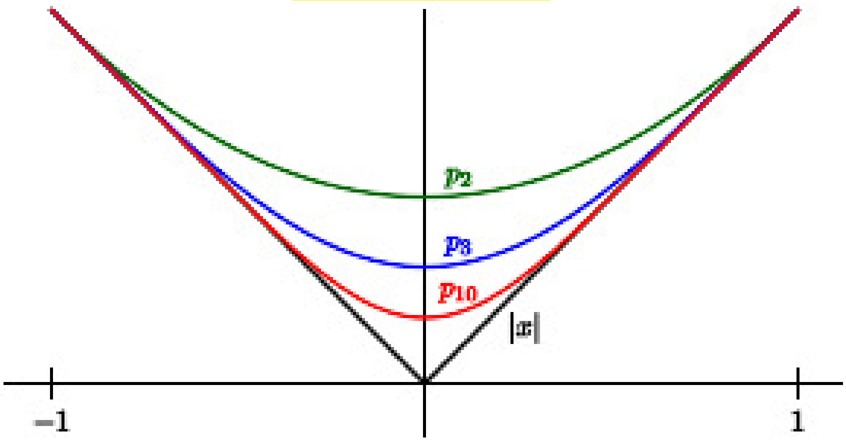
ΘΕΩΡΙΑ ΠΡΟΣΕΓΓΙΣΗΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

ΜΙΧΑΛΗΣ ΚΟΛΟΥΝΤΖΑΚΗΣ

ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ, ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

19 Δεκεμβρίου 2023



Περιεχόμενα

1. Νόρμες σε γραμμικούς χώρους

1.1. **Γραμμικοί χώροι συναρτήσεων.** Η πιο σημαντική ιδιότητα ενός συνόλου συναρτήσεων, με τις οποίες επιχειρούμε να προσεγγίσουμε άλλες, γενικότερες συναρτήσεις, είναι το σύνολο αυτό να αποτελεί γραμμικό χώρο. Λέμε ότι ένα σύνολο συναρτήσεων V αποτελεί **γραμμικό** ή **διανυσματικό χώρο** με συντελεστές από το $\mathbb{R}$ ή το $\mathbb{C}$ αντίστοιχα αν

- Οποτεδήποτε $f_1, f_2 \in V$ τότε ισχύει και $f_1 + f_2 \in V$, και
- Οποτεδήποτε $\lambda \in \mathbb{R}$ ή $\lambda \in \mathbb{C}$ και $f \in V$ τότε ισχύει και $\lambda \cdot f \in V$.

Η έννοια του γραμμικού χώρου συναρτήσεων είναι φυσικά μια εξειδίκευση της γενικής έννοιας του γραμμικού χώρου με συντελεστές από το $\mathbb{R}$ ή το $\mathbb{C}$ (τις περισσότερες φορές θα χρησιμοποιούμε μιγαδικούς συντελεστές σε αυτό το μάθημα αν και το να σκεφτόμαστε τους συντελεστές σα πραγματικούς αριθμούς μας βοηθάει στο να αντιλαμβανόμαστε πιο διαισθητικά τα πράγματα). Οι πρώτοι γραμμικοί χώροι που συναντάει κανείς είναι φυσικά οι πεπερασμένης διάστασης χώροι $\mathbb{R}^n$ ή $\mathbb{C}^n$.

Κάποια σημαντικά παραδείγματα γραμμικών χώρων (με συντελεστές από το $\mathbb{R}$ ή το $\mathbb{C}$) είναι τα:

(1) Τα **πολυώνυμα** με συντελεστές από το $\mathbb{C}$ και βαθμού $\le n$:
$$\mathcal{P}^n = \{p(x) = p_0 + p_1 x + \cdots + p_n x^n : \ p_j \in \mathbb{C}\}.$$
(2) Οι **συνεχείς συναρτήσεις** στο κλειστό διάστημα $[a, b]$:
$$C([a, b]) = \{f : [a, b] \to \mathbb{C} : \ f \text{ συνεχής}\}.$$
Ο ορισμός μπορεί να επεκταθεί σε κάθε σύνολο A πάνω στο οποίο είναι ορισμένες οι συναρτήσεις: $C(A)$ είναι το σύνολο όλων των συναρτήσεων $f : A \to \mathbb{C}$ που είναι συνεχείς σε κάθε σημείο του A.
(3) Οι **συνεχώς παραγωγίσιμες συναρτήσεις** στο ανοιχτό διάστημα (a, b):
$$C^1((a, b)) = \{f : (a, b) \to \mathbb{C} : \ f \text{ παραγωγίσιμη και με συνεχή παράγωγο στο } (a, b)\}.$$

Ομοίως ορίζεται ο χώρος $C^k((a,b))$ (υπάρχει k-τάξης παράγωγος της συνάρτησης και είναι συνεχής – με $C^0((a,b))$ εννοούμε και πάλι το χώρο $C((a,b))$) και ο χώρος

$$C^\infty((a,b)) = \bigcap_{k=1}^{\infty} C^k((a,b))$$

των συναρτήσεων που έχουν παράγωγο κάθε τάξης (*ομαλές συναρτήσεις*). Μπορεί κανείς να ορίσει τους χώρους αυτούς και σε κλειστά ή ημιανοιχτά διαστήματα. Σε αυτή την περίπτωση οι παράγωγοι στα άκρα εννοούνται ως πλευρικές παράγωγοι.

(4) Οι **συνεχείς συναρτήσεις πάνω στο $\mathbb{R}$ με φραγμένο (συμπαγή) φορέα**: μια συνάρτηση $f : \mathbb{R} \to \mathbb{C}$ λέμε ότι έχει φραγμένο φορέα αν υπάρχει ένα φραγμένο διάστημα $[a,b]$ (με $a,b \in \mathbb{R}$ δηλαδή) τέτοιο ώστε η f να μηδενίζεται ταυτοτικά εκτός του $[a,b]$:

$$C_c(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{C} : \ f \text{ μηδενίζεται έξω από κάποιο φραγμένο διάστημα } [a,b]\}.$$

Ο πρώτος χώρος είναι πεπερασμένης διάστασης ενώ οι υπόλοιποι είναι απειροδιάστατοι.

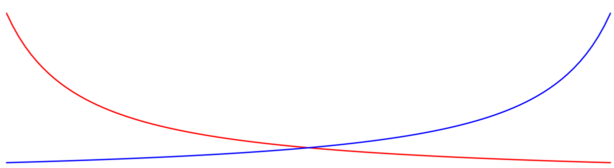
Πρόβλημα 1.1. *Δείξτε ότι το σύνολο $C_c(\mathbb{R})$ που ορίσαμε παραπάνω είναι γραμμικός χώρος.*

Πρόβλημα 1.2. *Δείξτε ότι ο γραμμικός χώρος $C([a,b])$ των συνεχών συναρτήσεων $f : [a,b] \to \mathbb{C}$ είναι απειροδιάστατος.*

Αρκεί να δείξετε ότι για κάθε n υπάρχουν συναρτήσεις $f_1, f_2, \ldots, f_n$ που είναι γραμμικώς ανεξάρτητες πάνω από το $\mathbb{C}$. Μπορείτε επίσης να δείξετε ότι υπάρχει μια άπειρη ακολουθία $g_n \in C([a,b])$ που είναι γραμμικώς ανεξάρτητη πάνω από το $\mathbb{C}$. (Παρ' ότι άπειρο το σύνολο $g_1, g_2, \ldots$ το να πούμε ότι είναι γραμμικώς ανεξάρτητο σημαίνει ότι οποιοσδήποτε πεπερασμένος *γραμμικός συνδυασμός των g_n δε μπορεί να είναι η μηδενική συνάρτηση παρά μόνο όταν όλοι οι συντελεστές είναι 0.)*

Πρόβλημα 1.3. *Ποια η διάσταση του χώρου $\mathcal{P}^n$ που ορίσαμε παραπάνω πάνω από το $\mathbb{C}$ και ποια η διαστασή του πάνω από το $\mathbb{R}$;*

1.2. **Νόρμες σε γραμμικούς χώρους.** Όπως και με τους αριθμούς έτσι και με τις συναρτήσεις ή τα διανύσματα έχουμε την ανάγκη να μετρήσουμε το μέγεθός τους. Σε αντίθεση με ένα πραγματικό ή μιγαδικό αριθμό x όμως, του οποίου το μέγεθος είναι πάντα το $|x|$ (και στην πραγματική και στη μιγαδική περίπτωση αυτό είναι η απόστασή του από το 0) δεν υπάρχει μοναδικός τρόπος να ορίσουμε το μέγεθος ενός διανύσματος ή μιας συνάρτησης όπως φαίνεται και στο παράδειγμα του Σχήματος 1. Ομοίως και



Σχήμα 1: Ποια συνάρτηση από τις δύο είναι πιο «μεγάλη»΄

για διανύσματα στο $\mathbb{R}^n$ ή στο $\mathbb{C}^n$ το μέγεθός τους μπορεί να οριστεί με διάφορους τρόπους. Η κεντρική έννοια εδώ είναι η έννοια της *νόρμας* σε ένα γραμμικό χώρο.

Το να ορίσουμε μια νόρμα σε ένα γραμμικό χώρο V σημαίνει το να μπορούμε για κάθε στοιχείο $x \in V$ να ορίσουμε τη νόρμα του $\|x\|$, που είναι ένας πραγματικός αριθμός, θετικός ή 0, που σκοπό έχει να μετρήσει το μέγεθος του στοιχείου x. Για να αποκαλέσουμε την απεικόνιση

$$x \to \|x\|$$

μια **νόρμα στο χώρο** V πρέπει αυτή η απεικόνιση να έχει τις παρακάτω ιδιότητες.

Αξιώματα Νόρμας

- $\|x\| \geq 0$ για κάθε $x \in V$. Αν $x \neq 0$ τότε $\|x\| > 0$.
- Αν $\lambda \in \mathbb{R}$ ή $\lambda \in \mathbb{C}$ (ανάλογα με το αν ο γραμμικός χώρος μας είναι πραγματικός ή μιγαδικός) και $x \in V$ τότε
$$\|\lambda x\| = |\lambda| \cdot \|x\|.$$
- Αν $x, y \in V$ τότε (**Τριγωνική Ανισότητα**)

$$\|x + y\| \leq \|x\| + \|y\|. \tag{1}$$

Ας δούμε πρώτα μερικές νόρμες ορισμένες στο γραμμικό χώρο $\mathbb{C}^n$ ή $\mathbb{R}^n$. Η πιο απλή είναι η λεγόμενη ℓ^1 νόρμα

$$\|x\|_1 = \sum_{j=1}^{n} |x_j|, \text{ για } x = (x_1, \ldots, x_n) \in \mathbb{C}^n \text{ ή } \mathbb{R}^n.$$

Για τη νόρμα αυτή οι δύο πρώτες ιδιότητες της νόρμας είναι προφανείς και η τρίγωνική ανισότητα είναι απλή συνέπεια της τριγωνικής ανισότητας $|x + w| \leq |z| + |w|$ για μιγαδικούς αριθμούς z, w.

Η νόρμα με την οποία έχουμε εφοδιάσει το χώρο μας επηρεάζει την έννοια της απόστασης στο χώρο μας αφού ως απόσταση δύο σημείων x και y του χώρου μας θεωρούμε την ποσότητα

$$\|x - y\|.$$

Με άλλα λόγια μια νόρμα $\|\cdot\|$ ορίζει με φυσιολογικό τρόπο μια μετρική στο γραμμικό μας χώρο (έννοια απόστασης) τη μετρική

$$d(x, y) = \|x - y\|.$$

Έχοντας λοιπόν μια νόρμα έχουμε αυτόματα και ένα μετρικό χώρο πάνω στο γραμμικό μας χώρο και αυτό σημαίνει ότι ορίζονται μέσω της νόρμας οι έννοιες της σύγκλισης, του ανοιχτού και του κλειστού συνόλου, κλπ. Η νόρμα δηλ. ενός γραμμικού χώρου μας ορίζει αυτόματα μια **τοπολογία** πάνω στο γραμμικό μας χώρο της οποίας οι ιδιότητες είναι άρρηκτα συνδεδεμένες με τις ιδιότητες της νόρμας. Από δω και πέρα όταν αναφερόμαστε σε αναλυτικές ιδιότητες του γραμμικού αυτού χώρου πάντα θα εννοείται σύμφωνα με τη μετρική που επάγεται από τη νόρμα.

Η **σφαίρα** ακτίνας R γύρω από ένα σημείο x_0 του χώρου μας γίνεται έτσι το σύνολο σημείων

$$S_{x_0}(R) = \{y : \|x_0 - y\| = R\}$$

ενώ η **ανοιχτή μπάλα** ακτίνας R γύρω από το x_0 είναι το σύνολο

$$B_{x_0}(R) = \{y : \|x_0 - y\| < R\}$$

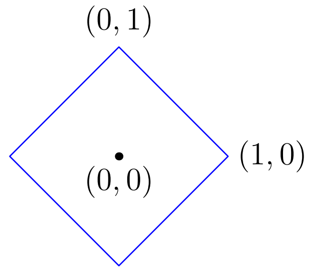
και η **κλειστή μπάλα** $\overline{B_{x_0}}(R)$ είναι το σύνολο

$$\overline{B_{x_0}}(R) = B_{x_0}(R) \cup S_{x_0}(R) = \{y : \|x_0 - y\| \leq R\}.$$

Στο Σχήμα 2 βλέπουμε τη μοναδιαία σφαίρα $S_{(0,0)}(1)$ του $\mathbb{R}^2$ με τη νόρμα ℓ^1.

Πρόβλημα 1.4. *Αποδείξτε ότι η ποσότητα (για κάθε $z = (z_1, z_2, \ldots, z_n) \in \mathbb{C}^n$)*

$$\|z\|_\infty = \max_{j=1,2,\ldots,n} |z_j| \tag{2}$$



Σχήμα 2: Η μοναδιαία σφαίρα του $\mathbb{R}^2$ με την ℓ^1 νόρμα

είναι νόρμα στο γραμμμικό χώρο $\mathbb{C}^n$ (λέγεται ℓ^∞ νόρμα). Με όμοιο τρόπο αποδείξτε ότι στο γραμμικό χώρο $C([a,b])$ η ποσότητα (για $f \in C([a,b])$)

$$\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|$$

είναι επίσης νόρμα.

Πρόβλημα 1.5. *Σχεδιάστε τις μοναδιαίες σφαίρες στο $\mathbb{R}^2$ και $\mathbb{R}^3$ στην ℓ^∞ νόρμα με κέντρο ένα τυχόν σημείο.*

Η πιο σημαντική και συνηθισμένη νόρμα που χρησιμοποιούμε στο $\mathbb{R}^n$ και στο $\mathbb{C}^n$ είναι η Ευκλείδια ή ℓ^2 νόρμα

$$\|x\|_2 = \sqrt{\sum_{j=1}^n |x_j|^2}, \text{ για } x = (x_1, \ldots, x_n) \in \mathbb{C}^n.$$

Και σε αυτή την περίπτωση οι δύο πρώτες ιδιότητες της νόρμας είναι προφανείς. Η απόδειξη της τριγωνικής ανισότητας όμως απαιτεί χρήση της πολύ χρήσιμης ανισότητας **Cauchy-Schwarz** (δείτε παρακάτω για την απόδειξή της):

$$\sum_{j=1}^n |z_j w_j| \le \left(\sum_{j=1}^n |z_j|^2\right)^{1/2} \cdot \left(\sum_{j=1}^n |w_j|^2\right)^{1/2}, \text{ για } z_j, w_j \in \mathbb{C}. \tag{3}$$

1.3. **Εσωτερικά γινόμενα.** Για να δείξουμε την τριγωνική ανισότητα για την ℓ^2 νόρμα στο $\mathbb{C}^n$ χρησιμοποιώντας την ανισότητα Cauchy-Schwarz (3) ορίζουμε πρώτα, και για μελλοντική χρήση, την έννοια του **εσωτερικού γινομένου** στο $\mathbb{C}^n$

$$\langle x, y\rangle = \sum_{j=1}^n x_j \overline{y_j}, \tag{4}$$

όπου $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$ και με $\overline{z} = a - bi$ συμβολίζουμε το μιγαδικό συζυγή του $z = a + bi$. Μπορεί κανείς να ορίσει το εσωτερικό γινόμενο και σε άλλους χώρους. Για παράδειγμα στο χώρο $C([a,b])$ μπορούμε να ορίσουμε το εσωτερικό γινόμενο δύο στοιχείων του με τον τύπο

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx.$$

Το εσωτερικό γινόμενο είναι μια **διγραμμική μορφή**, είναι δηλ. χωριστά γραμμικό ως προς το πρώτο και το δεύτερο μέρος του (με μια μικρή διαφορά ανάμεσα στο πρώτο και στο δεύτερο όρισμα του γινομένου στην περίπτωση του $\mathbb{C}^n$ που οφείλεται στην ύπαρξη του μιγαδικού συζυγούς στο δεύτερο όρισμα):

$$\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle,$$

και

$$\langle z, \lambda x + \mu y\rangle = \overline{\lambda}\langle z, x\rangle + \overline{\mu}\langle y, z\rangle,$$

για $x, y, z \in \mathbb{C}^n$ και $\lambda, \mu \in \mathbb{C}$. Επίσης ισχύει

$$\langle x, y\rangle = \overline{\langle y, x\rangle}.$$

Η σχέση εσωτερικού γινομένου και ℓ^2 νόρμας είναι πολύ απλά η

$$\|x\|_2^2 = \langle x, x\rangle. \tag{5}$$

Πρόβλημα 1.6. *Αποδείξτε τον* **κανόνα του παραλληλογράμμου***:*

$$2\|x\|_2^2 + 2\|y\|_2^2 = \|x-y\|_2^2 + \|x+y\|_2^2, \quad (\text{για } x, y \in \mathbb{C}^n) \tag{6}$$

και ερμηνεύστε τον γεωμετρικά.

Χρησιμοποιήστε την (5).

