσης κάθε κλασική λύση είναι και ασθενής λύση, και κάθε αρκετά ομαλή ασθενής λύση είναι και κλασική λύση.

4.2.1 Διακριτοποίηση με πεπερασμένα στοιχεία

Θεωρούμε το πρόβλημα (4.25) και υποθέτουμε ότι ο τελεστής είναι ελλειπτικός, δηλαδή ότι ικανοποιείται η συνθήκη (4.24), και ότι η συνάρτηση a_0 λαμβάνει μόνο μη αρνητικές τιμές. Η διγραμμική μορφή a, που δίνεται στην (4.27), είναι στον $H_0^1(\Omega)$ τόσο φραγμένη, δηλαδή

$$\exists C \in \mathbb{R} \ \ \forall v, w \in H_0^1 \quad |a(v,w)| \le C\|v\|_1 \, \|w\|_1, \tag{4.31}$$

όσο και ελλειπτική, δηλαδή

$$\exists c > 0 \ \ \forall v \in H_0^1 \quad a(v,v) \ge c\|v\|_1^2, \tag{4.32}$$

βλ. την Άσκηση 4.6. Υποθέτουμε ότι τα δεδομένα του προβλήματος (4.25) είναι τέτοια ώστε η μοναδική ασθενής λύση του $u \in H_0^1$, βλ. την (4.26), να ανήκει στον χώρο $H^2(\Omega) \cap H_0^1(\Omega)$ και να ισχύει η εκτίμηση ελλειπτικής ομαλότητας

$$\|u\|_2 \le \tilde{C}\|f\|, \tag{4.33}$$

με μια σταθερά $\tilde{C}$ ανεξάρτητη της f.

Έστω τώρα S_h^r ένας υπόχωρος του $H_0^1(\Omega)$, πεπερασμένης διάστασης. Η προσέγγιση πεπερασμένων στοιχείων (ή προσέγγιση Galerkin) $u_h \in S_h^r$ της λύσης u ορίζεται τώρα από τις σχέσεις

$$a(u_h, \chi) = (f, \chi) \quad \forall \chi \in S_h^r, \tag{4.34}$$

παράβαλε με τη μεταβολική διατύπωση (4.26). Αυτή η διατύπωση είναι η μέθοδος των πεπερασμένων στοιχείων (ή μέθοδος του Galerkin) για το πρόβλημα (4.25). Υπό τις υποθέσεις μας, το λήμμα των Lax–Milgram (ή, στην προκειμένη περίπτωση που η διγραμμική μορφή είναι συμμετρική, και το θεώρημα αναπαράστασης του Riesz) συνεπάγεται ύπαρξη και μοναδικότητα της u_h καθώς και την εκτίμηση

$$\|u-u_h\|_1 \le C \inf_{\chi\in S_h^r} \|u-\chi\|_1, \tag{4.35}$$

με μια κατάλληλη σταθερά C, ανεξάρτητη της f και του S_h^r.

Υποθέτουμε τώρα ότι μας δίνεται μια οικογένεια $\left(S_h^r\right)_{0<h<1}$ υποχώρων του $H_0^1(\Omega)$ με $r\ge 2$ με την *προσεγγιστική ιδιότητα*

$$\exists c>0 \;\; \forall v\in H^s\cap H_0^1 \;\; \exists\chi\in S_h^r \quad \|v-\chi\|+h\|v-\chi\|_1 \le Ch^s\|v\|_s, \quad s=2,r. \tag{4.36}$$

Η ανάλυση είναι εντελώς ανάλογη με την αντίστοιχη για το πρόβλημα δύο σημείων, ως εκ τούτου την παρουσιάζουμε εντελώς συνοπτικά. Κατ' αρχάς η (4.35) συνδυαζόμενη με την (4.36) δίνει αμέσως τη *βέλτιστης τάξης* εκτίμηση στην H^1-νόρμα

$$\|u-u_h\|_1 \le Ch^{r-1}\|v\|_r. \tag{4.37}$$

Για να οδηγηθούμε στην αντίστοιχη εκτίμηση στην L^2-νόρμα, θα καταφύγουμε στο τέχνασμα του Nitsche. Στην προκειμένη περίπτωση που η διγραμμική μορφή είναι συμμετρική (αφού ο τελεστής L είναι αυτοσυζυγής), θεωρούμε το πρόβλημα

$$a(w,v)=(e,v) \quad \forall v\in H_0^1, \tag{4.38}$$

με e το σφάλμα, $e:=u-u_h$. Επιλέγοντας $v=e$ στην (4.38), λαμβάνουμε

$$\|e\|^2 = a(w,e) = a(e,w).$$

Όμως, αφαιρώντας κατά μέλη την (4.34) από την (4.26) με $v=\chi$, έχουμε $a(e,\chi)=0$, για κάθε $\chi\in S_h^r$. Συνεπώς, η ανωτέρω σχέση δίνει, για κάθε $\chi\in S_h^r$,

$$\|e\|^2 = a(e,w-\chi).$$

Εκτιμούμε το δεξιό μέλος χρησιμοποιώντας την (4.31) και λαμβάνουμε

$$\|e\|^2 \le C\|e\|_1\,\|w-\chi\|_1,$$

οπότε, με κατάλληλη επιλογή του χ, βλ. την (4.36),

$$\|e\|^2 \le C\|e\|_1\, h\|w\|_2.$$

Όμως, σύμφωνα με την (4.33), $\|w\|_2 \le \tilde{C}\|e\|$, οπότε

$$\|e\|^2 \le C\|e\|_1\, h\|e\|$$

ή

$$\|e\| \le Ch\|e\|_1. \tag{4.39}$$

Συνδυάζοντας αυτή την εκτίμηση με την (4.37), οδηγούμαστε στην επιθυμητή *βέλτιστης τάξης* εκτίμηση στην L^2-νόρμα

$$\|u-u_h\| \le Ch^r\|u\|_r. \tag{4.40}$$

4.3 Συνεχείς κατά τμήματα γραμμικές (ομοπαραλληλικές) συναρτήσεις σε τριγωνισμούς πολυγώνων

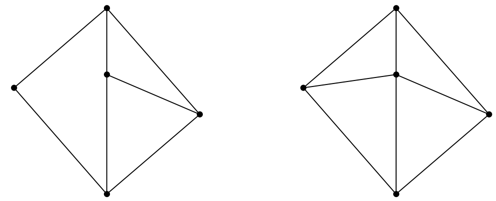
Έστω Ω ένα κυρτό πολύγωνο. Θεωρούμε έναν διαμερισμό του Ω σε τρίγωνα $\tau_i, i = 1,\dots,M$, που σχηματίζουν έναν *τριγωνισμό* του Ω, έννοια που εξηγείται στη συνέχεια. Υποθέτουμε ότι τα τ_i είναι ξένα μεταξύ τους και ότι είναι τέτοια ώστε το εσωτερικό της ένωσης των θηκών τους να δίνει το Ω,

$$\Omega = \Big(\bigcup_{i=1}^{M} \bar{\tau}_i\Big)^\circ.$$

Συμβολίζουμε με h τη μέγιστη διάμετρο, δηλαδή τη μέγιστη πλευρά, των τριγώνων τ_i,

$$h := \max_{1\le i\le M}(\operatorname{diam}\tau_i),$$

και με $\mathcal{T}_h = \{\tau_1,\dots,\tau_M\}$ τον τριγωνισμό. (Προφανώς, το h δεν προσδιορίζει πλήρως τον τριγωνισμό, στο ίδιο h μπορούν να αντιστοιχούν πολλοί τριγωνισμοί. Απλώς το h δείχνει ότι όλα τα τρίγωνα του τριγωνισμού $\mathcal{T}_h$ έχουν πλευρές όχι μεγαλύτερες του h. Όπως θα δούμε αργότερα, οι εκτιμήσεις μας εξαρτώνται κατά ουσιαστικό τρόπο από το h.) Οι κορυφές των τριγώνων τ_i καλούνται *κόμβοι* του τριγωνισμού. Μια σημαντική για τη συνέχεια υπόθεση για τον τριγωνισμό $\mathcal{T}_h$ είναι ότι *δεν* υπάρχουν



Σχήμα 4.2: Μη αποδεκτός και αποδεκτός, αριστερά και δεξιά, αντίστοιχα, τριγωνισμός.

κόμβοι στα εσωτερικά των πλευρών των τριγώνων. Φερ' ειπείν, ο τριγωνισμός στο αριστερό μέρος του Σχήματος 4.2 δεν είναι αποδεκτός, ενώ εκείνος στο δεξιό μέρος είναι αποδεκτός.

Θεωρούμε τώρα τον γραμμικό χώρο S_h, που αποτελείται από τις συνεχείς συναρτήσεις στο $\bar{\Omega}$ που μηδενίζονται στο σύνορο $\partial\Omega$ και σε κάθε τρίγωνο τ_i είναι πολυώνυμα βαθμού το πολύ ένα,

$$S_h := \{\chi \in C(\bar{\Omega}) \,:\, \chi(x) = \alpha_i + \beta_i x_1 + \gamma_i x_2, \;\; x = (x_1, x_2) \in \tau_i, \;\; \chi|_{\partial\Omega} = 0\}.$$

Έστω $N = N(h)$ το πλήθος των *εσωτερικών* κόμβων P_i του τριγωνισμού, δηλαδή των κόμβων που *δεν* ανήκουν στο σύνορο $\partial\Omega$. Κάθε στοιχείο χ του S_h ορίζεται μονοσήμαντα από τις τιμές του $\chi(P_j)$, $j = 1, \dots, N$, στους εσωτερικούς κόμβους, αφού στους μη εσωτερικούς κόμβους η χ μηδενίζεται και ένα πολυώνυμο δύο μεταβλητών, βαθμού το πολύ ένα, ορίζεται μονοσήμαντα από τις τιμές του σε τρία μη συνευθειακά σημεία, όπως οι κορυφές κάθε τριγώνου τ_i, αφού από τρία μη συνευθειακά σημεία διέρχεται ακριβώς ένα επίπεδο. Ιδιαίτερα, συμπεραίνουμε ότι η διάσταση του S_h είναι N, $\dim S_h = N$. Οι συναρτήσεις $\varphi_i \in S_h$, $i = 1, \dots, N$, με την ιδιότητα

$$\varphi_i(P_j) = \delta_{ij}, \quad i, j = 1, \dots, N,$$

αποτελούν βάση του S_h, όπως θα δούμε αμέσως μετά. Ο φορέας κάθε συνάρτησης φ_i είναι προφανώς η κλειστή θήκη της ένωσης εκείνων των τριγώνων του τριγωνισμού $\mathcal{T}_h$ που έχουν το P_i ως κορυφή τους. Ας δούμε κατ' αρχάς ότι οι συναρτήσεις φ_i είναι γραμμικώς ανεξάρτητες. Πράγματι, αν $\sum_{i=1}^N c_i \varphi_i = 0$, τότε $\sum_{i=1}^N c_i \varphi_i(P_j) = 0$, δηλαδή $c_j = 0$, $j = 1, \dots, N$. Επί πλέον, οι $\varphi_1, \dots, \varphi_N$ παράγουν τον S_h· πράγματι, κάθε $\chi \in S_h$ μπορεί να γραφεί στη μορφή

$$\chi = \sum_{i=1}^N \chi(P_i)\varphi_i,$$

αφού και τα δύο μέλη αυτής της σχέσης είναι στοιχεία του S_h και οι τιμές τους συμπίπτουν στους εσωτερικούς κόμβους $P_j, j = 1, \dots, N$. Επομένως, $\{\varphi_1, \dots, \varphi_N\}$ είναι μια βάση του S_h. Αργότερα θα δούμε ότι η εν λόγω βάση έχει πολλά πλεονεκτήματα, που αφορούν το υπολογιστικό μέρος των προσεγγίσεων.

Παρατήρηση 4.2 (Ο S_h είναι υπόχωρος του H_0^1.) Θα δούμε τώρα ότι ο S_h είναι όντως υπόχωρος του $H_0^1(\Omega)$. Κατ' αρχάς, προφανώς, ο S_h είναι υπόχωρος του $L^2(\Omega)$ και τα στοιχεία του S_h μηδενίζονται στο $\partial\Omega$. Απομένει, συνεπώς, να αποδείξουμε ότι, αν $\chi \in S_h$, τότε υπάρχουν συναρτήσεις $g_1, g_2 \in L^2(\Omega)$ τέτοιες ώστε

$$\int_\Omega \chi\varphi_{x_i}\,dx = -\int_\Omega g_i\varphi\,dx \quad \forall\varphi \in C_0^\infty(\Omega),\ \ i = 1, 2, \tag{4.41}$$

οπότε, φυσικά, οι g_1, g_2 θα είναι οι γενικευμένες παράγωγοι χ_{x_1} και χ_{x_2}, αντίστοιχα. (Για ευκολία στον συμβολισμό γράφουμε $x = (x_1, x_2)$ για τα στοιχεία του $\mathbb{R}^2$.)

Συμβολίζουμε με $\chi^{(\tau)}$ τον περιορισμό της χ σε κάθε τρίγωνο τ του τριγωνισμού $\mathcal{T}_h$. Ισχυριζόμαστε τώρα ότι οι συναρτήσεις g_i, που ορίζονται σε κάθε τρίγωνο τ του τριγωνισμού $\mathcal{T}_h$ ως

$$g_i := \big(\chi^{(\tau)}\big)_{x_i}, \quad i = 1, 2,\ \ \text{στο}\ \tau, \tag{4.42}$$

ικανοποιούν την (4.41), δηλαδή ότι παραγωγίζοντας ως προς x_i (με την κλασική έννοια) τη χ σε κάθε τρίγωνο του τριγωνισμού οδηγούμαστε στη γενικευμένη παράγωγο χ_{x_i} του χ. Σημειώστε ότι κάθε g_i είναι σταθερή σε κάθε τρίγωνο τ, αφού η χ είναι εκεί πολυώνυμο το πολύ πρώτου βαθμού. Τώρα, για $\varphi \in C_0^\infty(\Omega)$, έχουμε, σύμφωνα με τον τύπο του Green, ο οποίος ισχύει και σε τρίγωνα,

$$\int_\tau \chi^{(\tau)}\varphi_{x_i}\,dx = -\int_\tau g_i\varphi\,dx + \int_{\partial\tau} \chi^{(\tau)}\varphi n_i^{(\tau)}\,dx,$$

με $n^{(\tau)} = \big(n_1^{(\tau)}, n_2^{(\tau)}\big)$ το μοναδιαίο κάθετο εξωτερικό διάνυσμα στο σύνορο $\partial\tau$ του τ. Αθροίζοντας αυτές τις σχέσεις ως προς όλα τα τρίγωνα του τριγωνισμού $\mathcal{T}_h$, λαμβάνουμε

$$\sum_{\tau\in\mathcal{T}_h}\int_\tau \chi^{(\tau)}\varphi_{x_i}\,dx = -\sum_{\tau\in\mathcal{T}_h}\int_\tau g_i\varphi\,dx + \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau} \chi^{(\tau)}\varphi n_i^{(\tau)}\,dy,$$

δηλαδή

$$\int_\Omega \chi\varphi_{x_i}\,dx = -\int_\Omega g_i\varphi\,dx + \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau} \chi^{(\tau)}\varphi n_i^{(\tau)}\,dy \quad \forall\varphi \in C_0^\infty(\Omega). \tag{4.43}$$

Θα αποδείξουμε τώρα ότι το άθροισμα στην (4.43) μηδενίζεται. Αν υποθέσουμε, προς το παρόν, ότι αυτό έχει αποδειχθεί, διαπιστώνουμε αμέσως ότι η (4.43) συνεπάγεται την (4.41). Τώρα, οι όροι του αθροίσματος στην (4.43) διασπώνται σε ολοκληρώματα της μορφής

$$\int_S \chi^{(\tau)} \varphi n_i^{(\tau)} \, dy, \tag{4.44}$$

όπου S είναι πλευρά ενός τριγώνου τ του τριγωνισμού $\mathcal{T}_h$. Διακρίνουμε τώρα δύο περιπτώσεις: η S να είναι μέρος του συνόρου ή να είναι εσωτερική πλευρά. Στην πρώτη περίπτωση, ο αντίστοιχος όρος (4.44) εξαφανίζεται, αφού η φ ως συνάρτηση δοκιμής μηδενίζεται στο $\partial\Omega$. Στη δεύτερη περίπτωση, η S είναι πλευρά δύο παρακείμενων τριγώνων τ_ℓ και τ_m του τριγωνισμού $\mathcal{T}_h$. Έτσι, στο άθροισμα στην (4.43) εμφανίζονται ακριβώς δύο ολοκληρώματα πάνω στη συγκεκριμένη πλευρά,

$$\int_S \chi^{(\tau_\ell)} \varphi n_i^{(\tau_\ell)} \, dy \quad \text{και} \quad \int_S \chi^{(\tau_m)} \varphi n_i^{(\tau_m)} \, dy. \tag{4.45}$$

Όμως, οι περιορισμοί στο S των $\chi^{(\tau_\ell)}$ και $\chi^{(\tau_m)}$ συμπίπτουν, αφού η χ είναι συνεχής στο $\bar{\Omega}$. Περαιτέρω, $n_i^{(\tau_\ell)} = -n_i^{(\tau_m)}$, συνεπώς τα δύο ολοκληρώματα στην (4.45) είναι αντίθετοι αριθμοί, οπότε οι συνεισφορές των αντίστοιχων δύο όρων στο άθροισμα στην (4.43) αλληλοαναιρούνται. Επομένως, το άθροισμα στην (4.43) μηδενίζεται πράγματι, άρα ισχύει η (4.41) με τις $g_i \in L^2(\Omega)$ που ορίζονται τμηματικά στην (4.42). □

Σκοπός μας σε αυτή την παράγραφο είναι να αποδείξουμε ότι οι χώροι S_h έχουν, υπό κατάλληλες συνθήκες που θα διατυπώσουμε, την προσεγγιστική ιδιότητα (4.36) με $r = 2$. Τον καθοριστικό ρόλο σε αυτή την απόδειξη τον παίζει το λήμμα των Bramble–Hilbert, στη μελέτη του οποίου αφιερώνεται η ενότητα που ακολουθεί.

4.3.1 Το λήμμα των Bramble–Hilbert

Αντικείμενο μελέτης αυτής της ενότητας αποτελεί το λήμμα των Bramble–Hilbert, που αποτελεί θεμελιώδες εργαλείο στη μελέτη προσεγγιστικών ιδιοτήτων χώρων πεπερασμένων στοιχείων, βλ. και την Πρόταση 3.1. Το αποτέλεσμα ισχύει σε πολύ γενικότερη μορφή, εμείς αρκούμαστε εδώ σε μια απλή και αρκετή για τους στόχους μας.

Πρόταση 4.6 (Το λήμμα των Bramble–Hilbert.) *Συμβολίζουμε με τ ένα τρίγωνο. Έστω h_τ η μεγαλύτερη πλευρά του τ, ρ_τ η ακτίνα του εγγεγραμμένου κύκλου στο τ, και $\sigma_\tau := h_\tau/\rho_\tau$. Τότε υπάρχει μια σταθερά C, ανεξάρτητη του τ, τέτοια ώστε*

$$\|u - Qu\| \le C(1+\sigma_\tau)^2 h_\tau^2 |u|_2 \quad \forall u \in H^2(\tau) \tag{4.46}$$

και

$$|u - Qu|_1 \le C(1+\sigma_\tau)^2 h_\tau |u|_2 \quad \forall u \in H^2(\tau) \tag{4.47}$$

όπου Qu κατάλληλο πολυώνυμο, βαθμού το πολύ ένα. Επί πλέον, έστω $F : H^2(\tau) \to [0,\infty)$ ένα συναρτησιακό τέτοιο ώστε

$$\begin{aligned} &F(v+w) \le F(v) + F(w) && \forall v, w \in H^2,\\ &F(v) \le C\|v\|_2 && \forall v \in H^2,\\ &F(p) = 0 && \forall p \in \mathbb{P}_1, \end{aligned}$$

με μια σταθερά C ανεξάρτητη της v, όπου $H^2 := H^2(\tau)$ και $\mathbb{P}_1$ το σύνολο των πολυωνύμων δύο μεταβλητών στο τ βαθμού το πολύ ένα. Τότε υπάρχει μια σταθερά C_τ, τέτοια ώστε

$$F(v) \le C_\tau |v|_2 \quad \forall v \in H^2. \tag{4.48}$$

Η ημινόρμα $|\cdot|_2$ του H^2 δίνεται ως $|v|_2 := \big(\|v_{x_1x_1}\|^2 + \|v_{x_1x_2}\|^2 + \|v_{x_2x_2}\|^2\big)^{1/2}$.

Απόδειξη. Ας υποθέσουμε προς το παρόν ότι οι (4.46) και (4.47) έχουν ήδη αποδειχθεί. Τότε, λαμβάνοντας υπ' όψιν και το προφανές γεγονός ότι $|v - Qv|_2 = |v|_2$, έχουμε

$$F(v) = F(v - Qv + Qv) \le F(v - Qv) + F(Qv) = F(v - Qv) \le C\|v - Qv\| \le C_\tau |v|_2.$$

Απομένει, συνεπώς, να αποδείξουμε τις (4.46) και (4.47). Έστω B το εσωτερικό του εγγεγραμμένου κύκλου στο τ. Συμβολίζουμε με $x^\star$ το κέντρο του B, δηλαδή το έγκεντρο του τ, και με ρ_τ την ακτίνα του B. Για μια ομαλή συνάρτηση $u : \tau \to \mathbb{R}$ και $y \in \tau$, θεωρούμε το πολυώνυμο Taylor $T_y u$ της u ως προς το y, βαθμού το πολύ ένα,

$$T_y u(x) := (T_y u)(x) = u(y) + u_{x_1}(y)(x_1 - y_1) + u_{x_2}(y)(x_2 - y_2), \quad x \in \tau.$$

Αυτό το πολυώνυμο δεν μπορεί, δυστυχώς, να ορισθεί για συναρτήσεις $u \in H^2(\tau)$, αφού οι τιμές $u_{x_1}(y)$ και $u_{x_2}(y)$ δεν έχουν νόημα για τέτοιες συναρτήσεις. Σκοπός μας λοιπόν, κατ' αρχάς, είναι να τροποποιήσουμε κατάλληλα τον ορισμό, ώστε αυτός να έχει νόημα και για μη ομαλές συναρτήσεις. Θεωρούμε προς τούτο τη συνάρτηση φ,

$$\varphi(x) := \begin{cases} c\,\mathrm{e}^{-\left(1-\left(\frac{2|x-x^\star|}{\rho_\tau}\right)^2\right)^{-1}}, & |x-x^\star| < \frac{1}{2}\rho_\tau, \\ 0, & |x-x^\star| \ge \frac{1}{2}\rho_\tau, \end{cases}$$

με σταθερά c τέτοια ώστε το ολοκλήρωμα της φ (στον B ή στον $\mathbb{R}^2$) να ισούται με τη μονάδα. Προφανώς $\varphi \in C_0^\infty(B)$ και $\operatorname{supp}\varphi = B(x^\star; \rho_\tau/2)$. Φυσικά, η φ θα μπορούσε να είχε επιλεγεί έτσι ώστε ο φορέας της να είχε οποιαδήποτε ακτίνα μικρότερη του ρ_τ. Ορίζουμε τώρα το *σταθμισμένο* ως προς τον κύκλο B πολυώνυμο Taylor $Qu \in \mathbb{P}_1$ της u ως εξής

$$Qu(x) := (Qu)(x) := \int_B T_y u(x)\varphi(y)\,dy, \quad x \in \tau, \tag{4.49}$$

υποθέτοντας ακόμα ότι η συνάρτηση u είναι αρκετά ομαλή. Σημειώστε ότι το Qu στην (4.49) έχει νόημα και για $u \in H^1(\tau)$. Περαιτέρω, αμέσως διαπιστώνουμε ότι

$$\begin{aligned} Qu(x) = &\int_B u(y)\varphi(y)\,dy - \int_B \big[u_{x_1}(y)y_1 + u_{x_2}(y)y_2\big]\varphi(y)\,dy \\ &+ x_1 \int_B u_{x_1}(y)\varphi(y)\,dy + x_2 \int_B u_{x_2}(y)\varphi(y)\,dy, \end{aligned}$$

ιδιαίτερα έχουμε όντως $Qu \in \mathbb{P}_1$. Εφαρμόζοντας στα ολοκληρώματα που εμφανίζονται σε αυτή τη σχέση τον τύπο του Green, μπορούμε να τη γράψουμε στη μορφή

$$\begin{aligned} Qu(x) = &\int_B u(y)\big[3\varphi(y) + y_1\varphi_{x_1}(y) + y_2\varphi_{x_2}(y)\big]\,dy \\ &- x_1 \int_B u(y)\varphi_{x_1}(y)\,dy - x_2 \int_B u(y)\varphi_{x_2}(y)\,dy. \end{aligned} \tag{4.50}$$

Σημειώστε ότι το Qu στην (4.50) έχει πλέον νόημα και για $u \in L^1_{\mathrm{loc}}(B)$.

Σκοπός μας τώρα είναι να μελετήσουμε το σφάλμα $u - Qu$. Υποθέτουμε κατ' αρχάς ότι η u είναι αρκετά ομαλή· στην περίπτωση $u \in H^2(\tau)$ θα επανέλθουμε αργότερα. Χρησιμοποιώντας το γεγονός ότι το ολοκλήρωμα της φ είναι ένα, έχουμε

$$u(x) - Qu(x) = \int_B \big[u(x) - T_y u(x)\big]\varphi(y)\,dy, \quad x \in \tau. \tag{4.51}$$

Κατ' αρχάς, για μια συνάρτηση $f \in C^2[0,1]$, έχουμε, αναπτύσσοντας κατά Taylor ως προς το μηδέν και γράφοντας το υπόλοιπο σε ολοκληρωτική μορφή,

$$f(1) = f(0) + f'(0) + \int_0^1 s f''(1-s)\, ds. \tag{4.52}$$

Για $x \in \tau$ και $y \in B$, με $f(s) := u\big(y + s(x-y)\big)$, έχουμε, σύμφωνα με την (4.52),

$$\begin{aligned} u(x) - T_y u(x) = {} & (x_1 - y_1)^2 \int_0^1 s u_{x_1x_1}\big(x + s(y-x)\big)\, ds \\ & + 2(x_1-y_1)(x_2-y_2) \int_0^1 s u_{x_1x_2}\big(x + s(y-x)\big)\, ds \\ & + (x_2-y_2)^2 \int_0^1 s u_{x_2x_2}\big(x + s(y-x)\big)\, ds. \end{aligned} \tag{4.53}$$

Συνεπώς, η (4.51) γράφεται στη μορφή

$$u(x) - Qu(x) = I_1(x) + I_2(x) + I_3(x), \quad x \in \tau, \tag{4.54}$$

με

$$\begin{aligned} I_1(x) &:= \int_B (x_1-y_1)^2 \Big[\int_0^1 s u_{x_1x_1}\big(x + s(y-x)\big)\, ds\Big] \varphi(y)\, dy, \\ I_2(x) &:= 2\int_B (x_1-y_1)(x_2-y_2) \Big[\int_0^1 s u_{x_1x_2}\big(x + s(y-x)\big)\, ds\Big] \varphi(y)\, dy, \\ I_3(x) &:= \int_B (x_2-y_2)^2 \Big[\int_0^1 s u_{x_2x_2}\big(x + s(y-x)\big)\, ds\Big] \varphi(y)\, dy. \end{aligned}$$

Θα επικεντρώσουμε κατ' αρχάς το ενδιαφέρον μας στο $I_1(x)$, για τα $I_2(x)$ και $I_3(x)$ ισχύουν εντελώς παρόμοια πράγματα. Θα χρησιμοποιήσουμε την αλλαγή μεταβλητής

$$z := x + s(y-x) = sy + (1-s)x.$$

Προφανώς, καθώς το s μεταβάλλεται στο $[0,1]$, το z, ως κυρτός συνδυασμός των y και x, διαγράφει το ευθύγραμμο τμήμα που συνδέει τα y και x. Στον χώρο (y,s) η ολοκλήρωση γίνεται στο $B \times (0,1]$, στον χώρο (z,s) η ολοκλήρωση γίνεται στο A,

$$A := \{(z,s) : s \in (0,1],\ \ |\frac{1}{s}(z-x) + x - x^\star| < \rho_\tau\}, \tag{4.55}$$

αφού $y = \frac{1}{s}(z-x)+x$. Σημειώνουμε ότι, για $(z,s) \in A$, $\frac{1}{s}|z-x| < |x-x^\star| + \rho_\tau$, συνεπώς

$$s > \frac{|z-x|}{|x-x^\star|+\rho_\tau} =: t. \tag{4.56}$$

Χρησιμοποιώντας τη χαρακτηριστική συνάρτηση χ_A του A, γράφουμε το $I_1(x)$ στη μορφή

$$I_1(x) = \iint_{\mathbb{R}^3} \chi_A(z,s)\varphi\big(x+\frac{1}{s}(z-x)\big)(x_1-z_1)^2 u_{x_1x_1}(z)s^{-3}\,dzds,$$

αφού $dz = s^2 dy$. Συμβολίζουμε τώρα με C_x την κυρτή θήκη του συνόλου $\{x\}\cup B$, δηλαδή το ελάχιστο κυρτό σύνολο που περιέχει το x και τα σημεία του B. Σημειώνουμε ότι το C_x είναι υποσύνολο του τ. Προφανώς, το A είναι υποσύνολο του $C_x\times[0,1]$, οπότε μπορούμε να γράψουμε το $I_1(x)$ ως

$$I_1(x) = \int_{C_x}(x_1-z_1)^2 u_{x_1x_1}(z)\Big[\int_0^1 \varphi\big(x+\frac{1}{s}(z-x)\big)\chi_A(z,s)s^{-3}\,ds\Big]dz. \tag{4.57}$$

Τώρα, με

$$k(x,z) := \int_0^1 \varphi\big(x+\frac{1}{s}(z-x)\big)\chi_A(z,s)s^{-3}\,ds, \tag{4.58}$$

η (4.57) γράφεται στη μορφή

$$I_1(x) = \int_{C_x}(x_1-z_1)^2 k(x,z)u_{x_1x_1}(z)\,dz. \tag{4.59}$$

Χρησιμοποιώντας την (4.56) διαπιστώνουμε αμέσως ότι

$$|k(x,z)| \le \int_t^1 \varphi\big(x+\frac{1}{s}(z-x)\big)s^{-3}\,ds,$$

οπότε

$$|k(x,z)| \le \frac{1}{2}\,\|\varphi\|_{L^\infty(B)}\,\frac{1}{t^2}. \tag{4.60}$$

Τώρα,

$$\frac{1}{t} = \frac{|x-x^\star|+\rho_\tau}{|z-x|} \le \frac{h_\tau+\rho_\tau}{|z-x|} \le \rho_\tau(1+\sigma_\tau)|z-x|^{-1},$$

συνεπώς η (4.60) δίνει

$$|k(x,z)| \le \frac{1}{2}\,\|\varphi\|_{L^\infty(B)}\,\rho_\tau^2(1+\sigma_\tau)^2|z-x|^{-2}. \tag{4.61}$$

Όμως, έχουμε

$$\frac{1}{2}\,\|\varphi\|_{L^\infty(B)}\,\rho_\tau^2 \le C, \tag{4.62}$$

με μια απόλυτη σταθερά C, ανεξάρτητη από το τρίγωνο τ, βλ. τη σχετική συζήτηση μετά την (4.7), συνεπώς η (4.61) δίνει

$$|k(x,z)| \le C(1+\sigma_\tau)^2|z-x|^{-2}. \tag{4.63}$$

Τώρα $|z-x|^{-2}(x_1-z_1)^2 \le 1$, συνεπώς οι (4.59) και (4.63) δίνουν

$$|I_1(x)| \le C(1+\sigma_\tau)^2\int_{C_x}|u_{x_1x_1}(z)|\,dz,$$

ή, αφού $C_x \subset \tau$,

$$|I_1(x)| \le C(1+\sigma_\tau)^2\int_\tau|u_{x_1x_1}(z)|\,dz. \tag{4.64}$$

Επομένως, αφού το εμβαδόν του τ είναι μικρότερο του h_τ^2, εκτιμώντας το ολοκλήρωμα στο δεξιό μέλος της (4.64) με την ανισότητα των Cauchy–Schwarz, λαμβάνουμε

$$|I_1(x)| \le C(1+\sigma_\tau)^2 h_\tau\|u_{x_1x_1}\|_{L^2(\tau)}. \tag{4.65}$$

Από την (4.65) παίρνουμε αμέσως

$$\int_\tau |I_1(x)|^2\,dx \le C^2(1+\sigma_\tau)^4 h_\tau^4\|u_{x_1x_1}\|_{L^2(\tau)}^2. \tag{4.66}$$

Αντίστοιχες εκτιμήσεις ισχύουν και για τα I_2 και I_3, οπότε η (4.54) δίνει

$$\|u-Qu\| \le C(1+\sigma_\tau)^2 h_\tau^2|u|_2. \tag{4.67}$$

Προχωρούμε τώρα στην εκτίμηση των πρώτων μερικών παραγώγων. Κατ' αρχάς έχουμε

$$u_{x_i}(x)-\big(Qu\big)_{x_i}(x) = \int_B\big[u_{x_i}(x)-u_{x_i}(y)\big]\varphi(y)\,dy. \tag{4.68}$$

Τώρα, με $f(s) := u_{x_i}\big(y + s(x-y)\big)$, η σχέση

$$f(1) - f(0) = \int_0^1 f'(1-s)\,ds,$$

δίνει

$$\begin{aligned} u_{x_i}(x) - u_{x_i}(y) = (x_1 - y_1) &\int_0^1 u_{x_i x_1}\big(y + s(x-y)\big)\,ds \\ + (x_2 - y_2) &\int_0^1 u_{x_i x_2}\big(y + s(x-y)\big)\,ds. \end{aligned} \tag{4.69}$$

Επομένως, έχουμε

$$u_{x_i}(x) - \big(Qu\big)_{x_i}(x) = J_{i1}(x) + J_{i2}(x), \quad x \in \tau, \tag{4.70}$$

με

$$J_{ij}(x) := \int_B (x_j - y_j)\Big[\int_0^1 u_{x_i x_j}\big(y + s(x-y)\big)\,ds\Big]\varphi(y)\,dy, \quad i,j = 1,2.$$

Προχωρώντας όπως προηγουμένως λαμβάνουμε

$$J_{ij}(x) = \int_{C_x} (x_j - y_j) k(x,z) u_{x_i x_j}(z)\,dz, \tag{4.71}$$

με τον πυρήνα k που δίνεται στην (4.58). Επομένως, βλ. τις (4.63) και (4.64),

$$|J_{ij}(x)| \le C(1+\sigma_\tau)^2 \int_\tau \frac{|x_j - y_j|}{|x-z|^2}\,|u_{x_i x_j}(z)|\,dz,$$

οπότε

$$|J_{ij}(x)| \le C(1+\sigma_\tau)^2 \int_\tau |x-z|^{-1}\,|u_{x_i x_j}(z)|\,dz. \tag{4.72}$$

Τώρα, για να προχωρήσουμε στην εκτίμηση του $|J_{ij}(x)|$, σημειώνουμε πρώτα ότι, για $x \in \tau$, χρησιμοποιώντας πολικές συντεταγμένες έχουμε

$$\int_\tau |x-z|^{-1}\,dz \le \int_{B(x;h_\tau)} |x-z|^{-1}\,dz = \int_0^{h_\tau} \Big(\int_{\partial B(x;r)} \frac{1}{r}\,dS\Big)\,dr = \int_0^{h_\tau} 2\pi\,dr,$$

συνεπώς

$$\int_\tau |x-z|^{-1}\,dz \le 2\pi h_\tau. \tag{4.73}$$

Τώρα, με $g_{ij}(x) := \int_\tau |x-z|^{-1}\,|u_{x_ix_j}(z)|\,dz$, χρησιμοποιώντας την ανισότητα των Cauchy–Schwarz και την (4.73), έχουμε

$$\begin{aligned}
\|g_{ij}\|^2 &= \int_\tau \Big(\int_\tau |x-z|^{-1}\,|u_{x_ix_j}(z)|\,dz\Big)^2\,dx\\
&\le \int_\tau \Big[\int_\tau |x-z|^{-1}\,dz \int_\tau |x-z|^{-1}|u_{x_ix_j}(z)|^2\,dz\Big]\,dx\\
&\le 2\pi h_\tau \int_\tau\int_\tau |x-z|^{-1}|u_{x_ix_j}(z)|^2\,dz\,dx\\
&= 2\pi h_\tau \int_\tau \Big(\int_\tau |x-z|^{-1}\,dx\Big)|u_{x_ix_j}(z)|^2\,dz\\
&\le 4\pi^2 h_\tau^2 \|u_{x_ix_j}\|^2.
\end{aligned}$$

Επομένως, από την (4.72) παίρνουμε

$$\|J_{ij}\| \le C(1+\sigma_\tau)^2 h_\tau \|u_{x_ix_j}\|, \quad i,j=1,2. \tag{4.74}$$

Οι (4.70) και (4.74) δίνουν αμέσως

$$|u-Qu|_1 \le C(1+\sigma_\tau)^2 h_\tau |u|_2. \tag{4.75}$$

Ο επόμενος στόχος μας τώρα είναι να αποδείξουμε ότι οι εκτιμήσεις (4.67) και (4.75) ισχύουν και για συναρτήσεις $u \in H^2(\tau)$. Σημειώνουμε κατ' αρχάς ότι, σύμφωνα με την (4.50), ισχύει

$$\|Qu\|_{W^{1,\infty}(\tau)} \le C\|u\|_{L^1(B)}, \tag{4.76}$$

με μια σταθερά C ανεξάρτητη της u. Έστω τώρα $u \in H^2(\tau)$. Τότε, υπάρχει μια ακολουθία $(u_n)_{n\in\mathbb{N}} \subset C^\infty(\bar\tau)$, η οποία συγκλίνει στη u στον $H^2(\tau)$ (βλ. το Θεώρημα 4.9 για το αντίστοιχο αποτέλεσμα για τον $H^1(\tau)$, που ισχύει και για τρίγωνα). Τώρα,

$$\|Q(u-u_n)\|_{H^1(\tau)} \le C\|u-u_n\|_{L^1(B)} \to 0, \quad n\to\infty. \tag{4.77}$$

Επομένως, εφαρμόζοντας τις (4.67) και (4.75) στις u_n και αφήνοντας το n να τείνει στο άπειρο, οδηγούμαστε στα επιθυμητά αποτελέσματα (4.46) και (4.47), αντίστοιχα, και η απόδειξη είναι πλήρης. □

Παρατήρηση 4.3 (Γενίκευση της Πρότασης 4.6.) Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^2$, με ομαλό σύνορο $\partial\Omega$. Το Ω λέγεται *αστροειδές*, αν υπάρχει κύκλος B, τέτοιος ώστε, για κάθε $x \in \Omega$, η κυρτή θήκη C_x του συνόλου $\{x\} \cup \Omega$ να περιέχεται στο Ω. Από την απόδειξη της Πρότασης 4.6 προκύπτει αμέσως ότι τα αποτελέσματα της Πρότασης ισχύουν και για αστροειδή χωρία Ω. Το ρόλο του h_τ τον παίζει τότε το μήκος της διαμέτρου του Ω, δηλαδή του μεγαλυτέρου ευθυγράμμου τμήματος που περιέχεται στο $\bar{\Omega}$, και τον ρόλο του ρ_τ η ακτίνα του κύκλου B. Ας σημειωθεί ακόμα ότι η Πρόταση 4.6 γενικεύεται προς πολλές κατευθύνσεις, όπως στην περίπτωση που τον ρόλο του $\mathbb{P}_1$ αναλαμβάνει ένας χώρος $\mathbb{P}_m$, $m \geq 2$, πολυωνύμων βαθμού το πολύ m, ή όταν το Ω περιέχεται στον $\mathbb{R}^d$, με $d \geq 3$, και ακόμα όταν εργαζόμαστε στους χώρους $W^{k,p}(\Omega)$. □

Παρατήρηση 4.4 (Η ανισότητα του Sobolev.) Με τους συμβολισμούς στην Πρόταση 4.6 καθώς και στην Απόδειξή της, από την (4.54) και την (4.65), καθώς και τις αντίστοιχες της (4.65) εκτιμήσεις για τις $I_2(x)$ και $I_3(x)$, λαμβάνουμε

$$\|u - Qu\|_{L^\infty(\tau)} \leq C(1+\sigma_\tau)^2 h_\tau |u|_2. \tag{4.78}$$

Επί πλέον, από την (4.49) προκύπτει αμέσως, αφού $\|\varphi\| \leq C/\rho_\tau$,

$$\|Qu\|_{L^\infty(\tau)} \leq C(1+h_\tau)\frac{1}{\rho_\tau}\|u\|_{H^1(\tau)}. \tag{4.79}$$

Από τις (4.78) και (4.79) λαμβάνουμε την ανισότητα του Sobolev

$$\|u\|_{L^\infty(\tau)} \leq C\big[(1+\sigma_\tau)^2 h_\tau + (1+h_\tau)\frac{1}{\rho_\tau}\big]\|u\|_{H^2(\tau)} \quad u \in H^2(\tau). \tag{4.80}$$

Σημειώνουμε ότι από την ανισότητα του Sobolev μπορεί κανείς να οδηγηθεί εύκολα στο συμπέρασμα ότι κάθε $u \in H^2(\tau)$ είναι σ.π. ίση με μια συνεχή στο $\bar{\tau}$ συνάρτηση.
□

4.3.2 Η προσεγγιστική ιδιότητα (4.36)

Σκοπός μας σε αυτή την ενότητα είναι να αποδείξουμε ότι η οικογένεια των χώρων $(S_h)_{0<h\leq 1}$ έχει την προσεγγιστική ιδιότητα (4.36), με $r = 2$. Με κίνητρο τις εκτιμήσεις (4.46) και (4.47), καλούμε μια οικογένεια τριγωνισμών $(\mathcal{T}_h)_{0<h\leq 1}$ ενός κυρτού πολυγώνου Ω *ομαλή*, αν υπάρχει μια σταθερά $\sigma^\star$, ανεξάρτητη του h, τέτοια ώστε, για κάθε τρίγωνο τ του $\mathcal{T}_h$, να ισχύει

$$\sigma_\tau \leq \sigma^\star, \tag{4.81}$$

βλ. Πρόταση 4.6. Εναλλακτικοί χαρακτηρισμοί ομαλών οικογενειών τριγωνισμών είναι είτε ότι κάθε τρίγωνο τ, με μέγιστη πλευρά h_τ, έχει εμβαδόν μεγαλύτερο ή ίσο του ch_τ^2, με $c > 0$ ανεξάρτητο του h, είτε ότι όλες οι γωνίες των τριγώνων του $\mathcal{T}_h$ είναι μεγαλύτερες από μια θετική γωνία α, ανεξάρτητη του h, βλ. την Άσκηση 4.9.

Επιχείρημα ομοιογένειας. Έστω τ ένα τρίγωνο με μέγιστη πλευρά h_τ. Συμβολίζουμε με $\tilde{\tau}$ το τρίγωνο που προκύπτει από το τ ως εξής

$$\tilde{\tau} := \{\frac{1}{h_\tau} x : x \in \tau\}.$$

Τα τρίγωνα τ και $\tilde{\tau}$ είναι προφανώς όμοια. Ιδιαίτερα, $\sigma_\tau = \sigma_{\tilde{\tau}}$. Για $u \in H^2(\tau)$, θεωρούμε τη συνάρτηση $\tilde{u} : \tilde{\tau} \to \mathbb{R}$, $\tilde{u}(x) := u(h_\tau x)$. Προφανώς, $\tilde{u} \in H^2(\tilde{\tau})$. Επί πλέον, ισχυριζόμαστε ότι

$$|\tilde{u}|_{H^\ell(\tilde{\tau})} = h_\tau^{\ell-1} |u|_{H^\ell(\tau)}, \quad \ell = 0, 1, 2. \tag{4.82}$$

Πράγματι, για $\ell = 0$, με $x := \frac{1}{h_\tau} y$,

$$\|\tilde{u}\|_{L^2(\tilde{\tau})}^2 = \int_{\tilde{\tau}} |\tilde{u}(x)|^2 \, dx = \frac{1}{h_\tau^2} \int_\tau |\tilde{u}(\frac{1}{h_\tau} y)|^2 \, dy = \frac{1}{h_\tau^2} \int_\tau |u(y)|^2 \, dy,$$

δηλαδή

$$\|\tilde{u}\|_{L^2(\tilde{\tau})} = \frac{1}{h_\tau} \|u\|_{L^2(\tau)}.$$

Συνεπώς ισχύει η (4.82) για $\ell = 0$. Εξ άλλου, $\tilde{u}_{x_i}(x) = h_\tau u_{x_i}(h_\tau x)$ και $\tilde{u}_{x_i x_j}(x) = h_\tau^2 u_{x_i x_j}(h_\tau x)$, και όπως προηγουμένως διαπιστώνουμε ότι η (4.82) ισχύει και για $\ell = 1$ και $\ell = 2$.

Θεωρούμε στο $\tilde{\tau}$ τον τελεστή παρεμβολής $I_{\tilde{\tau}} : H^2(\tilde{\tau}) \to \mathbb{P}_1$,

$$(I_{\tilde{\tau}} \tilde{v})(\tilde{P}_i) = \tilde{v}(\tilde{P}_i), \quad i = 1, 2, 3,$$

όπου $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$ οι κορυφές του $\tilde{\tau}$. Θεωρούμε επίσης τις συναρτήσεις $\tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{\varphi}_3$, πολυώνυμα βαθμού το πολύ ένα στο $\tilde{\tau}$, με την ιδιότητα $\tilde{\varphi}_i(\tilde{P}_j) = \delta_{ij}$, $i, j = 1, 2, 3$. Προφανώς έχουμε

$$I_{\tilde{\tau}} \tilde{v} = \sum_{i=1}^3 \tilde{v}(\tilde{P}_i) \tilde{\varphi}_i. \tag{4.83}$$

Επομένως έχουμε

$$\|I_{\tilde{\tau}} \tilde{v}\|_{H^1(\tilde{\tau})} \le \beta(\tilde{\tau}) \sup_{\tilde{x} \in \tilde{\tau}} |\tilde{v}(\tilde{x})| \tag{4.84}$$

με

$$\beta(\tilde{\tau}) := \sum_{i=1}^{3} \|\tilde{\varphi}_i\|_{H^1(\tilde{\tau})}. \tag{4.85}$$

Όσον αφορά το σφάλμα της παρεμβολής, για $\tilde{v} \in H^2(\tilde{\tau})$, χρησιμοποιώντας το γεγονός ότι $h_{\tilde{\tau}} = 1$, και τις (4.84) και (4.80), έχουμε

$$\begin{aligned}\|\tilde{v} - I_{\tilde{\tau}}\tilde{v}\|_{H^1(\tilde{\tau})} &\le \|\tilde{v} - Q\tilde{v}\|_{H^1(\tilde{\tau})} + \|I_{\tilde{\tau}}(\tilde{v} - Q\tilde{v})\|_{H^1(\tilde{\tau})}\\ &\le \|\tilde{v} - Q\tilde{v}\|_{H^1(\tilde{\tau})} + \beta(\tilde{\tau}) \sup_{\tilde{x}\in\tilde{\tau}} |(\tilde{v} - Q\tilde{v})(\tilde{x})|\\ &\le \|\tilde{v} - Q\tilde{v}\|_{H^1(\tilde{\tau})} + C\big(\sigma_\tau, \beta(\tilde{\tau})\big)\|\tilde{v} - Q\tilde{v}\|_{H^2(\tilde{\tau})}.\end{aligned}$$

Επομένως, σύμφωνα με τις (4.46) και (4.47), αν λάβουμε υπ' όψιν μας και τη συνθήκη (4.81) , έχουμε

$$\|\tilde{v} - I_{\tilde{\tau}}\tilde{v}\|_{H^2(\tilde{\tau})} \le C\big(\sigma^\star, \beta(\tilde{\tau})\big)|\tilde{v}|_{H^2(\tilde{\tau})}. \tag{4.86}$$

Έστω τώρα P_1, P_2, P_3 οι κορυφές του τ. Θεωρούμε στο τ τον τελεστή παρεμβολής $I_\tau : H^2(\tau) \to \mathbb{P}_1$,

$$(I_\tau v)(P_i) = v(P_i), \quad i = 1, 2, 3.$$

Αμέσως διαπιστώνει κανείς ότι $\widetilde{I_\tau v} = I_{\tilde{\tau}}\tilde{v}$. Πράγματι στις κορυφές των τριγώνων έχουμε, με κατάλληλη αρίθμηση των κορυφών του τ,

$$(\widetilde{I_\tau v})(\tilde{P}_i) = (I_\tau v)(P_i) = v(P_i) = \tilde{v}(\tilde{P}_i) = (I_{\tilde{\tau}}\tilde{v})(\tilde{P}_i), \quad i = 1, 2, 3.$$

Το τελευταίο σημαντικό βήμα για να οδηγηθούμε στην απόδειξη της επιθυμητής προσεγγιστικής ιδιότητας της οικογένειας S_h είναι να αποδείξουμε ότι οι ποσότητες $\beta(\tilde{\tau})$ είναι ομοιόμορφα φραγμένες, υπό την προϋπόθεση (4.81), για όλα τα τρίγωνα $\tilde{\tau}$ με μέγιστη πλευρά ίση με τη μονάδα.

Θεωρούμε δύο τρίγωνα $\hat{\tau}$ και $\tilde{\tau}$ στο επίπεδο. Συμβολίζουμε με $\hat{P}_i$ και $\tilde{P}_i, i = 1, 2, 3$, τις κορυφές του $\hat{\tau}$ και του $\tilde{\tau}$, αντίστοιχα. Τότε υπάρχει ακριβώς μία *ομοπαραλληλική* απεικόνιση $A : \mathbb{R}^2 \to \mathbb{R}^2$, $A(\hat{x}) = \mathcal{A}\hat{x} + b$, με $\mathcal{A} \in \mathbb{R}^{2,2}$, τέτοια ώστε $A(\hat{P}_i) = \tilde{P}_i, i = 1, 2, 3$. Πράγματι, με $\mathcal{A} = \big(a_{ij}\big)_{i,j=1,2}$, $b = (b_1, b_2)^T$, $\hat{P}_i = (\hat{x}_{i1}, \hat{x}_{i2})$ και $\tilde{P}_i = (\tilde{x}_{i1}, \tilde{x}_{i2})$, $i = 1, 2, 3$, οι σχέσεις $A(\hat{P}_i) = \tilde{P}_i, i = 1, 2, 3$, γράφονται ισοδύναμα στη μορφή

$$a_{j1}\hat{x}_{i1} + a_{j2}\hat{x}_{i2} + b_j = \tilde{x}_{ij}, \quad i = 1, 2, 3, \quad j = 1, 2. \tag{4.87}$$

Για κάθε j, το (4.87) είναι ένα γραμμικό σύστημα με τρεις εξισώσεις και τρεις αγνώστους, τους a_{j1}, a_{j2} και b_j. Η ορίζουσα D του πίνακα αυτών των συστημάτων είναι

$$D = \begin{vmatrix} \hat{x}_{11} & \hat{x}_{12} & 1 \\ \hat{x}_{21} & \hat{x}_{22} & 1 \\ \hat{x}_{31} & \hat{x}_{32} & 1 \end{vmatrix} = \begin{vmatrix} \hat{x}_{11} & \hat{x}_{12} & 1 \\ \hat{x}_{21} - \hat{x}_{11} & \hat{x}_{22} - \hat{x}_{12} & 0 \\ \hat{x}_{31} - \hat{x}_{11} & \hat{x}_{32} - \hat{x}_{12} & 0 \end{vmatrix} = \begin{vmatrix} \hat{x}_{21} - \hat{x}_{11} & \hat{x}_{22} - \hat{x}_{12} \\ \hat{x}_{31} - \hat{x}_{11} & \hat{x}_{32} - \hat{x}_{12} \end{vmatrix}.$$

Η ορίζουσα αυτή είναι διάφορη του μηδενός, αφού έχει απόλυτη τιμή ίση με το διπλάσιο του εμβαδού του $\hat{\tau}$, βλ. την Άσκηση 4.11. Επομένως, τα a_{j1}, a_{j2} και b_j, $j = 1, 2$, ορίζονται μονοσήμαντα από τις (4.87).

Κάθε τρίγωνο (το εσωτερικό του και οι πλευρές του) αποτελεί την κυρτή θήκη των κορυφών του, συνεπώς η απεικόνιση A απεικονίζει το $\hat{\tau}$ στο $\tilde{\tau}$. Ιδιαίτερα, ο πίνακας $\mathcal{A}$ είναι αντιστρέψιμος και η A αποτελεί αμφιμονοσήμαντη απεικόνιση από το $\hat{\tau}$ στο $\tilde{\tau}$. Στη συνέχεια συγκεκριμενοποιούμε το *τρίγωνο αναφοράς* $\hat{\tau}$ να είναι το τρίγωνο με κορυφές τα σημεία $(0,0)$, $(1,0)$ και $(0,1)$. Σημειώστε ότι και η απεικόνιση που μας πηγαίνει από ένα τρίγωνο τ στο τρίγωνο $\tilde{\tau}$ είναι ομοπαραλληλική, πολύ απλής όμως μορφής. Θα μπορούσαμε να απεικονίσουμε το τρίγωνο αναφοράς $\hat{\tau}$ απ' ευθείας στο τ, αλλά απλουστεύει κάπως την ανάλυση να γίνεται αυτή η απεικόνιση σε δύο στάδια, παρεμβάλλοντας και το τρίγωνο $\tilde{\tau}$.

Συμβολίζουμε με $\tilde{B}$ τον εγγεγραμμένο κύκλο στο τρίγωνο $\tilde{\tau}$. Τότε, συμβολίζοντας με $|\Omega|$ το εμβαδόν ενός χωρίου Ω, έχουμε αφ' ενός

$$|\tilde{B}| \le |\tilde{\tau}| = \int_{\tilde{\tau}} d\tilde{x} = \int_{\hat{\tau}} |\det \mathcal{A}|\, d\hat{x} = |\det \mathcal{A}|\, |\hat{\tau}| < |\det \mathcal{A}|,$$

και αφ' ετέρου, λαμβάνοντας υπ' όψιν και την (4.81),

$$|\tilde{B}| = \pi\big(\rho_{\tilde{\tau}}\big)^2 = \frac{\pi}{\big(\sigma_{\tilde{\tau}}\big)^2} \ge \frac{\pi}{\big(\sigma^\star\big)^2}.$$

Συνεπώς, με $\varepsilon := \pi/(\sigma^\star)^2$ εξαρτώμενο μόνο από το $\sigma^\star$,

$$|\det \mathcal{A}| \ge \varepsilon > 0. \tag{4.88}$$

Τώρα, το σύνολο

$$\{\mathcal{B} \in \mathbb{R}^{2,2} : \ |\det \mathcal{B}| \ge \varepsilon\} \tag{4.89}$$

είναι κλειστό. Γράφουμε τις κορυφές του $\hat{\tau}$, σημεία του $\mathbb{R}^2$, ως στήλες $\hat{P}_1 = (1,0)^T = e^1$, $\hat{P}_0 = (0,1)^T = e^2$, $\hat{P}_3 = (0,0)^T = e^0$. Με κατάλληλη αρίθμηση, οι κορυφές του $\tilde{\tau}$

είναι $\tilde{P}_i = A(P_i)$, $i = 1, 2, 3$. Επομένως έχουμε

$$A(e^i) - A(e^0) = \mathcal{A}e^i = \tilde{P}_i - \tilde{P}_0, \quad i = 1, 2,$$

οπότε, αφού τα μήκη των πλευρών του $\tilde{\tau}$ δεν ξεπερνούν τη μονάδα, θα έχουμε

$$|\mathcal{A}e^i| \le 1, \quad i = 1, 2,$$

δηλαδή

$$(a_{i1})^2 + (a_{i2})^2 \le 1, \quad i = 1, 2.$$

Συνεπώς, τα στοιχεία του $\mathcal{A}$ έχουν απόλυτο τιμή μικρότερη ή ίση της μονάδας. (Ιδιαίτερα, $|\det \mathcal{A}| \le 2$.) Όλοι οι πίνακες $\mathcal{A}$ ανήκουν επομένως, σύμφωνα και με την (4.89), στο κλειστό και φραγμένο, δηλαδή συμπαγές, σύνολο $\mathcal{G}$,

$$\mathcal{G} := \{\mathcal{B} \in \mathbb{R}^{2,2} : \ |\det \mathcal{B}| \ge \varepsilon, \ |b_{ij}| \le 1, \ i, j = 1, 2\}. \tag{4.90}$$

Θα προσδιορίσουμε τώρα μια σχέση μεταξύ των $\beta(\tilde{\tau})$ και $\beta(\hat{\tau})$. Όπως αναφέραμε ήδη, η A είναι αμφιμονοσήμαντη απεικόνιση από το $\hat{\tau}$ στο $\tilde{\tau}$, $\hat{x} \mapsto \tilde{x} = \mathcal{A}\hat{x} + b$. Επομένως, η αντίστροφη απεικόνιση είναι $A^{-1} : \tilde{\tau} \to \hat{\tau}$, $\tilde{x} \mapsto \hat{x} = \mathcal{A}^{-1}\tilde{x} - \mathcal{A}^{-1}b$. Στην (4.85) οι συναρτήσεις $\tilde{\varphi}_i$ είναι πολυώνυμα πρώτου βαθμού στο $\tilde{\tau}$ και παίρνουν την τιμή ένα σε μία κορυφή του $\tilde{\tau}$, διαφορετική για κάθε i, ενώ μηδενίζονται στις άλλες δύο. Οι συναρτήσεις $\hat{\varphi}_i$,

$$\hat{\varphi}_i(\hat{x}) := \tilde{\varphi}_i(\tilde{x}), \ \ \hat{x} \in \hat{\tau}, \quad i = 1, 2, 3,$$

έχουν ακριβώς αντίστοιχες ιδιότητες, αλλά φυσικά στο τρίγωνο $\hat{\tau}$. Προφανώς έχουμε

$$\tilde{\varphi}_i(\tilde{x}) = \hat{\varphi}_i(\mathcal{A}^{-1}\tilde{x} - \mathcal{A}^{-1}b), \ \ \tilde{x} \in \tilde{\tau}, \quad i = 1, 2, 3. \tag{4.91}$$

Τώρα,

$$\begin{aligned}(I_{\tilde{\tau}}\tilde{v})(\tilde{x}) = (I_{\hat{\tau}}\hat{v})(\hat{x}) &= \sum_{i=1}^{3} \hat{v}(\hat{P}_i)\hat{\varphi}_i(\hat{x}) = \sum_{i=1}^{3} \hat{v}(\hat{P}_i)\hat{\varphi}_i(\mathcal{A}^{-1}\tilde{x} - \mathcal{A}^{-1}b)\\ &= \sum_{i=1}^{3} \tilde{v}(\tilde{P}_i)\hat{\varphi}_i(\mathcal{A}^{-1}\tilde{x} - \mathcal{A}^{-1}b),\end{aligned}$$

βλ. την (4.83). Τώρα, έχουμε

$$\begin{aligned}\|\tilde{\varphi}_i\|^2_{L^2(\tilde{\tau})} &= \int_{\tilde{\tau}} |\tilde{\varphi}_i(\tilde{x})|^2\, d\tilde{x} = \int_{\hat{\tau}} |\tilde{\varphi}_i(\mathcal{A}\hat{x}+b)|^2\, |\det \mathcal{A}|\, d\hat{x}\\ &= \int_{\hat{\tau}} |\hat{\varphi}_i(\hat{x})|^2\, |\det \mathcal{A}|\, d\hat{x},\end{aligned}$$

συνεπώς

$$\|\tilde{\varphi}_i\|_{L^2(\tilde{\tau})} = |\det \mathcal{A}|^{1/2}\, \|\hat{\varphi}_i\|_{L^2(\hat{\tau})}, \quad i = 1,2,3. \tag{4.92}$$

Επομένως,

$$\|\tilde{\varphi}_i\|_{L^2(\tilde{\tau})} \le \sqrt{2}\, \|\hat{\varphi}_i\|_{L^2(\hat{\tau})}, \quad i = 1,2,3. \tag{4.93}$$

Τώρα, $\hat{x} = \mathcal{A}^{-1}\tilde{x} - \mathcal{A}^{-1}b$, οπότε, αν συμβολίσουμε με a^{-1}_{ij} τα στοιχεία του $\mathcal{A}^{-1}$, $\hat{x}_i = a^{-1}_{i1}(\tilde{x}_1 - b_1) + a^{-1}_{i2}(\tilde{x}_2 - b_2)$, $i = 1,2$. Συνεπώς,

$$\frac{\partial \hat{x}_i}{\partial \tilde{x}_j} = a^{-1}_{ij}, \quad i,j = 1,2. \tag{4.94}$$

Παραγωγίζοντας στην (4.91) έχουμε επομένως

$$\frac{\partial \tilde{\varphi}_\ell}{\partial \tilde{x}_j}(\tilde{x}) = \sum_{i=1}^{2} a^{-1}_{ij}\, \frac{\partial \hat{\varphi}_\ell}{\partial \hat{x}_i},$$

δηλαδή

$$\begin{pmatrix} \frac{\partial \tilde{\varphi}_\ell}{\partial \tilde{x}_1}(\tilde{x}) \\ \frac{\partial \tilde{\varphi}_\ell}{\partial \tilde{x}_2}(\tilde{x}) \end{pmatrix} = (\mathcal{A}^{-1})^T \begin{pmatrix} \frac{\partial \hat{\varphi}_\ell}{\partial \hat{x}_1}(\hat{x}) \\ \frac{\partial \hat{\varphi}_\ell}{\partial \hat{x}_2}(\hat{x}) \end{pmatrix}, \quad \ell = 1,2,3. \tag{4.95}$$

Τώρα

$$(\mathcal{A}^{-1})^T = \frac{1}{\det \mathcal{A}} \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix},$$

επομένως, για την Ευκλείδεια νόρμα $|(\mathcal{A}^{-1})^T|$ του $(\mathcal{A}^{-1})^T$, έχουμε

$$|(\mathcal{A}^{-1})^T| \le \frac{1}{|\det \mathcal{A}|}\big(a_{11}^2 + a_{21}^2 + a_{12}^2 + a_{22}^2\big)^{1/2},$$

βλ. την Άσκηση 4.12. Λαμβάνοντας εδώ υπ' όψιν το γεγονός ότι $\mathcal{A} \in \mathcal{G}$, συμπεραίνουμε ότι

$$|(\mathcal{A}^{-1})^T| \le \frac{2}{\varepsilon}. \tag{4.96}$$

Για μια εναλλακτική απόδειξη μίας παρόμοιας εκτίμησης παραπέμπουμε στην Άσκηση 4.13. Από τις (4.95) και (4.96) προκύπτει ότι

$$\Big(\frac{\partial\tilde\varphi_\ell}{\partial\tilde x_1}(\tilde x)\Big)^2+\Big(\frac{\partial\tilde\varphi_\ell}{\partial\tilde x_2}(\tilde x)\Big)^2\le\frac{4}{\varepsilon^2}\Big[\Big(\frac{\partial\hat\varphi_\ell}{\partial\hat x_1}(\hat x)\Big)^2+\Big(\frac{\partial\hat\varphi_\ell}{\partial\hat x_2}(\hat x)\Big)^2\Big],\quad \ell=1,2,3. \tag{4.97}$$

Κατ' αναλογίαν προς την (4.93) έχουμε τώρα

$$|\tilde\varphi_i|_{H^1(\tilde\tau)}\le\frac{2\sqrt2}{\varepsilon}\,|\hat\varphi_i|_{H^1(\hat\tau)},\quad i=1,2,3. \tag{4.98}$$

Από τις (4.93) και (4.98) συμπεραίνουμε ότι

$$\beta(\tilde\tau)\le C\,\beta(\hat\tau) \tag{4.99}$$

με την ίδια σταθερά C για όλα τα τρίγωνα $\tilde\tau$ με μέγιστη πλευρά ίση με τη μονάδα, που ικανοποιούν την (4.81).

Επομένως, η (4.86) λαμβάνει τη μορφή

$$\|\tilde v-I_{\tilde\tau}\tilde v\|_{H^2(\tilde\tau)}\le C|\tilde v|_{H^2(\tilde\tau)}. \tag{4.100}$$

Σύμφωνα με τις (4.82) και (4.100) έχουμε τώρα για κάθε τρίγωνο τ του τριγωνισμού $\mathcal{T}_h$,

$$\|v-I_\tau v\|_{L^2(\tau)}=h_\tau\|\tilde v-I_{\tilde\tau}\tilde v\|_{L^2(\tilde\tau)}\le Ch_\tau|\tilde v|_{H^2(\tilde\tau)}=Ch_\tau^2|v|_{H^2(\tau)}$$

και

$$|v-I_\tau v|_{H^1(\tau)}=|\tilde v-I_{\tilde\tau}\tilde v|_{H^1(\tilde\tau)}\le C|\tilde v|_{H^2(\tilde\tau)}=Ch_\tau|v|_{H^2(\tau)},$$

δηλαδή

$$\|v-I_\tau v\|_{L^2(\tau)}\le Ch^2|v|_{H^2(\tau)} \tag{4.101}$$

και

$$|v-I_\tau v|_{H^1(\tau)}\le Ch|v|_{H^2(\tau)}. \tag{4.102}$$

Ύστερα από όλη αυτή την προεργασία, εύκολα τώρα μπορούμε να αποδείξουμε ότι, υπό την προϋπόθεση (4.81), οι χώροι S_h έχουν την προσεγγιστική ιδιότητα (4.36) με $r=2$. Για $v\in H^2(\Omega)$, συμβολίζουμε με I_hv το στοιχείο του S_h που παρεμβάλλεται στη v στους κόμβους του τριγωνισμού $\mathcal{T}_h$, δηλαδή την παρεμβάλλουσα της v. Τότε, σύμφωνα με την (4.101),

$$\|v-I_hv\|^2_{L^2(\Omega)}=\sum_{\tau\in\mathcal{T}_h}\|v-I_hv\|^2_{L^2(\tau)}=\sum_{\tau\in\mathcal{T}_h}\|v-I_\tau v\|^2_{L^2(\tau)}\le C^2h^4\sum_{\tau\in\mathcal{T}_h}|v|_{H^2(\tau)},$$

δηλαδή

$$\|v - I_h v\|_{L^2(\Omega)} \le Ch^2 |v|_{H^2(\Omega)}. \tag{4.103}$$

Ακριβώς αντίστοιχα, χρησιμοποιώντας τώρα την (4.102), λαμβάνουμε

$$|v - I_h v|_{H^1(\Omega)} \le Ch |v|_{H^2(\Omega)}. \tag{4.104}$$

Παρατήρηση 4.5 (Κυρτά επίπεδα χωρία.) Έστω $\Omega \subset \mathbb{R}^2$ ένα κυρτό χωρίο με ομαλό σύνορο $\partial\Omega$. Εγγράφουμε στο $\partial\Omega$ ένα κυρτό πολύγωνο Ω_h, με μέγιστη πλευρά το πολύ h. Θεωρούμε έναν τριγωνισμό $\mathcal{T}_h$ του Ω_h, με τις ιδιότητες που περιγράψαμε στην περίπτωση πολυγώνων, και σχηματίζουμε τους αντίστοιχους χώρους S_h, επεκτείνοντας τα στοιχεία τους στο $\Omega \setminus \Omega_h$ ως μηδενικές συναρτήσεις. Προφανώς οι S_h είναι υπόχωροι του $H_0^1(\Omega)$. Η απόσταση κάθε σημείου του $\partial\Omega_h$ από το $\partial\Omega$ είναι της τάξης του h^2, συνεπώς το μέρος του $\Omega \setminus \Omega_h$ που περικλείεται από μια πλευρά του Ω_h και μέρος του $\partial\Omega$ είναι της τάξης του h^3. Αφού το πλήθος των πλευρών του Ω_h είναι της τάξης h^{-1}, συμπεραίνουμε ότι το εμβαδόν του $\Omega \setminus \Omega_h$ είναι της τάξης h^2. Αποδεικνύεται μετά ότι οι χώροι αυτοί έχουν την προσεγγιστική ιδιότητα (4.36). □

Ασκήσεις

4.1 Έστω $\Omega := \{(x,y) \in \mathbb{R}^2 :\ -1 < x, y < 1\}$ και $u, v : \Omega \to \mathbb{R}$, $u(x,y) := \operatorname{sgn} x$, $v(x,y) := \operatorname{sgn} x - \operatorname{sgn} y$. Αποδείξτε ότι η γενικευμένη μερική παράγωγος $D^{(1,0)}u$ δεν υπάρχει, ενώ η $D^{(0,1)}u$ υπάρχει και είναι μάλιστα η μηδενική συνάρτηση, και ότι οι $D^{(1,0)}v$ και $D^{(0,1)}v$ δεν υπάρχουν, ενώ η $D^{(1,1)}v$ υπάρχει και είναι μάλιστα η μηδενική συνάρτηση.