Πρόβλημα 1.7. *Από τον τύπο $\|x\|_2^2 = \langle x, x\rangle$ φαίνεται ότι αν ξέρουμε να υπολογίζουμε εσωτερικά γινόμενα τότε μπορούμε να υπολογίζουμε και τις αντίστοιχες νόρμες. Ισχύει και το αντίστροφο (***polarization identities***). Δείξτε ότι αν $x, y \in \mathbb{R}^n$ τότε έχουμε*

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|_2^2 - \|x-y\|_2^2). \tag{7}$$

Δείξτε επίσης ότι αν $x, y \in \mathbb{C}^n$ τότε ισχύει

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|_2^2 - \|x-y\|_2^2 + i\|x+iy\|_2^2 - i\|x-iy\|_2^2). \tag{8}$$

Εκφράστε τα πάντα με εσωτερικά γινόμενα μέσω της (5).

Αποδεικνύουμε τώρα την ανισότητα Cauchy-Schwarz (3). Την αποδεικνύουμε πρώτα για το χώρο $\mathbb{R}^n$ όπου η απόδειξη είναι πιο απλή.

Ας είναι λοιπόν $x, y \in \mathbb{R}^n$ και $t \in \mathbb{R}$. Ισχύει τότε

$$\begin{aligned} 0 &\le \|x+ty\|_2^2 \\ &= \langle x+ty, x+ty\rangle \\ &= \langle x, x\rangle + t\langle y, x\rangle + t\langle x, y\rangle + t^2\langle y, y\rangle \text{ (επιμερίζοντας)} \\ &= \|x\|_2^2 + 2\langle x, y\rangle t + \|y\|_2^2 t^2. \end{aligned}$$

Επειδή αυτό το τετραγωνικό πολυώνυμο του t είναι μη αρνητικό για κάθε $t \in \mathbb{R}$ έπεται ότι η διακρίνουσά του είναι ≤ 0

$$4\langle x, y\rangle^2 - 4\|x\|_2^2\|y\|_2^2 \le 0,$$

και αναδιατάσσοντας τους όρους προκύπτει η ανισότητα

$$|\langle x, y\rangle| \le \|x\|_2\|y\|_2,$$

που είναι η ανισότητα Cauchy-Schwarz. Μπορούμε φυσικά και να παραλείψουμε την απόλυτη τιμή από το αριστερό μέλος (για πραγματικά διανύσματα μόνο) και να πάρουμε και πάλι μια σωστή ανισότητα.

Για να αποδείξουμε την ανισότητα στο $\mathbb{C}^n$ παρατηρούμε ότι αν αντικαταστήσουμε τις συντεταγμένες δύο μιγαδικών διανυσμάτων x, y με τις απόλυτες τιμές τους (το μέτρο τους, σωστότερα, μια και πρόκειται για μιγαδικούς αριθμούς) τότε το δεξί μέλος της ανισότητας Cauchy-Schwarz δεν αλλάζει ενώ το αριστερό μέλος μεγαλώνει, οπότε η μιγαδική Cauchy-Schwarz είναι συνέπεια της πραγματικής.

Η απόδειξη της τριγωνικής ανισότητας για την ℓ^2 νόρμα στο $\mathbb{C}^n$ είναι τώρα εύκολη:

$$\begin{aligned}
\|x+y\|_2^2 &= \langle x+y, x+y\rangle \\
&= \|x\|_2^2 + \|y\|_2^2 + 2\mathrm{Re}\,[\langle x, y\rangle] \quad \text{(επιμερίζοντας)} \\
&\le \|x\|_2^2 + \|y\|_2^2 + 2\|x\|\|y\| \ \text{(από Cauchy-Schwarz)} \\
&= (\|x\|_2 + \|y\|_2)^2.
\end{aligned}$$

Πρόβλημα 1.8. *Είδαμε την απόδειξη της ανισότητας Cauchy-Schwarz*

$$|\langle x, y\rangle|^2 \le \langle x, x\rangle\langle y, y\rangle, \ \ \text{για } x, y \in \mathbb{C}^n,$$

αναγοντάς την πρώτα στην περίπτωση που θα διανύσματα x, y έχουν συντεταγμένες που είναι μη αρνητικοί πραγματικοί αριθμοί. Με αυτό τον τρόπο όμως δε μπορούμε να βγαλουμε συμπέρασμα για το πότε ισχύει η παραπάνω ανισότητα ως ισότητα. Δώστε λοιπόν μια παρόμοια απόδειξη αλλά χωρίς να εγκαταλείψετε τους μιγαδικούς αριθμούς και συμπεράνετε ότι η ανισότητα Cauchy-Schwarz ισχύει ως ισότητα ακριβως όταν τα δυο διανύσματα είναι γραμμικώς εξαρτημένα (με μιγαδικούς συντελεστές).

Δουλέψτε όπως και στην περίπτωση των μη αρνητικών συντεταγμένων, παρατηρώντας δηλ. ότι για κάθε πραγματικό t ισχύει

$$0 \le \langle x + ty, x + ty\rangle$$

και παρατηρώντας κάνοντας πράξεις ότι η παράσταση στο δεξί μέλος είναι ένα πολυώνυμο ως προς t βαθμού το πολύ 2 με πραγματικούς συντελεστές. Κοιτώντας τη διακρίνουσα του πολυωνύμου αυτού συμπεράνετε ότι

$$(\mathrm{Re}\,[\langle x, y\rangle])^2 \le \langle x, x\rangle\langle y, y\rangle.$$

Για να βγάλετε το τελικό σας συμπέρασμα εφαρμόστε την ανισότητα αυτή στο διάνυσμα $e^{i\theta}x$ για κατάλληλο $\theta \in \mathbb{R}$.

1.4. **Ισοδυναμία νορμών.** Δύο νόρμες $\|\cdot\|_a$ και $\|\cdot\|_b$ σε ένα γραμμικό χώρο V ονομάζονται ισοδύναμες αν υπάρχουν δύο θετικές σταθερές C_1, C_2 τέτοιες ώστε για κάθε $x \in V$ να έχουμε

$$C_1\|x\|_b \le \|x\|_a \le C_2\|x\|_b.$$

Με άλλα λόγια όταν οι τιμές τους είναι πάντα «συγκρίσιμες», δε διαφέρει δηλαδή η μια νόρμα του x από την άλλη παρά το πολύ κατά ένα πολλαπλάσιο, ανεξάρτητα από το x.

Δεν είναι δύσκολο να βρει κανείς παραδείγματα νορμών που είναι ισοδύναμες. Για παράδειγμα, ας δουλέψουμε στο χώρο $\mathbb{C}^n$ και ας πάρουμε τη μια νόρμα να είναι η $\|\cdot\|_1$. Ας πάρουμε την άλλη νόρμα να είναι η

$$\|z\|_w = \sum_{j=1}^{\infty} |z_j| w_j$$

όπου $z = (z_1, z_2, \ldots, z_n) \in \mathbb{C}^n$ και $w_j > 0$ είναι μια σταθερή ακολουθία **βαρών** (η απόδειξη ότι η $\|\cdot\|_w$ είναι νόρμα είναι ταυτόσημη με την περίπτωση της $\|\cdot\|_1$). Αν ορίσουμε

$$w = \min_{j=1,2,\ldots,n} w_j > 0, \quad W = \max_{j=1,2,\ldots,n} w_j$$

τότε εύκολα βλέπουμε ότι για κάθε $z \in \mathbb{C}^n$ έχουμε

$$w\|z\|_1 \le \|z\|_w \le W\|z\|_1$$

άρα οι δύο αυτές νόρμες είναι ισοδύναμες όποια και να είναι η επιλογή των βαρών w_j.

Πρόβλημα 1.9. *Δίξτε ότι η ισοδυναμία νορμών είναι μια σχέση ισοδυναμίας.*

Θα δείξουμε παρακάτω ότι σε κάθε χώρο πεπερασμένης διάστασης, όπως ο $\mathbb{C}^n$, όλες οι νόρμες είναι ισοδύναμες.

Πρόβλημα 1.10. *Αν $z \in \mathbb{C}^n$ δείξτε τις παρακάτω σχέσεις ανάμεσα σε διαφορετικές νόρμες:*

$$\|z\|_1 \le \sqrt{n}\|z\|_2, \tag{9}$$

$$\|z\|_2 \le \|z\|_1, \tag{10}$$

$$\|z\|_1 \le n\|z\|_\infty, \tag{11}$$

$$\|z\|_\infty \le \|z\|_1, \tag{12}$$

$$\|z\|_\infty \le \|z\|_2, \tag{13}$$

$$\|z\|_2 \le \sqrt{n}\|z\|_\infty. \tag{14}$$

Άρα και οι τρεις νόρμες που αναφέρονται είναι ισοδύναμες στο χώρο $\mathbb{C}^n$ και οι καλύτερες σταθερές C_1 και C_2 είναι αυτές που δίνονται στις παραπάνω ανισότητες δηλ. πρέπει ακόμη να δείξετε ότι οι παραπάνω ανισότητες δε βελτιώνονται, δε μπορούν δηλ. να ισχύσουν με καλύτερες σταθερές από αυτές που δίδονται παραπάνω.

Πρόβλημα 1.11. *Αποδείξτε ότι στο χώρο $C([a, b])$ κανένα ζεύγος από τις νόρμες*

$$\|f\|_1 = \int_a^b |f(x)|\,dx,$$

$$\|f\|_2 = \left(\int_a^b |f(x)|^2\,dx\right)^{1/2},$$

$$\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|,$$

δεν αποτελείται από ισοδύναμες νόρμες. Για να δείξει κανείς ότι ένα ζεύγος νορμών $\|\cdot\|_a$ και $\|\cdot\|_b$ δεν είναι ισοδύναμες αρκεί να δείξει ότι το πηλίκο

$$\frac{\|v\|_a}{\|v\|_b}$$

μπορεί, με κατάλληλη επιλογή του στοιχείου v του γραμμικού χώρου, να γίνει είτε οσοδήποτε μεγάλο είτε οσοδήποτε μικρό (κοντά στο 0).

1.5. **Χώροι και υπόχωροι πεπερασμένης διάστασης.** Όταν η διάσταση ενός χώρου με νόρμα είναι άπειρη (π.χ. ο πολύ συνηθισμένος χώρος $C([a, b])$) τότε πολλά ασυνήθιστα και περίεργα φαινόμενα μπορούν να συμβούν και η ανάλυση δυσκολεύει (δηλ. γίνεται πιο ενδιαφέρουσα). Όταν η διάσταση του γραμμικού χώρου όμως είναι πεπερασμένη τότε πολλά πράγματα γίνονται απλούστερα. Το περιεχόμενο του επόμενου θεωρήματος είναι ότι η τοπολογία (δηλ. η έννοια της σύγκλισης, τα ανοιχτά και τα κλειστά σύνολα, οι συνεχείς συναρτήσεις κλπ) ενός γραμμικού χώρου πεπερασμένης διάστασης δεν αλλάζει αν αλλάξουμε τη νόρμα γιατί όλες οι νόρμες είναι ισοδύναμες.

Θεώρημα 1.1. *Αν V είναι χώρος πεπερασμένης διάστασης (πάνω από τους μιγαδικούς ή τους πραγματικούς αριθμούς) τότε οποιεσδήποτε δύο νόρμες στον V είναι μεταξύ τους ισοδύναμες.*

Απόδειξη: Ας είναι $e_1, \ldots, e_n$ μια βάση του V. Κάθε στοιχείο $z \in V$ γράφεται με μοναδικό τρόπο ως $z = z_1 e_1 + \cdots + z_n e_n$, όπου $z_j \in \mathbb{C}$ ($j = 1, 2, \ldots, n$). Ορίζουμε την ποσότητα $\|z\|_1 = \sum_{j=1}^n |z_j|$ που εύκολα αποδεικνύεται ότι είναι νόρμα στον V (ουσιαστικά ο V είναι ισόμορφος με τον $\mathbb{C}^n$ και η $\|\cdot\|_1$ είναι αντίστοιχη με την γνωστή μας ℓ_1 νόρμα στον $\mathbb{C}^n$). Θα δείξουμε ότι κάθε άλλη νόρμα $\|\cdot\|$ είναι ισοδύναμη προς την $\|\cdot\|_1$ και άρα οποιεσήποτε δύο νόρμες είναι μεταξύ τους ισοδύναμες αφού είναι ισοδύναμες προς την $\|\cdot\|_1$.

Έχουμε χρησιμοποιώντας την τριγωνική ανισότητα

$$\|z\| = \left\|\sum z_j e_j\right\| \leq \sum \|z_j e_j\| = \sum |z_j| \|e_j\| \leq C_2 \|z\|_1,$$

όπου $C_2 = \max\{\|e_1\|, \ldots, \|e_n\|\}$. Μένει να δείξουμε την ανισότητα $C_1\|z\|_1 \leq \|z\|$ για κάποιο $C_1 > 0$.

Ορίζουμε την μοναδιαία σφαίρα ως προς τη νόρμα $\|\cdot\|_1$

$$S = \{x \in V : \|x\|_1 = 1\}.$$

Αυτό είναι ένα μη κενό (γιατί;), κλειστό και φραγμένο σύνολο και αφού ο V είναι πεπερασμένης διάστασης προκύπτει ότι το S είναι συμπαγές σύνολο. Όμως η απεικόνιση $x \to \|x\|$ είναι μια συνεχής απεικόνιση (δείτε Πρόβλημα 1.12 παρακάτω) και άρα αφού κάθε συνεχής συνάρτηση σε συμπαγές σύνολο «πιάνει» το ελάχιστό της (αποδείξτε το!) έπεται ότι υπάρχει $x_0 \in S$ τέτοιο ώστε

$$\|x_0\| \leq \|x\| \quad \forall x \in S.$$

Για κάθε $x \in V$ έχουμε (εφόσον $\frac{x}{\|x\|_1} \in S$):

$$\|x\| = \left\| \|x\|_1 \frac{x}{\|x\|_1} \right\| = \|x\|_1 \left\| \frac{x}{\|x\|_1} \right\| \geq \|x_0\| \|x\|_1$$

που είναι η ανισότητα που ψάχνουμε με $C_1 = \|x_0\|$.
□

Πρόβλημα 1.12. *Δείξτε ότι σε κάθε χώρο με νόρμα V η απεικόνιση $x \to \|x\|$ είναι μια συνεχής απεικόνιση από το V στο $\mathbb{R}$.*

Από την τριγωνική ανισότητα $\|x + y\| \leq \|x\| + \|y\|$ δείξτε πρώτα την ανισότητα

$$\|x - y\| \geq |\|x\| - \|y\||, \quad \forall x, y \in V.$$

Πόρισμα 1.1. *Κάθε χώρος πεπερασμένης διάστασης με νόρμα είναι πλήρης.*

Απόδειξη: Αν ο χώρος μας έχει διάσταση n τότε είναι ισόμορφος ως γραμμικός χώρος με το χώρο $\mathbb{C}^n$. Ας είναι $T : V \to \mathbb{C}^n$ η ισομορφική απεικόνιση. Τότε η T ορίζει μια νόρμα στο $\mathbb{C}^n$

$$\|x\| = \left\|T^{-1} x\right\|_V,$$

όπου $\|\cdot\|_V$ είναι η νόρμα του V (η απόδειξη ότι η $\|\cdot\|$ είναι νόρμα είναι σχεδόν προφανής). Με αυτή τη νόρμα στο $\mathbb{C}^n$ οι ακολουθίες Cauchy του V αντιστοιχούν ακριβώς σε ακολουθίες Cauchy στον $\mathbb{C}^n$. Αλλά ο $\mathbb{C}^n$ έχει πεπερασμένη διάσταση και άρα με οποιαδήποτε νόρμα είναι πλήρης χώρος, άρα κάθε ακολουθία Cauchy στον $\mathbb{C}^n$ έχει όριο και παίρνοντας το T^{-1} του ορίου βρίσκουμε το όριο της αντίστοιχης ακολουθίας Cauchy στον V.
□

Πόρισμα 1.2. *Σε γραμμικό χώρο πεπερασμένης διάστασης κάθε κλειστή μπάλα είναι συμπαγής ως προς οποιαδήποτε νόρμα (όχι μόνο ως προς τη νόρμα που χρησιμοποιείται στον ορισμό της μπάλας).*

Απόδειξη: Η έννοια της συμπάγειας δείξαμε ότι (όπως και κάθε άλλη τοπολογική έννοια) δεν εξαρτάται από τη νόρμα σε χώρους πεπερασμένης διάστασης, άρα μπορούμε απλά να χρησιμοποιήσουμε το θεώρημα Heine-Borel που μας λέει ότι σε ένα χώρο με νόρμα πεπερασμένης διάστασης τα συμπαγή σύνολα είναι ακριβώς τα κλειστά και φραγμένα σύνολα.
□

Η Θεωρία Προσέγγισης μελετάει κυρίως το πρόβλημα της προσέγγισης από υπόχωρο. Στη γενική του μορφή το πρόβλημα αυτό έχει ως εξής. Έχουμε ένα γραμμικό χώρο X με νόρμα $\|\cdot\|$ και ένα υπόχωρό του Y. Έχουμε επίσης ένα $x \in X$ το οποίο θέλουμε να προσεγγίσουμε από ένα στοιχείο $y \in Y$ ώστε η απόσταση $\|x - y\|$ να είναι όσο το δυνατόν μικρή. Το κύριο θεώρημα που δείξαμε σήμερα είναι ότι όταν ο χώρος Y είναι πεπερασμένης διάστασης τότε πάντα υπάρχει ένα $y^* \in Y$ το οποίο αποτελεί *βέλτιστη προσέγγιση* του x:

$$\|x - y^*\| \le \|x - y\|, \quad \forall y \in Y.$$

Η βέλτιστη αυτή προσέγγιση δεν είναι εν γένει μοναδική (το αν είναι ή όχι εξαρτάται κυρίως από το για ποια νόρμα μιλάμε). Μια σημαντική εφαρμογή αυτού είναι ότι αν $f \in C([a, b])$ τότε αν ψάχνουμε να ελαχιστοποιήσουμε τη διαφορά

$$\|f(x) - p(x)\|_\infty, \quad p(x) \in \mathcal{P}^n,$$

υπάρχει τουλάχιστον ένα πολυώνυμο που υλοποιεί το infimum αυτής της διαφοράς. Υπάρχει δηλ. βέλτιστη πολυωνυμική προσέγγιση της f. Ο λόγος είναι ότι ο υπόχωρος $\mathcal{P}^n$ του $C([a, b])$ είναι πεπερασμένης διάστασης

$$\dim \mathcal{P}^n = n + 1.$$

Αυτό ισχύει και αν στη θέση της νόρμας $\|\cdot\|_\infty$ παραπάνω χρησιμοποιήσουμε μια οποιαδήποτε άλλη νόρμα, π.χ. μια από τις ολοκληρωτικές νόρμες $\|\cdot\|_1$ ή $\|\cdot\|_2$.