4.2 (Διάσταση του χώρου των πολυωνύμων d μεταβλητών βαθμού το πολύ k.) Αποδείξτε ότι η διάσταση του χώρου των πολυωνύμων d μεταβλητών βαθμού το πολύ k (και, φυσικά, και το πλήθος των πολυδεικτών $\alpha \in \mathbb{N}_0^d$ τάξης το πολύ k) είναι $\binom{d+k}{k}$.
[*Υπόδειξη*: Επαγωγικά ως προς d και ως προς k: Για $d = 1$ η σχέση είναι προφανής. Έστω ότι ισχύει για $d-1$, για όλα τα k. Στην περίπτωση d μεταβλητών το επαγωγικό βήμα $k-1 \to k$ είναι τότε: Κάθε πολυώνυμο d μεταβλητών βαθμού το πολύ k γράφεται κατά μοναδικό τρόπο στη μορφή

$$p_k(x_1, \dots, x_d) = q_k(x_1, \dots, x_{d-1}) + x_d p_{k-1}(x_1, \dots, x_d),$$

όπου το q_k είναι πολυώνυμο $d-1$ μεταβλητών βαθμού το πολύ k και το p_{k-1} πολυώνυμο d μεταβλητών βαθμού το πολύ $k-1$. Επομένως, σύμφωνα με την επαγωγική υπόθεση, η ζητούμενη διάσταση είναι $\binom{d-1+k}{k} + \binom{d+k-1}{k-1}$. Χρησιμοποιήστε την προφανή σχέση

$$\binom{m}{n-1} + \binom{m}{n} = \binom{m+1}{n}$$

για να οδηγηθείτε στο αποτέλεσμα.]

4.3 Χρησιμοποιήστε τον τύπο του Green (4.16) για να αποδείξετε τον τύπο (4.17).

4.4 (Αρχή του Dirichlet.) Με τα δεδομένα του προβλήματος (4.20), ορίζουμε στον $H_0^1 := H_0^1(\Omega)$ το συναρτησιακό J ως εξής $J(v) := \frac{1}{2}(v,v)_1 - (f,v)$. Αποδείξτε ότι η λύση u του προβλήματος (4.21) είναι το μόνο στοιχείο του H_0^1, στο οποίο το J λαμβάνει την ελάχιστη τιμή του.

4.5 (Ομαλές ασθενείς λύσεις είναι κλασικές λύσεις.) Υποθέτουμε ότι $f \in C(\Omega)$ και ότι η λύση u του προβλήματος (4.21) ανήκει στον χώρο $C^2(\bar{\Omega})$. Αποδείξτε τότε ότι αυτή η λύση αποτελεί και κλασική λύση του προβλήματος (4.20).
[*Υπόδειξη*: Κατ' αρχάς, όσον αφορά τη συνοριακή συνθήκη, αφού $u \in H_0^1(\Omega) \cap C^2(\bar{\Omega})$, θα έχουμε $u|_{\partial\Omega} = 0$ με την κλασική έννοια. Όσον αφορά τη διαφορική εξίσωση, εφαρμόζοντας τον τύπο του Green, από την (4.21) λαμβάνουμε

$$\int_\Omega (-\Delta u + u)v\, dx = \int_\Omega fv\, dx \quad \forall v \in C_0^\infty(\Omega).$$

Χρησιμοποιήστε τώρα την πυκνότητα του $C_0^\infty(\Omega)$ στον $L^2(\Omega)$ για να οδηγηθείτε στο συμπέρασμα ότι $-\Delta u + u = f$ σ.π. στο Ω. Τέλος, αφού και τα δύο μέλη αυτής της ισότητας είναι συνεχείς συναρτήσεις, θα έχουμε $-\Delta u + u = f$ στο Ω.]

4.6 Θεωρούμε το πρόβλημα (4.25). Υποθέτουμε ότι ισχύει η (4.24) και ότι η συνάρτηση a_0 λαμβάνει μη αρνητικές τιμές. Αποδείξτε ότι η διγραμμική μορφή a είναι ελλειπτική στον $H_0^1(\Omega)$, δηλαδή αποδείξτε την (4.32).
[*Υπόδειξη*: Χρησιμοποιήστε την ανισότητα των Poincaré–Friedrichs, βλ. την Πρόταση 4.5.]

4.7 Θεωρούμε το πρόβλημα (4.34). Με τη βοήθεια μιας βάσης του χώρου S_h^r, γράψτε αυτό το πρόβλημα ως ένα γραμμικό σύστημα με ίδιο πλήθος εξισώσεων και αγνώστων. Αποδείξτε ότι ο πίνακας του γραμμικού συστήματος είναι συμμετρικός και θετικά ορισμένος.

4.8 Θεωρούμε ένα τρίγωνο τ και συμβολίζουμε με h τη μεγαλύτερη πλευρά του και με ρ την ακτίνα του εγγεγραμμένου κύκλου του. Έστω E το εμβαδόν του τ. Αποδείξτε ότι $E \geq \rho h$, οπότε με $\sigma := h/\rho$, $E \geq h^2/\sigma$.
[*Υπόδειξη*: Το E ισούται με το γινόμενο του ρ επί την ημιπερίμετρο του τριγώνου.]

4.9 Θεωρούμε μια οικογένεια τριγώνων τ και συμβολίζουμε με h τη μεγαλύτερη πλευρά τους και με ρ την ακτίνα του εγγεγραμμένου κύκλου τους. Έστω E το εμβαδόν του τ. Αποδείξτε ότι οι ακόλουθες συνθήκες είναι ισοδύναμες:

i. Υπάρχει μια θετική σταθερά $\sigma^\star$, τέτοια ώστε, για κάθε τρίγωνο της οικογένειας, να ισχύει $\sigma \le \sigma^\star$, όπου $\sigma := h/\rho$.

ii. Υπάρχει μια θετική σταθερά c, τέτοια ώστε, για κάθε τρίγωνο της οικογένειας, να ισχύει $E \ge ch^2$.

iii. Υπάρχει μια θετική σταθερά α, τέτοια ώστε, όλες οι γωνίες των τριγώνων της οικογένειας να είναι μεγαλύτερες της α.

4.10 Έστω a, b, c και d πραγματικοί αριθμοί. Αποδείξτε ότι η ορίζουσα

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

έχει απόλυτο τιμή ίση με το εμβαδόν E του παραλληλογράμμου με πλευρές τα διανύσματα (a, b) και (c, d).
[*Υπόδειξη.* Η προβολή $(\tilde a, \tilde b)$ του (c, d) στο (a, b) είναι

$$(\tilde a, \tilde b) = \frac{1}{a^2 + b^2}(ac + bd)\,(a, b),$$

συνεπώς για το ύψος h του παραλληλογράμμου ως προς τη βάση (a, b) ισχύει, σύμφωνα με το Πυθαγόρειο θεώρημα,

$$h^2 = c^2 + d^2 - \frac{(ac + bd)^2}{a^2 + b^2}.$$

Τώρα

$$E^2 = h^2(a^2 + b^2) = (c^2 + d^2)(a^2 + b^2) - (ac + bd)^2 = a^2d^2 + b^2c^2 - 2abcd = (ad - bc)^2.\,]$$

4.11 Θεωρούμε ένα τρίγωνο με κορυφές (x_{i1}, x_{i2}), $i = 1, 2, 3$. Αποδείξτε ότι η ορίζουσα

$$\begin{vmatrix} x_{21} - x_{11} & x_{22} - x_{12} \\ x_{31} - x_{11} & x_{32} - x_{12} \end{vmatrix}$$

είναι διάφορη του μηδενός.
[*Υπόδειξη.* Θεωρήστε ένα σύστημα με αρχή των αξόνων την κορυφή (x_{11}, x_{12}). Τότε, ως προς αυτό το σύστημα, οι άλλες δύο κορυφές έχουν συντεταγμένες $(a, b) = (x_{21} - x_{11}, x_{22} - x_{12})$ και $(c, d) = (x_{31} - x_{11}, x_{32} - x_{12})$, αντίστοιχα. Χρησιμοποιήστε την Άσκηση 4.10.]

4.12 Έστω $A \in \mathbb{R}^{d,d}$, $A = (a_{ij})_{i,j=1,\dots,d}$. Θεωρούμε στον $\mathbb{R}^d$ την Ευκλείδεια νόρμα $|\cdot|$ και συμβολίζουμε με $|\cdot|$ και την επαγόμενη νόρμα στον $\mathbb{R}^{d,d}$. Αποδείξτε ότι

$$|A| \le \Big(\sum_{i,j=1}^{d} a_{ij}^2\Big)^{1/2}.$$

[*Υπόδειξη.* Για $x \in \mathbb{R}^d$ έχουμε

$$(Ax)_i = \sum_{j=1}^{d} a_{ij} x_j,$$

συνεπώς, χρησιμοποιώντας την ανισότητα των Cauchy–Schwarz,

$$|Ax|^2 = \sum_{i=1}^{d} \Big(\sum_{j=1}^{d} a_{ij} x_j \Big)^2 \leq \sum_{i=1}^{d} \sum_{j=1}^{d} a_{ij}^2 \sum_{j=1}^{d} x_j^2 = \sum_{i,j=1}^{d} a_{ij}^2 \, |x|^2 .]$$

4.13 Έστω $\hat{\Omega}$ και $\tilde{\Omega}$ δύο ομοπαραλληλικά ισοδύναμα χωρία στον $\mathbb{R}^d$, δηλαδή τέτοια ώστε να υπάρχει μια αμφιμονοσήμαντη ομοπαραλληλική απεικόνιση $A : \hat{\Omega} \to \tilde{\Omega}$, $A(\hat{x}) = \mathcal{A}\hat{x} + b$. Υποθέτουμε ότι στο Ω περιέχεται μια σφαίρα ακτίνας $\hat{\rho}$ και συμβολίζουμε με $\tilde{h}$ τη διάμετρο του $\tilde{\Omega}$. Αποδείξτε την εκτίμηση

$$|\mathcal{A}| \leq \frac{\tilde{h}}{\hat{\rho}}$$

για την Ευκλείδεια νόρμα του πίνακα $\mathcal{A}$. Ποια είναι η αντίστοιχη εκτίμηση για την $|\mathcal{A}^{-1}|$;
[*Υπόδειξη.* Προφανώς

$$|\mathcal{A}| = \frac{1}{\hat{\rho}} \sup_{\substack{\xi \in \mathbb{R}^d \\ |\xi| = \hat{\rho}}} |\mathcal{A}\xi| .$$

Τώρα, κάθε $\xi \in \mathbb{R}^d$ τέτοιο ώστε $|\xi| = \hat{\rho}$, μπορεί να γραφεί ως διαφορά στοιχείων του $\hat{\Omega}$, $\xi = \hat{x} - \hat{y}$. Επομένως $\mathcal{A}\xi = A(\hat{x}) - A(\hat{y})$, οπότε $|\mathcal{A}\xi| = |A(\hat{x}) - A(\hat{y})| \leq \tilde{h}$, αφού$A(\hat{x}), A(\hat{y}) \in \tilde{\Omega}$.]

4.14 Βελτιώστε την εκτίμηση (4.96), αποδεικνύοντας ότι

$$|(\mathcal{A}^{-1})^T| \leq \frac{\sqrt{2}}{\varepsilon}.$$

5. Μέθοδοι πεπερασμένων στοιχείων για παραβολικές εξισώσεις

Μέχρι τώρα γνωρίσαμε μεθόδους πεπερασμένων στοιχείων για *στατικές* εξισώσεις, δηλαδή εξισώσεις ανεξάρτητες του χρόνου· συγκεκριμένα, ασχοληθήκαμε με μεθόδους πεπερασμένων στοιχείων για ελλειπτικές εξισώσεις σε μία διάσταση, δηλαδή για το πρόβλημα δύο σημείων, και σε δύο διαστάσεις. Σε αυτό το κεφάλαιο θα μελετήσουμε μεθόδους πεπερασμένων στοιχείων για μια *δυναμική* εξίσωση, την εξίσωση της θερμότητας σε μία διάσταση. Πρώτα θα αναφέρουμε ορισμένες βασικές ιδιότητες της λύσης του συνεχούς προβλήματος. Στη συνέχεια θα ασχοληθούμε κατ' αρχάς με το ημιδιακριτό πρόβλημα, το οποίο είναι διακριτό ως προς τον χώρο και συνεχές ως προς τον χρόνο. Για να οδηγηθούμε σε μεθόδους οι οποίες δίνουν προσεγγίσεις που μπορούν να υπολογισθούν, πρέπει να διακριτοποιήσουμε και ως προς τον χρόνο, παίρνοντας έτσι πλήρως διακριτά σχήματα. Η διακριτοποίηση στον χρόνο θα γίνει με τρεις μεθόδους, τις μεθόδους του Euler, άμεση και πεπλεγμένη, και με τη μέθοδο των Crank–Nicolson. 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.

5.1 Μια παραβολική εξίσωση

Έστω $T > 0$. Θεωρούμε το εξής πρόβλημα αρχικών και συνοριακών συνθηκών για μια παραβολική εξίσωση: Ζητείται μια συνάρτηση $u : [0,1] \times [0,T] \to \mathbb{R}$, δύο φορές συνεχώς παραγωγίσιμη ως προς x και μία φορά ως προς t, τέτοια ώστε

$$(5.1)\qquad \begin{cases} u_t = (au_x)_x - cu + f & \text{στο } [0,1]\times[0,T],\\ u(0,\cdot) = u(1,\cdot) = 0 & \text{στο } [0,T],\\ u(\cdot,0) = u_0 & \text{στο } [0,1], \end{cases}$$

όπου $c, f \in C([0,1]\times[0,T])$ και $a \in C^1([0,1]\times[0,T])$, $a(x,t) > 0$, για $x \in [0,1]$ και $t \in [0,T]$. Στη συνέχεια θα υποθέτουμε επί πλέον συχνά ότι

$$(5.2)\qquad c(x,t) \geq 0 \qquad \forall (x,t) \in [0,1]\times[0,T].$$

Σε αντίθεση με το πρόβλημα δύο σημείων, στην προκειμένη περίπτωση η συνθήκη (5.2) 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 (5.1) γράφεται ισοδύναμα στη μορφή

$$(5.3)\qquad \begin{cases} \tilde{u}_t = (a\tilde{u}_x)_x - (\lambda + c)\tilde{u} + \mathrm{e}^{-\lambda t} f & \text{στο } [0,1]\times[0,T],\\ \tilde{u}(0,\cdot) = \tilde{u}(1,\cdot) = 0 & \text{στο } [0,T],\\ \tilde{u}(\cdot,0) = u_0 & \text{στο } [0,1]. \end{cases}$$

Αν το λ επιλεγεί αρκετά μεγάλο, τότε η αντίστοιχη της συνθήκης (5.2) για το πρόβλημα (5.3) προφανώς ικανοποιείται. Στο εξής θα υποθέτουμε συχνά ότι έχει ήδη γίνει ένας τέτοιος μετασχηματισμός, θα θεωρούμε το πρόβλημα στη μορφή (5.1), και θα υποθέτουμε ότι ισχύει η (5.2).

Είναι γνωστό ότι αν τα δεδομένα είναι αρκετά ομαλά και συμβατά, τότε το πρόβλημα (5.1) έχει ακριβώς μια λύση. Αναγκαία συνθήκη για τη συνέχεια της u είναι η $u_0(0) = u_0(1) = 0$, ενώ για τη συνέχεια της u_t απαιτείται η παράσταση

$$(au_0')_x - cu_0 + f$$

να μηδενίζεται στα σημεία $(0,0)$ και $(1,0)$.

Συνεχής εξάρτηση της λύσης από τα αρχικά δεδομένα. Έστω u_1 μια λύση του προβλήματος (5.1), 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 τη $u := u_1 - u_2$ έχουμε

$$(5.4)\qquad \begin{cases} u_t = (au_x)_x - cu & \text{στο } [0,1]\times[0,T],\\ u(0,\cdot) = u(1,\cdot) = 0 & \text{στο } [0,T],\\ u(\cdot,0) = u_0 - \tilde{u}_0 & \text{στο } [0,1]. \end{cases}$$

Πολλαπλασιάζοντας τη Δ.Ε. στο (5.4) επί u, ολοκληρώνοντας στο $[0, 1]$, ολοκληρώνοντας κατά μέρη και χρησιμοποιώντας τις συνοριακές συνθήκες παίρνουμε, λαμβάνοντας υπ' όψιν και την (5.2),

$$\int_0^1 u_t(x,t)u(x,t)dx \le -\int_0^1 a(x,t)[u_x(x,t)]^2dx.$$

Τώρα

$$\int_0^1 u_t(x,t)u(x,t)dx = \frac{1}{2}\frac{d}{dt}\int_0^1 [u(x,t)]^2dx = \frac{1}{2}\frac{d}{dt}\|u(\cdot,t)\|^2,$$

οπότε η προηγούμενη σχέση γράφεται στη μορφή

$$\frac{d}{dt}\|u(\cdot,t)\|^2 \le -2\int_0^1 a(x,t)[u_x(x,t)]^2dx. \tag{5.5}$$

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

$$\|u(\cdot,t)\| \le \|u(\cdot,0)\| \qquad \forall t \in [0,T] \tag{5.6}$$

ή

$$\max_{0\le t\le T}\|u_1(\cdot,t) - u_2(\cdot,t)\| \le \|u_0 - \tilde{u}_0\|, \tag{5.7}$$

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 (5.6), αν εκμεταλλευθούμε καλύτερα το δεξιό μέλος της (5.5). Σύμφωνα με την ανισότητα των Poincaré–Friedrichs, από την (5.5) λαμβάνουμε

$$\frac{d}{dt}\|u(\cdot,t)\|^2 \le -2\,\min_{x,t} a(x,t)\,\|u_x(\cdot,t)\|^2 \le -2\mu\,\|u(\cdot,t)\|^2,$$

όπου $\mu := \min_{x,t} a(x,t)$. Συνεπώς

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

οπότε

$$\|u(\cdot,t)\| \le \mathrm{e}^{-\mu t}\|u(\cdot,0)\| \qquad \forall t \in [0,T]. \tag{5.8}$$

Συνεχής εξάρτηση από τον μη ομογενή όρο. Πολλαπλασιάζοντας τη Δ.Ε. στο (5.1) επί u, ολοκληρώνοντας στο $[0,1]$, ολοκληρώνοντας κατά μέρη και χρησιμοποιώντας τις συνοριακές συνθήκες και την (5.2) λαμβάνουμε

$$\frac{1}{2}\frac{d}{dt}\|u(\cdot,t)\|^2 \leq -\min_{x,t} a(x,t)\|u_x(\cdot,t)\|^2 + (f(\cdot,t),u(\cdot,t)). \tag{5.9}$$

Συνεπώς, σύμφωνα με την ανισότητα των Poincaré–Friedrichs,

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

ή

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

οπότε

$$\|u(\cdot,t)\|^2 \leq \|u(\cdot,0)\|^2 + \frac{1}{2\mu}\int_0^t \|f(\cdot,s)\|^2 ds \qquad \forall t \in [0,T], \tag{5.10}$$

όπου $\mu = \min_{x,t} a(x,t)$. Αν u_1 και u_2 είναι λύσεις των προβλημάτων της μορφής (5.1) 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 την (5.10) ότι

$$\|u_1(\cdot,t) - u_2(\cdot,t)\|^2 \leq \frac{1}{2\mu}\int_0^t \|f_1(\cdot,s) - f_2(\cdot,s)\|^2 ds \qquad \forall t \in [0,T]. \tag{5.11}$$

Παρατήρηση 5.1 (i). Τόσο από την (5.7) όσο και από την (5.11) έπεται ότι το πρόβλημα (5.1) έχει το πολύ μία 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 (5.10) 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 (5.1) 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

$$\begin{cases} u_t = (au_x)_x - cu + f & \text{στο } [0,1]\times[0,T], \\ u_x(0,\cdot) = u_x(1,\cdot) = 0 & \text{στο } [0,T], \\ u(\cdot,0) = u_0 & \text{στο } [0,1]. \end{cases} \tag{5.12}$$

Οι συνθήκες ομαλότητας των a, c και f καθώς και του προσήμου της a είναι οι ίδιες με εκείνες για το πρόβλημα (5.1), για τη c όμως, πέραν της (5.2), υποθέτουμε τώρα ότι

$$c^\star := \min_{0\le t\le T} \int_0^1 c(x,t)dx > 0. \tag{5.13}$$

Όπως στην περίπτωση του προβλήματος (5.1), έτσι και τώρα ο περιορισμός αυτός στη c δεν είναι ουσιαστικός, βλ. την (5.3). Πολλαπλασιάζοντας τη Δ.Ε. στο (5.12) επί u, ολοκληρώνοντας στο $[0,1]$ και ολοκληρώνοντας κατά μέρη λαμβάνουμε

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

Σύμφωνα με τις υποδείξεις στις Ασκήσεις 3.9 και 3.10, έχουμε λοιπόν

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

με μια θετική σταθερά C, και συμπεραίνουμε εύκολα ότι

$$\|u(\cdot,t)\|^2 \le \|u(\cdot,0)\|^2 + C_1\int_0^t \|f(\cdot,s)\|^2 ds \qquad \forall t\in[0,T]. \tag{5.14}$$

Από την (5.14) έπεται η συνεχής εξάρτηση της λύσεως του προβλήματος (5.12), τόσο από τα αρχικά δεδομένα όσο και από τον μη ομογενή όρο.

(iv). Θεωρούμε τώρα το αντίστοιχο του προβλήματος (5.1) για μια πιο γενική Δ.Ε.

$$\begin{cases} u_t = (au_x)_x - bu_x - cu + f & \text{στο } [0,1]\times[0,T],\\ u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T],\\ u(\cdot,0) = u_0 & \text{στο } [0,1],\end{cases} \tag{5.15}$$

με τις ίδιες συνθήκες ομαλότητας για τις a, c, f και $a(x,t) > 0$ και υποθέτουμε τώρα ότι η b είναι συνεχώς παραγωγίσιμη ως προς την πρώτη μεταβλητή και

$$c(x,t) - \frac{1}{2}b_x(x,t) \ge 0 \quad \forall (x,t)\in[0,1]\times[0,T], \tag{5.16}$$

βλ. την (3.2). Χρησιμοποιώντας το γεγονός ότι

$$\begin{aligned}\int_0^1 b(x,t)u_x(x,t)u(x,t)dx &= \frac{1}{2}\int_0^1 b(x,t)\{[u(x,t)]^2\}_x dx\\ &= -\frac{1}{2}\int_0^1 b_x(x,t)[u(x,t)]^2 dx\end{aligned}$$

μπορούμε εύκολα να αποδείξουμε και γι' αυτό το πρόβλημα εκτιμήσεις της μορφής (5.7), (5.8) και (5.11). □

5.2 Ημιδιακριτοποίηση

Έστω $T > 0$. Για να απλοποιήσουμε κάπως τον συμβολισμό, θεωρούμε το εξής πρόβλημα αρχικών και συνοριακών συνθηκών

$$\begin{cases} u_t = u_{xx} + f & \text{στο } [0,1] \times [0,T], \\ u(0,\cdot) = u(1,\cdot) = 0 & \text{στο } [0,T], \\ u(\cdot,0) = u^0 & \text{στο } [0,1]. \end{cases} \tag{5.17}$$

Στη διακριτοποίηση αυτού του προβλήματος με τη μέθοδο των πεπερασμένων στοιχείων εμφανίζονται όλα τα ουσιαστικά σημεία της ανάλυσης· η γενίκευση για το πρόβλημα (5.15), φερ' ειπείν, δεν παρουσιάζει επιπρόσθετες δυσκολίες.

Εδώ $f : [0,1] \times [0,T] \to \mathbb{R}$ είναι μια δεδομένη συνάρτηση και $u^0 : [0,1] \to \mathbb{R}$ είναι μια δεδομένη αρχική τιμή. Υποθέτουμε ότι τα δεδομένα είναι ομαλά και συμβατά, ώστε το πρόβλημα (5.17) να έχει μια λύση, η οποία μάλιστα είναι αρκετά ομαλή ώστε να έχουν έννοια οι νόρμες της u οι οποίες θα χρησιμοποιηθούν στη συνέχεια.

Έστω $v \in H_0^1(0,1)$. Παίρνοντας στη διαφορική εξίσωση του (5.17) το L^2 εσωτερικό γινόμενο και των δύο μελών με τη v, ολοκληρώνοντας στον πρώτο όρο στο δεξιό μέλος κατά μέρη και χρησιμοποιώντας τις συνοριακές συνθήκες, γράφουμε το πρόβλημα (5.17) σε *μεταβολική* μορφή ως εξής: Ζητείται $u(\cdot,t) \in H_0^1(0,1), t \in [0,T]$, τέτοια ώστε

$$\begin{cases} \big(u_t(\cdot,t),v\big) + \big(u_x(\cdot,t),v'\big) = \big(f(\cdot,t),v\big) \quad \forall v \in H_0^1 \quad \forall t \in [0,T], \\ u(\cdot,0) = u^0. \end{cases} \tag{5.18}$$

Η διακριτοποίηση του προβλήματος (5.17) με τη μέθοδο των πεπερασμένων στοιχείων βασίζεται στη μεταβολική διατύπωση (5.18).

Σε αυτή την παράγραφο θα μελετήσουμε τη διακριτοποίηση στον χώρο του προβλήματος (5.17) 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 (5.17). 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 (5.17), πρέπει να διακριτοποιήσουμε ως προς τον χρόνο το ημιδιακριτό πρόβλημα. Αυτό γίνεται με τα λεγόμενα *πλήρως διακριτά σχήματα*. Στην επόμενη παράγραφο θα μελετήσουμε τρία πλήρως διακριτά σχήματα για το πρόβλημα (5.17): 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.

Έστω S_h^r μια οικογένεια υποχώρων του $H_0^1, N_h := \dim S_h^r < \infty$, με την προσεγγιστική ιδιότητα (3.21). Κατ' αναλογίαν προς το πρόβλημα (5.18) ορίζουμε την *ημιδιακριτή λύση* $u_h, u_h(\cdot,t) \in S_h^r, t \in [0,T]$, ως

$$\begin{cases} \big(u_{ht}(\cdot,t),\chi\big) + \big(u_{hx}(\cdot,t),\chi'\big) = \big(f(\cdot,t),\chi\big) \quad \forall \chi \in S_h^r \quad \forall t \in [0,T], \\ u_h(\cdot,0) = u_h^0, \end{cases} \tag{5.19}$$

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. Αφού $u_h(\cdot,t) \in S_h^r, t \in [0,T]$, η u_h γράφεται στη μορφή

$$u_h(x,t) = \sum_{j=1}^{N_h} \alpha_j(t)\varphi_j(x)$$

και το πρόβλημα (5.19) γράφεται ως

$$\begin{cases} \displaystyle\sum_{j=1}^{N_h}(\varphi_j,\varphi_i)\alpha_j'(t) + \sum_{j=1}^{N_h}(\varphi_j',\varphi_i')\alpha_j(t) = (f(\cdot,t),\varphi_i), \quad i = 1,\dots,N_h, \\ \alpha_j(0) = \gamma_j, \quad j = 1,\dots,N_h, \end{cases} \tag{5.20}$$

όπου τα γ_j είναι τέτοια ώστε $u_h^0 = \sum_{j=1}^{N_h} \gamma_j\varphi_j$.

Το (5.20) είναι ένα πρόβλημα αρχικών τιμών για ένα σύστημα συνήθων διαφορικών εξισώσεων. Ο πίνακας μάζας $(\varphi_j,\varphi_i)_{j,i=1,\dots,N_h}$ είναι αντιστρέψιμος, ως πίνακας του Gram, συνεπώς το (5.20) 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 τις δύο πρώτες σχέσεις των (5.18) και (5.19) καταλήγουμε στην

$$\big((u-u_h)_t(\cdot,t),\chi\big) + \big((u-u_h)_x(\cdot,t),\chi'\big) = 0 \quad \forall \chi \in S_h^r \quad \forall t \in [0,T].$$

Τονίζουμε ότι αυτή η σχέση ισχύει για $\chi \in S_h^r$, αφού ο S_h^r είναι υπόχωρος του H_0^1. 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, αφού η $u(\cdot,t)$ δεν είναι γενικά στοιχείο του S_h^r. Αντί λοιπόν να συγκρίνουμε τη u_h απ' ευθείας με τη u, θα τη συγκρίνουμε με μια "κατάλληλη" προβολή της u στον S_h^r. Το κατάλληλη εδώ αναφέρεται σε δύο πράγματα: Αφ' ενός η απόσταση της προβολής από τη u πρέπει να είναι της τάξης $O(h^r)$ στη νόρμα L^2, αφ' ετέρου δε πρέπει να ικανοποιεί μια σχέση αντίστοιχη της (5.19) 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, αφού 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 είναι ούτε αναγκαίος. Ας σημειωθεί ακόμα ότι ο συγκεκριμένος ορισμός της ελλειπτικής προβολής υπαγορεύεται από τη μορφή της εξίσωσης στο πρόβλημα (5.17) και πρέπει να τροποποιείται κατάλληλα για την αντίστοιχη κάθε φορά εξίσωση.

Λήμμα 5.1 (Εκτίμηση για το σφάλμα της ελλειπτικής προβολής.) *Έστω $w \in H^r \cap H_0^1$ και $R_h w$ η ελλειπτική προβολή του w στον S_h^r, η οποία ορίζεται ως*

$$\big((R_h w)', \chi'\big) = (w', \chi') \quad \forall \chi \in S_h^r. \tag{5.21}$$

Η $R_h w$ ορίζεται μονοσήμαντα και ισχύει

$$\|w - R_h w\| + h\|w' - (R_h w)'\| \le C h^r \|w\|_r. \tag{5.22}$$

Απόδειξη. Θέτουμε $\varphi := -w''$, οπότε η w είναι λύση του ελλειπτικού προβλήματος

$$\begin{cases} -w'' = \varphi & \text{στο } [0,1], \\ w(0) = w(1) = 0. \end{cases}$$

Εύκολα διαπιστώνουμε ότι η $R_h w \in S_h^r$ είναι η προσεγγιστική λύση της w, όπως αυτή ορίστηκε στη μελέτη της μεθόδου των πεπερασμένων στοιχείων για το πρόβλημα δύο σημείων, δηλαδή για ένα ελλειπτικό πρόβλημα, στο Κεφάλαιο 3. Από αυτή τη μελέτη γνωρίζουμε ότι η $R_h w$ είναι καλώς ορισμένη και ότι ισχύει η (5.22). □

Ευστάθεια του ημιδιακριτού προβλήματος. Έστω u_h η λύση του προβλήματος (5.19) 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

$$\big(v_{ht}(\cdot,t),\chi\big) + \big(v_{hx}(\cdot,t),\chi'\big) = 0. \tag{5.23}$$

Ευστάθεια στην L^2-νόρμα. Θέτοντας $\chi := v_h$ στην (5.23) 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

$$\|v_h(\cdot,t)\| \le \|v_h(\cdot,0)\|, \quad t \in [0,T]. \tag{5.24}$$

Ευστάθεια στην H^1-νόρμα. Θέτοντας $\chi := v_{ht}$ στην (5.23) 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

$$\|v_{hx}(\cdot,t)\| \le \|v_{hx}(\cdot,0)\|, \quad t \in [0,T]. \tag{5.25}$$

Συνέπεια. Έστω $W(\cdot,t) := R_h u(\cdot,t)$ η ελλειπτική προβολή της λύσης του προβλήματος (5.17). Θα αποδείξουμε συνέπεια του ημιδιακριτού σχήματος (5.18) για τη W. Έχουμε, για $\chi \in S_h^r$,

$$\begin{aligned}
&\big(W_t(\cdot,t),\chi\big) + \big(W_x(\cdot,t),\chi'\big) - \big(f(\cdot,t),\chi\big) = \\
&= \big(W_t(\cdot,t),\chi\big) + \big(u_x(\cdot,t),\chi'\big) - \big(f(\cdot,t),\chi\big) = \\
&= \big(W_t(\cdot,t),\chi\big) - \big(u_t(\cdot,t),\chi\big),
\end{aligned}$$

δηλαδή

$$(5.26)\qquad \big(W_t(\cdot,t),\chi\big) + \big(W_x(\cdot,t),\chi'\big) = \big(f(\cdot,t),\chi\big) - \big(u_t(\cdot,t) - W_t(\cdot,t),\chi\big) \quad \forall\chi\in S_h^r.$$

Η ποσότητα $-(u_t(\cdot,t) - W_t(\cdot,t)$ είναι το *σφάλμα συνέπειας* της W για το ημιδιακριτό πρόβλημα (5.19): ένα μέγεθος της αποτυχίας της W να ικανοποιεί την πρώτη σχέση στην (5.19).

Τώρα ο τελεστής της ελλειπτικής προβολής και η παραγώγιση ως προς t αντιμετατίθενται. Πράγματι έχουμε, για $\chi\in S_h^r$,

$$\begin{aligned}\big((R_hu_t)_x,\chi'\big) &= \big((u_t)_x,\chi'\big) = \big(\frac{\partial}{\partial t}u_x,\chi'\big) = \\ &= \frac{d}{dt}(u_x,\chi') = \frac{d}{dt}\big((R_hu)_x,\chi'\big) = \big(\frac{\partial}{\partial t}((R_hu)_x),\chi'\big) = \\ &= \big(((R_hu)_t)_x,\chi'\big),\end{aligned}$$

δηλαδή $R_hu_t = (R_hu)_t$ ή $R_h\frac{\partial}{\partial t}u = \frac{\partial}{\partial t}R_hu$. Συνεπώς, σύμφωνα με την (5.22), έχουμε

$$(5.27)\qquad \|u_t(\cdot,t) - W_t(\cdot,t)\| \le Ch^r, \quad t\in[0,T].$$

Σύγκλιση. Συνδυάζοντας τώρα ευστάθεια και συνέπεια αποδεικνύουμε σύγκλιση της ημιδιακριτής λύσης προς την ακριβή λύση του προβλήματος (5.17).

Σύγκλιση στην L^2-νόρμα.

Θεώρημα 5.1 (Εκτίμηση του σφάλματος στην L^2-νόρμα.) *Έστω ότι η λύση u του προβλήματος* (5.17) *είναι αρκετά ομαλή. Τότε ισχύει*

$$(5.28)\qquad \max_{0\le t\le T}\|u(\cdot,t) - u_h(\cdot,t)\| \le C\big[\|u^0 - u_h^0\| + h^r\big].$$

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 (5.22),

$$(5.29)\qquad \max_{0\le t\le T}\|\rho(\cdot,t)\| \le Ch^r.$$

Απομένει συνεπώς να εκτιμήσουμε την $\|\vartheta(\cdot,t)\|$. Αφαιρώντας την πρώτη σχέση της (5.19) από την (5.26) έχουμε

$$(5.30)\qquad \big(\vartheta_t(\cdot,t),\chi\big) + \big(\vartheta_x(\cdot,t),\chi'\big) = -\big(\rho_t(\cdot,t),\chi\big) \quad \forall\chi\in S_h^r \quad \forall t\in[0,T].$$

Θέτοντας $\chi := \vartheta(\cdot,t)$ στην (5.30) λαμβάνουμε

$$\frac{1}{2}\frac{d}{dt}\|\vartheta(\cdot,t)\|^2 + \|\vartheta_x(\cdot,t)\|^2 = -\big(\rho_t(\cdot,t),\vartheta(\cdot,t)\big)$$

ή

$$\frac{d}{dt}\|\vartheta(\cdot,t)\|^2 \le \|\rho_t(\cdot,t)\|^2 + \|\vartheta(\cdot,t)\|^2$$

ή

$$\frac{d}{dt}\{\mathrm{e}^{-t}\|\vartheta(\cdot,t)\|^2\} \le \mathrm{e}^{-t}\|\rho_t(\cdot,t)\|^2$$

ή

$$\mathrm{e}^{-t}\|\vartheta(\cdot,t)\|^2 \le \|\vartheta(\cdot,0)\|^2 + \int_0^t \mathrm{e}^{-s}\|\rho_t(\cdot,s)\|^2 ds$$

ή, σύμφωνα με την (5.27),

$$\|\vartheta(\cdot,t)\|^2 \le C\big[\|\vartheta(\cdot,0)\|^2 + h^{2r}\big].$$

Επομένως

$$\|\vartheta(\cdot,t)\| \le C\big[\|\vartheta(\cdot,0)\| + h^r\big]. \tag{5.31}$$

Τώρα $\vartheta(\cdot,0) = W(\cdot,0) - u_h^0 = (u^0 - u_h^0) - \big(u^0 - W(\cdot,0)\big)$, οπότε σύμφωνα με την (5.22),

$$\|\vartheta(\cdot,0)\| \le C\|u^0 - u_h^0\| + Ch^r. \tag{5.32}$$

Από τις (5.31) και (5.32) έπεται ότι

$$\max_{0\le t\le T}\|\vartheta(\cdot,t)\| \le C\big[\|u^0 - u_h^0\| + Ch^r\big]. \tag{5.33}$$

Οι (5.29) και (5.33) δίνουν αμέσως την (5.28). □

Σύγκλιση στην H^1-νόρμα.

Θεώρημα 5.2 (Εκτίμηση του σφάλματος στην H^1-νόρμα.) *Έστω ότι η λύση u του προβλήματος* (5.17) *είναι αρκετά ομαλή. Τότε ισχύει*

$$\max_{0\le t\le T}\|u_x(\cdot,t) - u_{hx}(\cdot,t)\| \le \|(u^0 - u_h^0)'\| + Ch^{r-1}. \tag{5.34}$$

Απόδειξη. Χρησιμοποιούμε τους συμβολισμούς που χρησιμοποιήσαμε στην απόδειξη του Θεωρήματος 5.1. Σύμφωνα με την (5.22) έχουμε κατ' αρχάς

$$\max_{0\le t\le T}\|\rho_x(\cdot,t)\| \le Ch^{r-1}. \tag{5.35}$$

Θέτοντας τώρα $\chi := \vartheta_t(\cdot,t)$ στην (5.30) λαμβάνουμε

$$\|\vartheta_t(\cdot,t)\|^2 + \frac{1}{2}\frac{d}{dt}\|\vartheta_x(\cdot,t)\|^2 = -\big(\rho_t(\cdot,t),\vartheta_t(\cdot,t)\big)$$

ή

$$\|\vartheta_t(\cdot,t)\|^2 + \frac{1}{2}\frac{d}{dt}\|\vartheta_x(\cdot,t)\|^2 \le \frac{1}{2}\|\rho_t(\cdot,t)\|^2 + \frac{1}{2}\|\vartheta_t(\cdot,t)\|^2,$$

οπότε

$$\frac{d}{dt}\|\vartheta_x(\cdot,t)\|^2 \le \|\rho_t(\cdot,t)\|^2,$$

συνεπώς

$$\|\vartheta_x(\cdot,t)\|^2 \le \|\vartheta_x(\cdot,0)\|^2 + \int_0^t \|\rho_t(\cdot,s)\|^2 ds.$$

Χρησιμοποιώντας εδώ την (5.27) έχουμε

$$\|\vartheta_x(\cdot,t)\|^2 \le \|\vartheta_x(\cdot,0)\|^2 + C^2h^{2r}$$

ή

$$\|\vartheta_x(\cdot,t)\| \le \|\vartheta_x(\cdot,0)\| + Ch^r. \tag{5.36}$$

Τώρα,

$$\|\vartheta_x(\cdot,0)\| \le \|(u^0-u_h^0)'\| + \|u_x^0 - W_x(\cdot,0)\|,$$

οπότε, σύμφωνα με την (5.22),

$$\|\vartheta_x(\cdot,0)\| \le \|(u^0-u_h^0)'\| + Ch^{r-1}. \tag{5.37}$$

Από τις (5.36) και (5.37) έπεται ότι

$$\max_{0\le t\le T}\|\vartheta_x(\cdot,t)\| \le \|(u^0-u_h^0)'\| + Ch^{r-1}. \tag{5.38}$$

Οι (5.35) και (5.38) δίνουν αμέσως την (5.34). □

Παρατήρηση 5.2 (Υπερσύγκλιση.) Για $u_h^0 := W(\cdot,0)$ έχουμε $\vartheta(\cdot,0) = 0$, και η (5.36) δίνει το εξής αποτέλεσμα *"υπερσύγκλισης"*

$$\max_{0\le t\le T} \|\vartheta_x(\cdot,t)\| \le Ch^r. \qquad (5.39)$$ □

Παρατήρηση 5.3 (Μεταβλητοί συντελεστές.) Ας υποθέσουμε ότι θέλουμε να μελετήσουμε τη διακριτοποίηση του προβλήματος

$$\begin{cases} u_t = (au_x)_x + f & \text{στο } [0,1]\times[0,T],\\ u(\cdot,0) = u(\cdot,1) = 0 & \text{στο } [0,T],\\ u(\cdot,0) = u^0 & \text{στο } [0,1], \end{cases} \qquad (5.40)$$

με τη μέθοδο των πεπερασμένων στοιχείων. Η συνάρτηση $a : [0,1]\times[0,T]$ υποτίθεται ομαλή και θετική, $a(x,t) \ge \alpha > 0$. Στην περίπτωση που η a είναι ανεξάρτητη του t, τα μέχρι τώρα αποτελέσματα γενικεύονται απρόσκοπτα. Στη γενική περίπτωση απαιτούνται κάποιες τροποποιήσεις σε όσα αναφέρονται στην ελλειπτική προβολή.

Ορίζουμε τώρα τον τελεστή ελλειπτικής προβολής $R_h(t) : H_0^1(0,1) \to S_h^r$, $t \in [0,T]$, ως εξής

$$\big(a(\cdot,t)(R_h(t)v)_x, \chi'\big) = \big(a(\cdot,t)v_x, \chi'\big) \quad \forall \chi \in S_h^r,\ \ t\in[0,T]. \qquad (5.41)$$

Τότε, προφανώς,

$$\|v - R_h(t)v\| + h\|v' - (R_h(t)v)_x\| \le ch^r. \qquad (5.42)$$

Παραγωγίζοντας στην (5.41) ως προς t παίρνουμε, για $t \in [0,T]$,

$$\big(a(\cdot,t)(R_h'(t)v)_x, \chi'\big) + \big(a_t(\cdot,t)(R_h(t)v)_x, \chi'\big) = \big(a_t(\cdot,t)v_x, \chi'\big) \quad \forall \chi \in S_h^r, \qquad (5.43)$$

δηλαδή

$$\big(a(\cdot,t)(R_h'(t)v)_x, \chi'\big) = \big(a_t(\cdot,t)(v - R_h(t)v)_x, \chi'\big) \quad \forall \chi \in S_h^r,\ \ t\in[0,T]. \qquad (5.44)$$

Συνδυάζοντας αυτή τη σχέση με την (5.42) λαμβάνουμε αμέσως

$$\|(R_h'(t)v)_x\| \le ch^{r-1}. \qquad (5.45)$$

Προκειμένου να οδηγηθούμε στην αντίστοιχη εκτίμηση βέλτιστης τάξης στην L^2-νόρμα, θεωρούμε το πρόβλημα δύο σημείων

$$\begin{cases} -\big(a(\cdot,t)\varphi_x\big)_x = R_h'(t)v & \text{στο } [0,1],\\ \varphi(0) = \varphi(1) = 0. \end{cases} \qquad (5.46)$$

Τότε

$$\big(a(\cdot,t)\varphi_x, z'\big) = \big(R_h'(t)v, z\big) \quad \forall z \in H_0^1(0,1). \tag{5.47}$$

Με $z := R_h'(t)v$ έχουμε συνεπώς, χρησιμοποιώντας και την (5.44), για οποιοδήποτε $\chi \in S_h^r$,

$$\begin{aligned}\|R_h'(t)v\|^2 &= \big(a(\cdot,t)\varphi_x, (R_h'(t)v)_x\big)\\ &= \big(a(\cdot,t)(R_h'(t)v)_x, \varphi_x - \chi'\big) + \big(a(\cdot,t)(R_h'(t)v)_x, \chi'\big)\\ &= \big(a(\cdot,t)(R_h'(t)v)_x, (\varphi-\chi)_x\big) + \big(a_t(\cdot,t)(v - R_h(t)v)_x, \chi'\big),\end{aligned}$$

δηλαδή

$$\begin{aligned}\|R_h'(t)v\|^2 =&\big(a(\cdot,t)(R_h'(t)v)_x, (\varphi-\chi)_x\big) - \big(a_t(\cdot,t)(v - R_h(t)v)_x, (\varphi-\chi)_x\big)\\ &+ \big(a_t(\cdot,t)(v - R_h(t)v)_x, \varphi_x\big).\end{aligned} \tag{5.48}$$

Εύκολα διαπιστώνουμε ότι και οι τρεις όροι στο δεξιό μέλος της (5.48) φράσσονται από $Ch^r\|\varphi\|_2$, οπότε, σύμφωνα με την εκτίμηση ελλειπτικής ομαλότητας, από $Ch^r\|R_h'(t)v\|$. Για τους δύο πρώτους όρους αυτό έπεται από τις (5.45) και (5.42), αντίστοιχα, καθώς και την προσεγγιστική ιδιότητα (3.21). Για τον τρίτο όρο απαιτείται πρώτα μία ολοκλήρωση κατά μέρη ώστε ο όρος $v - R_h(t)v$ να εμφανίζεται χωρίς την παράγωγο ως προς x και εν συνεχεία το αποτέλεσμα έπεται από την (5.42). Συμπεραίνουμε επομένως ότι

$$\|R_h'(t)v\| \le ch^r. \tag{5.49}$$

Ορίζουμε τώρα την ελλειπτική προβολή W της λύσης u της εξίσωσης (5.40) ως $W(\cdot,t) := R_h(t)u(\cdot,t)$. Τότε

$$W_t(\cdot,t) = R_h(t)u_t(\cdot,t) + R_h'(t)u(\cdot,t),$$

συνεπώς

$$u_t(\cdot,t) - W_t(\cdot,t) = \big[u_t(\cdot,t) - R_h(t)u_t(\cdot,t)\big] - R_h'(t)u(\cdot,t). \tag{5.50}$$

Σύμφωνα με τις (5.42) και (5.49) και οι δύο όροι στο δεξιό μέλος είναι στην L^2-νόρμα της τάξης h^r, συνεπώς

$$\|u_t(\cdot,t) - W_t(\cdot,t)\| \le ch^r, \quad t \in [0,T], \tag{5.51}$$

όπως και στην περίπτωση σταθερών συντελεστών, βλ. την (5.27). □

5.3 Πλήρως διακριτά σχήματα

Σε αυτή την παράγραφο θα μελετήσουμε τρία πλήρως διακριτά σχήματα για το πρόβλημα (5.17): την πεπλεγμένη μέθοδο του Euler, τη μέθοδο των Crank–Nicolson και τη μέθοδο του Euler.

5.3.1 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 (5.19) 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

$$\text{(5.52)} \qquad \begin{cases} (\partial U^n, \chi) + (U_x^{n+1}, \chi') = \big(f(\cdot, t^{n+1}), \chi\big) & \forall \chi \in S_h^r, \quad n = 0, \dots, N-1, \\ U^0 := u_h^0. \end{cases}$$

Έστω ότι έχουμε υπολογίσει το U^n. Αν χρησιμοποιήσουμε μια βάση του S_h^r, τότε διαπιστώνουμε 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 (5.52).

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 (5.52), δηλαδή είναι 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

$$\text{(5.53)} \qquad (\partial V^n, \chi) + (V_x^{n+1}, \chi') = 0 \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1.$$

Ευστάθεια στην L^2−νόρμα. Θέτοντας $\chi := V^{n+1}$ στην (5.53) λαμβάνουμε

$$(\partial V^n, V^{n+1}) + \|V_x^{n+1}\|^2 = 0,$$

οπότε

$$(\partial V^n, V^{n+1}) \le 0$$

ή

$$\|V^{n+1}\|^2 \le (V^n, V^{n+1}),$$

συνεπώς

$$\|V^{n+1}\|^2 \le \|V^n\|\, \|V^{n+1}\|,$$

οπότε

$$\|V^{n+1}\| \le \|V^n\|, \quad n = 0, \dots, N-1. \tag{5.54}$$

Ιδιαίτερα λοιπόν έχουμε

$$\max_{1\le n\le N} \|V^n\| \le \|V^0\|. \tag{5.55}$$

Παρατήρηση 5.4 (Ύπαρξη και μοναδικότητα.) Για δεδομένο $n,$ αν θεωρήσουμε το αντίστοιχο ομογενές γραμμικό σύστημα του (5.52), θέσουμε δηλαδή $U^n = 0$ και $f = 0,$ τότε σύμφωνα με την (5.54) θα έχουμε $U^{n+1} = 0.$ Σημειώνουμε ότι πρόκειται για τετραγωνικά γραμμικά συστήματα, με ίδιο πλήθος εξισώσεων και αγνώστων, και ότι σε τέτοιες περιπτώσεις η μοναδικότητα της λύσης συνεπάγεται και ύπαρξη της λύσης. Έτσι διαπιστώνουμε πάλι ότι τα $U^1, \dots, U^N$ ορίζονται καλώς. □

Ευστάθεια στην H^1-νόρμα. Θέτοντας $\chi := \partial V^n$ στην (5.53) λαμβάνουμε

$$\|\partial V^n\|^2 + (V_x^{n+1}, \partial V_x^n) = 0$$

και συμπεραίνουμε εύκολα ότι

$$\|V_x^{n+1}\| \le \|V_x^n\|, \quad n = 0, \dots, N-1. \tag{5.56}$$

Ιδιαίτερα λοιπόν έχουμε

$$\max_{1\le n\le N} \|V_x^n\| \le \|V_x^0\|. \tag{5.57}$$

Συνέπεια. Όπως και στην περίπτωση του ημιδιακριτού προβλήματος, θα αποδείξουμε συνέπεια του σχήματος (5.52) για την ελλειπτική προβολή W της λύσης u του προβλήματος (5.17), $W(\cdot,t) := R_h u(\cdot,t).$ Θέτουμε $W^n := W(\cdot,t^n), n = 0, \dots, N,$ και έχουμε, για $\chi \in S_h^r,$ σύμφωνα με την (5.21),

$$\begin{aligned}(\partial W^n,\chi) &+ (W_x^{n+1}, \chi') - \big(f(\cdot,t^{n+1}),\chi\big)\\ &= (\partial W^n,\chi) + (u_x^{n+1}, \chi') - \big(f(\cdot,t^{n+1}),\chi\big).\end{aligned}$$

Λαμβάνοντας τώρα υπ' όψιν την (5.18) έχουμε

$$(5.58)\qquad \begin{aligned}(\partial W^n,\chi)+(W_x^{n+1},\chi')&=\big(f(\cdot,t^{n+1}),\chi\big)\\&+\big(\partial W^n-u_t(\cdot,t^{n+1}),\chi\big)\quad\forall\chi\in S_h^r,\quad n=0,\dots,N-1.\end{aligned}$$

Θα πρέπει λοιπόν να εκτιμήσουμε το *σφάλμα συνέπειας* E^n της W για το σχήμα (5.52), $E^n:=\partial W^n-u_t(\cdot,t^{n+1})$. Το σφάλμα συνέπειας είναι ένα μέγεθος της αποτυχίας της W να ικανοποιεί την πρώτη σχέση του προσεγγιστικού σχήματος (5.52). Θέτουμε $E_1^n:=\partial W^n-\partial u^n$ και $E_2^n:=\partial u^n-u_t(\cdot,t^{n+1})$ και έχουμε $E^n=E_1^n+E_2^n$. 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 (5.27), έχουμε

$$(5.59)\qquad \max_{0\le n\le N-1}\|E_1^n\|\le Ch^r.$$

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

$$(5.60)\qquad \max_{0\le n\le N-1}\|E_2^n\|\le Ck.$$

Από τις (5.59) και (5.60) συμπεραίνουμε αμέσως ότι ισχύει η ακόλουθη εκτίμηση για το σφάλμα συνέπειας

$$(5.61)\qquad \max_{0\le n\le N-1}\|E^n\|\le C(k+h^r).$$

Σύγκλιση. Συνδυάζοντας τώρα ευστάθεια και συνέπεια αποδεικνύουμε σύγκλιση των προσεγγίσεων, που δίνει το πεπλεγμένο σχήμα του Euler, προς την ακριβή λύση του προβλήματος (5.17).

Θεώρημα 5.3 (Εκτίμηση του σφάλματος.) *Έστω ότι η λύση u του προβλήματος* (5.17) *είναι αρκετά ομαλή. Τότε για τις προσεγγίσεις $U^0,\dots,U^N$, που δίνονται από την* (5.52), *ισχύει*

$$(5.62)\qquad \max_{0\le n\le N}\|u^n-U^n\|\le\|u^0-u_h^0\|+C(k+h^r).$$

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 (5.22),

$$\max_{0\le n\le N} \|\rho^n\| \le Ch^r. \tag{5.63}$$

Απομένει συνεπώς να εκτιμήσουμε τη $\|\vartheta^n\|$. Αφαιρώντας την πρώτη σχέση της (5.52) από την (5.58) λαμβάνουμε

$$(\partial\vartheta^n,\chi) + (\vartheta_x^{n+1},\chi') = (E^n,\chi) \quad \forall\chi\in S_h^r, \quad n = 0,\dots,N-1. \tag{5.64}$$

Θέτοντας $\chi = \vartheta^{n+1}$ στην (5.64) και πολλαπλασιάζοντας επί k έχουμε

$$\|\vartheta^{n+1}\|^2 - (\vartheta^n,\vartheta^{n+1}) + k\|\vartheta_x^{n+1}\|^2 = k(E^n,\vartheta^{n+1}),$$

οπότε

$$\|\vartheta^{n+1}\|^2 \le \{\|\vartheta^n\| + k\|E^n\|\}\|\vartheta^{n+1}\|$$

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 (5.61),

$$\max_{0\le n\le N}\|\vartheta^n\| \le \|\vartheta^0\| + C(k+h^r). \tag{5.65}$$

Τώρα

$$\|\vartheta^0\| \le \|u^0 - u_h^0\| + \|u^0 - W^0\| \le \|u^0 - u_h^0\| + Ch^r,$$

και η (5.65) δίνει

$$\max_{0\le n\le N}\|\vartheta^n\| \le \|u^0 - u_h^0\| + C(k+h^r). \tag{5.66}$$

Από τις (5.63) και (5.66) έπεται αμέσως η (5.62). □

5.3.2 Η μέθοδος των Crank–Nicolson

Όπως βλέπουμε από την εκτίμηση (5.62), η πεπλεγμένη μέθοδος του Euler έχει τάξη ακρίβειας ένα ως προς την παράμετρο διακριτοποίησης στον 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 τη μέθοδο των Crank–Nicolson για παραβολικές εξισώσεις, με την έννοια ότι οι προσεγγίσεις που δίνει η μέθοδος του Euler μιμούνται καλύτερα τις ιδιότητες εξομάλυνσης (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 $v^0, \dots, v^N \in H_0^1$, και, ανάλογα, $t^{n+\frac{1}{2}} := \frac{1}{2}(t^n + t^{n+1})$. Η $v^{n+\frac{1}{2}}$ είναι δηλαδή ο μέσος όρος των v^n και v^{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

$$\begin{cases} (\partial U^n, \chi) + (U_x^{n+\frac{1}{2}}, \chi') = \big(f(\cdot, t^{n+\frac{1}{2}}), \chi\big) \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1, \\ U^0 := u_h^0. \end{cases} \tag{5.67}$$

Ευστάθεια. Κατ' αναλογίαν προς την (5.53) οδηγούμαστε στη μελέτη του σχήματος

$$(\partial V^n, \chi) + (V_x^{n+\frac{1}{2}}, \chi') = 0 \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1, \tag{5.68}$$

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 (5.68) λαμβάνουμε

$$(\partial V^n, V^{n+\frac{1}{2}}) + \|V_x^{n+\frac{1}{2}}\|^2 = 0,$$

συνεπώς

$$(V^{n+1} - V^n, V^{n+1} + V^n) \le 0$$

ή

$$\|V^{n+1}\|^2 \le \|V^n\|^2,$$

οπότε

$$\|V^{n+1}\| \le \|V^n\|, \quad n = 0, \dots, N-1. \tag{5.69}$$

Ιδιαίτερα λοιπόν έχουμε

$$\max_{1\le n\le N} \|V^n\| \le \|V^0\|. \tag{5.70}$$

Από την (5.69) 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 (5.68) λαμβάνουμε

$$\|\partial V^n\|^2 + (V_x^{n+\frac{1}{2}}, \partial V_x^n) = 0,$$

οπότε

$$\big((V^n + V^{n+1})_x, (V^{n+1} - V^n)_x\big) \le 0$$

ή

$$\|V_x^{n+1}\| \le \|V_x^n\|, \quad n = 0, \dots, N-1. \tag{5.71}$$

Ιδιαίτερα λοιπόν έχουμε

$$\max_{1\le n\le N} \|V_x^n\| \le \|V_x^0\|. \tag{5.72}$$

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 (5.17). Χρησιμοποιώντας την (5.21), για $\chi \in S_h^r$, έχουμε

$$\begin{aligned}(\partial W^n, \chi) + (W_x^{n+\frac{1}{2}}, \chi') - \big(f(\cdot, t^{n+\frac{1}{2}}), \chi\big) &= \\ = (\partial W^n, \chi) + (u_x^{n+\frac{1}{2}}, \chi') - \big(f(\cdot, t^{n+\frac{1}{2}}), \chi\big)&,\end{aligned}$$

οπότε λαμβάνοντας υπ' όψιν την (5.18) έχουμε, για $n = 0, \dots, N-1$,

$$(\partial W^n, \chi) + (W_x^{n+\frac{1}{2}}, \chi') = \big(f(\cdot, t^{n+\frac{1}{2}}), \chi\big) + (E^n, \chi) \quad \forall \chi \in S_h^r, \tag{5.73}$$

με *σφάλμα συνέπειας* E^n της W για το σχήμα (5.67),

$$E^n := (\partial W^n - \partial u^n) + [\partial u^n - u_t(\cdot, t^{n+\frac{1}{2}})] + [u_{xx}(\cdot, t^{n+\frac{1}{2}}) - u_{xx}^{n+\frac{1}{2}}].$$

Το σφάλμα συνέπειας είναι και πάλι ένα μέγεθος της αποτυχίας της W να ικανοποιεί την πρώτη σχέση του σχήματος (5.67). Θέτουμε

$$\begin{aligned} E_1^n &:= \partial W^n - \partial u^n, \\ E_2^n &:= \partial u^n - u_t(\cdot, t^{n+\frac12}), \\ E_3^n &:= u_{xx}(\cdot, t^{n+\frac12}) - u_{xx}^{n+\frac12}, \end{aligned}$$

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 σύμφωνα με την (5.59) έχουμε

$$\max_{1\le n\le N-1} \|E_1^n\| \le Ch^r. \tag{5.74}$$

Επίσης, αναπτύσσοντας κατά Taylor ως προς το σημείο $t^{n+\frac12}$, έχουμε

$$E_2^n = \frac{1}{2k}\Big[\int_{t^n}^{t^{n+\frac12}} (s-t^n)^2 u_{ttt}(\cdot, s)ds + \int_{t^{n+\frac12}}^{t^{n+1}} (s-t^{n+1})^2 u_{ttt}(\cdot, s)ds\Big],$$

επομένως

$$\max_{1\le n\le N-1} \|E_2^n\| \le Ck^2. \tag{5.75}$$

Επί πλέον έχουμε

$$2\big[u(\cdot, t^{n+\frac12}) - u^{n+\frac12}\big] = \int_{t^{n+\frac12}}^{t^{n+1}} (s-t^{n+1})u_{tt}(\cdot, s)ds - \int_{t^n}^{t^{n+\frac12}} (s-t^n)u_{tt}(\cdot, s)ds,$$

και ακριβώς αντίστοιχη σχέση για τη u_{xx}, συνεπώς

$$\max_{1\le n\le N-1} \|E_3^n\| \le Ck^2. \tag{5.76}$$

Από τις (5.74), (5.75) και (5.76) συμπεραίνουμε αμέσως ότι ισχύει η ακόλουθη εκτίμηση για το σφάλμα συνέπειας

$$\max_{1\le n\le N-1} \|E^n\| \le C(k^2 + h^r). \tag{5.77}$$

Σύγκλιση. Συνδυάζοντας τώρα ευστάθεια και συνέπεια αποδεικνύουμε σύγκλιση.

Θεώρημα 5.4 (Εκτίμηση του σφάλματος.) *Έστω ότι η λύση u του προβλήματος* (5.17) *είναι αρκετά ομαλή. Τότε για τις προσεγγίσεις $U^0, \dots, U^N$, που δίνονται από την* (5.67)*, ισχύει*

$$\max_{1\le n\le N} \|u^n - U^n\| \le \|u^0 - u_h^0\| + C(k^2 + h^r). \tag{5.78}$$

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 (5.22),

$$\max_{1\le n\le N} \|\rho^n\| \le Ch^r. \tag{5.79}$$

Απομένει συνεπώς να εκτιμήσουμε τη $\|\vartheta^n\|$. Αφαιρώντας την πρώτη σχέση της (5.67) από την (5.73) λαμβάνουμε

$$(\partial\vartheta^n, \chi) + (\vartheta_x^{n+\frac{1}{2}}, \chi') = (E^n, \chi) \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1. \tag{5.80}$$

Θέτοντας $\chi = \vartheta^{n+\frac{1}{2}}$ στην (5.80) και πολλαπλασιάζοντας επί $2k$ έχουμε

$$\|\vartheta^{n+1}\|^2 - \|\vartheta^n\|^2 + 2k\|\vartheta_x^{n+\frac{1}{2}}\|^2 = k(E^n, \vartheta^n + \vartheta^{n+1}),$$

οπότε

$$\|\vartheta^{n+1}\|^2 - \|\vartheta^n\|^2 \le k\|E^n\|(\|\vartheta^n\| + \|\vartheta^{n+1}\|)$$

ή

$$\|\vartheta^{n+1}\| \le \|\vartheta^n\| + k\|E^n\|, \quad n = 0, \dots, N-1. \tag{5.81}$$

Προχωρώντας τώρα όπως στην απόδειξη του Θεωρήματος 5.3 και λαμβάνοντας υπ' όψιν την (5.77) αποδεικνύουμε ότι

$$\max_{1\le n\le N} \|\vartheta^n\| \le \|u^0 - u_h^0\| + C(k^2 + h^r). \tag{5.82}$$

Από τις (5.79) και (5.82) έπεται αμέσως η (5.78). □

5.3.3 Η μέθοδος του Euler

Οι δύο μέθοδοι που μελετήσαμε μέχρι τώρα, δηλαδή η πεπλεγμένη μέθοδος του Euler και η μέθοδος των Crank–Nicolson, είναι πεπλεγμένες. Αποδείξαμε, μεταξύ άλλων, ότι αμφότερες είναι *απεριόριστα ευσταθείς*, δηλαδή 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. Μελετάται εδώ κυρίως για παιδαγωγικούς λόγους.

Διακριτοποιούμε το ημιδιακριτό πρόβλημα (5.18) ως προς τον 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

$$
\begin{cases} (\partial U^n, \chi) + (U_x^n, \chi') = (f(\cdot, t^n), \chi) \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1, \\ U^0 := u_h^0. \end{cases} \tag{5.83}
$$

Ευστάθεια. Κατ' αναλογίαν προς την (5.53) οδηγούμαστε στη μελέτη του σχήματος

$$
(\partial V^n, \chi) + (V_x^n, \chi') = 0 \quad \forall \chi \in S_h^r, \quad n = 0, \dots, N-1, \tag{5.84}
$$

με $V^n \in S_h^r$, $n = 0, \dots, N$.

Πέραν της προσεγγιστικής ιδιότητας (3.21) υποθέτουμε στο εξής και την ακόλουθη *αντίστροφη ανισότητα* για την οικογένεια $\{S_h^r\}_{0<h\le 1}$

$$
\|\chi'\| \le C_1 h^{-1} \|\chi\| \qquad \forall \chi \in S_h^r, \tag{5.85}
$$

με μια σταθερά C_1 ανεξάρτητη του h, βλ. την Άσκηση 3.27.

Θέτοντας $\chi := V^n$ στην (5.84) λαμβάνουμε

$$
(V^{n+1}, V^n) - \|V^n\|^2 + k\|V_x^n\|^2 = 0
$$

ή

$$
\frac{1}{2}\{\|V^{n+1}\|^2 - \|V^n\|^2 - \|V^{n+1} - V^n\|^2\} + k\|V_x^n\|^2 = 0. \tag{5.86}
$$

Θέτοντας εξ άλλου $\chi := \partial V^n$ στην (5.84) λαμβάνουμε

$$
\|\partial V^n\|^2 + (V_x^n, \partial V_x^n) = 0,
$$

ή

$$
\|V^{n+1} - V^n\|^2 + k(V_x^n, V_x^{n+1} - V_x^n) = 0
$$

οπότε

$$
\|V^{n+1} - V^n\|^2 \le k\|V_x^n\| \, \|V_x^{n+1} - V_x^n\|.
$$

Χρησιμοποιώντας συνεπώς την (5.85) έχουμε

$$
\|V^{n+1} - V^n\|^2 \le C_1 \frac{k}{h} \|V_x^n\| \, \|V^{n+1} - V^n\|,
$$

οπότε

$$\|V^{n+1} - V^n\| \le C_1 \frac{k}{h} \|V_x^n\|. \tag{5.87}$$

Από τις (5.87) και (5.87) λαμβάνουμε

$$\|V^{n+1}\|^2 - \|V^n\|^2 \le (C_1^2 \frac{k}{h^2} - 2)k\|V_x^n\|^2. \tag{5.88}$$

Υποθέτουμε τώρα ότι ικανοποιείται η εξής συνθήκη

$$\frac{k}{h^2} \le \frac{2}{C_1^2}. \tag{5.89}$$

Από τις (5.88) και (5.89) λαμβάνουμε

$$\|V^{n+1}\| \le \|V^n\|, \quad n = 0, \dots, N-1, \tag{5.90}$$

ιδιαίτερα επομένως έχουμε

$$\max_{1 \le n \le N} \|V^n\| \le \|V^0\|. \tag{5.91}$$

Παρατήρηση 5.5 (Πολύ περιοριστική συνθήκη.) Η συνθήκη (5.89), 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

$$\begin{aligned}(\partial W^n, \chi) + (W_x^n, \chi') = (f(\cdot, t^n), \chi) + (E^n, \chi) \quad & \forall \chi \in S_h^r,\\ & n = 0, \dots, N-1,\end{aligned} \tag{5.92}$$

με σφάλμα συνέπειας E^n τέτοιο ώστε

$$\max_{0 \le n \le N-1} \|E^n\| \le C(k + h^r). \tag{5.93}$$

Σύγκλιση. Συνδυάζοντας τώρα ευστάθεια και συνέπεια αποδεικνύουμε σύγκλιση.