Θεώρημα 1.2. *(Υπαρξη βέλτιστης προσέγγισης) Αν X είναι γραμμικός χώρος με νόρμα και $Y \subseteq X$ είναι γραμμικός υπόχωρος πεπερασμένης διάστασης τότε για κάθε $x \in X$ υπάρχει $x^* \in Y$ τέτοιο ώστε*

$$\|x - x^*\| \le \|x - y\| \quad \forall y \in Y.$$

Απόδειξη: Ας ορίσουμε κατ' αρχήν το σύνολο

$$K = \{y \in Y : \|x - y\| \le \|x\|\}.$$

Παρατηρείστε ότι $0 \in K \subseteq Y$ και, αφού το 0 έχει απόσταση $\|x\|$ από το x, δε χρειάζεται να ψάξουμε για τη βέλτιστη προσέγγιση x^* στο σύνολο $Y \setminus K$, και ότι αρκεί συνεπώς να ψάξουμε στο K. Με άλλα λόγια αρκεί να δείξουμε ότι η συνάρτηση $y \to \|x - y\|$ πιάνει το ελάχιστό της στο σύνολο K. Όμως η συνάρτηση αυτή είναι συνεχής (ακριβώς όπως στο πρόβλημα 1.12) και το σύνολο K είναι φραγμένο (αφού $K \subseteq B_{\|x\|}(x)$) και κλειστό αφού αν $K \ni y_n \to y \in Y$ τότε $\|x - y_n\| \to \|x - y\|$ (από τη συνέχεια της απεικόνισης $z \to \|z\|$) και άρα $y \in K$. Αφού ο χώρος Y είναι πεπερασμένης διάστασης έπεται ότι το σύνολο K είναι συμπαγές (Heine-Borel) και άρα η συνάρτηση $y \to \|x - y\|$ πιάνει το ελάχιστό της στο σύνολο K, το οποίο είναι ακριβώς η βέλτιστη προσέγγιση x^*.
□

Η βέλτιστη προσέγγιση του x από το χώρο Y μπορεί να μην είναι μοναδική.

Πρόβλημα 1.13. *Στο χώρο $X = \mathbb{C}^3$ με την ℓ_∞ νόρμα βρείτε όλες τις βέλτιστες προσεγγίσεις του στοιχείου $x = (0, 0, 1)$ από τον υπόχωρο $Y = \{(z, w, 0) : z, w \in \mathbb{C}\}$.*

Αν όμως είναι τότε η απεικόνιση $x \to x^*$ είναι συνεχής. Και αυτό είναι μια πολύ σημαντική ιδιότητα και πολύ επιθυμητή σε προβλήματα προσέγγισης. Όταν προσεγγίζουμε π.χ. μια συνεχή συνάρτηση στο διάστημα $[a, b]$ από ένα πολυώνυμο βαθμού έως n είναι πολύ επιθυμητό να έχουμε αυτό που ονομάζουμε *ευστάθεια*: αν αλλάξουν τα δεδομένα μας λίγο (τα δεδομένα μας είναι η συνάρτηση που προσεγγίζουμε και μπορεί αυτή να αλλάξει λίγο λόγω σφαλμάτων στη μέτρηση ή στον υπολογισμό) τότε θέλουμε και η προσέγγισή μας επίσης να αλλάζει με ελεγχόμενο τρόπο.

Θεώρημα 1.3. *(Συνέχεια βέλτιστης προσέγγισης) Αν X είναι γραμμικός χώρος με νόρμα και $Y \subseteq X$ είναι γραμμικός υπόχωρος πεπερασμένης διάστασης τότε για κάθε $x \in X$ η βέλτιστη προσέγγισή του από τον Y, $x^* \in Y$, είναι μοναδική, τότε η συνάρτηση $x \to x^*$ είναι μια συνεχής συνάρτηση $X \to Y$.*

Απόδειξη: Έστω $X \ni x_n \to x_0$. Τότε, από την τριγωνική ανισότητα έχουμε

$$\|x_n^*\| = \|x_n^* - x_n + x_n\| \le \|x_n^* - x_n\| + \|x_n\| \le \|x_n\| + \|x_n\| = 2\|x_n\|$$

άρα η ακολουθία $\|x_n^*\|$ είναι φραγμένη αφού $\|x_n\| \to \|x\|$ (αυτό το τελευταίο από την ανισότητα $|\|x_n\| - \|x\|| \le \|x_n - x\|$). Αφού η ακολουθία $x_n^* \in Y$ είναι φραγμένη και ο Y έχει πεπερασμένη διάσταση (οι μπάλες στον Y είναι άρα συμπαγείς) έπεται ότι η x_n^* έχει ακολουθία που συγκλίνει σε κάποιο $y \in Y$.

Θα δείξουμε ότι $y = x_0^*$. Αυτό συνεπάγεται ότι οποιαδήποτε συγκλίνουσα υπακολουθία της x_n^* έχει το ίδιο όριο και άρα $x_n^* \to x_0^*$. Για ευκολία συμβολισμού (για να μην έχουμε «διόρωφους» δείκτες στις ακολουθίες μας) ονομάζουμε την υπακολουθία της x_n^* που συγκλίνει στο y ξανά x_n^*, πετάμε δηλ. τους υπόλοιπους όρους της ακολουθίας. Έχουμε

$$\begin{aligned} \|y - x_0\| &= \lim_n \|x_n^* - x_n\| \quad \text{(από τη συνέχεια της νόρμας)} \\ &\le \lim_n \|x_0^* - x_n\| \quad \text{(γιατί το } x_n^* \text{ είναι η βέλτ. προσέγγιση του } x_n\text{)} \\ &= \|x_0^* - x_0\|. \quad \text{(από τη συνέχεια της νόρμας)} \end{aligned}$$

Αφού δείξαμε $\|y - x_0\| \le \|x_0^* - x_0\|$ και η βέλτιστη προσέγγιση είναι μοναδική έπεται ότι $y = x_0^*$.
□

2. Το Θεώρημα του Weierstrass

2.1. **Γενικά.** Το πολύ σημαντικό θεώρημα του Weierstrass είναι το παρακάτω.

Θεώρημα 2.1 (Weierstrass)**.** *Αν $a < b$ είναι πραγματικοί αριθμοί και η $f : [a, b] \to \mathbb{C}$ είναι συνεχής τότε υπάρχει ακολουθία πολυωνύμων που συγκλίνει ομοιόμορφα στην f στο διάστημα $[a, b]$.*

Ένας διαφορετικός τρόπος να πουμε το ίδιο πράγμα είναι να πούμε ότι υπάρχει ακολουθία πολυωνύμων p_n τ.ώ. $\|p_n - f\|_\infty \to 0$, όπου η $\|\cdot\|_\infty$ είναι αναφορικά με το διάστημα $[a, b]$, ισχύει δηλαδή

$$\|g\|_\infty = \sup_{x \in [a,b]} |g(x)|.$$

Τη δεύτερη αυτή διατύπωση μπορούμε και να την πάρουμε ως τον ορισμό της ομοιόμορφης σύγκλισης.

Ορισμός 2.1. Αν $f, f_n : K \to \mathbb{C}$ είναι συναρτήσεις ορισμένες πάνω σε ένα σύνολο K λέμε ότι η ακολουθία f_n συγκλίνει στην f ομοιόμορφα στο K αν

$$\|f - f_n\|_\infty = \sup_{x \in K} |f(x) - f_n(x)| \to 0, \quad (n \to \infty).$$

Παρατήρηση 2.1. Υπάρχει λόγος που στο Θεώρημα του Weierstrass περιοριζόμαστε σε συνεχείς συναρτήσεις, δηλ. το Θέωρημα 2.2 παρακάτω.

Θεώρημα 2.2. *Αν $K \subseteq \mathbb{C}$ και $f, f_n : K \to \mathbb{C}$, οι f_n είναι συνεχείς στο K και η ακολουθία f_n συγκλίνει ομοιόμορφα στην f πάνω στο K τότε και η f είναι συνεχής πάνω στο K.*

Απόδειξη: Έστω $\epsilon > 0$ και $x \in K$. Θα δείξουμε ότι υπάρχει $\delta > 0$ τέτοιο ώστε αν $y \in K$ να ισχύει η συνεπαγωγή

$$|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon.$$

Για κάθε n γράφουμε

$$\begin{aligned} |f(x) - f(y)| &= |f(x) - f_n(x) + f_n(x) - f_n(y) + f_n(y) - f(y)| \\ &\le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &= A_1 + A_2 + A_3. \end{aligned}$$

Από την ομοιόμορφη σύγκλιση $f_n \to f$ υπάρχει ένα n τέτοιο ώστε $\|f - f_n\|_\infty < \epsilon/3$. Για αυτό το n οι όροι A_1, A_3 παραπάνω είναι και οι δύο $< \epsilon/3$. Για αυτό το n επίσης η συνάρτηση f_n είναι συνεχής στο x, άρα υπάρχει ένα $\delta > 0$ ώστε να ισχύει η συνεπαγωγή

$$|x - y| < \delta \Longrightarrow |f_n(x) - f_n(y)| < \epsilon/3.$$

Άρα, με την προϋπόθεση ότι $|x - y| < \delta$ και οι τρεις όροι A_1, A_2, A_3 παραπάνω είναι μικρότεροι του $\epsilon/3$ και συνεπώς $|f(x) - f(y)| < \epsilon$ πως είχαμε να αποδείξουμε.
□

Παρατήρηση 2.2. Έχει σημασία ότι το πεδίο ορισμού της f είναι ένα κλειστό και φραγμένο διάστημα. Εύκολα μπορούμε να βρούμε συνεχείς συναρτήσεις πάνω σε φραγμένα ανοιχτά διαστήματα ή άφρακτα διαστήματα που δε μπορούν να προσεγγισθούν ομοιόμορφα στο πεδίο ορισμού τους. Για παράδειγμα η συνάρτηση $f_1(x) = e^x$, με πεδίο ορισμού το $\mathbb{R}$, είναι συνεχής παντού αλλά για κάθε πολυώνυμο $p(x)$ ισχύει φυσικά

$$|f_1(x) - p(x)| \to +\infty, \quad (x \to +\infty),$$

αφού η εκθετική συνάρτηση αυξάνει πιο γρήγορα από οποιαδήποτε πολυωνυμική συνάρτηση. Ένα άλλο παράδειγμα είναι η συνάρτηση $f_2(x) = 1/x$, με πεδίο ορισμού το $(0, 1)$. Η συνάρτηση αυτή είναι συνεχής στο πεδίο ορισμού της αλλά δεν μπορεί να προσεγγισθεί ομοιόμορφα από πολυώνυμο, μια και οποιοδήποτε πολυώνυμο $q(x)$ έχει

$$\lim_{x \to 0+} q(x) = q(0),$$

ενώ για την $f_2(x)$ το από δεξιά όριο στο 0 είναι $+\infty$.

2.2. **Η απόδειξη του Landau.**

Η πρώτη απόδειξη του θεωρήματος του Weierstrass που θα δούμε οφείλεται στον Landau. Έστω $f : [a, b] \to \mathbb{C}$ συνεχής.

2.2.1. *Περιορισμός στο διάστημα* $[0, 1]$. Κάνουμε κατ' αρχήν την παρατήρηση ότι μπορούμε να υποθέσουμε ότι το $[a, b]$ είναι όποιο συγκεκριμένο διάστημα μας βολεύει, για παράδειγμα το $[0, 1]$. Ο λόγος γι' αυτό είναι ότι μπορούμε να αντιστοιχίσουμε τα σημεία $t \in [0, 1]$ με τα στοιχεία $x \in [a, b]$ με μια *αφφινική απεικόνιση* (μια απεικόνιση δηλ. που είναι της μορφής $x \to Ax + B$), την

$$x = a + t(b - a) \ \text{ της οποίας η αντίστροφη είναι η } \ t = \frac{x - a}{b - a}.$$

Αν τώρα $f : [a, b] \to \mathbb{C}$ είναι συνεχής τότε και η συνάρτηση

$$g(t) = f(a + t(b - a))$$

είναι μια συνεχής συνάρτηση στο $[0,1]$. Αν υπάρχει πολυώνυμο $p(t)$ τ.ώ. $|p(t)-g(t)| \le \epsilon$ για $t \in [0,1]$ τότε η συνάρτηση

$$q(x) = p((x-a)/(b-a))$$

είναι κι αυτή πολυώνυμο και ισχύει φυσικά

$$|q(x) - f(x)| = |p(t) - g(t)| \le \epsilon.$$

Από δω και πέρα λοιπόν υποθέτουμε ότι $f : [0,1] \to \mathbb{C}$ είναι συνεχής.

2.2.2. *Η f μηδενίζεται στα άκρα του διαστήματος.* Μπορούμε επίσης χωρίς βλάβη της γενικότητας να υποθέσουμε ότι $f(0) = f(1) = 0$. Αυτό συμβαίνει γιατί μπορούμε να αφαιρέσουμε από την f ένα κατάλληλο πρωτοβάθμιο πολυώνυμο ώστε να πετύχουμε αυτή τη συνθήκη. Ακριβέστερα, αν η $f : [0,1] \to \mathbb{C}$ είναι οποιαδήποτε συνεχής συνάρτηση τότε θέτουμε

$$g(x) = f(x) - \ell(x), \ \text{ όπου } \ \ell(x) = f(0) + x(f(1) - f(0)),$$

και παρατηρούμε ότι (α) η g είναι συνεχής συνάρτηση που μηδενίζεται στα άκρα του διαστήματος και (β) αν μπορούμε να προσεγγίσουμε την g με ένα πολυώνυμο p

$$\|g(x) - p(x)\|_\infty \le \epsilon$$

τότε το πολυώνυμο $q(x) = p(x) + \ell(x)$ επίσης προσεγγίζει την $f(x)$

$$\|f(x) - q(x)\|_\infty = \|g(x) + \ell(x) - (p(x) + \ell(x))\|_\infty \le \epsilon.$$

Από δω και πέρα λοιπόν υποθέτουμε ότι η συνάρτησή μας, $f(x)$, μηδενίζεται στα άκρα του διαστήματος. Για ευκολία μας την επεκτείνουμε (με μηδενικές τιμές) στο υπόλοιπο της πραγματικής ευθείας, και η νέα μας συνάρτηση είναι συνεχής σε όλο το $\mathbb{R}$ μια και τα μόνα σημεία στα οποία γεννάται αμφιβολία γι' αυτό είναι τα άκρα του $[0,1]$, όμως εκεί τα πλευρικά όρια της f είναι και τα δύο 0.

2.2.3. *Τα πολυώνυμα προσέγγισης.* Ορίζουμε τώρα (θυμόμαστε ότι η f μηδενίζεται έξω από το διάστημα $[0,1]$)

$$L_n(x) = \int_{-\infty}^{\infty} f(x+t) K_n(t)\, dx = \int_{-x}^{1-x} f(x+t) K_n(t)\, dx,$$

όπου

$$K_n(t) = \begin{cases} c_n(1-t^2)^n, & |t| \le 1 \\ 0, & |t| > 1, \end{cases}$$

και η σταθερά c_n επιλέγεται με τρόπο τέτοιο ώστε να ισχύει

$$\int_{-\infty}^{\infty} K_n(t)\, dt = 1.$$

Για να δούμε ότι οι συναρτήσεις $L_n(x)$ είναι όντως πολυώνυμα του x αν $x \in [0,1]$[1] κάνουμε την αλλαγή μεταβλητής $u = x + t$ και παίρνουμε έτσι την έκφραση

$$L_n(x) = c_n \int_0^1 f(u)(1-(u-x)^2)^n\, du,$$

[1]Εύκολα βλέπει κανείς ότι οι συναρτήσεις $L_n(x)$ έχουν φραγμένο φορέα άρα δεν μπορούν να είναι πολυώνυμα σε ολόκληρο το $\mathbb{R}$. Ταυτίζονται όμως με κάποιο πολυώνυμο στο διάστημα $[0,1]$ που μας ενδιαφέρει.

η οποία εύκολα φαίνεται ότι είναι πολυώνυμο του x, μια και η συνάρτηση $(1-(u-x)^2)^n$ είναι πολυώνυμο του x με συντελεστές που εξαρτώνται από το u

$$(1-(u-x)^2)^n = \sum_{j=0}^{2n} q_j(u) x^j$$

και άρα

$$L_n(x) = c_n \sum_{j=0}^{2n} \left(\int_0^1 f(u) q_j(u)\, du \right) x^j.$$

2.2.4. *Προσέγγιση της μονάδας.* Η συνάρτηση K_n (για την ακρίβεια, η ακολουθία συναρτήσεων K_n) είναι αυτό που ονομάζεται *προσέγγιση της μονάδας*, έχει δηλ. τις παρακάτω ιδιότητες:

(1) Η $K_n(x)$ είναι κατά τμήματα συνεχής (αρκεί να υποθέσουμε ότι είναι Riemann ολοκληρώσιμη σε κάθε διάστημα),
(2) $K_n(x) \geq 0$ για κάθε $x \in \mathbb{R}$,
(3) $\int_{-\infty}^{\infty} K_n(x)\, dx = 1$,
(4) Για κάθε $\delta > 0$ ισχύει $\lim_{n\to\infty} \int_{|t|>\delta} K_n(t)\, dt = 0$.

(Το ολοκλήρωμα $\int_{|t|>\delta}$ είναι συντομογραφία του αθροίσματος $\int_{-\infty}^{-\delta} + \int_{\delta}^{\infty}$.) Η πιο καίρια ιδιότητα μιας προσέγγισης της μονάδας είναι η ιδιότητα 4 παραπάνω. Το νόημα αυτής της ιδιότητας είναι ότι, αφού τα ολοκληρώματα των K_n δεν αλλάζουν με το n και αφού το ολοκλήρωμα της K_n εκτός του διαστήματος $(-\delta, \delta)$ τείνει στο 0, τότε όλη η «μάζα» κάτω από το γράφημα του K_n μαζεύεται όλα και κοντύτερα στο 0 όσο το n μεγαλώνει. Αυτό έχει ως συνέπεια το ακόλουθο.

Θεώρημα 2.3. *Εστω $f : \mathbb{R} \to \mathbb{C}$ συνεχής και φραγμένη. Τότε για κάθε $x \in \mathbb{R}$ έχουμε*

$$\int_{-\infty}^{\infty} f(x+t) K_n(t)\, dt \to f(x). \tag{15}$$

Εάν η f είναι ομοιόμορφα συνεχής στο $\mathbb{R}$ τότε η σύγκλιση στην (15) δεν είναι απλά σύγκλιση κατά σημείο αλλά είναι και ομοιόμορφη σύγκλιση σε όλο το $\mathbb{R}$.

Απόδειξη. Γράφουμε $f(x) = \int_{-\infty}^{\infty} f(x)K_n(t)\,dt$ (αφού $\int_{-\infty}^{\infty} K_n(t)\,dt = 1$) και έτσι, αν $\delta > 0$ είναι οποιοδήποτε,

$$\begin{aligned}
&|f(x) - \int_{-\infty}^{\infty} f(x+t)K_n(t)\,dt| \\
&= \left|\int_{-\infty}^{\infty} (f(x+t) - f(x))K_n(t)\,dt\right| \\
&\le \int_{-\infty}^{\infty} |f(x+t) - f(x)|K_n(t)\,dt \\
&= \int_{-\delta}^{\delta} |f(x+t) - f(x)|K_n(t)\,dt + \\
&\quad \int_{|t|>\delta} |f(x+t) - f(x)|K_n(t)\,dt \\
&= I + II.
\end{aligned}$$

Έστω $\epsilon > 0$. Αρκεί να δείξουμε ότι, για αρκετά μεγάλο n, έχουμε $I < \epsilon$ και $II < \epsilon$.

Για τον πρώτο όρο έχουμε λόγω της συνέχειας της f στο x ότι υπάρχει $\delta > 0$ ώστε

$$|f(x+t) - f(x)| < \epsilon,$$

για $|t| \le \delta$. Αν η συνάρτηση είναι απλά συνεχής στο $\mathbb{R}$ τότε το δ αυτό εξαρτάται από το ϵ και το x ενώ αν η συνάρτηση είναι ομοιόμορφα συνεχής στο $\mathbb{R}$ τότε εξαρτάται μόνο από το ϵ. Έχουμε λοιπόν

$$I \le \int_{-\delta}^{\delta} \epsilon K_n(t)\,dt \le \epsilon \int_{-\infty}^{\infty} K_n(t)\,dt = \epsilon.$$

Για το δεύτερο όρο θα χρησιμοποιήσουμε την ιδιότητα 4 της προσέγγισης της μονάδας. Έχουμε

$$II \le \int_{|t|>\delta} 2\|f\|_\infty K_n(t)\,dt = 2\|f\|_\infty \int_{|t|>\delta} K_n(t)\,dt \to 0,$$

και άρα, για αρκετά μεγάλο n, έχουμε επίσης $II < \epsilon$. Στην παραπάνω ανισότητα χρησιμοποιήσαμε την τριγωνική ανισότητα για να φράξουμε την ποσότητα $|f(x+t) - f(x)| \le |f(x+t)| + |f(t)| \le 2\|f\|_\infty$, και η ποσότητα

$$\|f\|_\infty = \sup_{x\in\mathbb{R}} |f(x)| < \infty$$

από την υπόθεσή μας ότι η f είναι φραγμένη.

Στην εκτίμηση του II ο δείκτης n_0 πέρα από τον οποίο ισχύει $II < \epsilon$ δεν εξαρτάται από το x αλλά μόνο από το ϵ και το δ είτε η f είναι ομοιόμορφα συνεχής είτε όχι (αφού το μόνο που χρησιμοποιούμε από την f είναι ότι είναι φραγμένη).

Όμως η επιλογή του δ έγινε όταν φράξαμε το I και εκεί το δ εξαρτάται από το x, εκτός αν η συνάρτηση f υποτεθεί ομοιόμορφα συνεχής οπότε το δ εξαρτάται μόνο από το ϵ. Σε αυτή την περίπτωση έχουμε ταυτόχρονα $I + II < \epsilon$ για κάθε $x \in \mathbb{R}$ και άρα η σύγκλιση

$$\int_{-\infty}^{\infty} f(x+t)K_n(t)\,dt \to f(x)$$

είναι ομοιόμορφη στο $\mathbb{R}$.
□

Εφαρμόζοντας το Θεώρημα 2.3 παίρνουμε την ομοιόμορφη σύγκλιση των πολυωνύμων L_n στην f στο διάστημα $[0, 1]$.

Πρόβλημα N 2.1. *Ορίζουμε τη συνάρτηση* $G(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ *και υποθέτουμε ως γνωστό ότι* $\int_{-\infty}^{\infty} G(x)\, dx = 1$.

Ορίζουμε επίσης, για $\delta \to 0+$, *την οικογένεια συναρτήσεων*

$$G_\delta(x) = \frac{1}{\delta} G(x/\delta).$$

Δείξτε ότι η οικογένεια G_δ *είναι προσέγγιση της μονάδας. Ο ορισμός είναι ο ίδιος με την περίπτωση ακολουθίας συναρτήσεων μόνο που αντί για το όριο* $n \to \infty$ *έχουμε* $\delta \to 0+$.

Πρόβλημα 2.1. *Βρείτε όλες τις συναρτήσεις* $f : \mathbb{R} \to \mathbb{R}$ *που είναι ομοιόμορφα (στο* $\mathbb{R}$*) όρια πολυωνύμων.*

Ένα πολυώνυμο $p(x)$ *που είαι φραγμένο στο* $\mathbb{R}$ *είναι αναγκαστικά σταθερά.*

2.3. **Η απόδειξη του Bernstein.** Η δεύτερη απόδειξη του θεωρήματος του Weierstrass (Θεώρημα 2.1) που θα δούμε οφείλεται στον Bernstein και είναι πολύ ενδιαφέρουσα γιατί συνδέει τη θεωρία Πιθανοτήτων με την πολυωνυμική προσέγγιση που θέλουμε να πετύχουμε. Επίσης τα πολυώνυμα που δίνει ως προσέγγιση της f είναι πολύ απλό να περιγραφούν.

2.3.1. *Τα πολυώνυμα του Bernstein και η πιθανοθεωρητική τους ερμηνεία.* Υποθέτουμε και πάλι ότι η $f : [0, 1] \to \mathbb{C}$ είναι συνεχής (δε χρειάζεται αυτή τη φορά να κανονικοποιήσουμε την f ώστε να μηδενίζεται στα άκρα του διαστήματος). Ορίζουμε τα *πολυώνυμα Bernstein της* f να είναι η ακολουθία πολυωνύμων $B_n(f)$ που δίδεται από

$$B_n(f)(x) = \sum_{k=0}^{n} f(k/n) \binom{n}{k} x^k (1-x)^{n-k}. \tag{16}$$

Θα δείξουμε ότι τα $B_n(f)(x) \to f(x)$ ομοιόμορφα στο διάστημα $[0, 1]$.

Για την απόδειξη θα χρειαστεί να ερμηνεύσουμε τα πολυώνυμα του Bernstein σε πιθανοθεωρητική γλώσσα.[2]

Ας υποθέσουμε ότι έχουμε ένα νόμισμα το οποίο φέρνει κορώνα με πιθανότητα $x \in [0, 1]$ και το ρίχνουμε n φορές. Η τυχαία μεταβλητή $B_{x,n}$ μετράει το πόσες κορώνες φέραμε σε αυτό το πείραμα. Προφανώς ισχύει πάντα $0 \le B_{x,n} \le n$ και εύκολα βλέπει κανείς ότι η $B_{x,n}$ ακολουθεί τη λεγόμενη *διωνυμική κατανομή*, έχουμε δηλαδή για $k \in \mathbb{Z}$

$$\mathbb{P}\left[B_{x,n} = k\right] = \begin{cases} \binom{n}{k} x^k (1-x)^{n-k} & \text{αν } 0 \le k \le n \\ 0 & \text{αλλιώς} \end{cases}.$$

Ο λόγος για τον παραπάνω τύπο είναι ότι το να φέρουμε ακριβώς k κορώνες στο πείραμα μπορεί να συμβεί με ακριβώς $\binom{n}{k}$ τρόπους (επιλέγουμε από τις n ρίψεις σε ποιες k θα έρθουν οι κορώνες) και κάθε ένας από αυτούς τους τρόπους έχει πιθανότητα $x^k(1-x)^{n-k}$ να συμβεί.

[2]Μπορείτε να αναφέρεστε στις σημειώσεις μου Διακριτή Πιθανότητα για τους διάφορους ορισμούς και ορισμένους υπολογισμούς.

Κάνουμε επίσης την παρατήρηση ότι, αν γράψουμε I_j για την *δείκτρια* τυχαία μεταβλητή που είναι 1 αν φέρουμε κορώνα στην j ρίψη και 0 αν φέρουμε γράμματα στην j ρίψη, ισχύει

$$B_{x,n} = I_1 + I_2 + \cdots + I_n.$$

Έχουμε $\mathbb{E}\left[I_j\right] = x$ και $\mathbf{Var}\left[I_j\right] = x(1-x)$ με ένα απλό υπολογισμό οπότε συμπεραίνουμε ότι

$$\mathbb{E}\left[B_{x,n}\right] = \mathbb{E}\left[I_1\right] + \cdots + \mathbb{E}\left[I_n\right] = nx$$

και από το γεγονός ότι οι I_j είναι ανεξάρτητες προκύπτει επίσης ότι

$$\begin{align}
\sigma^2(B_{x,n}) &= \mathbf{Var}\left[B_{x,n}\right] \tag{17}\\
&= \mathbf{Var}\left[I_1\right] + \cdots + \mathbf{Var}\left[I_n\right] \tag{18}\\
&= nx(1-x) \tag{19}\\
&\le n. \tag{20}
\end{align}$$

Τέλος, θα χρειαστούμε και την ανισότητα του Chebyshev που δίνει ένα άνω φράγμα για την πιθανότητα να αποκλίνει μια τυχαία μεταβλητή από τη μέση τιμή της.

Θεώρημα 2.4 (Chebyshev). *Αν $X \in \mathbb{Z}$ είναι μια τυχαία μεταβλητή με*

$$\mathbb{E}\left[|X|^2\right] < \infty,$$

και αν $\mu = \mathbb{E}\left[X\right]$ και $\sigma = \sqrt{\mathbf{Var}\left[X\right]}$ τότε

$$\mathbb{P}\left[|X-\mu| \ge \lambda\sigma\right] \le \frac{1}{\lambda^2}, \tag{21}$$

για κάθε $\lambda > 0$.

Πρόβλημα 2.2. *Θα χρειαστούμε επίσης αργότερα και την ανισότητα $\mathbb{E}\left[|X|\right] \le \mathbb{E}\left[|X|^2\right]^{1/2}$ για μια διακριτή τυχαία μεταβλητή $X \in \mathbb{Z}$. Αποδείξτε αυτή την ανισότητα.*

Οι δύο ποσότητες που μας ενδιαφέρουν είναι οι

$$\mathbb{E}\left[|X|\right] = \sum_{k\in\mathbb{Z}} |k| \mathbb{P}\left[X = k\right]$$

και

$$\mathbb{E}\left[|X|^2\right] = \sum_{k\in\mathbb{Z}} |k|^2 \mathbb{P}\left[X = k\right].$$

Χρησιμοποιείστε την ανισότητα Cauchy-Schwarz (3) για να δείξετε το ζητούμενο. Στην (3) η ανισότητα είναι διατυπωμένη για πεπερασμένες ακολουθίες αλλά ισχύει και για ακολουθίες όπου ο δείκτης παίρνει όλες τις ακέραιες τιμές και τα αθροίσματα γίνονται άπειρες σειρές.

Παρατηρούμε τώρα ότι υπάρχει η εξής σχέση ανάμεσα στα πολυώνυμα Bernstein, την τυχαία μεταβλητή $B_{x,n}$ και τη συνάρτηση f:

$$B_n(f)(x) = \mathbb{E}\left[f(B_{x,n}/n)\right]. \tag{22}$$

Πρόβλημα 2.3. *Βεβαιωθείτε ότι καταλαβαίνετε γιατί ισχύει αυτό. Θυμίζουμε ότι αν* $X \in \mathbb{Z}$ *είναι μια διακριτή τυχαία μεταβλητή και* $\phi : \mathbb{Z} \to \mathbb{C}$ *μια συνάρτηση τότε έχουμε*

$$\mathbb{E}\left[\phi(X)\right] = \sum_{k \in \mathbb{Z}} \phi(k) \mathbb{P}\left[X = k\right],$$

με την προϋπόθεση φυσικά ότι η σειρά συγκλίνει απόλυτα. Στη δικιά μας περίπτωση, όπου $X = B_{x,n}$ *δεν τίθεται θέμα σύγκλισης μια και όλα τα αθροίσματα έχουν πεπερασμένο πλήθος μη μηδενικών όρων αφού* $0 \le B_{x,n} \le n$.

Από την (22) γίνεται τώρα φανερό το γιατί πρέπει να περιμένουμε τη σύγκλιση

$$B_n(f)(x) \to f(x) \quad \text{για } n \to \infty.$$

Ο λόγος είναι ότι η τυχαία μεταβλητή $B_{x,n}/n$, η οποία έχει μέση τιμή x τείνει να συγκεντρώνεται γύρω από τη μέση της τιμή (νόμος των μεγάλων αριθμών). Έτσι λοιπόν, με πολύ μεγάλη πιθανότητα, οι τιμές $f(B_{x,n}/n)$ είναι πολύ κοντά στην $f(x)$ λόγω της συνέχειας της f και άρα είναι αναμενόμενο και η μέση τιμή της μεταβλητής αυτής να είναι κοντά στο $f(x)$. Αυτό το συλλογισμό ποσοτικοποιούμε παρακάτω στην απόδειξη.

Γράφουμε E για το ενδεχόμενο

$$E = \left\{ \left| \frac{B_{x,n}}{n} - x \right| \ge n^{-1/3} \right\},$$

και κάνουμε την παρατήρηση ότι η ανισότητα του Chebyshev μας δίνει το παρακάτω άνω φράγμα για την πιθανότητα του E:

$$\mathbb{P}\left[E\right] \le n^{-1/3}. \tag{23}$$

Πρόβλημα 2.4. *Αποδείξτε ότι η ανισότητα* (23) *προκύπτει από την ανισότητα Chebyshev* (21).

Χρησιμοποιείστε το φράγμα $\mathbf{Var}\left[B_{x,n}\right] \le n$ *και την τιμή* $\lambda = n^{1/6}$ *και εφαρμόστε την ανισότητα* (21) *για το ενδεχόμενο* E *γραμμένο στη μορφή*

$$|B_{x,n} - nx| \le \lambda n^{1/2}.$$

Φράσσουμε τώρα τη διαφορά της συνάρτησής μας και της προσέγγισής της από ένα πολυώνυμο Bernstein:

$$\begin{aligned}
&|f(x) - \mathbb{E}f(\frac{B_{x,n}}{n})| \\
&= \left| f(x) - \sum_{k=0}^{n} f(k/n)\mathbb{P}\left[B_{x,n} = k\right] \right| \\
&= \left| \sum_{k=0}^{n} f(x)\mathbb{P}\left[B_{x,n} = k\right] - \sum_{k=0}^{n} f(\frac{k}{n})\mathbb{P}\left[B_{x,n} = k\right] \right| \\
&\qquad (\text{αφού} \sum_{k} \mathbb{P}\left[B_{x,n} = k\right] = 1) \\
&= \left| \sum_{k=0}^{n} (f(x) - f(k/n))\mathbb{P}\left[B_{x,n} = k\right] \right| \\
&\le \sum_{k=0}^{n} |f(x) - f(k/n)| \cdot \mathbb{P}\left[B_{x,n} = k\right] \\
&\qquad (\text{τριγωνική ανισότητα}) \\
&= \sum_{\substack{k=0 \\ |x-k/n| \ge n^{-1/3}}}^{n} + \sum_{\substack{k=0 \\ |x-k/n| < n^{-1/3}}}^{n} \\
&\qquad (\text{διαχωρίζουμε τα } k \text{ σε δύο είδη}) \\
&= I + II.
\end{aligned}$$

Έστω $\epsilon > 0$. Θα δείξουμε ότι αν το n είναι αρκετά μεγάλο τότε τα αθροίσματα I και II φράσσονται από ϵ.