Θεώρημα 5.5 (Εκτίμηση του σφάλματος.) *Έστω ότι η λύση u του προβλήματος* (5.17) *είναι αρκετά ομαλή, και ότι ισχύουν οι* (5.85) *και*(5.89)*. Τότε για τις προσεγγίσεις $U^0, \dots,$ U^N, που δίνονται από την* (5.83)*, ισχύει*

$$\max_{0 \le n \le N} \|u^n - U^n\| \le C\|u^0 - u_h^0\| + (k + h^r). \tag{5.94}$$

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 (5.22),

$$\max_{0\le n\le N} \|\rho^n\| \le Ch^r. \tag{5.95}$$

Μένει συνεπώς να εκτιμήσουμε την $\|\vartheta^n\|$. Αφαιρώντας την πρώτη σχέση της (5.83) από την (5.92) λαμβάνουμε

$$(\partial\vartheta^n, \chi) + (\vartheta^n_x, \chi') = (E^n, \chi) \quad \forall \chi \in S^r_h, \quad n = 0, \dots, N-1. \tag{5.96}$$

Θέτοντας $\chi := \vartheta^n$ στην (5.96) λαμβάνουμε

$$(\vartheta^{n+1}, \vartheta^n) - \|\vartheta^n\|^2 + k\|\vartheta^n_x\|^2 = k(E^n, \vartheta^n)$$

ή

$$\frac{1}{2}\big[\|\vartheta^{n+1}\|^2 - \|\vartheta^n\|^2 - \|\vartheta^{n+1} - \vartheta^n\|^2\big] + k\|\vartheta^n_x\|^2 = k(E^n, \vartheta^n). \tag{5.97}$$

Θέτοντας εξ άλλου $\chi := \partial\vartheta^n$ στην (5.96) 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 (5.85),

$$\|\vartheta^{n+1} - \vartheta^n\|^2 \le C_1\frac{k}{h}\|\vartheta^n_x\|\,\|\vartheta^{n+1} - \vartheta^n\| + k\|E^n\|\,\|\vartheta^{n+1} - \vartheta^n\|$$

ή

$$\|\vartheta^{n+1} - \vartheta^n\| \le C_1\frac{k}{h}\|\vartheta^n_x\| + k\|E^n\|,$$

συνεπώς

$$\|\vartheta^{n+1} - \vartheta^n\|^2 \le C_1^2\frac{k^2}{h^2}\|\vartheta^n_x\|^2 + k^2\|E^n\|^2 + 2C_1\frac{k^2}{h}\|\vartheta^n_x\|\,\|E^n\|. \tag{5.98}$$

Από τις (5.97) και (5.98), χρησιμοποιώντας πάλι την (5.85), λαμβάνουμε

$$\begin{aligned}\|\vartheta^{n+1}\|^2 - \|\vartheta^n\|^2 \le{}& (C_1^2\frac{k^2}{h^2} - 2k)\|\vartheta^n_x\|^2 + k^2\|E^n\|^2 + 2k(E^n, \vartheta^n)\\ &+ 2C_1^2\frac{k^2}{h^2}\|\vartheta^n\|\,\|E^n\|\end{aligned}$$

οπότε

$$\|\vartheta^{n+1}\|^2 \le (1+ck)\|\vartheta^n\|^2 + Ck\|E^n\|^2.$$

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 (5.93) λαμβάνουμε

$$\max_{0\le n\le N}\|\vartheta^n\|^2 \le C\|\vartheta^0\|^2 + C(k+h^r)^2$$

ή

$$\max_{0\le n\le N}\|\vartheta^n\| \le C\|\vartheta^0\| + C(k+h^r). \tag{5.99}$$

Από τις (5.95) και (5.99) έπεται εύκολα η (5.94). □

Ασκήσεις

5.1 (Η ανισότητα του Gronwall σε ολοκληρωτική μορφή.) Έστω φ μια συνεχής συνάρτηση στο διάστημα $[0,T]$, και $\alpha,\beta\in\mathbb{R}$ με $\beta\ge 0$. Αν ισχύει

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

αποδείξτε ότι

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

[*Υπόδειξη*: Θεωρήστε ένα θετικό ε και αποδείξτε ότι η συνάρτηση $\psi, \psi(t) := (\alpha+\varepsilon)\,\mathrm{e}^{\beta t}$, $t\in[0,T]$, ικανοποιεί την ολοκληρωτική εξίσωση

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

Προφανώς $\varphi(0) < \psi(0)$. Έστω t_0 ο μικρότερος αριθμός στο διάστημα $[0,T]$ για τον οποίον $\varphi(t_0) = \psi(t_0)$. Αποδείξτε ότι αυτό δεν είναι δυνατόν, γιατί οδηγεί στη σχέση $\varphi(t_0) < \psi(t_0)$.]

5.2 (Γενίκευση της Άσκησης 5.1.) Με τους συμβολισμούς της προηγούμενης Άσκησης, υποθέτουμε ότι

$$\varphi(t) \le \alpha + \int_0^t h(s)\varphi(s)ds \quad \forall t \in [0,T],$$

όπου h μια συνεχής συνάρτηση στο $[0,T]$ η οποία λαμβάνει μόνο μη αρνητικές τιμές. Αποδείξτε ότι

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

[*Υπόδειξη*: Θεωρήστε ένα θετικό ε και αποδείξτε ότι η συνάρτηση ψ, $\psi(t) := (\alpha+\varepsilon)\, \mathrm{e}^{\int_0^t h(s)\, ds}$, $t \in [0,T]$, ικανοποιεί την ολοκληρωτική εξίσωση

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

Εργαστείτε όπως στην Άσκηση 5.2.]

5.3 (Η ανισότητα του Gronwall σε διαφορική μορφή.) Υποθέτουμε ότι η συνάρτηση φ της Άσκησης 5.1 είναι συνεχώς παραγωγίσιμη στο διάστημα $[0,T]$ και ικανοποιεί την ανισότητα

$$\varphi'(t) \le \beta\varphi(t) \quad \forall t \in [0,T].$$

Αποδείξτε ότι

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

[*Υπόδειξη*: Αποδείξτε ότι

$$\varphi(t) \le \varphi(0) + \beta \int_0^t \varphi(s)ds \quad \forall t \in [0,T]$$

και χρησιμοποιήστε την Άσκηση 5.1. Εναλλακτικός τρόπος απόδειξης: Γράψτε τη διαφορική ανίσωση στη μορφή

$$\big(\mathrm{e}^{-\beta s}\varphi(s)\big)' \le 0$$

και ολοκληρώστε από 0 μέχρι t.]

5.4 (Γενίκευση της Άσκησης 5.3.) Υποθέτουμε ότι μια συνάρτηση φ είναι συνεχώς παραγωγίσιμη στο διάστημα $[0,T]$ και ικανοποιεί την ανισότητα

$$\varphi'(t) \le h(t)\varphi(t) \quad \forall t \in [0,T],$$

όπου h μια συνεχής συνάρτηση στο $[0,T]$, η οποία λαμβάνει μόνο μη αρνητικές τιμές. Αποδείξτε ότι

$$\varphi(t) \le \varphi(0)\, \mathrm{e}^{\int_0^t h(s)ds} \quad \forall t \in [0,T].$$

5.5 (Διακριτή ανισότητα του Gronwall, αντίστοιχη της διαφορικής μορφής: ομοιόμορφος διαμερισμός.) 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

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

5.6 (Διακριτή ανισότητα του Gronwall, αντίστοιχη της διαφορικής μορφής: μη ομοιόμορφος διαμερισμός.) Έστω $T > 0,\ 0 = t^0 < t^1 < \cdots < t^N = T$ ένας διαμερισμός του $[0,T]$, $h_n := t^{n+1} - t^n$, $n = 0, \dots, N-1$. Αν $\alpha_0, \dots, \alpha_N$ και $\varepsilon_0, \dots, \varepsilon_{N-1}$ μη αρνητικοί αριθμοί τέτοιοι ώστε με $\gamma > 0$

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

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

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

και ότι

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

5.7 (Διακριτή ανισότητα του Gronwall, αντίστοιχη της ολοκληρωτικής μορφής.) Έστω $T > 0,\ N \in \mathbb{N}$ και $h := \frac{T}{N}$. Αν $\alpha_0, \dots, \alpha_N$ και E μη αρνητικοί αριθμοί τέτοιοι ώστε με $\gamma > 0$ να ισχύει

$$\alpha_{n+1} \le E + \gamma h \sum_{\ell=0}^{n} \alpha_\ell \qquad n = 0, \dots, N-1,$$

αποδείξτε ότι

$$\max_{1\le n\le N} \alpha_n \le C(h\alpha_0 + E)$$

με μια σταθερά C ανεξάρτητη του h.

[*Υπόδειξη*: Έστω $\varphi_n := \gamma h \sum_{\ell=0}^{n} \alpha_\ell$. Τότε $\varphi_{n+1} - \varphi_n = \gamma h \alpha_{n+1}$, συνεπώς

$$\varphi_{n+1} \leq \gamma h E + (1+\gamma k)\varphi_n, \quad n = 0, \dots, N-1.$$

Χρησιμοποιήστε την Άσκηση 5.6 για να εκτιμήσετε τα φ_n, και την

$$\alpha_{n+1} \leq E + \varphi_n$$

για να εκτιμήσετε τα α_n.]

5.8 Θεωρούμε το πρόβλημα

$$(\star) \qquad \begin{cases} u_t = (au_x)_x - bu_x - cu & \text{στο } [0,1]\times[0,T], \\ u(0,\cdot) = u(1,\cdot) = 0 & \text{στο } [0,T], \\ u(\cdot,0) = u_0 & \text{στο } [0,1], \end{cases}$$

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

$$(\bullet) \qquad \max_{x,t} |b(x,t)|^2 \leq 4 \min_{x,t} a(x,t)\ \min_{x,t} c(x,t),$$

αποδείξτε ότι

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

5.9 Θεωρούμε το πρόβλημα $(\star)$ της Άσκησης 5.8, χωρίς όμως τώρα τη συνθήκη $(\bullet)$. Αποδείξτε ότι

$$\frac{1}{2}\frac{d}{dt}\|u(\cdot,t)\|^2 \leq C\|u(\cdot,t)\|^2 \quad \forall t \in [0,T]$$

και οδηγηθείτε στο συμπέρασμα ότι

$$\|u(\cdot,t)\| \leq \mathrm{e}^{Ct}\,\|u_0\| \quad \forall t \in [0,T],$$

με μια θετική σταθερά C.

5.10 Θεωρούμε το πρόβλημα $(\star)$ της Άσκησης 5.8, χωρίς τη συνθήκη $(\bullet)$ και πάλι. Αποδείξτε ότι

$$\frac{1}{2}\frac{d}{dt}\int_a^b a(u_x)^2 dx \leq C\|u_x(\cdot,t)\|^2 \quad \forall t \in [0,T]$$

και

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

με σταθερές C και $\tilde{C}$.

[*Υπόδειξη*: Πολλαπλασιάστε τη Δ.Ε. στο $(\star)$ επί u_t, ολοκληρώστε στο $[0,1]$, ολοκληρώστε κατά μέρη, χρησιμοποιήστε την αριθμητική–γεωμετρική ανισότητα και την ανισότητα των Poincaré–Friedrichs.]

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

$$\begin{cases} u_t = \mathrm{i}\,\alpha u_{xx} + \mathrm{i}\,\beta(x,t)u & \text{στο } [0,1]\times[0,T],\\ u(a,\cdot) = u(b,\cdot) = 0 & \text{στο } [0,T],\\ u(\cdot,0) = u_0 & \text{στο } [0,1], \end{cases}$$

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.]

5.12 Θεωρούμε το ακόλουθο πρόβλημα αρχικών και συνοριακών συνθηκών

$$\begin{cases} u_t = u_{xx} - u + f & \text{στο } [0,1]\times[0,T],\\ u_x(a,\cdot) = u_x(b,\cdot) = 0 & \text{στο } [0,T],\\ u(\cdot,0) = u^0 & \text{στο } [0,1]. \end{cases}$$

Θεωρήστε έναν κατάλληλο χώρο πεπερασμένων στοιχείων για αυτό το πρόβλημα, βλ. την Άσκηση 3.18. Ποια μορφή παίρνουν για αυτό το πρόβλημα οι μέθοδοι πεπερασμένων στοιχείων Euler, πεπλεγμένη Euler, και Crank–Nicolson. Αποδείξτε εκτιμήσεις αντίστοιχες αυτών που δόθηκαν σε αυτό το Κεφάλαιο.

5.13 Θεωρούμε το πρόβλημα της Άσκησης 5.11. 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*.

5.14 Έστω $\vartheta \in [0,1]$. Θεωρούμε την εξής μέθοδο για τη διακριτοποίηση του ημιδιακριτού προβλήματος (5.19) στο χρόνο

$$\begin{cases} (\partial U^n,\chi) + (\vartheta U_x^{n+1} + (1-\vartheta)U_x^n,\chi') =\\ \qquad\qquad = (\vartheta f(\cdot,t^{n+1}) + (1-\vartheta)f(\cdot,t^n),\chi) \quad \forall \chi\in S_h^r,\\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad n=0,\dots,N-1,\\ U^0 := u_h^0. \end{cases}$$

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*.

5.15 Έστω Ω ένα φραγμένο χωρίο του $\mathbb{R}^d$. Θεωρούμε το πρόβλημα αρχικών και συνοριακών συνθηκών

$$\begin{cases} u_t = \Delta u + f & \text{στο } \Omega \times [0,T], \\ u = 0 & \text{στο } \partial\Omega \times [0,T], \\ u(\cdot,0) = u_0 & \text{στο } \Omega, \end{cases}$$

όπου $u = u(x,t) : \Omega \times [0,T] \to \mathbb{R}$, $\Delta u = u_{x_1 x_1} + \cdots + u_{x_d x_d}$. Οι $f : \Omega \times [0,T] \to \mathbb{R}$ και $u_0 : \Omega \to \mathbb{R}$ είναι εδώ δεδομένες συναρτήσεις. Υποθέτουμε ότι τα δεδομένα είναι ομαλά και συμβατά ώστε το πρόβλημα να έχει μια ομαλή λύση. Θεωρούμε μια οικογένεια $\big(S_h^r\big)_{0<h\le 1}$ υποχώρων του $H_0^1(\Omega)$, πεπερασμένης διάστασης, με την προσεγγιστική ιδιότητα (4.36).

$i.$ Μελετήστε την ημιδιακριτοποίηση στον χώρο του δοθέντος προβλήματος με τη μέθοδο των πεπερασμένων στοιχείων. Ιδιαίτερα, δώστε εκτιμήσεις βέλτιστης τάξης στην L^2-νόρμα για το σφάλμα της ημιδιακριτοποίησης.

$ii.$ Μελετήστε τη διακριτοποίηση του ημιδιακριτού προβλήματος του $i.$ στο χρόνο με την πεπλεγμένη μέθοδο του Euler και δώστε εκτιμήσεις βέλτιστης τάξης στην L^2-νόρμα για το σφάλμα του πλήρως διακριτού σχήματος.

6. Παραβολικές εξισώσεις σε αφηρημένη μορφή

Σε αυτό το κεφάλαιο, θα μελετήσουμε αριθμητικές μεθόδους για παραβολικές εξισώσεις. Θα αρχίσουμε με μερικές γενικές παρατηρήσεις για παραβολικές εξισώσεις σε αφηρημένη μορφή, τόσο γραμμικές όσο και μη γραμμικές. Στη συνέχεια θα μελετήσουμε αριθμητικές μεθόδους, και θα δώσουμε τόσο εκ των προτέρων όσο και εκ των υστέρων εκτιμήσεις των σφαλμάτων. Το ενδιαφέρον μας εδώ επικεντρώνεται στη διακριτοποίηση στο χρόνο. Ως εκ τούτου, δεν έχει ιδιαίτερη σημασία το ελλειπτικό μέρος της εξίσωσης, γιαυτό και τη δίνουμε σε αφηρημένη μορφή. Θα δούμε ορισμένα αποτελέσματα που τα γνωρίζουμε ήδη από το προηγούμενο κεφάλαιο, για μια ειδική περίπτωση· θα γνωρίσουμε όμως και κάποια άλλα πράγματα, φερ' ειπείν εκ των υστέρων εκτιμήσεις του σφάλματος.

6.1 Εισαγωγή

Έστω $(H,(\cdot,\cdot))$ ένας (πραγματικός) χώρος του Hilbert, $A : D(A) \to H$ ένας θετικά ορισμένος, αυτοσυζυγής, γραμμικός τελεστής με πεδίο ορισμού $D(A)$ πυκνό στον H, $f : [0,T] \to H$ μια συνάρτηση και $u^0 \in H$. Θεωρούμε το ακόλουθο πρόβλημα αρχικών τιμών: Να προσδιορισθεί συνάρτηση $u : [0,T] \to D(A)$ τέτοια ώστε

$$\text{(6.1)} \qquad \begin{cases} u'(t) + Au(t) = f(t), \quad 0 < t < T, \\ u(0) = u^0, \end{cases}$$

Υπό αυτές τις υποθέσεις ορίζεται ο τελεστής $A^{1/2}$, βλ. την Άσκηση 6.13 για τον ορισμό της ρίζας ενός τελεστή σε μια απλή περίπτωση, ορισμό που μπορεί να γενικευθεί κατάλληλα. Θέτουμε $V := D(A^{1/2})$, και συμβολίζουμε τις νόρμες στους H και V με $|\cdot|$ και $\|\cdot\|$, $\|v\| := |A^{1/2}v| = (Av,v)^{1/2}$, αντίστοιχα. Είναι γνωστό ότι ο χώρος $(V, \|\cdot\|)$ είναι χώρος του Hilbert. Ταυτίζουμε τώρα τον H με τον δυϊκό του (με το θεώρημα αναπαράστασης του Riesz), και συμβολίζουμε με V' τον δυϊκό του V $(V \subset H \subset V')$. Συμβολίζουμε με $(\cdot,\cdot)$ το δυϊκό ζεύγος μεταξύ των V' και V, δηλαδή χρησιμοποιούμε τον ίδιο συμβολισμό όπως και στο εσωτερικό γινόμενο του H, και με $\|\cdot\|_\star$ τη δυϊκή νόρμα στον V', $\|v\|_\star := |A^{-1/2}v| = (v, A^{-1}v)^{1/2}$. Θυμίζουμε ότι η δυϊκή νόρμα ορίζεται ως

$$\|v\|_\star := \sup_{\substack{w\in V\\ w\neq 0}} \frac{(v,w)}{\|w\|}.$$

Τονίζουμε ότι η $\|\cdot\|_\star$ είναι η μικρότερη δυνατή νόρμα που ικανοποιεί τη γενικευμένη ανισότητα των Cauchy–Schwarz,

$$|(v,w)| \le \|v\|_\star \|w\| \quad \forall v \in V' \quad \forall w \in V. \tag{6.2}$$

6.1.1 Ευστάθεια

Σε αυτή την ενότητα θα αποδείξουμε ευστάθεια των λύσεων του προβλήματος αρχικών τιμών (6.1), ιδιαίτερα συνεχή εξάρτηση των λύσεων από τα αρχικά δεδομένα και τον μη ομογενή όρο.

Λαμβάνοντας στη Δ.Ε. στην (6.1) το εσωτερικό γινόμενο με τη u, έχουμε

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

όπου, για να απλοποιήσουμε κάπως τον συμβολισμό, παραλείψαμε το t. Τώρα

$$(u', u) = \frac{1}{2}\frac{d}{dt}|u|^2 \quad \text{και} \quad (f,u) \le \|f\|_\star \|u\| \le \frac{1}{2}\big(\|f\|_\star^2 + \|u\|^2\big),$$

οπότε συμπεραίνουμε ότι

$$\frac{1}{2}\frac{d}{dt}|u|^2 + \|u\|^2 \le \frac{1}{2}\big(\|f\|_\star^2 + \|u\|^2\big),$$

ή

$$\frac{d}{dt}|u|^2 + \|u\|^2 \le \|f\|_\star^2.$$

Ολοκληρώνοντας αυτή τη σχέση από 0 έως t, $t \le T$, παίρνουμε

$$|u(t)|^2 - |u(0)|^2 + \int_0^t \|u(s)\|^2 ds \le \int_0^t \|f(s)\|_\star^2 ds$$

ή

$$|u(t)|^2 + \int_0^t \|u(s)\|^2 ds \le |u^0|^2 + \int_0^t \|f(s)\|_\star^2 ds, \quad t \in [0,T]. \tag{6.3}$$

Αν τώρα θεωρήσουμε και το πρόβλημα αρχικών τιμών

$$\begin{cases} v'(t) + Av(t) = g(t), & 0 < t < T, \\ v(0) = v^0, \end{cases} \tag{6.4}$$

τότε η διαφορά $u - v$ των λύσεων των προβλημάτων (6.1) και (6.4) αποτελεί, λόγω της γραμμικότητας των εξισώσεων, λύση του προβλήματος αρχικών τιμών

$$\begin{cases} (u-v)'(t) + A(u-v)(t) = (f-g)(t), & 0 < t < T, \\ (u-v)(0) = u^0 - v^0. \end{cases} \tag{6.5}$$

Επομένως, σύμφωνα με την (6.3), ισχύει η ακόλουθη εκτίμηση ευστάθειας αναφορικά τόσο με τις αρχικές τιμές όσο και με τους μη ομογενείς όρους

$$|(u-v)(t)|^2 + \int_0^t \|(u-v)(s)\|^2 ds \le |u^0 - v^0|^2 + \int_0^t \|(f-g)(s)\|_\star^2 ds, \quad t \in [0,T]. \tag{6.6}$$

Ιδιαίτερα από αυτή τη σχέση έπεται η *μοναδικότητα* ομαλών λύσεων του προβλήματος (6.1).

6.1.2 Ευστάθεια: Μη γραμμικές εξισώσεις

Σε αυτή την ενότητα θα αποδείξουμε ευστάθεια των λύσεων του προβλήματος αρχικών τιμών για μη γραμμικές εξισώσεις, βλ. την (6.7) στη συνέχεια, ιδιαίτερα συνεχή εξάρτηση των λύσεων από τα αρχικά δεδομένα.

Με τους συμβολισμούς που χρησιμοποιήσαμε προηγουμένως και με έναν (μη γραμμικό, γενικά) τελεστή $B(t,\cdot) : D(A) \to H$, ο οποίος μπορεί να επεκταθεί και ως τελεστής $B(t,\cdot) : V \to V'$, για $t \in [0,T]$, (ουσιαστική για τα επόμενα είναι η δεύτερη αυτή ιδιότητα, η πρώτη τίθεται μόνο χάριν ευκολίας) θεωρούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} u'(t) + Au(t) = B\big(t, u(t)\big), & 0 < t < T, \\ u(0) = u^0. \end{cases} \tag{6.7}$$

Δίχως περιορισμούς στον B, το πρόβλημα αυτό δεν είναι γενικά παραβολικό, φερ' ειπείν, με $B = 2A$ οδηγούμαστε στην εξίσωση $u'(t) - Au(t) = 0$ και ο τελεστής $-A$ είναι *αρνητικά* ορισμένος. Θα χρειαστεί επομένως κάποια συνθήκη. Φυσιολογικές συνθήκες θεωρούνται οι συνθήκες του Lipschitz, είτε η αμφίπλευρη είτε η μονόπλευρη. Οι συνθήκες αυτές, για να μην είναι πολύ περιοριστικές, τίθενται συνήθως τοπικά, σε μια περιοχή της λύσης. Για απλότητα εμείς εδώ θα υποθέσουμε, προς το παρόν, την ολική μονόπλευρη συνθήκη του Lipschitz,

$$\big(B(t,v) - B(t,w), v - w\big) \le \lambda \|v - w\|^2 + \mu |v - w|^2 \quad \forall v, w \in V, \tag{6.8}$$

με μια σταθερά λ *μικρότερη* της μονάδας, $\lambda < 1$, και μια σταθερά μ. Σημειώνουμε ότι αυτή η σχέση συσχετίζει τους τελεστές A και B, παρά το ότι ο A δεν εμφανίζεται άμεσα, αφού η νόρμα $\|\cdot\|$ ορίζεται με βάση τον τελεστή A. Ο περιορισμός στη σταθερά λ τίθεται για να αποκλείσουμε περιπτώσεις όπως η $B = 2A$, που αναφέρθηκε προηγουμένως. Εύκολα διαπιστώνει κανείς ότι η συνθήκη (6.8) μπορεί να γραφεί ισοδύναμα και ως συνθήκη τύπου Gårding,

$$\big(F(t,v) - F(t,w), v - w\big) \ge (1-\lambda) \|v - w\|^2 - \mu |v - w|^2 \quad \forall v, w \in V, \tag{6.9}$$

με $F(t,v) := Av - F(t,v)$. Συγκρίνετε αυτή τη σχέση με τις σχέσεις στις Ασκήσεις 6.1, 6.2 και 6.3· παρατηρήστε ότι η F εδώ παίζει τον ρόλο της $-f$ σε αυτές τις ασκήσεις.

Παράλληλα με το (6.7) θεωρούμε και το πρόβλημα αρχικών τιμών

$$\begin{cases} v'(t) + Av(t) = B\big(t, v(t)\big), \quad 0 < t < T, \\ v(0) = v^0. \end{cases} \tag{6.10}$$

Αφαιρώντας κατά μέλη τις διαφορικές εξισώσεις στα προβλήματα (6.7) και (6.10) λαμβάνουμε

$$(u - v)' + A(u - v) = B(t,u) - B(t,v).$$

Παίρνοντας το εσωτερικό γινόμενο με $u - v$ έχουμε, όπως και στη γραμμική περίπτωση,

$$\frac{1}{2}\frac{d}{dt}|u - v|^2 + \|u - v\|^2 = \big(B(t,u) - B(t,v), u - v\big),$$

οπότε, σύμφωνα με την (6.8),

$$\frac{1}{2}\frac{d}{dt}|u - v|^2 + \|u - v\|^2 \le \lambda \|u - v\|^2 + \mu |u - v|^2,$$

συνεπώς

$$\frac{d}{dt}|u-v|^2 - 2\mu|u-v|^2 + 2(1-\lambda)\|u-v\|^2 \le 0,$$

οπότε

$$\frac{d}{dt}\left(\mathrm{e}^{-2\mu t}|u-v|^2\right) + 2(1-\lambda)\mathrm{e}^{-2\mu t}\|u-v\|^2 \le 0.$$

Ολοκληρώνοντας αυτή τη σχέση από 0 έως t, $t \le T$, παίρνουμε

$$\mathrm{e}^{-2\mu t}|(u-v)(t)|^2 - |(u-v)(0)|^2 + 2(1-\lambda)\int_0^t \mathrm{e}^{-2\mu s}\|(u-v)(s)\|^2 ds \le 0$$

ή

$$|(u-v)(t)|^2 + 2(1-\lambda)\int_0^t \mathrm{e}^{2\mu(t-s)}\|(u-v)(s)\|^2 ds \le \mathrm{e}^{2\mu t}|u^0-v^0|^2, \quad t\in[0,T]. \tag{6.11}$$

Ιδιαίτερα από αυτή τη σχέση έπεται η *μοναδικότητα* ομαλών λύσεων του προβλήματος (6.7).

6.2 Διακριτοποίηση στον χρόνο

Έστω $N \in \mathbb{N}, k := T/N$, και $t^n := nk, n = 0,\dots,N$. Οι αριθμητικές μέθοδοι, με τις οποίες θα ασχοληθούμε εδώ, δίνουν προσεγγίσεις U^n των τιμών u^n της ακριβούς λύσης στο σημείο t^n, $u^n := u(t^n)$.

6.2.1 Πεπλεγμένη μέθοδος του Euler

Κατά τα γνωστά, η πεπλεγμένη μέθοδος του Euler για το πρόβλημα (6.1) είναι

$$\begin{cases} U^0 := u^0, \\ U^{n+1} + kAU^{n+1} = U^n + kf(t^{n+1}), \quad n = 0,\dots,N-1. \end{cases} \tag{6.12}$$

Η ύπαρξη και η μοναδικότητα των λύσεων μπορούν να αποδειχθούν εύκολα με το θεώρημα αναπαράστασης του Riesz, αφού ο $\big(V,\|\cdot\|\big)$ είναι χώρος του Hilbert.

Η πεπλεγμένη μέθοδος του Euler για το πρόβλημα (6.7) είναι

$$\begin{cases} U^0 := u^0, \\ U^{n+1} + kAU^{n+1} = U^n + kB\big(t^{n+1},U^{n+1}\big), \quad n = 0,\dots,N-1. \end{cases} \tag{6.13}$$

Η μοναδικότητα των λύσεων, για $k < 1/\mu$, αποδεικνύεται εύκολα ως εξής. Δοθείσης της προσέγγισης U^n, έστω ότι υπάρχουν δύο προσεγγίσεις U^{n+1} και V^{n+1}, οι οποίες ικανοποιούν την (6.13) και την

$$V^{n+1} + kAV^{n+1} = U^n + kB\big(t^{n+1}, V^{n+1}\big),$$

αντίστοιχα. Για τη διαφορά τους ϑ^{n+1}, $\vartheta^{n+1} := U^{n+1} - V^{n+1}$, έχουμε

$$\vartheta^{n+1} + kA\vartheta^{n+1} = k\big[B\big(t^{n+1}, U^{n+1}\big) - B\big(t^{n+1}, V^{n+1}\big)\big].$$

Παίρνοντας εδώ το εσωτερικό γινόμενο με ϑ^{n+1} και χρησιμοποιώντας την υπόθεση (6.8), λαμβάνουμε

$$|\vartheta^{n+1}|^2 + k\|\vartheta^{n+1}\|^2 \le \lambda k\|\vartheta^{n+1}\|^2 + \mu k|\vartheta^{n+1}|^2,$$

οπότε

$$(1-\mu k)|\vartheta^{n+1}|^2 + k(1-\lambda)\|\vartheta^{n+1}\|^2 \le 0.$$

Λόγω των υποθέσεων $\lambda < 1$ και $\mu k < 1$ συμπεραίνουμε ότι $\vartheta^{n+1} = 0$.

Για την ύπαρξη θα αποδείξουμε ότι ο τελεστής $G : V \to V'$, $G(v) := v + kAv - kB(t^{n+1}, v) - U^n$, έχει δύο ιδιότητες: είναι *ισχυρά μονότονος* και ικανοποιεί τη συνθήκη του Lipschitz. Οι δύο αυτές ιδιότητες δίνουν μαζί ύπαρξη και μοναδικότητα της προσέγγισης U^{n+1}, για $k < 1/\mu$. Για την πρώτη ιδιότητα αρκούν οι συνθήκες που υποθέσαμε μέχρι τώρα, για τη δεύτερη υποθέτουμε επί πλέον ότι ο B ικανοποιεί την εξής συνθήκη του Lipschitz

$$\|B(t,v) - B(t,w)\|_\star \le \tilde{L}\|v-w\| \quad \forall v, w \in V, \tag{6.14}$$

με μια σταθερά $\tilde{L}$ που επιτρέπεται να ξεπερνάει τη μονάδα. Αναφέρουμε ότι αυτή η συνθήκη δεν είναι περιοριστική.

Κατ' αρχάς, όπως αποδείξαμε προηγουμένως μοναδικότητα, διαπιστώνουμε ότι

$$\big(G(v) - G(w), v - w\big) \ge (1-\mu k)|v-w|^2 + k(1-\lambda)\|v-w\|^2,$$

οπότε, με μια θετική σταθερά α, που εξαρτάται και από το k,

$$\big(G(v) - G(w), v - w\big) \ge \alpha\|v-w\|^2 \quad \forall v, w \in V. \tag{6.15}$$

Επίσης, με μια σταθερά L, που εξαρτάται και από το k, διαπιστώνουμε ότι

$$\|G(t,v) - G(t,w)\|_\star \le L\|v-w\| \quad \forall v, w \in V. \tag{6.16}$$

Τώρα, ο τελεστής $\tilde{G}$, $\tilde{G}(v) := A^{-1/2}G\big(A^{-1/2}v\big)$, απεικονίζει τον H στον εαυτό του και ικανοποιεί τις συνθήκες

$$\big(\tilde{G}(v) - \tilde{G}(w), v - w\big) \ge \alpha|v - w|^2, \quad \forall v, w \in H \tag{6.17}$$

και

$$|G(t, v) - G(t, w)| \le L|v - w| \quad \forall v, w \in H. \tag{6.18}$$

(Η συνθήκη (6.17) αναφέρεται στη βιβλιογραφία ως *ισχυρή μονοτονία* του τελεστή $\tilde{G}$.) Από τις (6.17) και (6.18) έπεται ότι ο τελεστής $\mathcal{F}$,

$$\mathcal{F}(v) := v - \frac{\alpha}{L^2}\tilde{G}(v),$$

είναι συστολή στον $\big(H, (\cdot, \cdot)\big)$, επομένως έχει ακριβώς ένα σταθερό σημείο, το οποίο είναι η μόνη ρίζα του $\tilde{G}$ και, συνεπώς, και του G, βλ. την Άσκηση 6.24.

Συνέπεια. Αποδεικνύουμε συνέπεια της μεθόδου κατ' ευθείαν για το πρόβλημα (6.7), οπότε έπεται ως ειδική περίπτωση και η συνέπεια για το πρόβλημα (6.1). Σημειώνουμε ότι η απόδειξη είναι εξ ίσου απλή και στις δύο περιπτώσεις. Κατά τα γνωστά, ορίζουμε το σφάλμα συνέπειας E^n ως

$$kE^n = u(t^{n+1}) - u(t^n) + kAu(t^{n+1}) - kB\big(t^{n+1}, u(t^{n+1})\big). \tag{6.19}$$

Χρησιμοποιώντας εδώ τη διαφορική εξίσωση στο (6.7), έχουμε

$$kE^n = u(t^{n+1}) - u(t^n) - ku'(t^{n+1}). \tag{6.20}$$

Τώρα, αναπτύσσοντας κατά Taylor,

$$u(t^n) = u(t^{n+1}) - ku'(t^{n+1}) + \int_{t^n}^{t^{n+1}} (s - t^n)u''(s)\, ds,$$

συνεπώς

$$kE^n = \int_{t^n}^{t^{n+1}} (s - t^n)u''(s)\, ds,$$

οπότε

$$\max_{0\le n\le N-1} \|E^n\|_\star \le Ck, \tag{6.21}$$

με προφανή την απαιτούμενη ομαλότητα της u και την έννοια της σταθεράς C.

Σύγκλιση. Δεν θα αποδείξουμε ξεχωριστά ευστάθεια της μεθόδου αλλά θα προχωρήσουμε κατ' ευθείαν στην εκτίμηση του σφάλματος· η απόδειξη της ευστάθειας θα είναι μετά προφανής. Η απόδειξη είναι απλούστερη στη γραμμική περίπτωση, (6.1), ας δούμε όμως την απόδειξη για το γενικότερο μη γραμμικό πρόβλημα (6.7).

Θέτοντας $e^m := u^m - U^m$ και αφαιρώντας το βήμα της μεθόδου που δίνεται στην (6.13) από την (6.19), παίρνουμε

$$e^{n+1} + kAe^{n+1} = e^n + k\big[B\big(t^{n+1}, u(t^{n+1})\big) - B\big(t^{n+1}, U^{n+1}\big)\big] + kE^n. \tag{6.22}$$

Λαμβάνοντας εδώ το εσωτερικό γινόμενο με e^{n+1} και χρησιμοποιώντας την (6.8), έχουμε

$$|e^{n+1}|^2 + k\|e^{n+1}\|^2 \le (e^n, e^{n+1}) + \lambda k\|e^{n+1}\|^2 + \mu k|e^{n+1}|^2 + k(E^n, e^{n+1}). \tag{6.23}$$

Τώρα,

$$(e^n, e^{n+1}) \le |e^n|\,|e^{n+1}| \le \frac{1}{2}|e^n|^2 + \frac{1}{2}|e^{n+1}|^2$$

και

$$(E^n, e^{n+1}) \le \|E^n\|_\star\,\|e^{n+1}\| \le \frac{1}{4\varepsilon}\|E^n\|_\star^2 + \varepsilon\|e^{n+1}\|^2,$$

για κάθε θετική σταθερά ε (η οποία θα επιλεγεί κατάλληλα στη συνέχεια), και η (6.23) δίνει

$$(1-\mu k)|e^{n+1}|^2 + 2(1-\lambda-\varepsilon)k\|e^{n+1}\|^2 \le |e^n|^2 + \frac{1}{2\varepsilon}k\|E^n\|_\star^2. \tag{6.24}$$

Επιλέγουμε τώρα το ε, έτσι ώστε $1-\lambda-\varepsilon > 0$, φερ' ειπείν $\varepsilon := (1-\lambda)/2$, και έχουμε

$$(1-\mu k)|e^{n+1}|^2 + (1-\lambda)k\|e^{n+1}\|^2 \le |e^n|^2 + \frac{1}{1-\lambda}k\|E^n\|_\star^2. \tag{6.25}$$

Επομένως, για $k \le 1/(2\mu)$,

$$|e^{n+1}|^2 + 2(1-\lambda)k\|e^{n+1}\|^2 \le (1+2\mu k)|e^n|^2 + \frac{2}{1-\lambda}k\|E^n\|_\star^2. \tag{6.26}$$

Από αυτή τη σχέση έπεται εύκολα επαγωγικά (διακριτή ανισότητα του Gronwall) ότι

$$|e^n|^2 + 2(1-\lambda)k\sum_{\ell=1}^{n}\|e^\ell\|^2 \le \mathrm{e}^{2\mu t^n}\frac{2}{1-\lambda}k\sum_{\ell=1}^{n-1}\|E^\ell\|_\star^2. \tag{6.27}$$

Συνδυάζοντας αυτή την εκτίμηση με τη συνέπεια (6.21), οδηγούμαστε στην επιθυμητή εκτίμηση του σφάλματος

$$\max_{1\le n\le N}\Big(|e^n|^2+2(1-\lambda)k\sum_{\ell=1}^{n}\|e^\ell\|^2\Big)^{1/2}\le \tilde{C}k, \tag{6.28}$$

με σταθερά $\tilde{C}$ ανεξάρτητη του k.

Η (6.28) λέγεται *εκτίμηση εκ των προτέρων*, γιατί το φράγμα στο δεξιό της μέλος δεν εξαρτάται από τις προσεγγιστικές λύσεις, μας δίνει δηλαδή πληροφορίες για τις προσεγγιστικές λύσεις χωρίς να τις χρησιμοποιεί. Τέτοιες εκτιμήσεις είναι χρήσιμες, γιατί μας εξασφαλίζουν ότι όταν η λύση έχει μια συγκεκριμένη ομαλότητα, τότε το σφάλμα έχει μια συγκεκριμένη ασυμπτωτική συμπεριφορά, στην προκειμένη περίπτωση ότι φθίνει γραμμικά με το k. Από τις εκ των προτέρων εκτιμήσεις δεν μπορούμε να πάρουμε πληροφορίες της μορφής ότι το σφάλμα δεν ξεπερνάει έναν προκαθορισμένο αριθμό, γιατί οι σταθερές που εμφανίζονται σε αυτές εξαρτώνται από νόρμες της (άγνωστης) λύσης. Για τέτοιες πληροφορίες καταφεύγουμε σε *εκ των υστέρων εκτιμήσεις*, στις οποίες θα επανέλθουμε αναλυτικότερα αργότερα.

6.2.2 Μέθοδος των Crank–Nicolson

Η μέθοδος του μέσου, στην περίπτωση Μερικών Διαφορικών Εξισώσεων, αναφέρεται ως μέθοδος των Crank–Nicolson. Για να απλοποιήσουμε λίγο τα πράγματα εισάγουμε λίγο συμβολισμό. Για δεδομένα $v^0,\dots,v^N$, θέτουμε

$$v^{n+\frac12}:=\frac12(v^n+v^{n+1}),\quad n=0,\dots,N-1.$$

Για το πρόβλημα (6.1) η μέθοδος είναι

$$\begin{cases} U^0:=u^0,\\ U^{n+1}+kAU^{n+\frac12}=U^n+kf(t^{n+\frac12}),\quad n=0,\dots,N-1.\end{cases} \tag{6.29}$$

Η μέθοδος των Crank–Nicolson για το πρόβλημα (6.7) είναι

$$\begin{cases} U^0:=u^0,\\ U^{n+1}+kAU^{n+\frac12}=U^n+kB\big(t^{n+\frac12},U^{n+\frac12}\big),\quad n=0,\dots,N-1.\end{cases} \tag{6.30}$$

Η ύπαρξη και η μοναδικότητα των λύσεων αποδεικνύεται όπως και στην περίπτωση της πεπλεγμένης μεθόδου του Euler, παραδείγματος χάριν, στη μη γραμμική περίπτωση ο κατάλληλος τελεστής είναι τώρα $G : V \to V'$,

$$G(v) := v + \frac{1}{2}kAv - kB\big(t^{n+\frac{1}{2}}, \frac{1}{2}(v + U^n)\big) - \frac{1}{2}kAU^n - U^n.$$

Συνέπεια. Όπως και στην περίπτωση της πεπλεγμένης μεθόδου του Euler, αποδεικνύουμε συνέπεια της μεθόδου των Crank–Nicolson κατ' ευθείαν για το πρόβλημα (6.7), οπότε έπεται ως ειδική περίπτωση και η συνέπεια για το πρόβλημα (6.1). Κατά τα γνωστά, ορίζουμε το σφάλμα συνέπειας E^n ως

$$kE^n = u^{n+1} - u^n + kAu^{n+\frac{1}{2}} - kB\big(t^{n+\frac{1}{2}}, u^{n+\frac{1}{2}}\big). \tag{6.31}$$

Προσθαφαιρώντας προφανείς όρους γράφουμε αυτή τη σχέση στη μορφή

$$\begin{aligned} kE^n =& u^{n+1} - u^n + kA\big[u^{n+\frac{1}{2}} - u(t^{n+\frac{1}{2}})\big] - k\big[B\big(t^{n+\frac{1}{2}}, u^{n+\frac{1}{2}}\big) \\ &- B\big(t^{n+\frac{1}{2}}, u(t^{n+\frac{1}{2}})\big)\big] + k\big[Au(t^{n+\frac{1}{2}}) - B\big(t^{n+\frac{1}{2}}, u(t^{n+\frac{1}{2}})\big)\big], \end{aligned} \tag{6.32}$$

οπότε, χρησιμοποιώντας τη διαφορική εξίσωση στο (6.7),

$$\begin{aligned} kE^n =& u^{n+1} - u^n + ku'(t^{n+\frac{1}{2}}) + kA\big[u^{n+\frac{1}{2}} - u(t^{n+\frac{1}{2}})\big] \\ &- k\big[B\big(t^{n+\frac{1}{2}}, u^{n+\frac{1}{2}}\big) - B\big(t^{n+\frac{1}{2}}, u(t^{n+\frac{1}{2}})\big)\big] \end{aligned} \tag{6.33}$$

Τώρα, αναπτύσσοντας κατά Taylor ως προς το σημείο $t^{n+\frac{1}{2}}$, έχουμε

$$u^{n+\frac{1}{2}} = u(t^{n+\frac{1}{2}}) + \frac{1}{2}\Big[\int_{t^n}^{t^{n+\frac{1}{2}}} (s - t^n)u''(s)\,ds + \int_{t^{n+\frac{1}{2}}}^{t^{n+1}} (t^{n+1} - s)u''(s)\,ds\Big]. \tag{6.34}$$

Επί πλέον, αναπτύσσοντας πάλι κατά Taylor ως προς το σημείο $t^{n+\frac{1}{2}}$,

$$\begin{aligned} &u(t^{n+1}) - u(t^n) - ku'(t^{n+\frac{1}{2}}) = \\ &\qquad \frac{1}{2}\Big[\int_{t^n}^{t^{n+\frac{1}{2}}} (s - t^n)^2 u'''(s)\,ds + \int_{t^{n+\frac{1}{2}}}^{t^{n+1}} (t^{n+1} - s)^2 u'''(s)\,ds\Big]. \end{aligned} \tag{6.35}$$

οπότε, η (6.33) δίνει, αν χρησιμοποιήσουμε και τη συνθήκη του Lipschitz,

$$\max_{0\le n\le N-1} \|E^n\|_\star \le Ck^2, \tag{6.36}$$

με προφανή την απαιτούμενη ομαλότητα της u και την έννοια της σταθεράς C.

Σύγκλιση. Θα προχωρήσουμε κατ' ευθείαν στην εκτίμηση του σφάλματος της μεθόδου των Crank–Nicolson· η απόδειξη της ευστάθειας θα είναι μετά προφανής. Προκειμένου να αποφύγουμε επαναλήψεις, θα δώσουμε την απόδειξη για το μη γραμμικό πρόβλημα (6.7).

Θέτοντας $e^m := u^m - U^m$ και αφαιρώντας το βήμα της μεθόδου που δίνεται στην (6.30) από την (6.31), παίρνουμε

$$e^{n+1} - e^n + kAe^{n+\frac{1}{2}} = k\big[B\big(t^{n+\frac{1}{2}}, u^{n+\frac{1}{2}}\big) - B\big(t^{n+\frac{1}{2}}, U^{n+\frac{1}{2}}\big)\big] + kE^n. \tag{6.37}$$

Λαμβάνοντας εδώ το εσωτερικό γινόμενο με $e^{n+\frac{1}{2}}$ και χρησιμοποιώντας την (6.8), έχουμε

$$\frac{1}{2}\big(|e^{n+1}|^2 - |e^n|^2\big) + k\|e^{n+\frac{1}{2}}\|^2 \le \lambda k\|e^{n+\frac{1}{2}}\|^2 + \mu k|e^{n+\frac{1}{2}}|^2 + k(E^n, e^{n+\frac{1}{2}})$$

ή

$$\frac{1}{2}\big(|e^{n+1}|^2 - |e^n|^2\big) + k\|e^{n+\frac{1}{2}}\|^2 \le \lambda k\|e^{n+\frac{1}{2}}\|^2 + \mu k|e^{n+\frac{1}{2}}|^2 + \varepsilon k\|e^{n+\frac{1}{2}}\|^2 + \frac{1}{4\varepsilon}k\|E^n\|_\star^2 \tag{6.38}$$

για κάθε θετική σταθερά ε (η οποία θα επιλεγεί κατάλληλα στη συνέχεια), βλ. την αντίστοιχη απόδειξη στην περίπτωση της πεπλεγμένης μεθόδου του Euler. Τώρα,

$$|e^{n+\frac{1}{2}}|^2 \le \frac{1}{4}\big(|e^n| + |e^{n+1}|\big)^2 \le \frac{1}{2}\big(|e^n|^2 + |e^{n+1}|^2\big)$$

και η (6.38) δίνει

$$|e^{n+1}|^2 + 2(1-\lambda-\varepsilon)k\|e^{n+\frac{1}{2}}\|^2 \le |e^n|^2 + \mu k|e^{n+1}|^2 + \mu k|e^n|^2 + \frac{1}{2\varepsilon}k\|E^n\|_\star^2$$

ή

$$(1-\mu k)|e^{n+1}|^2 + 2(1-\lambda-\varepsilon)k\|e^{n+\frac{1}{2}}\|^2 \le (1+\mu k)|e^n|^2 + \frac{1}{2\varepsilon}k\|E^n\|_\star^2.$$

Επιλέγουμε τώρα τη σταθερά ε έτσι ώστε $1-\lambda-\varepsilon > 0$, π.χ., $\varepsilon := (1-\lambda)/2$, και έχουμε

$$(1-\mu k)|e^{n+1}|^2 + (1-\lambda)k\|e^{n+\frac{1}{2}}\|^2 \le (1+\mu k)|e^n|^2 + \frac{1}{1-\lambda}k\|E^n\|_\star^2. \tag{6.39}$$

Υποθέτουμε τώρα ότι το k είναι αρκετά μικρό, $k \le 1/(2\mu)$, οπότε $\mu k \le 1/2$. Σημειώνουμε ότι η άλλη περίπτωση είναι εύκολη. Τώρα

$$\frac{1}{1-\mu k} \le \frac{1}{2} \quad \text{και} \quad \frac{1+\mu k}{1-\mu k} \le 1 + 4\mu k,$$

οπότε η (6.39) δίνει , για $k \leq 1/(2\mu)$,

$$|e^{n+1}|^2 + 2(1-\lambda)k\|e^{n+\frac{1}{2}}\|^2 \leq (1+4\mu k)|e^n|^2 + \frac{2}{1-\lambda}k\|E^n\|_\star^2. \tag{6.40}$$

Από αυτή τη σχέση έπεται εύκολα επαγωγικά (διακριτή ανισότητα του Gronwall) ότι

$$|e^n|^2 + 2(1-\lambda)k\sum_{\ell=0}^{n-1}\|e^{\ell+\frac{1}{2}}\|^2 \leq \mathrm{e}^{4\mu t^n}\frac{2}{1-\lambda}k\sum_{\ell=1}^{n-1}\|E^\ell\|_\star^2. \tag{6.41}$$

Συνδυάζοντας αυτή την εκτίμηση με τη συνέπεια (6.36), οδηγούμαστε στην επιθυμητή εκτίμηση του σφάλματος

$$\max_{1\leq n\leq N}\Big(|e^n|^2 + 2(1-\lambda)k\sum_{\ell=0}^{n-1}\|e^{\ell+\frac{1}{2}}\|^2\Big)^{1/2} \leq \tilde{C}k^2, \tag{6.42}$$

με σταθερά $\tilde{C}$ ανεξάρτητη του k.

6.2.3 Η διβηματική μέθοδος ανάδρομων διαφορών

Θέτουμε $U^0 := u^0$ και ορίζουμε τη U^1 ως την προσέγγιση που δίνει η πεπλεγμένη μέθοδος του Euler για το πρόβλημα (6.7),

$$U^1 + kAU^1 = kB(t^1, U^1).$$

Έχοντας ήδη στη διάθεσή μας δύο αρχικές προσεγγίσεις, υπολογίζουμε τις προσεγγίσεις $U^2, \dots, U^N$ με τη διβηματική μέθοδο ανάδρομων διαφορών, για το πρόβλημα (6.7),

$$\frac{3}{2}U^{n+2} - 2U^{n+1} + \frac{1}{2}U^n + kAU^{n+2} = kB(t^{n+2}, U^{n+2}), \quad n = 0, \dots, N-2. \tag{6.43}$$

Η ύπαρξη και μοναδικότητα των προσεγγίσεων, για αρκετά μικρό k, αποδεικνύεται εντελώς ανάλογα με την περίπτωση της πεπλεγμένης μεθόδου του Euler· προς αποφυγήν επαναλήψεων, η απόδειξη παραλείπεται.

Συνέπεια. Κατά τα γνωστά, ορίζουμε το σφάλμα συνέπειας E^n της μεθόδου, για το πρόβλημα (6.7), ως

$$kE^n = \frac{3}{2}u^{n+2} - 2u^{n+1} + \frac{1}{2}u^n + k\big[Au^{n+2} - B(t^{n+2}, u^{n+2})\big]. \tag{6.44}$$

Χρησιμοποιώντας εδώ τη διαφορική εξίσωση στο (6.7), έχουμε

$$kE^n = \frac{3}{2}u^{n+2} - 2u^{n+1} + \frac{1}{2}u^n - ku'(t^{n+2}). \tag{6.45}$$

Τώρα, αναπτύσσοντας κατά Taylor, ως προς το σημείο t^{n+2}, έχουμε

$$u^{n+1} = u^{n+2} - ku'(t^{n+2}) + \frac{1}{2}k^2u''(t^{n+2}) - \frac{1}{2}\int_{t^{n+1}}^{t^{n+2}} (s - t^{n+1})^2 u'''(s)\, ds$$

και

$$u^n = u^{n+2} - 2ku'(t^{n+2}) + 2k^2u''(t^{n+2}) - \frac{1}{2}\int_{t^n}^{t^{n+2}} (s - t^n)^2 u'''(s)\, ds.$$

Επομένως, η (6.45) γράφεται στη μορφή

$$kE^n = \int_{t^{n+1}}^{t^{n+2}} (s - t^{n+1})^2 u'''(s)\, ds - \frac{1}{4}\int_{t^n}^{t^{n+2}} (s - t^n)^2 u'''(s)\, ds, \tag{6.46}$$

οπότε, με προφανή την απαιτούμενη ομαλότητα της u και την έννοια της σταθεράς C, έχουμε

$$\max_{0\le n\le N-2} \|E^n\|_\star \le Ck^2. \tag{6.47}$$

Σύγκλιση. Θέτοντας $e^m := u^m - U^m$ και αφαιρώντας κατά μέλη την (6.43) από την (6.44), παίρνουμε

$$\frac{3}{2}e^{n+2} - 2e^{n+1} + \frac{1}{2}e^n + kAe^{n+2} + k\big[B(t^{n+2}, u^{n+2}) - B(t^{n+2}, U^{n+2})\big] + kE^n. \tag{6.48}$$

Λαμβάνοντας εδώ το εσωτερικό γινόμενο με e^{n+2} και χρησιμοποιώντας την (6.8), έχουμε

$$\big(\frac{3}{2}e^{n+2} - 2e^{n+1} + \frac{1}{2}e^n, e^{n+2}\big) + k\|e^{n+2}\|^2 \le \lambda k\|e^{n+2}\|^2 + \mu k|e^{n+2}|^2 + k(E^n, e^{n+2}).$$

οπότε, κατά τα γνωστά,

$$\big(\frac{3}{2}e^{n+2} - 2e^{n+1} + \frac{1}{2}e^n, e^{n+2}\big) + (1-\lambda-\varepsilon)k\|e^{n+2}\|^2 \le \mu k|e^{n+2}|^2 + \frac{1}{4\varepsilon}k\|E^n\|_\star^2, \tag{6.49}$$

για οποιαδήποτε θετική σταθερά ε. Επιλέγουμε τώρα $\varepsilon := (1-\lambda)/2$, και έχουμε

$$\big(\frac{3}{2}e^{n+2} - 2e^{n+1} + \frac{1}{2}e^n, e^{n+2}\big) + \frac{1}{2}(1-\lambda)k\|e^{n+2}\|^2 \le \mu k|e^{n+2}|^2 + \frac{1}{2(1-\lambda)}k\|E^n\|_\star^2. \tag{6.50}$$

Χρησιμοποιώντας εδώ την ταυτότητα

$$\begin{aligned}(\frac{3}{2}\varepsilon^{n+2} - 2\varepsilon^{n+1} + \frac{1}{2}\varepsilon^n, \varepsilon^{n+2}) &= \frac{5}{4}|\varepsilon^{n+2}|^2 - |\varepsilon^{n+1}|^2 - \frac{1}{4}|\varepsilon^n|^2 \\ &\quad - \big[(\varepsilon^{n+2}, \varepsilon^{n+1}) - (\varepsilon^{n+1}, \varepsilon^n)\big] + \frac{1}{4}|\varepsilon^{n+2} - 2\varepsilon^{n+1} + \varepsilon^n|^2\end{aligned}$$

λαμβάνουμε

$$\begin{aligned}&\big[\frac{5}{4}|e^{n+2}|^2 - (e^{n+2}, e^{n+1}) + \frac{1}{4}|e^{n+1}|^2\big] + \frac{1}{2}(1-\lambda)k\|e^{n+2}\|^2 \leq \\ &\qquad \big[\frac{5}{4}|e^{n+1}|^2 - (e^{n+1}, e^n) + \frac{1}{4}|e^n|^2\big] + \mu k|e^{n+2}|^2 + \frac{1}{2(1-\lambda)}k\|E^n\|_\star^2.\end{aligned} \tag{6.51}$$

Τώρα

$$(e^{n+2}, e^{n+1}) \leq |e^{n+2}|^2 + \frac{1}{4}|e^{n+1}|^2,$$

οπότε

$$\big[\frac{5}{4}|e^{n+2}|^2 - (e^{n+2}, e^{n+1}) + \frac{1}{4}|e^{n+1}|^2\big] \geq \frac{1}{4}|e^{n+2}|^2, \tag{6.52}$$

και η (6.51) δίνει

$$\begin{aligned}&(1-4\mu k)\big[\frac{5}{4}|e^{n+2}|^2 - (e^{n+2}, e^{n+1}) + \frac{1}{4}|e^{n+1}|^2\big] + \frac{1}{2}(1-\lambda)k\|e^{n+2}\|^2 \leq \\ &\qquad \big[\frac{5}{4}|e^{n+1}|^2 - (e^{n+1}, e^n) + \frac{1}{4}|e^n|^2\big] + \frac{1}{2(1-\lambda)}k\|E^n\|_\star^2.\end{aligned} \tag{6.53}$$

Έστω τώρα $k \leq 1/(8\mu)$. Τότε

$$\frac{1}{1-4\mu k} \leq 2 \quad \text{και} \quad \frac{1}{1-4\mu k} \leq 1 + 8\mu k,$$

επομένως η (6.53) δίνει

$$\begin{aligned}&\big[\frac{5}{4}|e^{n+2}|^2 - (e^{n+2}, e^{n+1}) + \frac{1}{4}|e^{n+1}|^2\big] + (1-\lambda)k\|e^{n+2}\|^2 \leq \\ &\qquad (1+8\mu k)\big[\frac{5}{4}|e^{n+1}|^2 - (e^{n+1}, e^n) + \frac{1}{4}|e^n|^2\big] + \frac{1}{1-\lambda}k\|E^n\|_\star^2.\end{aligned} \tag{6.54}$$

Από αυτή τη σχέση έπεται εύκολα επαγωγικά (διακριτή ανισότητα του Gronwall) ότι

$$\begin{aligned}&\big[\frac{5}{4}|e^{n+2}|^2 - (e^{n+2}, e^{n+1}) + \frac{1}{4}|e^{n+1}|^2\big] + (1-\lambda)k\sum_{\ell=1}^{n+2}\|e^\ell\|^2 \leq \\ &\qquad \mathrm{e}^{2\mu t^{n+1}}\Big[\big[\frac{5}{4}|e^1|^2 - (e^1, e^0) + \frac{1}{4}|e^0|^2\big] + \frac{1}{1-\lambda}k\sum_{\ell=1}^{n}\|E^\ell\|_\star^2\Big].\end{aligned}$$

ή

$$
\begin{aligned}
&\Big[\frac{5}{4}|e^{n+2}|^2-(e^{n+2},e^{n+1})+\frac{1}{4}|e^{n+1}|^2\Big]+(1-\lambda)k\sum_{\ell=1}^{n+2}\|e^\ell\|^2\le\\
&\qquad \mathrm{e}^{2\mu t^{n+1}}\Big[\frac{5}{4}|e^1|^2+\frac{1}{1-\lambda}k\sum_{\ell=1}^{n}\|E^\ell\|_\star^2\Big].
\end{aligned}
\tag{6.55}
$$

Συνδυάζοντας αυτή την εκτίμηση με την (6.52), τη συνέπεια (6.47), και την εκτίμηση για την $|e^1|$, στην οποία οδηγείται κανείς πολύ εύκολα, όπως είδαμε στην πεπλεγμένη μέθοδο του Euler, οδηγούμαστε στην επιθυμητή εκτίμηση του σφάλματος

$$
\max_{1\le n\le N}\Big(|e^n|^2+(1-\lambda)k\sum_{\ell=1}^{n}\|e^\ell\|^2\Big)^{1/2}\le \tilde{C}k^2,
\tag{6.56}
$$

με σταθερά $\tilde{C}$ ανεξάρτητη του k.

6.3 Εκ των υστέρων εκτιμήσεις των σφαλμάτων

Για να οδηγηθούμε σε εκ των προτέρων εκτιμήσεις, συνδυάζουμε τη συνέπεια με την ευστάθεια της αριθμητικής μεθόδου. Όπως γνωρίζουμε ήδη, το σφάλμα συνέπειας αποτελεί ένα μέτρο για το μέγεθος της αποτυχίας της ακριβούς λύσης να είναι προσεγγιστική λύση, δηλαδή να ικανοποιεί την αριθμητική μέθοδο. Στις εκ των υστέρων εκτιμήσεις οδηγούμαστε συνδυάζοντας το λεγόμενο *υπόλοιπο* με την ευστάθεια του συνεχούς προβλήματος αρχικών τιμών. Το υπόλοιπο παίζει σε αυτές τις εκτιμήσεις ρόλο αντίστοιχο εκείνου του σφάλματος συνέπειας στις εκ των προτέρων εκτιμήσεις. Το υπόλοιπο αποτελεί ένα μέτρο για το μέγεθος της αποτυχίας της προσεγγιστικής λύσης να είναι ακριβής λύση, δηλαδή να ικανοποιεί τη διαφορική εξίσωση. Φυσικά, για να μπορεί κανείς καν να μιλήσει για το υπόλοιπο, πρέπει να έχει στη διάθεσή του μια προσεγγιστική λύση ορισμένη σε όλο το διάστημα $[0,T]$ και όχι μόνο στους κόμβους μιας διακριτοποίησης. Ο τρόπος με τον οποίο κατασκευάζουμε τέτοιες προσεγγίσεις, προκειμένου να οδηγηθούμε σε κατάλληλες εκ των υστέρων εκτιμήσεις, αποτελεί κεντρικό θέμα σ' αυτή τη μελέτη και εξαρτάται κάθε φορά από τη μέθοδο στην οποία αναφερόμαστε. Σε αντιδιαστολή με τις εκ των προτέρων εκτιμήσεις, όπου η έρευνα έχει προχωρήσει πάρα πολύ, η έρευνα σε θέματα εκ των υστέρων εκτιμήσεων για παραβολικές εξισώσεις βρίσκεται ακόμα σε αρχικό στάδιο.

6.3.1 Πεπλεγμένη μέθοδος του Euler

Σύμφωνα με την (6.28), για το σφάλμα των προσεγγίσεων που δίνει η μέθοδος του Euler, έχουμε

$$\max_{1\le n\le N} |u^n - U^n| \le \tilde{C}k, \tag{6.57}$$

υποθέτοντας κατάλληλη ομαλότητα της ακριβούς λύσης. Επομένως, μια λογική επιλογή για να οδηγηθούμε σε προσεγγίσεις $U(t)$ των τιμών $u(t)$ με τάξη ακρίβειας ένα, είναι να ορίσουμε τη U, $U(0) = U^0$, κατά τμήματα σταθερή στα διαστήματα $I_n := (t^{n-1}, t^n]$, $n = 1, \dots, N$, με τιμές $U(t) = U^n, t \in I_n$, $n = 1, \dots, N$. Είναι πράγματι εύκολο να πεισθούμε ότι ισχύει

$$\sup_{0\le t\le T} |u(t) - U(t)| \le \tilde{c}k. \tag{6.58}$$

Όπως θα δούμε στη συνέχεια, το υπόλοιπο της U *δεν* μας οδηγεί σε πρώτης τάξης εκ των υστέρων εκτιμήσεις. Για να οδηγηθούμε σε τέτοιες εκτιμήσεις, εισάγουμε και την προσέγγιση $\hat{U}$, $\hat{U}(0) = U^0$, ως την κατά τμήματα γραμμική (πολυώνυμο βαθμού το πολύ ένα) συνάρτηση που παρεμβάλλεται στις τιμές $U^0, \dots, U^N$, δηλαδή

$$\hat{U}(t) := U^{n-1} + \frac{1}{k}(U^n - U^{n-1})(t - t^{n-1}), \quad t \in I_n, \quad n = 1, \dots, N. \tag{6.59}$$

Η $\hat{U}$ αναφέρεται ως *ανακατασκευή* της προσέγγισης U.

Η γραμμική περίπτωση. Στην περίπτωση του γραμμικού προβλήματος (6.1), το υπόλοιπο R της προσέγγισης U είναι $R(t) := U'(t) + AU(t) - f(t)$, δηλαδή

$$R(t) = AU(t) - f(t). \tag{6.60}$$

Σημειώνουμε ότι το υπόλοιπο δεν ορίζεται στους κόμβους t^n, αλλά αυτό δεν είναι σημαντικό για όσα θα ακολουθήσουν. Χρησιμοποιώντας τη διαφορική εξίσωση του (6.1) στην (6.60) παίρνουμε

$$R(t) = -u'(t) - A[u(t) - U(t)]. \tag{6.61}$$

Όπως βλέπουμε ο δεύτερος όρος είναι πρώτης τάξης ως προς k, αλλά ο πρώτος όρος είναι μηδενικής τάξης, δηλαδή ανεξάρτητος του k. Για αυτόν ακριβώς τον λόγο, το εν λόγω υπόλοιπο δεν είναι κατάλληλο για να μας οδηγήσει σε εκ των υστέρων εκτιμήσεις του σφάλματος πρώτης τάξης, δηλαδή ίδιας τάξης με την εκ των προτέρων εκτίμηση. [Η βαθύτερη αιτία για αυτό το φαινόμενο έγκειται στο γεγονός ότι, όταν προσεγγίζουμε μια συνάρτηση με κατά τμήματα πολυωνυμικές συναρτήσεις βαθμού το πολύ q, και προσεγγίζουμε την παράγωγό της με την παράγωγο της προσέγγισης, τότε αυτή είναι πολυωνυμική συνάρτηση βαθμού το πολύ $q-1$, και επομένως χάνεται μία τάξη ακρίβειας.] Για να ανακτήσουμε τη βέλτιστη τάξη, καταφεύγουμε στην ανακατασκευή $\hat{U}$ της U, η οποία είναι τμηματικά πολυωνυμική συνάρτηση βαθμού το πολύ ένα. Το υπόλοιπο $\hat{R}$ της $\hat{U}$ είναι

$$\hat{R}(t) := \hat{U}'(t) + A\hat{U}(t) - f(t). \tag{6.62}$$

Τώρα, σύμφωνα με την (6.12),

$$\frac{1}{k}(U^n - U^{n-1}) = -AU^n + f(t^n),$$

οπότε, για $t \in (t^{n-1}, t^n)$,

$$\hat{U}'(t) = -AU^n + f(t^n) \tag{6.63}$$

ή

$$\hat{U}'(t) + AU(t) = f(t^n), \tag{6.64}$$

συνεπώς η (6.62) μπορεί να γραφεί και στις μορφές

$$\hat{R}(t) = A[\hat{U}(t) - U^n] - [f(t) - f(t^n)], \quad t \in (t^{n-1}, t^n), \tag{6.65}$$

ή

$$\hat{R}(t) = A[\hat{U}(t) - U(t)] - [f(t) - f(t^n)], \quad t \in (t^{n-1}, t^n). \tag{6.66}$$

Όμως, σύμφωνα με την (6.59), για $t \in (t^{n-1}, t^n)$,

$$\hat{U}(t) - U^n = \frac{1}{k}(U^n - U^{n-1})(t - t^n),$$

συνεπώς η (6.64) μπορεί να γραφεί και στη μορφή

$$\hat{R}(t) = \frac{1}{k}A(U^n - U^{n-1})(t - t^n) - [f(t) - f(t^n)], \quad t \in (t^{n-1}, t^n). \tag{6.67}$$

Το υπόλοιπο $\hat{R}(t)$ είναι πρώτης τάξης ως προς k, υπό κατάλληλες συνθήκες ομαλότητας.