Για να φράξουμε το II χρησιμοποιούμε την *ομοιόμορφη συνέχεια* της f (η οποία είναι συνέπεια της συνέχειας της f σε κλειστό διάστημα, δείτε το Πρόβλημα 2.5): για κάθε $\epsilon > 0$ υπάρχει $\delta > 0$ τ.ώ. αν $|x - y| < \delta$ να έπεται ότι $|f(x) - f(y)| < \epsilon$. Αν λοιπόν το n είναι αρκετά μεγάλο ώστε να ισχύει $n^{-1/3} < \delta$ τότε στο άθροισμα I η ποσότητα $|f(x) - f(k/n)|$ φράσσεται από ϵ οπότε ισχύει

$$II \le \epsilon \sum_{\substack{k=0 \\ |x-k/n| < n^{-1/3}}}^{n} \mathbb{P}\left[B_{x,n} = k\right] \le \epsilon \sum_{k=0}^{n} \mathbb{P}\left[B_{x,n} = k\right] = \epsilon.$$

Για τον όρο I φράσσουμε την ποσότητα $|f(x) - f(k/n)|$ από $2\|f\|_\infty$ (η συνάρτηση f είναι φραγμένη αφού είναι συνεχής σε κλειστό διάστημα, οπότε $\|f\|_\infty < \infty$) και έχουμε

$$\begin{aligned}
I &\le 2\|f\|_\infty \sum_{\substack{k=0 \\ |x-k/n| \ge n^{-1/3}}}^{n} \mathbb{P}\left[B_{x,n} = k\right] \\
&= 2\|f\|_\infty \mathbb{P}\left[E\right] \\
&\le 2\|f\|_\infty n^{-1/3}.
\end{aligned}$$

Επιλέγοντας και πάλι το n ακόμη μεγαλύτερο (ώστε να ισχύει και $2\|f\|_\infty n^{-1/3} < \epsilon$) πετυχαίνουμε να ισχύει και η ανισότητα $II < \epsilon$. Η απόδειξή μας είναι πλήρης.

Πρόβλημα 2.5. *Μια συνάρτηση $f : K \to \mathbb{C}$ λέγεται* ομοιόμορφα συνεχής *στο K αν για κάθε $\epsilon > 0$ υπάρχει $\delta > 0$ τ.ώ.*

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

Η διαφορά με την απλή συνέχεια στο σύνολο K είναι ότι στην περίπτωση της απλής συνέχειας το δ εξαρτάται όχι μόνο από το ϵ αλλά και από το x. Η f δηλ. είναι συνεχής στο K αν

$$\forall x \in K\ \forall \epsilon > 0\ \exists \delta > 0 : \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

Δείξτε ότι αν η συνάρτηση $f : [a, b] \to \mathbb{C}$ είναι συνεχής στο κλειστό και φραγμένο διάστημα $[a, b]$ τότε είναι και ομοιόμορφα συνεχής εκεί. Βρείτε επίσης μια συνάρτηση $g : \mathbb{R} \to \mathbb{R}$ και μια συνάρτηση $h : (0, 1) \to \mathbb{R}$ που να είναι συνεχείς αλλά όχι ομοιόμορφα συνεχείς στο πεδίο ορισμού τους.

Για να αποδείξετε την ομοιόμορφη συνέχεια της f πρέπει να χρησιμοποιήσετε τη συμπάγεια φραγμένου κλειστού διαστήματος $[a, b]$: κάθε ακολουθία $x_n \in [a, b]$ έχει υπακολουθία η οποία συγκλίνει (αναγκαστικά σε κάποιο σημείο του $[a, b]$).

Υποθέστε ότι η f δεν είναι ομοιόμορφα συνεχής. Αυτό σημαίνει ότι για κάποιο $\epsilon > 0$ τότε για υπάρχουν $x_n, y_n \in [a, b]$ τ.ώ. $|x_n - y_n| \to 0$ αλλά με

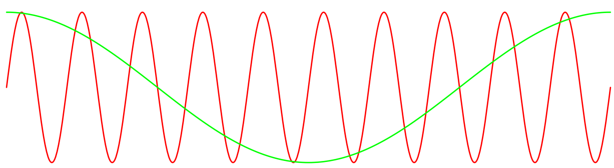
$$\inf_n |f(x_n) - f(y_n)| > 0.$$

Βρείτε μια ακολουθία δεικτών n_k τέτοια ώστε οι δύο ακολουθίες x_{n_k} και y_{n_k} να συγκλίνουν για $k \to \infty$ και χρησιμοποιείστε τη συνέχεια της f για να καταλήξετε σε άτοπο.

Όσον αφορά τις συναρτήσεις g και h που πρέπει να βρείτε μπορείτε να δοκιμάσετε τις $g(x) = e^x$ και $h(x) = 1/x$.

Πρόβλημα 2.6. *Αν $f : [1, \infty) \to \mathbb{C}$ είναι συνεχής και $\lim_{x\to+\infty} f(x)$ είναι πραγματικός αριθμός, δείξτε ότι η f προσεγγίζεται ομοιόμορφα στο διάστημα $[1, \infty)$ από συναρτήσεις της μορφής $p(1/x)$, όπου p πολυώνυμο.*

2.3.2. *Το μέτρο συνέχειας μιας συνάρτησης.* Το να πούμε ότι μια συνάρτηση είναι συνεχής σε κάποιο σύνολο είναι μια ποιοτική και όχι ποσοτική δήλωση. Είναι μια ιδιότητα που η συνάρτηση την έχει ή όχι. Κατά κάποιον τρόπο όμως υπάρχουν συναρτήσεις που είναι «πιο συνεχείς» από άλλες, όπως για παράδειγμα η πράσινη συνάρτηση στο Σχήμα 3 είναι πιο συνεχής από την κόκκινη συνάρτηση στο ίδιο Σχήμα, μεταβάλλεται δηλ. πιο αργά.



Σχήμα 3: Μια συνάρτηση που είναι «πιο συνεχής» από μια άλλη

Η έννοια του *μέτρου συνέχειας* μιας συνάρτησης παίζει ακριβώς αυτό το ρόλο της ποσοτικοποίησης του πόσο γρήγορα αλλάζει μια συνάρτηση όταν αλλάζει η μεταβλητή.

Ορισμός 2.2. Αν $f : K \to \mathbb{C}$ τότε το *μέτρο συνέχειας της f στο K* είναι η συνάρτηση

(24) $$\omega_f(\delta) = \sup\{|f(x) - f(y)| :\ x, y \in K,\ |x - y| \le \delta\}, \quad (\delta > 0).$$

Η ποσότητα $\omega_f(\delta)$, με άλλα λόγια, μας λέει πόσο πολύ μπορεί να μεταβληθεί η τιμή της συνάρτησης f αν η μεταβλητή της αλλάξει το πολύ κατά δ.

Η συνάρτηση $\omega_f(\delta)$ μπορεί να μην είναι πεπερασμένη ακόμη κι όταν η συνάρτηση f είναι συνεχής στο K. Για παράδειγμα, εύκολα μπορεί κανείς να δει ότι αν $f(x) = e^x$, ορισμένη για $x \in \mathbb{R}$, τότε για κάθε $\delta > 0$ ισχύει $\omega_f(\delta) = +\infty$. Δείτε όμως το Πρόβλημα 2.7.

Πρόβλημα 2.7. *Δείξτε ότι μια συνάρτηση $f : K \to \mathbb{C}$ είναι ομοιόμορφα συνεχής πάνω στο K αν και μόνο αν $\omega_f(\delta) \to 0$ για $\delta \to 0$.*

Πρόβλημα 2.8. *Αν $f : K \to \mathbb{C}$ και $f = u + iv$, όπου u, v πραγματικές συναρτήσεις δείξτε τότε ότι*

$$\omega_f(\delta) \le |\omega_u(\delta) + i\omega_v(\delta)| = \sqrt{\omega_u(\delta)^2 + \omega_v(\delta)^2}.$$

Πρόβλημα 2.9. *Αν η $f : [a, b] \to \mathbb{R}$ είναι παραγωγίσιμη στο $[a, b]$ (στα άκρα εννοείται ότι υπάρχουν οι πλευρικές παράγωγοι) και $|f'(x)| \le M$ για $x \in [a, b]$, δείξτε ότι $\omega_f(\delta) \le M\delta$, για κάθε $\delta > 0$.*

Χρησιμοποιείστε το θεώρημα μέσης τιμής για να φράξετε τη διαφορά $|f(x) - f(y)|$.

Το μέτρο συνέχειας της f είναι *υποπροσθετική συνάρτηση*.

Πρόβλημα 2.10. *Αν $f : I \to \mathbb{R}$, όπου $I \subseteq \mathbb{R}$ διάστημα, τότε $\omega_f(\delta_1 + \delta_2) \le \omega_f(\delta_1) + \omega_f(\delta_2)$.*

Αν $|x - y| \le \delta_1 + \delta_2$ τότε, αν υπάρχει z ανάμεσα στα x και y τ.ώ. $|x - z| \le \delta_1$ και $|z - y| \le \delta_2$.

Πρόβλημα 2.11. *Αν $f : I \to \mathbb{R}$, όπου $I \subseteq \mathbb{R}$ διάστημα, και $\lambda > 0$ τότε*

$$\omega_f(\lambda\delta) \le (1 + \lambda)\omega_f(\delta).$$

Χρησιμοποιείστε το Πρόβλημα 2.10.

2.3.3. *Εκτίμηση του σφάλματος για τα πολυώνυμα Bernstein.* Το Θεώρημα του Weierstrass δε μας δίνει κάποια εκτίμηση για το πόσο μεγάλο μπορεί να είναι το «σφάλμα» $\|f - p\|_\infty = \sup_{x\in[a,b]} |f(x) - p(x)|$ όταν η $f(x)$ είναι μια συνεχής συνάρτηση στο διάστημα $[a, b]$ και το $p(x)$ είναι πολυώνυμο βαθμού $\le n$. Κεντρικό ερώτημα της Θεωρίας Προσέγγισης είναι το πόσο μικρή μπορεί να γίνει αυτή η ποσότητα αν μας επιτρέπεται να διαλέξουμε κατάλληλα το πολυώνυμο $p(x)$. Με άλλα λόγια μας ενδιαφέρει η ποσότητα

$$E_n(f) = \inf\{\|f - p\|_\infty :\ \text{το } p(x) \text{ είναι πολυώνυμο βαθμού το πολύ } n\}.$$

Όπως αποδείξαμε στα εισαγωγικά μαθήματα, όταν μιλάγαμε για το πρόβλημα της προσέγγισης των στοιχείων ενός διανυσματικού χώρου από στοιχεία ενός υποχώρου, το παραπάνω infimum «πιάνεται» από κάποιο πολυώνυμο $p(x)$ βαθμού μέχρι n, μια και ο χώρος αυτός των πολυωνύμων έχει πεπερασμένη διάσταση και μάλιστα ίση με $n + 1$ (αφού κάθε τέτοιο πολυώνυμο μπορεί να προσδιορισθεί δίνοντας $n + 1$ παραμέτρους, τους συντελεστές του). Φυσικά μια εκτίμηση για την ποσότητα $E_n(f)$ θα πρέπει να επηρεάζεται από το πόσο «καλή» είναι η συνάρτηση f. Αυτή η ιδιότητα της f ποσοτικοποιείται από το μέτρο συνέχειας της f τη συνάρηση $\omega_f(\delta)$.

Θεώρημα 2.5. *Εστω $f \in C([0, 1])$. Τότε ισχύει*

$$E_n(f) \le 2\omega_f(1/\sqrt{n}).$$

Πιο συγκεκριμένα, αν $B_n(f)$ είναι το n-οστό πολυώνυμο Bernstein της f τότε $\|f - B_n(f)\|_\infty \le 2\omega_f(1/\sqrt{n})$.

Ξεκινάμε χρησιμοποιώντας της σχέση (22).

$$\begin{aligned}
|f(x)-B_n(f)(x)| &= |f(x)-\mathbb{E}\left[f(B_{x,n}/n)\right]| \\
&= |\mathbb{E}\left[f(x)-f(B_{x,n}/n)\right]| \\
&\leq \mathbb{E}\left[|f(x)-f(B_{x,n}/n)|\right] \quad (\text{τριγωνική ανισότητα: } |\mathbb{E}\left[X\right]| \leq \mathbb{E}\left[|X|\right]) \\
&= \sum_{k=0}^{n} |f(x)-f(k/n)| \mathbb{P}\left[B_{x,n}=k\right] \quad (\text{από το } \mathbb{E}\left[\phi(X)\right] = \textstyle\sum_k \phi(k)\mathbb{P}\left[X=k\right]) \\
&\leq \sum_{k=0}^{n} \omega_f(\left|x-\frac{k}{n}\right|)\mathbb{P}\left[B_{x,n}=k\right] \\
&= \sum_{k=0}^{n} \omega_f(\lambda\frac{1}{\sqrt{n}})\mathbb{P}\left[B_{x,n}=k\right] \quad (\text{με } \lambda = \sqrt{n}|x-\tfrac{k}{n}|) \\
&\leq \omega_f(1/\sqrt{n})\sum_{k=0}^{n}(1+\sqrt{n}\left|x-\frac{k}{n}\right|)\mathbb{P}\left[B_{x,n}=k\right] \quad (\text{από το Πρόβλημα 2.11}) \\
&= \omega_f(1/\sqrt{n})(1+\sqrt{n}\mathbb{E}\left[\left|x-\frac{B_{x,n}}{n}\right|\right]) \\
&= \omega_f(1/\sqrt{n})(1+\frac{1}{\sqrt{n}}\mathbb{E}\left[|B_{x,n}-nx|\right]) \\
&\leq \omega_f(1/\sqrt{n})(1+\frac{1}{\sqrt{n}}\sigma(B_{x,n})) \quad (\text{από το Πρόβλημα 2.2}) \\
&\leq 2\omega_f(1/\sqrt{n}) \quad (\text{από την (17)}).
\end{aligned}$$

Η απόδειξη του Θεωρήματος 2.5 είναι πλήρης.

Παρατήρηση 2.3. Υπάρχει πολύ καλύτερη εκτίμηση για το σφάλμα $E_n(f)$ από αυτή του Θεωρήματος 2.5. Το Θεώρημα του Jackson μας λέει ότι $E_n(f) \leq C\omega_f(1/n)$, όπου C είναι μια απόλυτη σταθερά.

3. Τριγωνομετρικά πολυώνυμα

3.1. **Τριγωνομετρικά Πολυώνυμα.** Ένα *τριγωνομετρικό πολυώνυμο* είναι ένας πεπερασμένος γραμμικός συνδυασμός, με μιγαδικούς συντελεστές, μιγαδικών εκθετικών συναρτήσεων, δηλ. των συναρτήσεων

$$e_n(x) = e^{inx}, \quad (n \in \mathbb{Z}).$$

Ένας άλλος τρόπος να πούμε το ίδιο πράγμα είναι να πούμε ότι τριγωνομετρικά πολυώνυμα είναι συναρτήσεις της μορφής

$$p(x) = \sum_{k=-N}^{N} p_k e_k(x), \tag{25}$$

όπου N είναι ένας μη αρνητικός ακέραιος. Ο αριθμός k ονομάζεται και *συχνότητα* του εκθετικού $e_k(x)$. Η μεγαλύτερη, κατ' απόλυτη τιμή, συχνότητα που εμφανίζεται σ' ένα τριγωνομετρικό πολυώνυμο ονομάζεται βαθμός του πολυωνύμου και συμοβολίζεται με $\deg p$. Για παραδειγμα, αν $p(x) = 3e^{-i4x}+1+e^{i2x}$ τότε $\deg p = 4$.

Τα p_k στην (25) ονομάζονται συντελεστές του $p(x)$ και καθορίζονται μοναδικά. Δε μπορεί δηλ. η ίδια συνάρτηση $p(x)$ να γραφεί με δυο διαφορετικούς τρόπους στη γραφή (25).

Πρόβλημα 3.1. *(Μοναδικότητα των συντελεστών)*
Αν $\sum_{k=-N}^{N} p_k e^{ikx} = \sum_{k=-N}^{N} q_k e^{ikx}$ για κάθε x σε ένα σύνολο $A \subseteq \mathbb{R}$ με $|A| \geq 2N+1$ τότε $p_k = q_k$ για κάθε k.

Αρκεί να δείξουμε ότι αν $\sum_{k=-N}^{N} p_k e^{ikx} = 0$ για κάθε $x \in A = \{a_1, a_2, \ldots, a_{2N+1}, \ldots\}$ τότε $p_k = 0$ για κάθε k. Δείξτε ότι αρκεί ο $(2N+1) \times (2N+1)$ πίνακας με στοιχεία τα e^{ika_j}, $k = -N, \ldots, N$, $j = 1, 2, \ldots, 2N+1$, είναι αντιστρέψιμος.

Αυτό ανάγεται σε ένα πίνακα Vandermonde A με $A_{jk} = x_j^k$,

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{bmatrix}$$

όπου $j = 0, 2, \ldots, n-1$, $k = 1, 2, \ldots, n$, και τα $x_j \in \mathbb{C}$. Δείξτε ότι αν όλα τα x_j είναι διαφορετικά τότε ο πίνακας είναι αντιστρέψιμος υπολογίζοντας την ορίζουσά του και δείχνοντας ότι αυτή ισούται με $\pm$ το γινόμενο όλων των διαφορών $x_r - x_s$, όπου $r \neq s$:

$$\det A = \pm \prod_{\substack{r,s=1,\ldots,n \\ r<s}} (x_r - x_s).$$

Αυτό μπορεί να αποδειχτεί με επαγωγή ως προς το n. New! Για παράδειγμα, κάνοντας πράξεις στηλών μετατρέψτε την πρώτη γραμμή σε $(1, 0, 0, \ldots, 0)$ αφαιρώντας από κάθε στήλη ένα κατάλληλο πολλαπλάσιο της πρώτης στήλης. Έπειτα αναπτύξτε την ορίζουσα ως προς την πρώτη γραμμή και χρησιμοποιήστε την επαγωγική υπόθεση.