Θεωρούμε τώρα τα σφάλματα e και $\hat{e}$ των προσεγγίσεων U και $\hat{U}$, αντίστοιχα,

$$e(t) := u(t) - U(t) \quad \text{και} \quad \hat{e}(t) := u(t) - \hat{U}(t).$$

Αφαιρώντας την (6.64) κατά μέλη από τη διαφορική εξίσωση στο πρόβλημα (6.1), παίρνουμε

$$\hat{e}'(t) + Ae(t) = R_f(t) \tag{6.68}$$

με

$$R_f(t) := f(t) - f(t^n), \quad t \in (t^{n-1}, t^n). \tag{6.69}$$

Τώρα

$$(Ae, \hat{e}) = \frac{1}{2}\Big(\|e\|^2 + \|\hat{e}\|^2 - \|\hat{e} - e\|^2\Big) = \frac{1}{2}\Big(\|e\|^2 + \|\hat{e}\|^2 - \|\hat{U} - U\|^2\Big),$$

οπότε, παίρνοντας στην (6.68) το εσωτερικό γινόμενο με $\hat{e}(t)$ λαμβάνουμε

$$\frac{d}{dt}|\hat{e}(t)|^2 + \Big(\|e(t)\|^2 + \|\hat{e}(t)\|^2\Big) = \|\hat{U}(t) - U(t)\|^2 + 2\big(R_f(t), \hat{e}(t)\big). \tag{6.70}$$

Τώρα,

$$2\big(R_f(t), \hat{e}(t)\big) \le 2\|R_f(t)\|_\star^2 + \frac{1}{2}\|\hat{e}(t)\|^2,$$

συνεπώς η (6.70) δίνει

$$\frac{d}{dt}|\hat{e}(t)|^2 + \Big(\|e(t)\|^2 + \frac{1}{2}\|\hat{e}(t)\|^2\Big) \le \|\hat{U}(t) - U(t)\|^2 + 2\|R_f(t)\|_\star^2.$$

Ολοκληρώνοντας αυτή τη σχέση από 0 έως $t \le T$, λαμβάνουμε

$$|\hat{e}(t)|^2 + \int_0^t \Big(\|e(s)\|^2 + \frac{1}{2}\|\hat{e}(s)\|^2\Big)\,ds \le \int_0^t \|\hat{U}(s) - U(s)\|^2\,ds + 2\int_0^t \|R_f(s)\|_\star^2\,ds \tag{6.71}$$

και εύκολα οδηγούμαστε στο συμπέρασμα ότι

$$\begin{aligned} \max_{0\le\tau\le t}\Big[|\hat{e}(\tau)|^2 + \int_0^\tau \Big(\|e(s)\|^2 + \frac{1}{2}\|\hat{e}(s)\|^2\Big)\,ds\Big] \\ \le \int_0^t \|\hat{U}(s) - U(s)\|^2\,ds + 2\int_0^t \|R_f(s)\|_\star^2\,ds. \end{aligned} \tag{6.72}$$

Σημειώνουμε ότι αυτή είναι μια εκ των υστέρων εκτίμηση, αφού το φράγμα στο δεξιό της μέλος μπορεί να υπολογισθεί.

Η μη γραμμική περίπτωση. Το υπόλοιπο $\hat{R}$ της ανακατασκευής $\hat{U}$, βλ. την (6.59), για το πρόβλημα (6.7), είναι

$$\hat{R}(t) := \hat{U}'(t) + A\hat{U}(t) - B\big(t, \hat{U}(t)\big). \tag{6.73}$$

Τώρα, σύμφωνα με την (6.13),

$$\frac{1}{k}(U^n - U^{n-1}) = -AU^n + B(t^n, U^n),$$

οπότε, για $t \in (t^{n-1}, t^n)$,

$$\hat{U}'(t) = -AU^n + B(t^n, U^n), \tag{6.74}$$

συνεπώς η (6.73) μπορεί να γραφεί και στις μορφές

$$\hat{R}(t) = A[\hat{U}(t) - U^n] - [B\big(t, \hat{U}(t)\big) - B(t^n, U^n)], \quad t \in (t^{n-1}, t^n), \tag{6.75}$$

ή

$$\hat{R}(t) = A[\hat{U}(t) - U(t)] - [B\big(t, \hat{U}(t)\big) - B\big(t^n, U(t)\big)], \quad t \in (t^{n-1}, t^n). \tag{6.76}$$

Γράφουμε τώρα την (6.74) στη μορφή

$$\hat{U}'(t) + AU(t) = B(t^n, U^n), \quad t \in (t^{n-1}, t^n). \tag{6.77}$$

Αφαιρώντας την (6.77) κατά μέλη από τη διαφορική εξίσωση στο πρόβλημα (6.7), παίρνουμε

$$\hat{e}'(t) + Ae(t) = B\big(t, u(t)\big) - B(t^n, U^n),$$

σχέση που τη γράφουμε στη μορφή

$$\hat{e}'(t) + Ae(t) = B\big(t, u(t)\big) - B\big(t, U(t)\big) + R_U(t) \tag{6.78}$$

με

$$R_U(t) := B\big(t, U^n\big) - B(t^n, U^n), \quad t \in (t^{n-1}, t^n). \tag{6.79}$$

Τώρα,

$$\begin{aligned}\big(B(t,u(t))-B(t,U(t)),\hat{e}(t)\big) =&\big(B(t,u(t))-B(t,U(t)),e(t)\big)\\ &+\big(B(t,u(t))-B(t,U(t)),U(t)-\hat{U}(t)\big)\end{aligned}$$

και, λόγω των (6.8) και (6.14), στοιχειώδεις υπολογισμοί δίνουν

$$\big(B(t,u(t))-B(t,U(t)),\hat{e}(t)\big)\le\lambda\|e(t)\|^2+\mu|e(t)|^2+\tilde{L}\|e(t)\|\,\|(\hat{U}-U)(t)\|. \tag{6.80}$$

Αντίστοιχα,

$$\begin{aligned}\big(B(t,u(t))-B(t,U(t)),\hat{e}(t)\big) =&\big(B(t,u(t))-B(t,\hat{U}(t)),\hat{e}(t)\big)\\ &+\big(B(t,\hat{U}(t))-B(t,U(t)),\hat{e}(t)\big)\end{aligned}$$

και

$$\big(B(t,u(t))-B(t,U(t)),\hat{e}(t)\big)\le\lambda\|\hat{e}(t)\|^2+\mu|\hat{e}(t)|^2+\tilde{L}\|\hat{e}(t)\|\,\|(\hat{U}-U)(t)\|. \tag{6.81}$$

Αθροίζοντας τις (6.80) και (6.81), λαμβάνουμε

$$\begin{aligned}2\big(B(t,u(t))-B(t,U(t)),\hat{e}(t)\big)&\le\lambda\big(\|\hat{e}(t)\|^2+\|e(t)\|^2\big)\\ &+\mu\big(|\hat{e}(t)|^2+|e(t)|^2\big)+\tilde{L}\big(\|\hat{e}(t)\|+\|e(t)\|\big)\|(\hat{U}-U)(t)\|.\end{aligned} \tag{6.82}$$

Τώρα,

$$|e(t)|^2\le\big(|\hat{e}(t)|+|(\hat{U}-U)(t)|\big)^2\le 2|\hat{e}(t)|^2+2|(\hat{U}-U)(t)|^2$$

και

$$\begin{aligned}&\tilde{L}\big(\|\hat{e}(t)\|+\|e(t)\|\big)\|(\hat{U}-U)(t)\|\le\\ &\qquad\le\frac{\varepsilon}{2}\big(\|\hat{e}(t)\|+\|e(t)\|\big)^2+\frac{\tilde{L}^2}{2\varepsilon}\|(\hat{U}-U)(t)\|^2\\ &\qquad\le\varepsilon\big(\|\hat{e}(t)\|^2+\|e(t)\|^2\big)+\frac{\tilde{L}^2}{2\varepsilon}\|(\hat{U}-U)(t)\|^2.\end{aligned}$$

Επομένως, από την (6.82) παίρνουμε

$$\begin{aligned}2\big(B(t,u(t))-B(t,U(t)),\hat{e}(t)\big)&\le\lambda\big(\|\hat{e}(t)\|^2+\|e(t)\|^2\big)+3\mu|e(t)|^2\\ &+2\mu|(\hat{U}-U)(t)|^2+\varepsilon\big(\|\hat{e}(t)\|^2+\|e(t)\|^2\big)+\frac{\tilde{L}^2}{2\varepsilon}\|(\hat{U}-U)(t)\|^2\end{aligned} \tag{6.83}$$

για οποιαδήποτε θετική σταθερά ε.

Παίρνοντας στη (6.78) το εσωτερικό γινόμενο με $\hat e(t)$ έχουμε

$$
\begin{aligned}
\frac{d}{dt}|\hat e(t)|^2 + \|e(t)\|^2 + \|\hat e(t)\|^2 &= \|\hat U(t) - U(t)\|^2 \\
&\quad + 2\big(B(t,u(t)) - B(t,U(t)), \hat e(t)\big) + 2\big(R_U(t), \hat e(t)\big),
\end{aligned}
\tag{6.84}
$$

οπότε, λόγω της (6.83),

$$
\begin{aligned}
\frac{d}{dt}|\hat e(t)|^2 + (1-\lambda-2\varepsilon)\big(\|\hat e(t)\|^2 + \|e(t)\|^2\big) &\le 3\mu|\hat e(t)|^2 \\
&\quad + \big(1 + \frac{\tilde L^2}{2\varepsilon}\big)\|(\hat U - U)(t)\|^2 + 2\mu|(\hat U - U)(t)|^2 + \frac{1}{\varepsilon}\|R_U(t)\|_\star^2.
\end{aligned}
\tag{6.85}
$$

Με τη διακριτή ανισότητα του Gronwall καταλήγουμε τώρα εύκολα στην επιθυμητή εκ των υστέρων εκτίμηση, επιλέγοντας, φερ' ειπείν, $\varepsilon := (1-\lambda)/4$,

$$
\begin{aligned}
|\hat e(t)|^2 &+ \frac{1-\lambda}{2}\int_0^t \mathrm{e}^{3\mu(t-s)}\Big(\|\hat e(s)\|^2 + \|e(s)\|^2\Big)ds \\
&\le \int_0^t \mathrm{e}^{3\mu(t-s)}\Big[\big(1 + \frac{2\tilde L^2}{1-\lambda}\big)\|(\hat U - U)(s)\|^2 \\
&\qquad + 2\mu|(\hat U - U)(s)|^2 + \frac{4}{1-\lambda}\|R_U(s)\|_\star^2\Big]ds.
\end{aligned}
\tag{6.86}
$$

6.3.2 Μέθοδος των Crank–Nicolson

Η γραμμική περίπτωση.

Θεωρούμε το πρόβλημα (6.1) και τη διακριτοποίησή του με τη μέθοδο των Crank–Nicolson, με έναν ομοιόμορφο διαμερισμό του διαστήματος $[0,T]$, με βήμα k, βλ. την (6.29).

Αφού τα σφάλματα $u^m - U^m$ είναι δεύτερης τάξης, για να πάρουμε μια προσέγγιση $U(t)$ της $u(t)$, δεύτερης τάξης, για κάθε $t\in[0,T]$, ορίζουμε την προσέγγιση των Crank–Nicolson U της u παρεμβάλλοντας γραμμικά μεταξύ των τιμών U^{n-1} και U^n,

$$
U(t) = U^{n+\frac12} + \frac{1}{k}(U^n - U^{n-1})(t - t^{n+\frac12}), \quad t\in I_n.
\tag{6.87}
$$

Συμβολίζουμε με $R(t)\in H$,

$$
R(t) := U'(t) + AU(t) - f(t), \quad t\in I_n,
\tag{6.88}
$$

το *υπόλοιπο* της U, το οποίο αποτελεί ένα μέτρο για το μέγεθος της αποτυχίας της προσεγγιστικής λύσης U να αποτελεί ακριβή λύση του (6.1). Τώρα

$$U'(t) + AU(t) = \frac{1}{k}(U^n - U^{n-1}) + AU^{n+\frac{1}{2}} + \frac{1}{k}(t - t^{n+\frac{1}{2}})A(U^n - U^{n-1}), \quad t \in I_n,$$

οπότε, λόγω της (6.29),

$$U'(t) + AU(t) = f(t^{n+\frac{1}{2}}) + \frac{1}{k}(t - t^{n+\frac{1}{2}})A(U^n - U^{n-1}), \quad t \in I_n.$$

Επομένως, το υπόλοιπο μπορεί επίσης να γραφεί στη μορφή

$$R(t) = \frac{1}{k}(t - t^{n+\frac{1}{2}})A(U^n - U^{n-1}) + [f(t^{n+\frac{1}{2}}) - f(t)], \quad t \in I_n. \tag{6.89}$$

Προφανώς, το $R(t)$ είναι μια εκ των υστέρων ποσότητα πρώτης τάξης, ακόμη και στην περίπτωση μιας βαθμωτής συνήθους διαφορικής εξίσωσης $u'(t) = f(t)$, μολονότι η μέθοδος των Crank–Nicolson δίνει προσεγγίσεις δεύτερης τάξης. Επειδή το σφάλμα $e := u - U$ ικανοποιεί την εξίσωση $e' + Ae = -R$, αν εφαρμόσουμε τη μέθοδο της ενέργειας στην εξίσωση σφάλματος, οδηγούμαστε αναγκαστικά σε φράγματα τα οποία δεν είναι βέλτιστης τάξης.

Ανακατασκευή της προσέγγισης των Crank–Nicolson.

Προκειμένου να ανακτήσουμε τη βέλτιστη τάξη, εισάγουμε την *ανακατασκευή* $\hat{U}$ της προσέγγισης των Crank–Nicolson U, ως μια τμηματικά πολυωνυμική συνάρτηση στον χρόνο, βαθμού το πολύ δύο, $\hat{U} : [0, T] \to H$ όπως στη συνέχεια. Κατ' αρχάς εισαγάγουμε τη συνάρτηση $\varphi : I_n \to H$ ως τη γραμμική παρεμβάλλουσα της f στους κόμβους t^{n-1} και $t^{n+\frac{1}{2}}$,

$$\varphi(t) := f(t^{n+\frac{1}{2}}) + \frac{2}{k}(t - t^{n+\frac{1}{2}})[f(t^{n+\frac{1}{2}}) - f(t^{n-1})], \quad t \in I_n, \tag{6.90}$$

και εν συνεχεία ορίζουμε μια τμηματικά πολυωνυμική συνάρτηση, το πολύ δευτέρου βαθμού, Φ ως $\Phi(t) := \int_{t^{n-1}}^{t} \varphi(s)ds, t \in I_n$, δηλαδή

$$\Phi(t) = (t - t^{n-1})f(t^{n+\frac{1}{2}}) - \frac{1}{k}(t - t^{n-1})(t^n - t)[f(t^{n+\frac{1}{2}}) - f(t^{n-1})]. \tag{6.91}$$

Όπως θα καταστεί σαφές στη συνέχεια, μια πολύ σημαντική ιδιότητα της Φ είναι ότι

$$\Phi(t^{n-1}) = 0, \quad \Phi(t^n) = kf(t^{n+\frac{1}{2}}) = \int_{I_n} f(t^{n+\frac{1}{2}})dt. \tag{6.92}$$

Ορίζουμε τώρα την ανακατασκευή $\hat U$ της προσέγγισης των Crank–Nicolson U ως

$$\hat U(t) := U^{n-1} - \int_{t^{n-1}}^{t} AU(s)\,ds + \Phi(t) \quad \forall t \in I_n. \tag{6.93}$$

Προφανώς,

$$\hat U'(t) + AU(t) = \varphi(t), \quad \forall t \in I_n.$$

Υπολογίζοντας το ολοκλήρωμα στην (6.93) με τον κανόνα του τραπεζίου, λαμβάνουμε

$$\hat U(t) = U^{n-1} - \frac{1}{2}(t - t^{n-1})A[U(t) + U^{n-1}] + \Phi(t) \quad \forall t \in I_n, \tag{6.94}$$

σχέση που μπορεί να γραφεί και στη μορφή

$$\hat U(t) = U^{n-1} - A[(t - t^{n-1})U^{n-1} + \frac{1}{2k}(t - t^{n-1})^2(U^n - U^{n-1})] + \Phi(t) \quad \forall t \in I_n.$$

Προφανώς, $\hat U(t^{n-1}) = U^{n-1}$. Επί πλέον, λόγω των (6.92) και (6.29), έχουμε

$$\begin{aligned}\hat U(t^n) &= U^{n-1} - kAU^{n+\frac12} + \Phi(t^n)\\ &= U^{n-1} + k[-AU^{n+\frac12} + f(t^{n+\frac12})] = U^{n-1} + (U^n - U^{n-1}) = U^n.\end{aligned}$$

Συνεπώς, οι $\hat U$ και U συμπίπτουν στους κόμβους$t^0, \dots, t^N$. Ιδιαίτερα, η $\hat U : [0,T] \to H$ είναι συνεχής.

Λαμβάνοντας υπ' όψιν τις (6.93) και (6.90), έχουμε

$$\hat U'(t) + AU(t) = f(t^{n+\frac12}) + \frac{2}{k}(t - t^{n+\frac12})[f(t^{n+\frac12}) - f(t^{n-1})], \quad t \in I_n. \tag{6.95}$$

Συνεπώς, το *υπόλοιπο* $\hat R(t)$ της $\hat U$,

$$\hat R(t) := \hat U'(t) + A\hat U(t) - f(t), \quad t \in I_n, \tag{6.96}$$

μπορεί να γραφεί στη μορφή

$$\begin{aligned}\hat R(t) &= A[\hat U(t) - U(t)]\\ &\quad + \Big[f(t^{n+\frac12}) + \frac{2}{k}(t - t^{n+\frac12})[f(t^{n+\frac12}) - f(t^{n-1})] - f(t)\Big], \quad t \in I_n.\end{aligned} \tag{6.97}$$

Η εκ των υστέρων ποσότητα $\hat R(t)$ είναι δεύτερης τάξης.

Εκτίμηση της διαφοράς $\hat U - U$. Θα δούμε τώρα χρήσιμες παραστάσεις, καθώς και εκτιμήσεις, της διαφοράς $\hat U - U$.

Από την (6.94) παίρνουμε

$$\begin{aligned}\hat U(t)-U(t)&=U^{n-1}-U(t)-\frac{1}{2}(t-t^{n-1})A[U(t)+U^{n-1}]+\varPhi(t)\\&=-\frac{1}{k}(t-t^{n-1})(U^n-U^{n-1})-\frac{1}{2}(t-t^{n-1})A[U(t)+U^{n-1}]+\varPhi(t).\end{aligned}$$

Συνεπώς, λόγω της (6.29),

$$\begin{aligned}\hat U(t)-U(t)&=(t-t^{n-1})\big[AU^{n+\frac12}-f(t^{n+\frac12})\big]-\frac{1}{2}(t-t^{n-1})A[U(t)+U^{n-1}]+\varPhi(t)\\&=-\frac{1}{2}(t-t^{n-1})A[U(t)-U^n]+\varPhi(t)-(t-t^{n-1})f(t^{n+\frac12}),\end{aligned}$$

οπότε, λόγω της (6.91), για $t\in I_n$,

$$\hat U(t)-U(t)=(t-t^{n-1})(t^n-t)\Big(\frac{1}{2k}A(U^n-U^{n-1})-\frac{1}{k}[f(t^{n+\frac12})-f(t^{n-1})]\Big), \tag{6.98}$$

και από αυτή τη σχέση διαπιστώνουμε αμέσως ότι $\max_{t\in I_n}|\hat U(t)-U(t)|=O(k^2)$.

Εκτίμηση του σφάλματος. Όπως και πριν, συμβολίζουμε με e και $\hat e$ τα σφάλματα των προσεγγίσεων U και $\hat U$, αντίστοιχα, $e:=u-U$ και $\hat e:=u-\hat U$. Αφαιρώντας την (6.95) από τη διαφορική εξίσωση στο πρόβλημα (6.1), λαμβάνουμε

$$\hat e'(t)+Ae(t)=R_f(t) \tag{6.99}$$

με

$$R_f(t):=f(t)-\big[f(t^{n+\frac12})+\frac{2}{k}(t-t^{n+\frac12})[f(t^{n+\frac12})-f(t^{n-1})]\big],\quad t\in I_n. \tag{6.100}$$

Παίρνοντας στην (6.99) το εσωτερικό γινόμενο με $\hat e(t)$, έχουμε

$$\frac{1}{2}\frac{d}{dt}|\hat e(t)|^2+\big(Ae(t),\hat e(t)\big)=\big(R_f(t),\hat e(t)\big). \tag{6.101}$$

Τώρα,

$$\big(Ae(t),\hat e(t)\big)=\frac{1}{2}\big(\|e(t)\|^2+\|\hat e(t)\|^2-\|\hat e(t)-e(t)\|^2\big)$$

και

$$\big(R_f(t),\hat e(t)\big)\le\|R_f(t)\|_\star^2+\frac{1}{4}\|\hat e(t)\|^2,$$

συνεπώς η (6.101) δίνει

$$\frac{d}{dt}|\hat e(t)|^2+\|e(t)\|^2+\frac{1}{2}\|\hat e(t)\|^2\le\|\hat U(t)-U(t)\|^2+2\|R_f(t)\|_\star^2. \tag{6.102}$$

Ολοκληρώνοντας την (6.102) από το 0 έως το $t \le T$ και λαμβάνοντας υπ' όψιν το γεγονός ότι $\hat{e}(0) = 0$, παίρνουμε

$$
\begin{aligned}
&|\hat{e}(t)|^2 + \int_0^t \Big(\|e(s)\|^2 + \frac{1}{2}\|\hat{e}(s)\|^2\Big)\, ds \\
&\qquad \le \int_0^t \|\hat{U}(s) - U(s)\|^2\, ds + 2\int_0^t \|R_f(s)\|_\star^2\, ds.
\end{aligned}
\tag{6.103}
$$

Από την (6.103) συμπεραίνουμε εύκολα την επιθυμητή εκ των υστέρων εκτίμηση

$$
\begin{aligned}
&\max_{0\le\tau\le t}\Big[|\hat{e}(\tau)|^2 + \int_0^\tau \Big(\|e(s)\|^2 + \frac{1}{2}\|\hat{e}(s)\|^2\Big)\, ds\Big] \\
&\qquad \le \int_0^t \|\hat{U}(s) - U(s)\|^2\, ds + 2\int_0^t \|R_f(s)\|_\star^2\, ds.
\end{aligned}
\tag{6.104}
$$

Η μη γραμμική περίπτωση.

Όπως και στη γραμμική περίπτωση, ορίζουμε την προσέγγιση των Crank–Nicolson

$$
U(t) = U^{n+\frac{1}{2}} + \frac{1}{k}(U^n - U^{n-1})(t - t^{n+\frac{1}{2}}), \quad t \in I_n. \tag{6.105}
$$

Έστω $b : I_n \to H$ η γραμμική παρεμβάλλουσα της $B(\cdot, U(\cdot))$ στους κόμβους t^{n-1} και $t^{n-\frac{1}{2}}$,

$$
b(t) := B(t^{n+\frac{1}{2}}, U^{n+\frac{1}{2}}) + \frac{2}{k}(t - t^{n+\frac{1}{2}})\big[B(t^{n+\frac{1}{2}}, U^{n+\frac{1}{2}}) - B(t^{n-1}, U^{n-1})\big], \tag{6.106}
$$

$t \in I_n$. Με κίνητρο τον ορισμό στη γραμμική περίπτωση, βλ. την (6.93), ορίζουμε την ανακατασκευή $\hat{U}$ της προσέγγισης των Crank–Nicolson U ως εξής

$$
\hat{U}(t) := U^{n-1} - \int_{t^{n-1}}^t AU(s)\, ds + + \int_{t^{n-1}}^t b(s)\, ds \quad \forall t \in I_n, \tag{6.107}
$$

δηλαδή

$$
\begin{aligned}
\hat{U}(t) =& U^{n-1} - \frac{1}{2}(t - t^{n-1})A[U(t) + U^{n-1}] + (t - t^{n-1})B(t^{n+\frac{1}{2}}, U^{n+\frac{1}{2}}) \\
&+ \frac{1}{k}(t - t^{n-1})(t - t^n)\big[B(t^{n+\frac{1}{2}}, U^{n+\frac{1}{2}}) - B(t^{n-1}, U^{n-1})\big]
\end{aligned}
\tag{6.108}
$$

$t \in I_n$. Σημειώστε ότι η (6.108) ανάγεται στην (6.94) στην περίπτωση που ο τελεστής B είναι ανεξάρτητος του u. Από την (6.105) έπεται αμέσως ότι

$$
\hat{U}'(t) + AU(t) = b(t) \quad \forall t \in I_n. \tag{6.109}
$$

Επί πλέον, αμέσως διαπιστώνουμε ότι, για $t \in I_n$,

$$U(t) - \hat{U}(t) =$$
$$-\frac{1}{2k}(t-t^{n-1})(t^n-t)\Big[A(U^n-U^{n-1}) - 2\big[B(t^{n+\frac12},U^{n+\frac12}) - B(t^{n-1},U^{n-1})\big]\Big].$$

Αφαιρώντας την (6.109) από τη διαφορική εξίσωση στο πρόβλημα (6.1), παίρνουμε

$$\hat{e}'(t) + Ae(t) = B(t,u(t)) - B(t,U(t)) + R_U(t), \tag{6.110}$$

με

$$R_U(t) := B(t,U(t)) - b(t). \tag{6.111}$$

Προχωρώντας όπως στην περίπτωση της πεπλεγμένης μεθόδου του Euler, οδηγούμαστε στην επιθυμητή εκ των υστέρων εκτίμηση

$$\begin{aligned} &|\hat{e}(t)|^2 + \frac{1-\lambda}{2}\int_0^t \mathrm{e}^{3\mu(t-s)}\Big(\|\hat{e}(s)\|^2 + \|e(s)\|^2\Big)ds \\ &\le \int_0^t \mathrm{e}^{3\mu(t-s)}\Big[\big(1+\frac{2\tilde{L}^2}{1-\lambda}\big)\|(\hat{U}-U)(s)\|^2 \\ &\qquad + 2\mu|(\hat{U}-U)(s)|^2 + \frac{4}{1-\lambda}\|R_U(s)\|_\star^2\Big]ds. \end{aligned} \tag{6.112}$$

Ασκήσεις

6.1 Έστω $f:[a,b]\times\mathbb{R}\to\mathbb{R}$ μια συνεχής συνάρτηση τέτοια ώστε

$$\forall t\in[a,b]\quad \forall y_1,y_2\in\mathbb{R}\quad \big(f(t,y_1)-f(t,y_2)\big)(y_1-y_2)\le 0.$$

(Δηλαδή, για κάθε σταθερή τιμή t της πρώτης μεταβλητής, η συνάρτηση $f(t,\cdot)$ είναι φθίνουσα.) Έστω y και z οι λύσεις των προβλημάτων αρχικών τιμών

$$\begin{cases} y' = f(t,y), & t\in[a,b],\\ y(a)=y_0, \end{cases} \qquad \text{και} \qquad \begin{cases} z' = f(t,z), & t\in[a,b],\\ z(a)=z_0, \end{cases}$$

αντίστοιχα. Αποδείξτε ότι, για κάθε $t\in[a,b]$, ισχύει

$$|y(t)-z(t)|\le|y_0-z_0|.$$

6.2 Θεωρούμε τα προβλήματα αρχικών τιμών της Άσκησης 6.1, υποθέτοντας αυτή τη φορά ότι η συνεχής συνάρτηση $f:[a,b]\times\mathbb{R}\to\mathbb{R}$ ικανοποιεί τη συνθήκη

$$\forall t\in[a,b]\quad \forall y_1,y_2\in\mathbb{R}\quad \big(f(t,y_1)-f(t,y_2)\big)(y_1-y_2)\le\mu(y_1-y_2)^2,$$

για κάποια σταθερά μ. (Σημειώνουμε ότι για $\mu = 0$ η παρούσα συνθήκη συμπίπτει με εκείνη στην προηγούμενη Άσκηση.) Συνθήκες αυτής της μορφής αναφέρονται στη βιβλιογραφία ως *μονόπλευρες* συνθήκες του Lipschitz. Αποδείξτε ότι, για κάθε $t \in [a, b]$, ισχύει

$$|y(t) - z(t)| \le \mathrm{e}^{\mu(t-a)}|y_0 - z_0|.$$

6.3 Οι Ασκήσεις 6.1 και 6.2 γενικεύονται και για συστήματα συνήθων διαφορικών εξισώσεων. Η γενίκευση της Άσκησης 6.1 είναι: Έστω $f : [a, b] \times \mathbb{R}^m \to \mathbb{R}^m$ μια συνεχής συνάρτηση τέτοια ώστε

$$\forall t \in [a, b] \quad \forall y_1, y_2 \in \mathbb{R}^m \quad \big(f(t, y_1) - f(t, y_2), y_1 - y_2\big) \le 0.$$

Έστω y και z οι λύσεις των προβλημάτων αρχικών τιμών

$$\begin{cases} y' = f(t, y), & t \in [a, b], \\ y(a) = y_0, \end{cases} \qquad \text{και} \qquad \begin{cases} z' = f(t, z), & t \in [a, b], \\ z(a) = z_0, \end{cases}$$

αντίστοιχα. Αποδείξτε ότι, για κάθε $t \in [a, b]$, ισχύει

$$\|y(t) - z(t)\| \le \|y_0 - z_0\|.$$

Συμβολίσαμε εδώ με $\big(\cdot, \cdot\big)$ και $\| \cdot \|$ το Ευκλείδειο εσωτερικό γινόμενο και την Ευκλείδεια νόρμα, αντίστοιχα, στον $\mathbb{R}^m$.

6.4 Θεωρούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} y' = f(t, y), & t \in [a, b], \\ y(a) = y_0, \end{cases}$$

και υποθέτουμε ότι η συνάρτηση f πληροί τη μονόπλευρη συνθήκη του Lipschitz που δίνεται στην Άσκηση 6.2. Βεβαιωθείτε ότι με την αλλαγή μεταβλητής

$$u(t) := \mathrm{e}^{-\mu(t-a)} y(t)$$

το πρόβλημα γράφεται στη μορφή

$$\begin{cases} u' = F(t, u), & t \in [a, b], \\ u(a) = y_0, \end{cases}$$

με

$$F(t, v) := \mathrm{e}^{-\mu(t-a)} f\big(t, \mathrm{e}^{\mu(t-a)} v\big) - \mu v$$

και αποδείξτε ότι η συνάρτηση F ικανοποιεί τη συνθήκη που δίνεται στην Άσκηση 6.1,

$$\forall t \in [a, b] \quad \forall y_1, y_2 \in \mathbb{R} \quad \big(F(t, y_1) - F(t, y_2)\big)(y_1 - y_2) \le 0.$$

6.5 Έστω $a \in \mathbb{R}$ και $f : [0, \infty) \to \mathbb{R}$ μια συνεχής συνάρτηση. Αποδείξτε ότι η λύση y του προβλήματος αρχικών τιμών

$$\begin{cases} y'(t) = ay(t) + f(t), \quad t \geq 0, \\ y(0) = y_0 \end{cases}$$

δίνεται από τη σχέση

$$y(t) = \mathrm{e}^{at} y_0 + \int_0^t \mathrm{e}^{a(t-s)} f(s)\, ds, \quad t \geq 0.$$

Σημειώνουμε ότι ο πρώτος όρος στο δεξιό μέλος παριστά τη λύση του προβλήματος για την ομογενή εξίσωση

$$\begin{cases} x'(t) = ax(t), \quad t \geq 0, \\ x(0) = y_0 \end{cases}$$

ενώ η ολοκληρωτέα ποσότητα $\mathrm{e}^{a(t-s)} f(s)$ παριστά την τιμή στο σημείο t της λύσης του προβλήματος

$$\begin{cases} x'(t) = ax(t), \quad t \geq s, \\ x(s) = f(s). \end{cases}$$

6.6 Έστω $M \in \mathbb{R}^{m,m}$ ένας πίνακας.

α) Κατ' αναλογίαν προς τον ορισμό της εκθετικής συνάρτησης e^x για πραγματικό αριθμό x, ορίζουμε τον πίνακα e^M ως

$$\mathrm{e}^M := \sum_{\ell=0}^{\infty} \frac{1}{\ell!} M^\ell.$$

Έστω $\|\cdot\|$ μια νόρμα πινάκων. Αποδείξτε ότι

$$\forall \varepsilon > 0 \quad \exists n \in \mathbb{N} \quad \forall k \in \mathbb{N} \quad \|\sum_{\ell=n}^{n+k} \frac{1}{\ell!} M^\ell\| \leq \varepsilon,$$

δηλαδή ότι η σειρά συγκλίνει και συνεπώς ο πίνακας e^M είναι καλά ορισμένος.

[*Υπόδειξη*: Χρησιμοποιήστε την εκτίμηση

$$\|\sum_{\ell=n}^{n+k} \frac{1}{\ell!} M^\ell\| \leq \sum_{\ell=n}^{n+k} \frac{1}{\ell!} \|M\|^\ell$$

και το γεγονός ότι η σειρά $\sum_{\ell=0}^{\infty} \frac{1}{\ell!} x^\ell$ συγκλίνει για κάθε $x \in \mathbb{R}$.]

β) Θεωρούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} y'(t) = My(t), \quad t \geq 0, \\ y(0) = y_0. \end{cases}$$

Αποδείξτε ότι η λύση του y δίνεται από τη σχέση

$$y(t) = \mathrm{e}^{tM} y_0, \quad t \geq 0.$$

[*Υπόδειξη*: Αποδείξτε ότι

$$(\mathrm{e}^{tM})' = \Big(\sum_{\ell=0}^{\infty} \frac{1}{\ell!} t^\ell M^\ell\Big)' = \sum_{\ell=1}^{\infty} \frac{1}{(\ell-1)!} t^{\ell-1} M^\ell = M\mathrm{e}^{tM}.]$$

γ) Συμβολίζουμε με $E(t)$ τον *τελεστή λύσης* του προβλήματος αρχικών τιμών του μέρους β), δηλαδή $E(t) = \mathrm{e}^{tM}$, οπότε $y(t) = E(t)y_0$. Κατ' αναλογίαν προς τη σχέση $\mathrm{e}^{x+y} = \mathrm{e}^x \mathrm{e}^y$ για πραγματικούς αριθμούς x και y, ισχύει και για πίνακες $A, B \in \mathbb{R}^{m,m}$ ότι $\mathrm{e}^{A+B} = \mathrm{e}^A \mathrm{e}^B$, υπό την προϋπόθεση ότι οι A και B *αντιμετατίθενται*, δηλαδή $AB = BA$. Δεχθείτε αυτό το γεγονός και αποδείξτε ότι ο τελεστής $E(t)$ έχει την *ιδιότητα της ημιομάδας*, δηλαδή ότι

$$E(\sigma + \tau) = E(\sigma)E(\tau) \quad \forall \sigma, \tau \geq 0.$$

Αυτό σημαίνει ότι αν λύσουμε διαδοχικά τα προβλήματα

$$\begin{cases} x'(t) = Mx(t), \quad 0 \leq t \leq \sigma, \\ x(0) = y_0. \end{cases}$$

και

$$\begin{cases} x'(t) = Mx(t), \quad \sigma \leq t \leq \sigma + \tau, \\ x(\sigma) = E(\sigma)y_0 \end{cases}$$

ή το αρχικό πρόβλημα στο διάστημα $[0, \sigma + \tau]$ παίρνουμε το ίδιο αποτέλεσμα, $y(\sigma + \tau) = x(\sigma + \tau)$. Με άλλα λόγια, στην τιμή της λύσης στη χρονική στιγμή $\sigma + \tau$ μπορούμε να οδηγηθούμε είτε λύνοντας το αρχικό πρόβλημα στο διάστημα $[0, \sigma + \tau]$, είτε λύνοντας το πρόβλημα κατ' αρχάς στο διάστημα $[0, \sigma]$ και εν συνεχεία στο διάστημα $[\sigma, \sigma + \tau]$, αλλά φυσικά με τη σωστή τώρα αρχική τιμή $E(\sigma)y_0$.