Ένας άλλος τρόπος να αποδείξετε ότι ένα τριγωνομετρικό πολυώνυμο βαθμού $\leq N$ το οποίο μηδενίζεται σε $2N+1$ σημεία έχει όλους τους συντελεστές του μηδενικούς είναι να χρησιμοποιήσετε την αντίστοιχη πρόταση για τα αλγεβρικά πολυώνυμα, ότι δηλ. ένα αλγεβρικό πολυώνυμο βαθμού $\leq M$ που μηδενίζεται σε $M+1$ σημεία είναι αναγκαστικά το μηδενικό πολυώνυμο, αυτό δηλ. με όλους τους συντελεστές ίσους με το μηδέν. Χρησιμοποιείστε το πολυώνυμο

$$q(z) = p_{-N} + p_{-N+1}z + p_{-N+2}z^2 + \cdots + p_N z^{2N} = z^N \sum_{k=-N}^{N} p_k z^k.$$

Ποια η σχέση των μηδενικών του τριγωνομετρικού πολυωνύμου $p(x)$ με τις ρίζες του αλγεβρικού πολυωνύμου $q(z)$ όταν κοιτάξετε μόνο τα z με $|z| = 1$ και θέσετε $z = e^{ix}$;

Άμεση συνέπεια του θεμελιώδους τύπου $e^{ix} = \cos x + i \sin x$, για $x \in \mathbb{R}$, είναι ότι

$$\cos x = (e^{-ix} + e^{ix})/2, \quad \text{και} \quad \sin x = (-e^{-ix} + e^{ix})/(2i),$$

πράγμα που σημαίνει ότι οι συναρτήσεις $\cos x$ και $\sin x$ είναι τριγωνομετρικά πολυώνυμα βαθμού 1.

Πρόβλημα 3.2. *(Τριγωνομετρικές συναρτήσεις αθροισμάτων και διαφορών)*
Βρείτε τύπους για τα $\cos(a \pm b), \sin(a \pm b)$ μέσω των τριγωνομετρικών αριθμών των a και b χρησιμοποιώντας τον τύπο $e^{ix} = \cos x + i \sin x$.

Ξεκινείστε γράφοντας το $e^{i(a\pm b)}$ με δύο τρόπους.

Μια συνάρτηση $f : \mathbb{R} \to \mathbb{C}$ ονομάζεται *περιοδική με περίοδο T* αν

$$f(x+T) = f(x), \quad (x \in \mathbb{R}).$$

Για παράδειγμα η συνάρτηση $\sin x$ είναι περιοδική με περίοδο 2π και η συνάρτηση $e^{2\pi i x}$ είναι περιοδική με περίοδο 1.

Πρόβλημα 3.3. *(Σύνολο των περιόδων μιας συνάρτησης)*
Η περίοδος μιας περιοδικής συνάρτησης δεν είναι ποτέ μοναδική. Αποδείξτε ότι το σύνολο των περιόδων μιας συνάρτησης αποτελεί ομάδα. *Αν δηλ. T_1 και T_2 είναι περίοδοι τότε και οι αριθμοί $-T_1, -T_2, T_1+T_2$ είναι επίσης περίοδοι. Δείξτε επίσης ότι κάθε μη σταθερή, συνεχής περιοδική συνάρτηση έχει μια ελάχιστη θετική περίοδο και κάθε άλλη περίοδος της συνάρτησης είναι πολλαπλάσιό της.*

Πρόβλημα 3.4. *Αν $f : \mathbb{R} \to \mathbb{C}$ είναι μια συνεχής (ή, γενικότερα, μια Riemann ολοκληρώσιμη) T-περιοδική συνάρτηση τότε αν $x, y \in \mathbb{R}$ έχουμε*

$$\int_x^{x+T} f(t)\,dt = \int_y^{y+T} f(t)\,dt.$$

Παρατήρηση 3.1. Οι μιγαδικές εκθετικές συναρτήσεις $e_n(x) = e^{inx}$, $n \in \mathbb{Z}$, έχουν όλες περίοδο 2π (αλλά μόνο οι $e_1(x)$ και $e_{-1}(x)$ έχουν ελάχιστη περίοδο 2π), και είναι ιδιαίτερα χρήσιμες στη μελέτη συναρτήσεων που είναι περιοδικές με αυτή την περίοδο. Είναι όμως αρκετά κοινό, αν θέλουμε να μελετήσουμε φαινόμενα που έχουν άλλη περίοδο, να χρησιμοποιούμε εκθετικές συναρτήσεις που είναι ελαφρώς διαφορετικές. Για παράδειγμα, εξίσου κοινές με τις παραπάνω συναρτήσεις είναι και οι εκθετικές συναρτήσεις $e^{2\pi i n x}$, $n \in \mathbb{Z}$, οι οποίες έχουν περίοδο 1 αντί για 2π. Όλα όσα θα πούμε παρακάτω ισχύουν, φυσικά με τις κατάλληλες ελάχιστες τροποποιήσεις, και σε κάθε τέτοια οικογένεια μιγαδικών εκθετικών συναρτήσεων και για τα αντίστοιχα τριγωνομετρικά πολυώνυμα. Η αναγωγή από τη μια περίπτωση στην άλλη γίνεται με μια απλή γραμμική αλλαγή μεταβλητής.

3.2. **Εσωτερικό γινόμενο.** Αν $f, g \in C([0, 2\pi])$ ορίζουμε το *εσωτερικό τους γινόμενο* να είναι η ποσότητα

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x)\overline{g(x)}\,dx.$$

Παρατήρηση 3.2. Εδώ $\overline{g(x)}$ είναι το μιγαδικό συζυγές του $g(x)$. Αν οι συναρτήσεις είναι πραγματικές τότε ο παραπάνω ορισμός του εσωτερικού γινομένου δίδεται συνήθως χωρίς το μιγαδικό συζυγές. Επίσης ο παράγοντας $\frac{1}{2\pi}$ χρησιμεύει στο να μετατρέψει το ολοκλήρωμα σε ένα μέσο όρο πάνω στο διάστημα $[0, 2\pi]$ και απλουστεύονται πολύ τύποι έτσι. Μπορεί όμως σε άλλα κείμενα να δείτε τον ορισμό χωρίς τον παράγοντα αυτό. Για παράδειγμα, αν τα εκθετικά που χρησιμοποιούνται είναι τα $e^{2\pi i n x}$, $n \in \mathbb{Z}$, τότε ο ορισμός του εσωτερικού γινομένου είναι $\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}\,dx$. Σε πολλά βιβλία επίσης θα δείτε το εσωτερικό γινόμενο να έχει το συζυγές στον πρώτο παράγοντα αντί για το δεύτερο.

Πρόβλημα 3.5. *(Αλγεβρικές ιδιότητες του εσωτερικού γινομένου)*
Το εσωτερικό γινόμενο είναι ουσιαστικά γραμμικό στους δύο παράγοντες, αν εξαιρέσουμε την μικρή επιπλοκή που δημιουργεί η ύπαρξη του μιγαδικού συζυγούς στο δεύτερο παράγοντα. Αποδείξτε τις παρακάτω

ιδιότητες ($f, g, h \in C([0, 2\pi])$):

$$\langle g, f\rangle = \overline{\langle f, g\rangle}$$

$$\langle \lambda f + \mu h, g\rangle = \lambda\langle f, g\rangle + \mu\langle h, g\rangle \qquad \text{(για } \lambda, \mu \in \mathbb{C}\text{)}$$

$$\langle f, \lambda g + \mu h\rangle = \overline{\lambda}\langle f, g\rangle + \overline{\mu}\langle f, h\rangle \qquad \text{(για } \lambda, \mu \in \mathbb{C}\text{)}$$

☞ $$\langle f, f\rangle = \|f\|_2^2 := \frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2\, dx \qquad \text{(παραλλαγή ορισμού της 2-νόρμας).}$$

Δύο συναρτήσεις $f, g \in C([0, 2\pi])$ ονομάζονται *ορθογώνιες* αν $\langle f, g\rangle = 0$.

Πρόβλημα 3.6. *(Πυθαγόρειο Θεώρημα)*
Αν $f_1, f_2, \ldots, f_n$ είναι ανά δύο ορθογώνιες τότε

(26) $$\|f_1 + \ldots + f_n\|_2^2 = \|f_1\|_2^2 + \cdots + \|f_n\|_2^2.$$

Χρησιμοποιείστε επαγωγή ως προς n. Για $n = 2$ γράψτε

$$\|f_1 + f_2\|_2^2 = \langle f_1 + f_2, f_1 + f_2\rangle = \langle f_1, f_1, +\rangle\langle f_1, f_2\rangle + \langle f_2, f_1\rangle + \langle f_2, f_2\rangle$$

και χρησιμοποιείστε την ορθογωνιότητα.

Είναι πολύ βασικό και πολύ χρήσιμο ότι οι μιγαδικές εκθετικές συναρτήσεις $e_n(x) = e^{inx}$, $n \in \mathbb{Z}$, είναι ανά δύο ορθογώνιες. Αν $m \neq n$ και θέτοντας $k = m - n \neq 0$ έχουμε

$$\begin{aligned}\langle e_m, e_n\rangle &= \frac{1}{2\pi}\int_0^{2\pi} e^{imx}e^{-inx}\, dx\\ &= \frac{1}{2\pi}\int_0^{2\pi} e^{ikx}\, dx\\ &= \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{e^{ikx}}{ik}\right)'\, dx\\ &= \frac{1}{ik}(e^{ik2\pi} - e^{ik\cdot 0})\\ &= 0.\end{aligned}$$

Ισχύει επίσης $\langle e_n, e_n\rangle = \|e_n\|_2^2 = 1$ για $n \in \mathbb{Z}$. Γι' αυτό το λόγο οι συναρτήσεις $e_n(x)$ λέμε ότι αποτελούν ένα *ορθοκανονικό σύστημα*.

Πρόβλημα 3.7. *(Ορθογωνιότητα των τριγωνομετρικών συναρτήσεων)*
Αποδείξτε ότι οι συναρτήσεις

$$1, \cos nx, \sin nx, \quad (n = 1, 2, \ldots)$$

είναι ανά δύο ορθογώνιες. Αν τις διαιρέσουμε όλες (εκτός από τη σταθερή συνάρτηση) με $\sqrt{2}$ τότε εκτός από ορθογώνιο σύστημα συναρτήσεων είναι και ορθοκανονικό.

Γενικά οι υπολογισμοί ολοκληρωμάτων με τις τριγωνομετρικές συναρτήσεις δεν είναι από τις πιο απολαυστικές ασχολίες. Είναι πολύ καλύτερο να τις μετατρέψουμε σε μιγαδικές εκθετικές συναρτήσεις και να *κάνουμε εκεί τις πράξεις μας μια και οι εκθετικές συναρτήσεις είναι* φτιαγμένες *για να πολλαπλασιάζονται. Για παράδειγμα, για να δείξετε την ορθογωνιότητα των* $\cos mx$ *και* $\cos nx$ *υπολογίστε το ολοκλήρωμα αφού πρώτα γράψετε*

$$\cos mx \cdot \cos nx = \frac{1}{4}(e^{imx} + e^{-imx})(e^{inx} + e^{-inx}).$$

Θυμίζουμε τον ορισμό της *γραμμικής ανεξαρτησίας* διανυσμάτων $v_1, \ldots, v_n$ σε ένα διανυσματικό χώρο V: θεωρούνται αυτά γραμμικώς ανεξάρτητα αν ισχύει η συνεπαγωγή

$$\sum_{k=1}^{n} c_k v_k = 0 \Longrightarrow c_1 = c_2 = \cdots = c_n = 0, \quad (\text{για κάθε επιλογή των συντελεστών } c_j \in \mathbb{C}).$$

Οι διανυσματικοί χώροι που μας απασχολούν σε αυτό το μάθημα είναι κατά κανόνα χώροι συναρτήσεων όπως ο $C([0, 2\pi])$ με τον οποίο ασχολούμαστε εδώ. Είναι πολύ βασικό ότι η ορθογωνιότητα συνεπάγεται τη γραμμική ανεξαρτησία. Αν οι μη μηδενικές $f_1, f_2, \ldots, f_n \in C([0, 2\pi])$ είναι ανά δύο ορθογώνιες τότε, υποθέτοντας ότι έχουμε ένα μηδενιζόμενο γραμμικό συνδυασμό τους

$$0 = \sum_{k=1}^{n} c_k f_k,$$

και παίρνοντας το εσωτερικό γινόμενο και τον δύο μελών με την f_1, έχουμε

$$0 = \sum_{k=1}^{n} c_k \langle f_k, f_1 \rangle = c_1 \langle f_1, f_1 \rangle = c_1 \|f_1\|_2^2$$

πράγμα που συνεπάγεται $c_1 = 0$ αφού $\|f_1\|_2^2 > 0$ αρκεί η f_1 να μην είναι η μηδενική συνάρτηση. Ομοίως αποδεικνύουμε ότι όλα τα c_j είναι μηδέν και άρα τα $f_1, \ldots, f_n$ είναι γραμμικώς ανεξάρτητα.

Πρόβλημα 3.8. *(Η 2-νόρμα ενός τριγωνομετρικού πολυωνύμου)*
Αν $p(x) = \sum_{k=-N}^{N} p_k e^{ikx}$ *δείξτε ότι*

$$☞ \qquad \|p\|_2^2 = \sum_{k=-N}^{N} |p_k|^2.$$

Χρησιμοποιείστε την ορθογωνιότητα των εκθετικών συναρτήσεων και το Πυθαγόρειο Θεώρημα.

Πρόβλημα 3.9. *Αποδείξτε ξανά τη μοναδικότητα των συντελεστών των τριγωνομετρικών πολυωνύμων χωρίς τον πίνακα Vandermonde (δείτε το Πρόβλημα 3.1). Αν* $p(x), q(x)$ *είναι δύο τριγωνομετρικά πολυώνυμα που ταυτίζονται* σε ολόκληρο *το διάστημα* $[0, 2\pi]$ *τότε έχουν τους ίδιους συντελεστές.*

Αποδείξτε ότι οι συντελεστές ενός τριγωνομετρικού πολυωνύμου δίδονται από τον τύπο

$$(27) \qquad ☞ \qquad p_k = \langle p(x), e^{ikx} \rangle = \frac{1}{2\pi} \int_0^{2\pi} p(x) e^{-ikx}\, dx, \quad (k \in \mathbb{Z}).$$

Πρόβλημα Ν 3.1. *Αποδείξαμε στο Πρόβλημα 3.9 ότι οποιοσδήποτε πεπερασμένος* $\mathbb{C}$*-γραμμικός συνδυασμός των εκθετικών συναρτήσεων* e^{inx}*, με* $n \in \mathbb{Z}$*, δε μπορεί να είναι η μηδενική συνάρτηση εκτός αν όλοι οι συντελεστές είναι* 0 *(γραμμική ανεξαρτησία των συναρτήσεων* e^{inx}*, για* $n \in \mathbb{Z}$*). Η μέθοδος* *που χρησιμοποιήσαμε είναι ότι δείξαμε πρώτα ότι οι συναρτήσεις αυτές είναι μεταξύ τους ορθογώνιες στο διάστημα* $[0, 2\pi]$*, ισχύει δηλαδή*

$$\langle e^{imx}, e^{inx} \rangle = \int e^{i(m-n)x}\, dx = 0$$

αν $m \neq n$*. Αυτό παύει να ισχύει αν οι συχνότητες δεν είναι ακέραια πολλαπλάσια του ίδιου αριθμού. Αν λοιπόν* $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n$ *είναι πραγματικοί αριθμοί τότε χρειαζόμαστε κάποια άλλη μέθοδο για να δείξουμε ότι οι εκθετικές συναρτήσεις*

$$e^{i\lambda_1 x}, e^{i\lambda_2 x}, \ldots, e^{i\lambda_n x}$$

είναι $\mathbb{C}$*-γραμμικώς ανεξάρτητες. Δείξτε το αυτό υποθέτοντας ότι για κάθε* $x \in \mathbb{R}$ *ισχύει*

$$f(x) = \sum_{j=1}^{n} c_j e^{i\lambda_j x} = 0$$

και παίρνοντας N*-οστές παραγώγους της* f *για* N *πολύ μεγάλο. Εξηγήστε γιατί δε μπορεί να μηδενίζεται ταυτοτικά η* $f^{(N)}(x)$*. Εξηγήστε επίσης γιατί η συνθήκη* $0 < \lambda_1$ *παραπάνω δε χρειάζεται και μπορούν τα* λ_j *να είναι οποιοιδήποτε διαφορετικοί* ***μιγαδικοί*** *αριθμοί.*

Παρατήρηση 3.3. Το δεξί μέλος της (27) μπορεί να εφαρμοστεί σε οποιαδήποτε συνεχή συνάρτηση $p(x)$ στο διάστημα $[0, 2\pi]$. Η ποσότητα αυτή ονομάζεται *συντελεστής Fourier* k *τάξης* της συνάρτησης $p(x)$.

Πρόβλημα 3.10. *Αν* $p(x) = \sum_{k=-N}^{N} p_k e^{ikx}$, $q(x) = \sum_{k=-N}^{N} q_k e^{ikx}$ *δείξτε ότι*

$$\langle p(x), q(x) \rangle = \sum_{k=-N}^{N} p_q \overline{q_k}.$$

Έστω $\mathcal{P}_N$ το σύνολο όλων των τριγωνομετρικών πολυωνύμων βαθμού $\le N$

$$\mathcal{P}_N = \left\{ p(x) = \sum_{k=-N}^{N} p_k e^{ikx} : \ \ p_k \in \mathbb{C} \right\}.$$

Αφού το άθροισμα δυο τέτοιων πολυωνύμων παραμένει στοιχείο του $\mathcal{P}_N$ και γινόμενο ενός μιγαδικού αριθμού με στοιχείο του $\mathcal{P}_N$ παραμένει στοιχείο του $\mathcal{P}_N$ προκύπτει ότι το σύνολο αυτό είναι ένας μιγαδικός διανυσματικός χώρος. Μάλιστα με την αντιστοίχιση

$$p(x) \to (p_{-N}, p_{-N+1}, \ldots, p_{N-1}, p_N)$$

που είναι καλώς ορισμένη (μοναδικότητα των συντελεστών πολυωνύμου), γραμμική και αντιστρέψιμη είναι φανερό πως ο χώρος $\mathcal{P}_N$ είναι ισομορφικός, ως γραμμικός χώρος, με το χώρο $\mathbb{C}^{2N+1}$. Μια βάση του χώρου $\mathcal{P}_N$ αποτελείται από τα $2N+1$ τριγωνομετρικά πολυώνυμα

$$e^{-iNx}, e^{-i(N-1)x}, \ldots, e^{-ix}, 1, e^{ix}, \ldots, e^{iNx}.$$

Αυτή μάλιστα η βάση έχουμε αποδείξει ότι είναι και ορθοκανονική, είναι δηλαδή τα στοιχεία της ανά δύο ορθογώνια και κάθε στοιχείο της έχει 2-νόρμα ίση με 1.

Υπάρχει μια ακόμη ορθοκανονική βάση του $\mathcal{P}_N$ η οποία είναι χρήσιμη στις εφαρμογές, ειδικά όταν πρόκειται για τριγωνομετρικά πολυώνυμα που παίρνουν πραγματικές τιμές.

Θεώρημα 3.1. *Οι συναρτήσεις*

$$1, \sqrt{2}\cos x, \sqrt{2}\cos 2x, \ldots, \sqrt{2}\cos Nx, \sqrt{2}\sin x, \sqrt{2}\sin 2x, \ldots, \sqrt{2}\sin Nx \tag{28}$$

είναι μια ορθοκανονική βάση του $\mathcal{P}_N$*. Επιπλέον, αν μια συνάρτηση* $p(x) \in \mathcal{P}_N$ *παίρνει πραγματικές τιμές τότε οι συντελεστές της ως προς αυτή τη βάση είναι πραγματικοί.*

Από το Πρόβλημα 3.7 έχουμε την ορθογωνιότητα των συναρτήσεων ανά δύο. Το ότι κάθε μια από αυτές έχει 2-νόρμα ίση με το 1 είναι ένας απλός υπολογισμός που στην περίπτωση της σταθερής συνάρτησης είναι προφανής και σε όλες τις άλλες περιπτώσεις ανάγεται στον τύπο

$$\int_0^{2\pi} \cos^2 x \, dx = \pi.$$

Για να αποδείξουμε το δεύτερο μέρος του Θεωρήματος θα υπολογίσουμε ακριβώς τους συντελεστές του πολυωνύμου

$$p(x) = \sum_{k=-N}^{N} p_k e^{ikx}$$

ως πρός τη βάση (28). Έστω λοιπόν ότι

$$p(x) = a_0 + \sum_{k=1}^{N} (a_k \cos kx + b_k \sin kx) \tag{29}$$

$$= a_0 + \sum_{k=1}^{N} \left(a_k \frac{e^{ikx} + e^{-ikx}}{2} + b_k \frac{a^{ikx} - e^{-ikx}}{2i} \right). \tag{30}$$

Εξισώνοντας (από τη μοναδικότητα των συντελεστών των τριγωνομετρικών πολυωνύμων) τους συντελεστές των δύο μελών, και παρατηρώντας ότι οι συναρτήσεις e^{ikx} και e^{-ikx} εμφανίζονται μόνο στον k προσθετέο του αθροίσματος (30), παίρνουμε ότι $p_0 = a_0$ και για $k = 1, 2, \ldots, N$ ότι

$$p_k e^{ikx} + p_{-k} e^{-ikx} = a_k \frac{e^{ikx} + e^{-ikx}}{2} + b_k \frac{a^{ikx} - e^{-ikx}}{2i}.$$

Κάνοντας πράξεις αυτό γράφεται

$$(p_k - a_k/2 - b_k/(2i))e^{ikx} + (p_{-k} - a_k/2 + b_k/(2i))e^{-ikx} = 0.$$

Από τη μοναδικότητα παίρνουμε τις δύο εξισώσεις

$$p_k = a_k/2 + b_k/(2i), \quad p_{-k} = a_k/2 - b_k/(2i). \tag{31}$$

Λύνοντας ως προς τις ποσότητες a_k, b_k παίρνουμε

$$a_k = p_k + p_{-k}, \quad b_k = i(p_k - p_{-k}). \tag{32}$$

Οι τύποι (31) και (32), μαζί με τον $a_0 = p_0$, μας λένε το πώς γράφουμε μια συνάρτηση $p(x)$ στην τριγωνομετρική βάση (28) αν την έχουμε γραμμένη ως προς την εκθετική βάση και το αντίστροφο.

Τέλος, ας υποθέσουμε ότι η συνάρτηση $p(x)$ παίρνει πραγματικές τιμές. Έχουμε τότε

$$\begin{aligned}
a_0 &= p(0) = \langle p(x), 1 \rangle \\
a_k &= p_k + p_{-k} = \langle (p(x), e^{ikx} + e^{-ikx} \rangle = 2\langle p(x), \cos kx \rangle \\
b_k &= i(p_k - p_{-k}) = i\langle p(x), e^{ikx} - e^{-ikx} \rangle = -2\langle p(x), \sin x \rangle.
\end{aligned}$$

Όλα τα εσωτερικά γινόμενα που εμφανίζονται παραπάνω έχουν και τους δύο παράγοντες πραγματικές συναρτήσεις, άρα είναι πραγματικά.

3.3. **Άρτιες και περιττές συναρτήσεις.** Μια συνάρτηση $f : \mathbb{R} \to \mathbb{C}$ λέγεται *άρτια* αν $f(-x) = f(x)$ για κάθε $x \in \mathbb{R}$ (οποιοδήποτε πεδίο ορισμού D μπορούμε να έχουμε εδώ το οποίο είναι συμμετρικό ως προς το 0, ισχύει δηλ. $x \in D \iff -x \in D$). Η $f : \mathbb{R} \to \mathbb{C}$ λέγεται *περιττή* αν ισχύει $f(-x) = -f(x)$ για κάθε $x \in \mathbb{R}$.

Είναι φανερό ότι το να είναι μια συνάρτηση άρτια ή περιττή είναι μια σχετικά σπάνια ιδιότητα; οι «πιο πολλές» συναρτήσεις δεν είναι ούτε το ένα ούτε το άλλο. Παρ' όλ' αυτά ισχύει το παρακάτω θεώρημα το οποίο κάποιες φορές μας επιτρέπει να περάσουμε ιδιότητες των αρτίων και των περιττών συναρτήσεων σε γενικές συναρτήσεις.

Θεώρημα 3.2. *Αν $f : \mathbb{R} \to \mathbb{C}$ τότε υπάρχει μια άρτια συνάρτηση $f_e : \mathbb{R} \to \mathbb{C}$ και μια περιττή συνάρτηση $f_o : \mathbb{R} \to \mathbb{C}$ τ.ώ. $f(x) = f_e(x) + f_o(x)$ για κάθε $x \in \mathbb{R}$. Μάλιστα αυτή η διάσπαση της f σε άθροισμα περιττής και άρτιας συνάρτησης είναι μοναδική.*

Η απόδειξη είναι εξαιρετικά απλή. Παίρνουμε
$$f_e(x) = \frac{f(x) + f(-x)}{2}, \quad f_o(x) = \frac{f(x) - f(-x)}{2}.$$
Είναι φανερό ότι $f = f_e + f_o$ και ότι η f_e είναι άρτια και η f_o περιττή.

Για να δείξουμε τη μοναδικότητα υποθέτουμε ότι υπάρχει και δεύτερη διάσπαση της f σε άθροισμα άρτιας και περιττής συνάρτησης
$$f = \widetilde{f}_e + \widetilde{f}_o.$$
Από την ισότητα $f_e + f_o = \widetilde{f}_e + \widetilde{f}_o$ παίρνουμε
$$f_e - \widetilde{f}_e = \widetilde{f}_o - f_o.$$
Το αριστερό μέλος είναι άρτια συνάρτηση και το δεξί περιττή (ως γραμμικός συνδυασμός αντιστοίχων συναρτήσεων). Όμως η μόνη συνάρτηση που υπάρχει που είναι ταυτόχρονα άρτια και περιττή είναι η μηδενική συνάρτηση, άρα $f_e = \widetilde{f}_e$ και $f_o = \widetilde{f}_o$.

Οι συναρτήσεις f_e και f_o ονομάζονται *άρτιο και περιττό μέρος της f*.

Πρόβλημα 3.11. *Αν $p(x) = a_0 + \sum_{k=1}^{N} (a_k \cos kx + b_k \sin kx)$ είναι ένα τριγωνομετρικό πολυώνυμο γραμμένο στην τριγωνομετρική του μορφή τότε βρείτε τις συναρτήσεις $p_e(x)$ και $p_o(x)$ γραμμένες επίσης στην τριγωνομετρική τους μορφή. Ίδιο ερώτημα αν το τριγ. πολυώνυμο $p(x)$ δίδεται στην εκθετική του μορφή*
$$p(x) = \sum_{k=-N}^{N} p_k e^{ikx}.$$

3.4. **Προσέγγιση από τριγωνομετρικά πολυώνυμα.** Αντίστοιχα με την προσέγγιση μιας συνεχούς συνάρτησης από αλγεβρικά (συνηθισμένα) πολυώνυμα έχουμε και προσέγγιση από τριγωνομετρικά πολυώνυμα. Ας είναι $f \in C([0, 2\pi])$. Μια προφανής αναγκαία συνθήκη για να μπορεί η f να προσεγγισθεί ομοιόμορφα στο $[0, 2\pi]$ από τριγωνομετρικά πολυώνυμα είναι η
$$f(0) = f(2\pi).$$
Αυτό ισχύει γιατί η συνθήκη αυτή ισχύει για κάθε τριγωνομετρικό πολυώνυμο (λόγω της 2π-περιοδικότητας) και άρα ισχύει και για κάθε συνάρτηση που είναι όριο (έστω και κατά σημείο όριο) τριγ. πολυωνύμων. Η συνθήκη αυτή πέρα από αναγκαία είναι και ικανή:

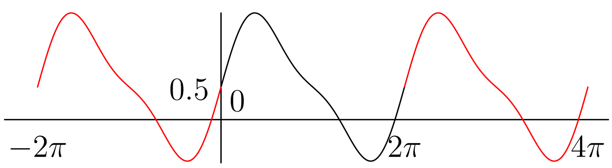
Θεώρημα 3.3. *Αν $f \in C([-\pi, \pi])$ έχει*

$$f(-\pi) = f(\pi)$$

τότε για κάθε $\epsilon > 0$ υπάρχει τριγ. πολυώνυμο $p(x)$ τ.ώ. $|f(x) - p(x)| \leq \epsilon$ για $x \in [-\pi, \pi]$.

Παρατήρηση 3.4. Το Θεώρημα 3.3 ισχύει αν στη θέση του $[-\pi, \pi]$ βάλουμε οποιοδήποτε άλλο διάστημα της μορφής $[s, s + 2\pi]$ και απαιτήσουμε η f να είναι συνεχής σε αυτό το κλειστό διάστημα και να έχει τις ίδιες τιμές στα δύο άκρα. Όλες οι αποδείξεις είναι ίδιες.

Υπάρχει και ένας διαφορετικός τρόπος να εκφράσουμε αυτή την υπόθεση για την f: η f πρέπει να είναι συνεχής σε όλο το $\mathbb{R}$ και 2π-περιοδική. Πράγματι, αν η f είναι συνεχής και 2π-περιοδική τότε ο περιορισμός της σε οποιοδήποτε διάστημα μήκους 2π, έστω στο $[s, s+2\pi]$, είναι συνεχής και η f παίρνει την ίδια τιμή στα δύο άκρα του διαστήματος. Αντίστροφα, αν η f είναι ορισμένη μόνο στο $[s, s+2\pi]$ και είναι συνεχής εκεί και με την ίδια τιμή στα δύο άκρα του διαστήματος τότε η f μπορεί να επεκταθεί σε όλο το $\mathbb{R}$ ως μια συνεχής και 2π-περιοδική συνάρτηση: επεκτείνουμε πρώτα στο διάστημα $[s+2\pi, s+4\pi]$ χρησιμοποιώντας τη σχέση $f(x) = f(x + 2\pi)$ που θέλουμε να ισχύει για κάθε x και παρατηρούμε ότι η συνάρτησή μας είναι συνεχής στο καινούργιο της πεδίο ορισμού. Το μόνο σημείο στο οποίο υπάρχει κάποια αμφιβολία για την συνέχεια της f είναι εκεί που κολλάνε τα δύο διαστήματα, το σημείο $s + 2\pi$ δηλαδή, αλλά εκεί η συνάρτησή μας είναι συνεχής και από τις δύο μεριές ως συνέπεια της υπόθεσής μας για ίδιες τιμές της f στα δύο άκρα του αρχικού διαστήματος. Με όμοιο τρόπο επεκτείνουμε την f σε διαστήματα μήκους 2π δεξιά κι αριστερά του πεδίου ορισμού της κι έτσι ορίζεται τελικά η f σε όλο το $\mathbb{R}$ με τις επιθυμητές ιδιότητες. Δείτε το Σχήμα 4 όπου φαίνεται η διαδικασία αυτή της *περιοδικοποίησης* για $s = 0$, ξεκινώντας δηλ. από το διάστημα $[0, 2\pi]$.



Σχήμα 4: Η διαδικασία της επέκτασης μιας συνάρτησης σε περιοδική (*περιοδικοποίηση*). Το γράφημα στα διαστήματα $[-2\pi, 0]$ και $[2\pi, 4\pi]$ είναι μεταφορές του γραφήματος στο διάστημα $[0, 2\pi]$

Με άλλα λόγια ορίζουμε $f(x)$ να είναι η τιμή $f(x - 2k\pi)$ όπου $k \in \mathbb{Z}$ είναι ο μοναδικός ακέραιος τέτοιος ώστε να ισχύει

$$x - 2k\pi \in [s, s + 2\pi).$$

Πρόβλημα 3.12. *Αν f_n, f είναι T-περιοδικές συναρτήσεις και $f_n \to f$ ομοιόμορφα στο διάστημα $[s, s+T]$ τότε $f_n \to f$ ομοιόμορφα σε ολόκληρο το $\mathbb{R}$.*

Πρόβλημα 3.13. *Αν η $f : \mathbb{R} \to \mathbb{C}$ είναι συνεχής και T-περιοδική τότε είναι ομοιόμορφα συνεχής σε ολόκληρο το $\mathbb{R}$.*

Από δω και στο εξής λοιπόν το να μιλάμε για μια συνάρτηση που είναι συνεχής σε κλειστό διάστημα μήκους T και έχει τις ίδιες τιμές στα δύο άκρα του διαστήματος είναι το ίδιο με το να μιλάμε για μια συνάρτηση που είναι ορισμένη και συνεχής σε ολόκληρο το $\mathbb{R}$ και είναι επίσης T-περιοδική.

Ας αποδείξουμε τώρα το Θ. 3.3. Θα θεωρούμε τη συνάρτηση f ορισμένη και συνεχή σε ολόκληρο το $\mathbb{R}$ μέσω 2π-περιοδικότητας όπως αναφέραμε στην Παρατήρηση 3.4.

Βήμα 1. Αν η f είναι άρτια τότε προσεγγίζεται ομοιόμορφα από άρτια τριγ. πολυώνυμα.

Αφού η f είναι άρτια αρκεί να την προσεγγίσουμε ομοιόμορφα από άρτιο τριγωνομετρικά πολυώνυμα στο διάστημα $[0, \pi]$ μόνο. Στο διάστημα αυτό η συνάρτηση $\cos x$ είναι μια συνεχής και 1 προς 1 απεικόνιση του διαστήματος $[0, \pi]$ στο διάστημα $[-1, 1]$. Έστω arccos («τόξο συνημιτόνου», που καμιά φορά το συμβολίζουμε και με $\cos^{-1}$) η αντίστροφή της, επίσης συνεχής, συνάρτηση που απεικονίζει το$[-1, 1]$ στο $[0, \pi]$:

$$y = \cos x \Leftrightarrow x = \arccos y \quad (x \in [0, \pi],\ y \in [-1, 1]).$$

Άρα η $g(y) = f(\arccos y)$ είναι μια συνεχής συνάρτηση ορισμένη πάνω στο $[-1, 1]$ και από το Θ. Weierstrass για αλγεβρικά πολυώνυμα (Θεώρημα 2.1) για κάθε $\epsilon > 0$ υπάρχει πολυώνυμο $p(y)$ τέτοιο ώστε

$$|g(y) - p(y)| \le \epsilon$$

για κάθε $y \in [-1, 1]$. Όμως αυτή η συνθήκη, μετά την αλλαγή μεταβλητής $y = \cos x$, δε λέει τίποτε άλλο παρά το ότι

$$|f(x) - p(\cos x)| \le \epsilon$$

για κάθε $x \in [0, \pi]$. Τέλος δε μένει παρά να παρατηρήσουμε ότι η συνάρτηση $p(\cos x)$ είναι άρτιο τριγωνομετρικό πολυώνυμο. Αυτό είναι συνέπεια του ότι η συνάρτηση $\cos x$ είναι άρτιο τριγωνομετρικό πολυώνυμο και του ότι τα τριγωνομετρικά πολυώνυμα είναι κλειστά (όπως και τα αλγεβρικά πολυώνυμα) στις πράξεις της πρόσθεσης και πολλαπλασιασμού συναρτήσεων.

Βήμα 2: Η συνάρτηση $f(x)\sin^2 x$ προσεγγίζεται ομοιόμορφα από τριγ. πολυώνυμα.

Έστω $\epsilon > 0$. Οι συναρτήσεις $f_1(x) = f(x) + f(-x)$ και $f_2(x) = (f(x) - f(-x))\sin x$ είναι άρτιες, 2π-περιοδικές και συνεχείς και άρα από το Βήμα 1 υπάρχουν άρτια τριγ. πολυώνυμα $T_1(x)$ και $T_2(x)$ τ.ώ. $\|f_1 - T_1\|_\infty \le \epsilon$ και $\|f_2 - T_2\|_\infty \le \epsilon$. Αυτό μπορούμε να το εκφράσουμε και ως εξής

$$f_1(x) = T_1(x) + E_1(x), \quad (|E_1(x)| \le \epsilon)$$

και

$$f_2(x) = T_2(x) + E_2(x), \quad (|E_2(x)| \le \epsilon).$$

Άρα

$$f(x)\sin^2(x) = f_1(x)\sin^2 x + f_2(x)\sin x = T_1(x)\sin^2 x + T_2(x)\sin x + E(x)$$

όπου $E(x) = E_1(x)\sin^2 x + E_2(x)\sin x$, οπότε $|E(x)| \le 2\epsilon$.