6.7 Έστω $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_m)$ ένας διαγώνιος πίνακας και $M \in \mathbb{R}^{m,m}$ ένας διαγωνιοποιήσιμος πίνακας, $M = U\Lambda U^{-1}$ με $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_m)$. Αποδείξτε ότι

α) $\mathrm{e}^\Lambda = \operatorname{diag}\left(\mathrm{e}^{\lambda_1}, \dots, \mathrm{e}^{\lambda_m}\right)$

β) $\mathrm{e}^M = U\mathrm{e}^\Lambda U^{-1}$, οπότε, φυσικά, $\mathrm{e}^{tM} = U\mathrm{e}^{t\Lambda} U^{-1}$, για $t \in \mathbb{R}$.

6.8 Έστω $f : [0, \infty) \to \mathbb{R}^m$ μια συνεχής συνάρτηση. Με τους συμβολισμούς της Άσκησης 6.6, θεωρούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} y'(t) = My(t) + f(t), \quad t \geq 0, \\ y(0) = y_0. \end{cases}$$

Αποδείξτε ότι

$$\big(\mathrm{e}^{-tM}y(t)\big)' = \mathrm{e}^{-tM}f(t)$$

και οδηγηθείτε στο συμπέρασμα ότι

$$y(t) = \mathrm{e}^{tM}y_0 + \int_0^t \mathrm{e}^{(t-s)M}f(s)\,ds, \quad t \geq 0.$$

(Συγκρίνετε με το αντίστοιχο αποτέλεσμα στην Άσκηση 6.5.) Η σχέση αυτή γράφεται και στη μορφή

$$y(t) = E(t)y_0 + \int_0^t E(t-s)f(s)\,ds,$$

όπου $E(t) = \mathrm{e}^{tM}$, η οποία συνήθως αναφέρεται ως *αρχή του Duhamel*. Σημειώνουμε ότι ο πρώτος όρος στο δεξιό μέλος είναι η τιμή στο σημείο t της λύσης του αντίστοιχου ομογενούς προβλήματος

$$\begin{cases} x'(t) = Mx(t), \quad t \geq 0, \\ x(0) = y_0, \end{cases}$$

ενώ η παράσταση $E(t-s)f(s)$ είναι η τιμή στο σημείο t της λύσης του προβλήματος

$$\begin{cases} x'(t) = Mx(t), \quad t \geq s, \\ x(s) = f(s). \end{cases}$$

6.9 Έστω $M \in \mathbb{C}^{m,m}$ ένας πίνακας με ιδιοτιμές $\lambda_1, \dots, \lambda_m$ με μη θετικό πραγματικό μέρος, $\operatorname{Re}\lambda_i \leq 0$, $i = 1, \dots, m$. Θεωρούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} y'(t) = My(t), \quad t \geq 0, \\ y(0) = y_0 \end{cases}$$

με $y_0 \neq 0$. Σκοπός μας στην παρούσα Άσκηση, καθώς και σε ορισμένες που ακολουθούν, είναι να μελετήσουμε τη συμπεριφορά του λόγου

$$\frac{\|y(t)\|}{\|y_0\|},$$

όπου $\|\cdot\|$ μια νόρμα στον $\mathbb{C}^m$.

α) Αν $m = 1$, αποδείξτε ότι
$$|y(t)| \le |y_0|, \quad t \ge 0.$$
β) Αν $m = 2$, και $M = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, οπότε, προφανώς, $\lambda_1 = \lambda_2 = 0$, αποδείξτε ότι η λύση $y(t)$ δίνεται από τη σχέση
$$y(t) = \begin{pmatrix} (y_0)_1 + (y_0)_2 t \\ (y_0)_2 \end{pmatrix}, \quad t \ge 0,$$
όπου $(y_0)_1$ και $(y_0)_2 \ne 0$ οι συνιστώσες του y_0, και οδηγηθείτε στο συμπέρασμα ότι
$$\frac{\|y(t)\|}{\|y_0\|} \to \infty, \quad t \to \infty,$$
σε αντιδιαστολή με την περίπτωση $m = 1$ που εξετάστηκε προηγουμένως.

6.10 Έστω $\mu \in \mathbb{R}$. Χρησιμοποιούμε τον συμβολισμό της Άσκησης 6.9 και υποθέτουμε τώρα ότι $M = \left(\begin{smallmatrix} -1 & \mu \\ 0 & 0 \end{smallmatrix}\right)$, οπότε οι ιδιοτιμές είναι $\lambda_1 = -1$ και $\lambda_2 = 0$. Η λύση $y(t)$ δίνεται από τη σχέση
$$y(t) = \begin{pmatrix} (y_0)_1 \mathrm{e}^{-t} + (y_0)_2 \mu (1 - \mathrm{e}^{-t}) \\ (y_0)_2 \end{pmatrix}, \quad t \ge 0.$$
Ιδιαίτερα, για $(y_0)_1 = 0$ και μια νόρμα $\|\cdot\|_p$ $(p \ge 1)$ στον $\mathbb{R}^2$, αποδείξτε ότι
$$\frac{\|y(t)\|_p}{\|y_0\|_p} \le \big(1 + |\mu|^p\big)^{1/p}, \quad t \ge 0.$$
Συγκρίνετε πάλι με το αποτέλεσμα στην περίπτωση $m = 1$ που εξετάστηκε στην Άσκηση 6.9α, όπου ο αντίστοιχος λόγος φράσσεται από τη μονάδα.

6.11 Έστω $\lambda \in \mathbb{C}$ ένας αριθμός με *αρνητικό* πραγματικό μέρος, $\mathrm{Re}\,\lambda < 0$. Χρησιμοποιούμε τον συμβολισμό της Άσκησης 6.9 και υποθέτουμε τώρα ότι ο πίνακας M έχει το λ στις θέσεις της διαγωνίου του, τη μονάδα στις θέσεις της υπερδιαγωνίου του και μηδενικά στις υπόλοιπες θέσεις,
$$M = \begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \lambda \end{pmatrix},$$
οπότε οι ιδιοτιμές είναι $\lambda_1 = \cdots = \lambda_m = \lambda$. Βεβαιωθείτε ότι η λύση $y(t) = \big(y_1(t), \dots, y_m(t)\big)^T$ δίνεται από τις σχέσεις
$$\begin{cases} y_n(t) = y_n(0)\mathrm{e}^{\lambda t}, \\ y_i(t) = y_i(0)\mathrm{e}^{\lambda t} + \displaystyle\int_0^t y_{i+1}(s)\mathrm{e}^{\lambda(t-s)} ds, \quad i = m-1, \dots, 1, \end{cases}$$

και χρησιμοποιώντας το γεγονός ότι η συνάρτηση $\varphi, \varphi(t) := \int_0^t |\mathrm{e}^{\lambda(t-s)}|\, ds$, είναι ομοιόμορφα φραγμένη, συμπεράνετε ότι

$$\|y(t)\|_\infty \le C\|y_0\|_\infty, \quad t \ge 0,$$

για κάποια σταθερά C. Συγκρίνετε με την περίπτωση που εξετάστηκε στην Άσκηση 6.9β.

6.12 Χρησιμοποιούμε τον συμβολισμό της Άσκησης 6.9 και υποθέτουμε τώρα ότι, πέραν της υπόθεσης $\operatorname{Re}\lambda_i \le 0, i = 1, \dots, m$, ισχύει $\operatorname{Re}\lambda_i < 0$, αν το λ_i είναι *πολλαπλή* ιδιοτιμή. Αποδείξτε ότι

$$\|y(t)\| \le C\|y_0\|, \quad t \ge 0,$$

για κάποια σταθερά C που εξαρτάται και από τη νόρμα $\|\cdot\|$.
[*Υπόδειξη*: Για $m = 1$ ο ισχυρισμός ισχύει προφανώς. Για $m > 1$, έστω $T \in \mathbb{C}^{m,m}$ ένας πίνακας τέτοιος ώστε $T^{-1}MT = J$ να είναι η κανονική μορφή του Jordan του πίνακα M. Με $x(t) := T^{-1}y(t)$ γράψτε το σύστημα των διαφορικών εξισώσεων στη μορφή

$$x'(t) = Jx(t).$$

Το πρόβλημα διασπάται τώρα σε προβλήματα της μορφής που μελετήθηκαν στην Άσκηση 6.8 και προβλήματα με $m = 1$. Επομένως

$$\|x(t)\|_\infty \le \tilde{C}\|x(0)\|_\infty, \quad t \ge 0,$$

και η ισοδυναμία των νορμών στον $\mathbb{C}^m$ οδηγεί στο αποτέλεσμα.]

6.13 Με τον συμβολισμό της Άσκησης 6.9, υποθέτουμε ότι ο πίνακας $M \in \mathbb{R}^{m,m}$ είναι *συμμετρικός*, και ότι οι (πραγματικές) ιδιοτιμές του είναι μη θετικές, $\lambda_i \le 0,\ 1 \le i \le m$.

α) Έστω $\varphi : (-\infty, 0] \to \mathbb{R}$ μια φραγμένη, συνεχής συνάρτηση. Ο τελεστής $\varphi(M) \in \mathbb{R}^{m,m}$ ορίζεται ως εξής: Θεωρούμε τα ορθομοναδιαία, ως προς το Ευκλείδειο εσωτερικό γινόμενο $(\cdot,\cdot)$ στον $\mathbb{R}^m$, ιδιοδιανύσματα του M, $v^{(i)}, i = 1, \dots, m$, τέτοια ώστε $Mv^{(i)} = \lambda_i v^{(i)},\ 1 \le i \le m$. Για κάθε $v \in \mathbb{R}^m$ ορίζουμε τη δράση του $\varphi(M)$ από τη σχέση ('φασματική' παράσταση του M)

$$\varphi(M)v = \sum_{i=1}^m \varphi(\lambda_i)(v, v^{(i)})v^{(i)} \ .$$

Αποδείξτε ότι

$$\|\varphi(M)\|_2 = \max_{1\le i\le m} |\varphi(\lambda_i)|,$$

όπου $\|\cdot\|_2$ η νόρμα πινάκων που παράγεται από την Ευκλείδεια νόρμα στον $\mathbb{R}^m$.

β) Χρησιμοποιώντας την παράσταση

$$y(t) = \mathrm{e}^{tM} y(0), \quad t \geq 0,$$

της λύσης y του προβλήματός μας, αποδείξτε ότι ισχύει

$$\|y(t)\|_2 \leq \mathrm{e}^{t(\max_i \lambda_i)} \|y(0)\|_2, \quad t \geq 0.$$

γ) Αποδείξτε ότι ο ορισμός του e^{tM} που χρησιμοποιήσαμε εδώ είναι συμβατός με εκείνον της Άσκησης 6.6α.

6.14 Έστω $M \in \mathbb{R}^{m,m}$ ένας *μη θετικά* ορισμένος πίνακας, $(Mx, x) \leq 0$ για κάθε $x \in \mathbb{R}^m$. Θεωρούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} y'(t) = My(t), & t \geq 0, \\ y(0) = y_0. \end{cases}$$

Αποδείξτε ότι η Ευκλείδεια νόρμα $\|y(\cdot)\|$ είναι φθίνουσα συνάρτηση.

6.15 Θεωρούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} x'(t) = -2x(t) + y(t), & t \geq 0, \\ y'(t) = 2x(t) - 2y(t), & t \geq 0, \\ x(0) = x_0, \\ y(0) = y_0. \end{cases}$$

Αποδείξτε ότι η συνάρτηση $[x(\cdot)]^2 + [y(\cdot)]^2$ είναι φθίνουσα.
[*Υπόδειξη*: Ο πίνακας $\left(\begin{smallmatrix} -2 & 1 \\ 2 & -2 \end{smallmatrix}\right)$ είναι αρνητικά ορισμένος. Βλ. την Άσκηση 6.14.]

6.16 Έστω $M \in \mathbb{R}^{m,m}$ ένας *αντισυμμετρικός* πίνακας, δηλαδή τέτοιος ώστε για τα στοιχεία του να ισχύει $M_{ij} = -M_{ji}$, $i, j = 1, \dots, m$. Θεωρούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} y'(t) = My(t), & t \geq 0, \\ y(0) = y_0. \end{cases}$$

Αποδείξτε ότι η Ευκλείδεια νόρμα $\|y(\cdot)\|$ είναι σταθερή συνάρτηση, $\|y(t)\| = \|y(0)\|$, για κάθε $t \geq 0$.
[*Υπόδειξη*: Αποδείξτε ότι $(Mx, x) = 0$ για κάθε $x \in \mathbb{R}^m$. Πάρτε τώρα στο σύστημα των διαφορικών εξισώσεων το εσωτερικό γινόμενο με $y(t)$ και χρησιμοποιήστε την προαναφερθείσα ιδιότητα.]

6.17 Έστω $M \in \mathbb{R}^{m,m}$ ένας συμμετρικός πίνακας. Θεωρούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} y'(t) = \mathrm{i}\, My(t), & t \geq 0, \\ y(0) = y_0 \end{cases}$$

για μια συνάρτηση $y : [0, \infty) \to \mathbb{C}^m$, όπου i η φανταστική μονάδα. Αποδείξτε ότι η Ευκλείδεια νόρμα $\|y(\cdot)\|$ είναι σταθερή συνάρτηση, $\|y(t)\| = \|y(0)\|$, για κάθε $t \geq 0$.
[*Υπόδειξη*: Πάρτε στο σύστημα διαφορικών εξισώσεων το εσωτερικό γινόμενο με $y(t)$ και χρησιμοποιήστε το γεγονός ότι στη μιγαδική περίπτωση ισχύει

$$\frac{d}{dt}\|y(t)\|^2 = \frac{d}{dt}\big(y(t), y(t)\big) = \big(y'(t), y(t)\big) + \big(y(t), y'(t)\big) = 2\mathrm{Re}\,\big(y'(t), y(t)\big),$$

δηλαδή $\mathrm{Re}\,\big(y'(t), y(t)\big) = \frac{1}{2}\frac{d}{dt}\|y(t)\|^2$. Επί πλέον ισχύει $(Mz, z) \in \mathbb{R}$, για κάθε $z \in \mathbb{C}^m$.]

6.18 Θεωρούμε μια συνάρτηση $f : \mathbb{R} \to \mathbb{R}$ η οποία είναι συνεχώς παραγωγίσιμη και τέτοια ώστε $f'(x) \leq 0$ για κάθε πραγματικό αριθμό $x, x \in \mathbb{R}$. Διακριτοποιούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} y'(t) = f\big(y(t)\big), & t \in [a, b], \\ y(a) = y_0, \end{cases}$$

με τη μέθοδο του τραπεζίου, θεωρώντας έναν ομοιόμορφο διαμερισμό του διαστήματος $[a, b]$. Αποδείξτε ότι οι προσεγγίσεις είναι καλά ορισμένες.

6.19 Έστω $\varphi : [0, 1] \to \mathbb{R}$ μια συνεχής συνάρτηση. Διακριτοποιούμε το πρόβλημα αρχικών τιμών

$$\begin{cases} y'(t) = -\big(y(t)\big)^3 + \varphi(t), & t \in [0, 1], \\ y(0) = 1, \end{cases}$$

με τη μέθοδο του τραπεζίου, θεωρώντας έναν ομοιόμορφο διαμερισμό του διαστήματος $[0, 1]$ με βήμα h. Αποδείξτε ότι οι προσεγγίσεις είναι καλά ορισμένες.

6.20 Σύμφωνα με το θεώρημα *σταθερού σημείου του Brouwer*, αν K είναι ένα μη κενό, κυρτό, κλειστό και φραγμένο υποσύνολο του $\mathbb{R}^m$ και $f : K \to K$ μια συνεχής συνάρτηση, τότε η f έχει στο K τουλάχιστον ένα σταθερό σημείο, δηλαδή υπάρχει $x^\star \in K$ τέτοιο ώστε $f(x^\star) = x^\star$. Για την απλούστερη περίπτωση, όπου $m = 1$ και $K = [a, b]$, βλ. την Πρόταση 2.2 στο [2]. Το θεώρημα αυτό μας είναι χρήσιμο πολλές φορές για να αποδείξουμε ύπαρξη αριθμητικών λύσεων.

Χρησιμοποιήστε το θεώρημα σταθερού σημείου του Brouwer για να αποδείξετε την εξής ιδιαίτερα εύχρηστη εκδοχή του: Έστω $g : \mathbb{R}^m \to \mathbb{R}^m$ μια συνεχής συνάρτηση τέτοια ώστε $(g(x), x) \geq 0$, για κάθε $x \in \mathbb{R}^m$ με $\|x\| = \alpha$, για έναν θετικό αριθμό α. Τότε υπάρχει $x^\star \in \mathbb{R}^m$ τέτοιο ώστε $g(x^\star) = 0$ και $\|x^\star\| \leq \alpha$. Τα $(\cdot, \cdot)$ και $\|\cdot\|$ συμβολίζουν εδώ το Ευκλείδειο εσωτερικό γινόμενο και την Ευκλείδεια νόρμα στον $\mathbb{R}^m$, αντίστοιχα.
[*Υπόδειξη*: Το σύνολο $K := \{x \in \mathbb{R}^m : \|x\| \leq \alpha\}$ είναι προφανώς μη κενό, κυρτό, κλειστό και φραγμένο. Αν $g(x) \neq 0$ για κάθε $x \in K$, τότε η συνάρτηση $f : K \to K$,

$$f(x) = -\alpha \frac{g(x)}{\|g(x)\|},$$

θα ήταν συνεχής και, σύμφωνα με το θεώρημα σταθερού σημείου του Brouwer, θα είχε ένα σταθερό σημείο $x^\star$ στο K. Οδηγηθείτε στο άτοπο ότι συγχρόνως $\|x^\star\| = \alpha$ και $\|x^\star\|^2 = (f(x^\star), x^\star) \leq 0$.]

6.21 Δώστε μια απόδειξη ύπαρξης και μοναδικότητας των προσεγγίσεων με την πεπλεγμένη μέθοδο του Euler για προβλήματα αρχικών τιμών που ικανοποιούν τη συνθήκη της Άσκησης 6.1, χωρίς περιορισμό στο h, η οποία να μπορεί να γενικευθεί και για συστήματα Σ.Δ.Ε., βλ. τις Ασκήσεις 6.22 και 6.23.
[*Υπόδειξη*: Θεωρούμε τη συνεχή συνάρτηση g, $g(x) := x - y^n - hf(t^{n+1}, x)$, $x \in \mathbb{R}$. Για $x \in \mathbb{R}$, έχουμε

$$\begin{aligned} g(x) \cdot x &= x^2 - y^n \cdot x - hf(t^{n+1}, x) \cdot x \\ &= x^2 - y^n \cdot x - h[f(t^{n+1}, x) - f(t^{n+1}, 0)] \cdot x - hf(t^{n+1}, 0) \cdot x\ , \end{aligned}$$

οπότε, χρησιμοποιώντας τη συνθήκη της Άσκησης 6.1, έχουμε

$$g(x) \cdot x \geq x^2 - [y^n + hf(t^{n+1}, 0)] \cdot x.$$

Συνεπώς, επειδή $[y^n + hf(t^{n+1}, 0)] \cdot x \leq \frac{1}{2}\big(x^2 + [y^n + hf(t^{n+1}, 0)]^2\big)$,

$$g(x) \cdot x \geq \frac{1}{2}\big(|x|^2 - |y^n + hf(t^{n+1}, 0)|^2\big) \quad \forall x \in \mathbb{R}.$$

Για $x \in \mathbb{R}, |x| = |y^n + hf(t^{n+1}, 0)| + 1$, έχουμε επομένως $g(x) \cdot x > 0$. Συμπεραίνουμε ότι η g έχει διαφορετικά πρόσημα στα άκρα του διαστήματος

$$[-|y^n + hf(t^{n+1}, 0)| - 1, |y^n + hf(t^{n+1}, 0)| + 1],$$

συνεπώς έχει τουλάχιστον μία ρίζα σε αυτό το διάστημα. Όσον αφορά τώρα τη μοναδικότητα, υποθέτοντας ότι $x, y \in \mathbb{R}$ είναι ρίζες της g, έχουμε

$$g(x) - g(y) = 0$$

ή

$$x - y = h[f(t^{n+1}, x) - f(t^{n+1}, y)].$$

Πολλαπλασιάζοντας επί $x - y$ και χρησιμοποιώντας τη (2.43) παίρνουμε $(x - y)^2 \leq 0$, οπότε $x = y$.]

6.22 Θεωρούμε ένα πρόβλημα αρχικών τιμών της μορφής

$$\begin{cases} y'(t) = f(t, y), \quad t \in [a, b], \\ y(a) = y_0, \end{cases}$$

όπου $f : [a, b] \times \mathbb{R}^m \to \mathbb{R}^m$ μια συνεχής συνάρτηση τέτοια ώστε

$$\forall t \in [a, b] \quad \forall y_1, y_2 \in \mathbb{R}^m \quad \big(f(t, y_1) - f(t, y_2), y_1 - y_2\big) \le 0,$$

με $(\cdot, \cdot)$ το Ευκλείδειο εσωτερικό γινόμενο στον $\mathbb{R}^m$. Αποδείξτε ότι οι προσεγγίσεις με την πεπλεγμένη μέθοδο του Euler, ως προς έναν ομοιόμορφο διαμερισμό με βήμα h, είναι καλά ορισμένες.
[*Υπόδειξη*: Θεωρήστε τη συνάρτηση $g : \mathbb{R}^m \to \mathbb{R}^m$, $g(x) := x - y^n - hf(t^{n+1}, x)$. Τότε

$$\begin{aligned}\big(g(x), x\big) &= \|x\|^2 - (y^n + hf(t^{n+1}, 0), x) - h(f(t^{n+1}, x) - f(t^{n+1}, 0), x)\\ &\ge \|x\|^2 - (y^n + hf(t^{n+1}, 0), x)\\ &\ge \frac{1}{2}\big[\|x\|^2 - \|y^n + hf(t^{n+1}, 0)\|^2\big].\end{aligned}$$

Για $x \in \mathbb{R}^m$ τέτοια ώστε $\|x\| = \|y^n + hf(t^{n+1}, 0)\| + 1$ το δεξιό μέλος είναι προφανώς θετικό. Χρησιμοποιήστε τώρα την Άσκηση 6.20 για να αποδείξετε ύπαρξη της y^{n+1}. Η μοναδικότητα αποδεικνύεται εύκολα.]

6.23 Με τον συμβολισμό της Άσκησης 6.22, υποθέτουμε τώρα ότι η $f : [a, b] \times \mathbb{R}^m \to \mathbb{R}^m$ είναι συνεχής και ικανοποιεί τη γενικότερη συνθήκη

$$\forall t \in [a, b] \quad \forall y_1, y_2 \in \mathbb{R}^m \quad \big(f(t, y_1) - f(t, y_2), y_1 - y_2\big) \le \nu\|y_1 - y_2\|^2,$$

με μια θετική σταθερά ν. Αποδείξτε ότι οι προσεγγίσεις με την πεπλεγμένη μέθοδο του Euler, ως προς έναν ομοιόμορφο διαμερισμό με βήμα h, είναι καλά ορισμένες, αν το γινόμενο νh είναι αρκετά μικρό.

6.24 Εκτός από το θεώρημα σταθερού σημείου του Brouwer, σημαντικό ρόλο σε αποδείξεις ύπαρξης και μοναδικότητας προσεγγιστικών λύσεων παίζει και το ακόλουθο αποτέλεσμα, το οποίο αναφέρεται στη βιβλιογραφία ως θεώρημα σταθερού σημείου του Zarantonello: Έστω $G : \mathbb{R}^m \to \mathbb{R}^m$ μια *ισχυρά μονότονη* απεικόνιση, δηλαδή τέτοια ώστε

$$\forall v, w \in \mathbb{R}^m \quad \big(G(v) - G(w), v - w\big) \ge c\|v - w\|^2,$$

με μια θετική σταθερά c, η οποία ικανοποιεί επί πλέον και τη συνθήκη του Lipschitz,

$$\forall v, w \in \mathbb{R}^m \quad \|G(v) - G(w)\| \le L\|v - w\|.$$

Αποδείξτε ότι η G μηδενίζεται σε ένα ακριβώς σημείο.
[*Υπόδειξη*: Αρκεί να αποδείξουμε ότι η απεικόνιση

$$F : \mathbb{R}^m \to \mathbb{R}^m, \quad F(v) := v - \frac{c}{L^2}G(v),$$

έχει ένα ακριβώς σταθερό σημείο. Παρατηρήστε κατ' αρχάς ότι

$$\|F(v) - F(w)\|^2 = \|v - w\|^2 - 2\frac{c}{L^2}\big(G(v) - G(w), v - w\big) + \frac{c^2}{L^4}\|G(v) - G(w)\|^2$$

και οδηγηθείτε στο συμπέρασμα ότι

$$\|F(v) - F(w)\| \leq \sqrt{1 - \frac{c^2}{L^2}}\,\|v - w\|,$$

δηλαδή ότι η F είναι συστολή.]

6.25 Με τον συμβολισμό της Άσκησης 6.22, υποθέτουμε τώρα ότι η $f : [a, b] \times \mathbb{R}^m \to \mathbb{R}^m$ ικανοποιεί τη συνθήκη

$$\forall t \in [a, b] \quad \forall y_1, y_2 \in \mathbb{R}^m \quad \big(f(t, y_1) - f(t, y_2), y_1 - y_2\big) \leq \nu\|y_1 - y_2\|^2,$$

με μια θετική σταθερά ν, καθώς και τη συνθήκη του Lipschitz

$$\forall t \in [a, b] \quad \forall y_1, y_2 \in \mathbb{R}^m \quad \|f(t, y_1) - f(t, y_2)\| \leq L\|y_1 - y_2\|.$$

Χρησιμοποιήστε την Άσκηση 6.24 για να αποδείξετε ότι οι προσεγγίσεις που δίνει η πεπλεγμένη μέθοδο του Euler, ως προς έναν ομοιόμορφο διαμερισμό με βήμα h, είναι καλά ορισμένες, αν το γινόμενο νh είναι αρκετά μικρό.

[Σημειώνουμε ότι ναι μεν υποθέσαμε ότι ισχύει η συνθήκη του Lipschitz, η σταθερά L όμως δεν υπεισέρχεται στον περιορισμό για το βήμα h. Με μια πρώτη ματιά στις Ασκήσεις 6.22 και 6.24 θα μπορούσε να πει κανείς ότι είναι προτιμότερο να χρησιμοποιήσουμε το θεώρημα σταθερού σημείου του Brouwer παρά το αντίστοιχο του Zarantonello, αφού στη δεύτερη περίπτωση απαιτείται και η συνθήκη του Lipschitz. Όμως, το πλεονέκτημα του θεωρήματος σταθερού σημείου του Zarantonello συνίσταται στο ότι αυτό μπορεί να μας οδηγήσει σε αποδείξεις ύπαρξης προσεγγιστικών λύσεων και σε απειροδιάστατους χώρους, αρκεί αυτοί να είναι πλήρεις, όπως απαιτείται στο θεώρημα της συστολής, ενώ στο θεώρημα σταθερού σημείου του Brouwer ο χώρος πρέπει να έχει πεπερασμένη διάσταση, όπως ο $\mathbb{R}^m$.]

Βιβλιογραφία

1. R. A. Adams, J. J. F. Fournier: *Sobolev Spaces*. 2nd ed., 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, β′ ανατύπωση, 2006.)
3. Γ. Δ. Ακρίβης, Β. Α. Δουγαλής: *Αριθμητικές Μέθοδοι για Συνήθεις Διαφορικές Εξισώσεις*. Πανεπιστημιακές Εκδόσεις Κρήτης, Ηράκλειο, 2006.
4. Γ. Δ. Ακρίβης, Β. Α. Δουγαλής: *Αριθμητικές Μέθοδοι για Διαφορικές Εξισώσεις*. Ιωάννινα, 2004.
5. Ν. Δ. Αλικάκος, Γ. Η. Καλογερόπουλος: *Συνήθεις Διαφορικές Εξισώσεις*. Σύγχρονη Εκδοτική, Αθήνα, 2003.
6. N. Bakaev: *Linear Discrete Parabolic Problems*. North–Holland Mathematical Studies v. 203, Elsevier, Amsterdam, 2006.
7. G. Bangerth, R. Rannacher: *Adaptive Finite Element Methods for Differential Equations*. Birkhäuser, Basel, 2003.
8. S. C. Brenner, L. R. Scott: *The Mathematical Theory of Finite Element Methods*. Springer–Verlag, New York, 1994. (2nd ed., 2002.)
9. J. Butcher: *The Numerical Analysis of Ordinary Differential Equations – Runge–Kutta and General Linear Methods*. Wiley, Chichester, 1987.
10. W. Cheney: *Analysis for Applied Mathematics*. Springer–Verlag, New York, 2001.
11. P. G. Ciarlet: *The Finite Element Method for Elliptic Problems*. North Holland, Amsterdam, 1978. (Ανατύπωση SIAM, Classics in Applied Mathematics v. 40, Philadelphia, 2002.)

12. 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.

13. P. G. Ciarlet, J. L. Lions (editors): *Handbook of Numerical Analysis, vol. II, Finite Element Methods* (*Part 1*). North Holland, Amsterdam, 1991.

14. 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.

15. 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.

16. P. G. Ciarlet, J. L. Lions (editors): *Handbook of Numerical Analysis, vol. V, Techniques of Scientific Computing* (*Part 2*). North Holland, Amsterdam, 1997.

17. 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.

18. 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.

19. R. A. DeVore, G. G. Lorentz: *Constructive Approximation*. Springer–Verlag, Berlin, 1993.

20. Β. Α. Δουγαλής: *Μέθοδοι Πεπερασμένων Στοιχείων για Μερικές Διαφορικές Εξισώσεις*. Αθήνα, 2004.

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*. North–Holland, Amsterdam, 2001.
25. E. Hairer, Ch. Lubich, G. Wanner: *Geometric Numerical Integration: Structure–Preserving Algorithms for Ordinary Differential Equations*. Springer–Verlag, Berlin, 2002.
26. 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.
27. E. Hairer, G. Wanner: *Solving Ordinary Differential Equations II – Stiff and Differential–Algebraic Problems*. Springer–Verlag, Berlin, 2nd revised ed., 2002.
28. P. Henrici: *Discrete Variable Methods in Ordinary Differential Equations*. Wiley, New York, 1962.
29. W. Hundsdorfer, J. G. Verwer: *Numerical Solution of Time–Dependent Advection–Diffusion–Reaction Equations*. Springer–Verlag, Berlin, 2003.
30. A. Iserles: *A First Course in the Numerical Analysis of Differential Equations*. Cambridge U. P., Cambridge, 1996.
31. C. Johnson: *Numerical Solutions of Partial Differential Equations by the Finite Element Method.* Cambridge University Press, Cambridge, 1987.
32. H. B. Keller: *Numerical Methods for Two–Point Boundary Value Problems*. Blaisdell, Waltham, Mass., 1968. (Ανατύπωση Dover, New York, 1992.)
33. J. D. Lambert: *Numerical Methods for Ordinary Differential Systems*. Wiley, Chichester, 1991.
34. S. Larsson, V. Thomée: *Partial Differential Equations with Numerical Methods*. Springer–Verlag, Berlin, 2003.
35. C. Moler: *Numerical Computing with MATLAB.* SIAM, Philadelphia, 2004.
36. K. W. Morton, D. F. Mayers: *Numerical Solution of Partial Differential Equations: An Introduction*. 2nd ed., Cambridge Univ. Press, Cambridge, 2005.
37. R. D. Richtmyer, K. W. Morton: *Difference Methods for Initial–Value Problems*. 2nd ed., Wiley, New York, 1967.
38. W. Rudin: *Principles of Mathematical Analysis*. 3rd ed., McGraw–Hill, Singapore, 1976. (Ελληνική μετάφραση με τίτλο *Αρχές Μαθηματικής Αναλύσεως*, Εκδόσεις Leader Books, Αθήνα, 2000.)

39. A. A. Samarskij: *Theorie der Differenzenverfahren*. Akademische Verlagsgesellschaft Geest und Portig, Leipzig, 1984.
40. L. L. Schumaker: *Spline Functions: Basic Theory*. Wiley, New York, 1981.
41. J. C. Strikwerda: *Finite Difference Schemes and Partial Differential Equations*. Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989.
42. V. Thomée: *Galerkin Finite Element Methods for Parabolic Problems*. Lecture Notes in Mathematics, v. 1054, Springer–Verlag, Berlin, 1984.
43. V. Thomée: *Galerkin Finite Element Methods for Parabolic Problems*. 2nd ed., Springer–Verlag, Berlin, 2006.
44. 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.
45. R. Verfürth: *A Review of A Posteriori Error Estimation and Adaptive Mesh–Refinement Techniques*. Teubner–Wiley, Stuttgart, 1996.

Ευρετήριο