Βήμα 3: Η συνάρτηση $f(x)\cos^2 x$ προσεγγίζεται ομοιόμορφα από τριγ. πολυώνυμα.

Αυτό είναι συνέπεια της εφαρμογής του Βήματος 2 στη συνάρτηση $F(x) = f(x - \frac{\pi}{2})$ που είναι φυσικά 2π-περιοδική και συνεχής. Από το Βήμα 2 έχουμε ότι η $F(x)\sin^2 x$ προσεγγίζεται ομοιόμορφα από τριγ. πολυώνυμα $p(x)$ το οποίο σημαίνει ότι η συνάρτηση

$$f(x)\cos^2 x = F(x + \frac{\pi}{2})\sin^2(x + \frac{\pi}{2})$$

προσεγγίζεται ομοιόμορφα από τα πολυώνυμα $p(x + \frac{\pi}{2})$.

Για να τελειώσουμε την απόδειξη του Θεωρήματος 3.3 παρατηρούμε απλά ότι

$$f(x) = f(x)\sin^2 x + f(x)\cos^2 x$$

και προσεγγίζουμε την f από το άθροισμα των δύο τριγωνομετρικών πολυωνύμων που προσεγγίζουν τις δύο συναρτήσεις του δεξιού μέλους αντίστοιχα.

Παρατήρηση 3.5. Έχουμε έτσι αποδείξει το Θεώρημα του Weierstrass για προσέγγιση από τριγωνομετρικά πολυώνυμα (Θεώρημα 3.3) χρησιμοποιώντας το Θεώρημα του Weierstrass για αλγεβρικά πολυώνυμα (Θεώρημα 2.1). Είναι ενδιαφέρον ότι και ο αντίστροφος συλλογισμός είναι δυνατός. Αν δηλ. υποθέσει κανείς το Θεώρημα 3.3 ως γνωστό τότε μπορεί να αποδείξει, χρησιμοποιώντας το, το Θεωρημα 2.1. Αυτό είναι ενδιαφέρον μια και υπάρχουν και άλλες αποδείξεις του Θεωρήματος 3.3, που δε χρησιμοποιούν το Θεώρημα 2.1, δίνοντάς μας έτσι νέες αποδείξεις του Θεωρήματος 2.1. Για παράδειγμα, ένα από τα βασικά θεωρήματα της Αρμονικής Ανάλυσης είναι το Θεώρημα του Fejér, που λέει ότι αν έχουμε μια συνεχή και 2π-περιοδική συνάρτηση f τότε αυτή είναι το ομοιόμορφο όριο των τριγωνομετρικών πολυωνύμων, που ονομάζονται Cesáro μέσοι της f:

$$\sigma_N(f)(x) = \sum_{j=-N}^{N} \widehat{f}(j) e^{ijx}.$$

Οι αριθμοί $\widehat{f}(j)$ ονομάζονται *συντελεστές Fourier* (τους είδαμε και στην Παρατήρηση 3.3) της f και δίνονται από τον τύπο

$$\widehat{f}(j) = \frac{1}{2\pi}\int_0^{2\pi} f(x) e^{-ijx}\,dx, \quad (j \in \mathbb{Z}). \tag{33}$$

Ας δούμε λοιπόν τώρα πώς αποδεικνύει κανείς το Θεώρημα 2.1 υποθέτοντας ως γνωστό το Θεώρημα 3.3.

Χωρίς βλάβη της γενικότητας μπορούμε να πάρουμε το διάστημα αναφοράς να είναι το $[-1, 1]$, και ας είναι $f \in C([-1, 1])$. Η συνάρτηση $g : \mathbb{R} \to \mathbb{C}$ ορίζεται ως

$$g(x) = f(\cos x)$$

και είναι συνεχής, 2π-περιοδική και άρτια. Για να δείξουμε ότι η f προσεγγίζεται ομοιόμορφα από ένα αλγεβρικό πολυώνυμο $q(x)$ στο διάστημα $[-1, 1]$ αρκεί να δείξουμε ότι η συνάρτηση $g(x)$ προσεγγίζεται ομοιόμορφα από το πολυώνυμο $q(\cos x)$ (το οποίο είναι τριγωνομετρικό πολυώνυμο) στο διάστημα $[0, \pi]$.

Έστω $\epsilon > 0$. Ας είναι $p(x)$ ένα τριγωνομετρικό πολυώνυμο τέτοιο ώστε $g(x) = p(x) + E(x)$, με $|E(x)| \le \epsilon$ (εδώ χρησιμοποιούμε το Θεώρημα 3.3). Έχουμε

$$g(x) = \frac{g(x) + g(-x)}{2} = \frac{p(x) + p(-x)}{2} + \frac{E(x) + E(-x)}{2}.$$

Προφανώς $\left|\frac{E(x)+E(-x)}{2}\right| \le \epsilon$ και η συνάρτηση $q(x) = \frac{p(x)+p(-x)}{2}$ είναι ένα άρτιο τριγωνομετρικό πολυώνυμο οπότε

$$q(x) = \sum_{k=0}^{N} q_k \cos kx$$

για κάποιο φυσικό αριθμό N και κάποια $q_k \in \mathbb{C}$ (αυτό φαίνεται πολύ εύκολα αν γράψει κανείς το τριγ. πολυώνυμο $q(x)$ σε τριγωνομετρική μορφή, κάνει το ίδιο για το $q(-x)$ και χρησιμοποιήσει το θεώρημα μοναδικότητας).

Για να τελειώσει η απόδειξη πρέπει να δείξουμε ότι η συνάρτηση $q(x)$ είναι αλγεβρικό πολυώνυμο του $\cos x$, για το οποίο προφανώς αρκεί να δείξουμε για κάθε $k = 0, 1, 2, \ldots$ ότι

(34) η συνάρτηση $\cos kx$ είναι αλγεβρικό πολυώνυμο του $\cos x$.

Όμως

$$
\begin{aligned}
\cos kx &= \mathrm{Re}\left[e^{ikx}\right] \\
&= \mathrm{Re}\left[(\cos x + i\sin x)^k\right] \\
&= \mathrm{Re}\left[\sum_{j=0}^{k}\binom{k}{j}(\cos x)^{k-j} i^j (\sin x)^j\right] && \text{(από το διωνυμικό θεώρημα)} \\
&= \sum_{\substack{j=0 \\ j \text{ άρτιο}}}^{k}\binom{k}{j}(\cos x)^{k-j}(-1)^{j/2}(\sin^2 x)^{j/2} && \text{(αφού } i^j = \pm i \text{ για } j \text{ περιττό)} \\
&= \sum_{\substack{j=0 \\ j \text{ άρτιο}}}^{k}\binom{k}{j}(\cos x)^{k-j}(-1)^{j/2}(1-\cos^2 x)^{j/2} && \text{(αφού } \sin^2 x = 1-\cos^2 x\text{)}
\end{aligned}
$$

το οποιο προφανώς είναι ένα αλγεβρικό πολυώνυμο της ποσότητας $\cos x$. Η απόδειξή μας είναι πλήρης.

4. Ιδιότητες της βέλτιστης προσέγγισης

Επανερχόμαστε τώρα στο πρόβλημα της προσέγγισης από αλγεβρικά πολυώνυμα. Έστω $f \in C([a,b])$ και έστω $p^* \in \mathcal{P}_N$ *μια* βέλτιστη προσέγγιση της f από το σύνολο $\mathcal{P}_N$ (αλγεβρικά πολυώνυμα βαθμού μεχρι N) στη νόρμα $\|\cdot\|_\infty$. Το πολυώνυμο p^* δηλ. έχει την ιδιότητα ότι για κάθε άλλο πολυώνυμο $p \in \mathcal{P}_N$ ισχύει

$$\|f - p^*\|_\infty \le \|f - p\|_\infty. \tag{35}$$

Η ύπαρξη του πολυωνύμου αυτού δεν είναι προφανής. Είναι όμως άμεση σχεδόν συνέπεια του γεγονότος ότι ο γραμμικός χώρος $\mathcal{P}_N$ έχει πεπερασμένη διάσταση (συγκεκριμένα $\dim \mathcal{P}_N = N+1$), όπως αποδείξαμε στις δύο πρώτες εισαγωγικές διαλέξεις (μπορείτε να το βρείτε στις σημειώσεις του Carothers, Κεφ. 1). Θα δούμε κάποιες ιδιότητες που έχει αυτό το πολυώνυμο p^* και οι οποίες, μεταξύ άλλων, θα μας επιτρέψουν να αποδείξουμε και τη μοναδικότητά του, ιδιότητα που δεν προκύπτει από κάποιο γενικό επιχείρημα όπως αυτό της πεπερασμένης διάστασης.

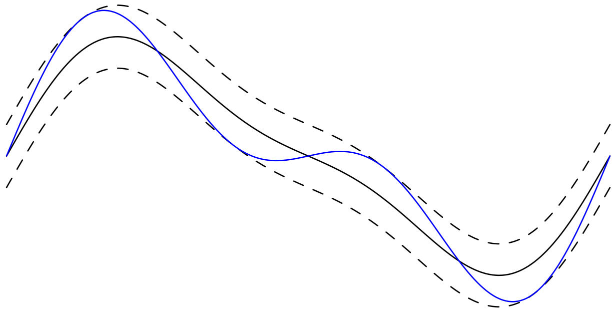
Περιοριζόμαστε από δω και πέρα σε πραγματικές συναρτήσεις $f[a,b] \to \mathbb{R}$ μια και το συμπέρασμα στο οποίο θέλουμε να καταλήξουμε δεν έχει νόημα για μιγαδικές συναρτήσεις. Είναι προφανές σε αυτή την περίπτωση ότι και η p^* είναι επίσης πραγματικό πολυώνυμο (γιατί;).

Αν $p : [a,b] \to \mathbb{R}$ είναι οποιαδήποτε συνεχής συνάρτηση τότε από τη συνέχεια της $|f-p|$ προκύπτει ότι αυτή «πιάνει» το supremum της για κάποιο $x \in [a,b]$. Ισχύει δηλ. για κάποιο x

$$f(x) - p(x) = \pm\|f-p\|_\infty.$$

Αν ζωγραφίσουμε γύρω από το γράφημα της f μια ζώνη με πλάτος $\rho = \|f-p\|_\infty$ από κάθε μεριά (βλέπε Σχήμα 5) τότε αυτό που μόλις είπαμε σημαίνει ότι το γράφημα της $p(x)$, το οποίο βρίσκεται εξ ολοκλήρου μέσα σε αυτή τη ζώνη (αφού $|f(x)-p(x)| \le \rho$ για κάθε x) σίγουρα ακουμπάει το άνω ή το κάτω σύνορο αυτής της ζώνης.

Στο Θεώρημα 4.1 που ακολουθεί δείχνουμε ότι ακουμπάει και πάνω και κάτω.



Σχήμα 5: Η βέλτιστη πολυωνυμική προσέγγιση (μπλε) της συνάρτησης πρέπει να ακουμπάει και το σύνορο της ρ-ζώνης (διακεκομένη γραμμή) και προς τα πάνω και προς τα κάτω. Εδώ $\rho = \|f - p^*\|_\infty$.

Θεώρημα 4.1. *Έστω $f : [a, b] \to \mathbb{R}$ συνεχής. Αν $p^* \in \mathcal{P}_N$ είναι ένα πολυώνυμο που ελαχιστοποιεί την ποσότητα $\|f - p\|_\infty$, $p \in \mathcal{P}_N$, τότε υπάρχουν σημεία $x, y \in [a, b]$ τ.ώ.*

$$f(x) - p^*(x) = p^*(y) - f(y) = \|f - p^*\|_\infty.$$

Η απόδειξη του Θεωρήματος 4.1 είναι πολύ εύκολη αν σκεφτεί κανείς λίγο γεωμετρικά. Τι θα γινόταν αν, π.χ., η μπλε γραμμή στο Σχήμα 5 ακουμπούσε μόνο στο κάτω σύνρο της ζώνης (κάτω διακεκομένη γραμμή); Πολύ απλά θα μπορούσαμε να μετατοπίσουμε λίγο προς τα πάνω τη μπλε γραμμή ώστε να παραμείνει εντός της ζώνης και να έχουμε τώρα μια καλύτερη προσέγγιση από πριν, πράγμα αδύνατο μια και υποθέσαμε ότι η p^* είναι βέλτιστη προσέγγιση.

Ας υποθέσουμε λοιπόν ότι για κάθε $x \in [a, b]$ ισχύει $p^*(x) - f(x) < \rho = \|f - p^*\|_\infty$ (δεν ακουμπάει επάνω). Αφού η συνεχής συνάρτηση $p^*(x) - f(x)$ είναι $< \rho$ παντού στο $[a, b]$ έπεται ότι υπάρχει κάποιος αριθμός $0 < \rho' < \rho$ ώστε να ισχύει

$$p^*(x) - f(x) \le \rho', \quad (x \in [a, b]).$$

(Αυτό είναι απλά συνέπεια του ότι κάθε συνεχής συνάρτηση σε κλειστό διάστημα «πιάνει» τα ακρότατα της σε κάποια σημεία του διαστηματος.) Έστω $\delta = (\rho - \rho')/2 > 0$ και ορίζουμε τη συνάρτηση $p_1(x) = p^*(x) + \delta$, η οποία επίσης είναι στο $\mathcal{P}_N$ αφού επηρεασαμε μόνο τη σταθερά του πολυωνύμου.

Ισχύει τώρα

$$\begin{aligned} p_1(x) - f(x) &= p^*(x) - f(x) + \delta \le \rho' + \delta < \rho \\ f(x) - p_1(x) &= f(x) - p^*(x) - \delta \le \rho - \delta = \rho' + \delta < \rho, \end{aligned}$$

ή, με άλλα λόγια, $|f(x) - p_1(x)| \le \rho' + \delta < \rho$ για κάθε $x \in [a, b]$. Αυτό αντιφάσκει όμως με τη σχέση (35), άρα κάπου μέσα στο διάστημα $[a, b]$ η συνάρτηση $p^*(x) - f(x)$ ισούται με ρ. Με όμοιο τρόπο αποδεικνύουμε ότι κάπου μέσα στο διάστημα η $f(x) - p^*(x)$ ισχούται με ρ, και η απόδειξη είναι πλήρης.

Πρόβλημα 4.1. *Αν $f : [a, b] \to \mathbb{R}$ δείξτε ότι η βέλτιστη προσέγγιση από σταθερά (η βέλτιστη προσέγγιση δηλ. από το χώρο $\mathcal{P}_0$) στην ∞-νόρμα είναι η σταθερά* $(\max f + \min f)/2$.

Το Θεώρημα 4.1 μπορεί να ισχυροποιηθεί πάρα πολύ, σε σημείο που να αποτελεί *χαρακτηρισμό* της βέλτιστης πολυωνυμικής προσέγγισης. Δηλ. η ιδιότητα που εμφανίζεται στο επόμενο Θεώρημα 4.2 όχι μόνο είναι αναγκαία για να είναι το πολυώνυμο p^* μια βέλτιστη προσέγγιση αλλά είναι και ικανή.

Ορισμός 4.1. Αν $F : [a, b] \to \mathbb{R}$ είναι μια συνάρτηση που δεν είναι ταυτοτικά μηδενική (ώστε $\|F\|_\infty > 0$) και $y_1 < y_2 < \cdots < y_k$ είναι κάποια σημεία του $[a, b]$ λέμε ότι αυτά είναι *εναλλασόμενα για την F* αν

(α) $|F(y_j)| = \|F\|_\infty$ για κάθε $j = 1, 2, \ldots, k$, και
(β) Αν τα πρόσημα των $F(y_j)$ εναλλάσσονται, ισχύει δηλ.

$$F(y_j) \cdot F(y_{j+1}) < 0, \quad j = 1, 2, \ldots, N-1.$$

Θεώρημα 4.2. *Έστω $f : [a,b] \to \mathbb{R}$ και N φυσικός αριθμός. Τότε υπάρχει ακριβώς μια βέλτιστη προσέγγιση $p^* \in \mathcal{P}_N$ της f. Ένα πολυώνυμο $p \in \mathcal{P}_N$ είναι το p^* αν και μόνο αν υπάρχει ένα σύνολο τουλάχιστον $N+2$ σημείων του $[a,b]$ που να είναι εναλασσόμενα για τη συνάρτηση $f-p$.*

Πρόβλημα 4.2. *Αποδείξτε ότι η βέλτιστη γραμμική (δηλ. πολυωνυμική βαθμού μέχρι 1) προσέγγιση της συνάρτησης $f(x) = x^2$ στο διάστημα $[0,1]$ για την ∞-νόρμα είναι η συνάρτηση $p(x) = x - \frac{1}{8}$.*

Εφαρμόστε το Θεώρημα 4.2 με $N = 1$.

Πρόβλημα 4.3. *Βρείτε τη βέλτιστη προσέγγιση από γραμμικά πολυώνυμα (από πολυώνυμα δηλ. στο $\mathcal{P}_1$) της συνάρτησης $f(x) = x^2$ στο διάστημα $[0,1]$ αλλά τώρα στην L^2 νόρμα.*

Δε χρειάζεται να χρησιμοποιήσετε εδώ κάποιο βαρύ θεώρημα. Πρέπει να βρείτε τα $a, b \in \mathbb{R}$ ώστε η ποσότητα $\left\|x^2 - ax - b\right\|_2^2$ να ελαχιστοποιείται. Η ποσότητα αυτή είναι μια συνάρτηση των $a, b \in \mathbb{R}$ την οποία μπορείτε να υπολογίσετε εύκολα. Οι μερικές της παράγωγοι ως προς a, b πρέπει να μηδενίζονται στο σημείο ελαχιστοποίησης.

4.1. **Απόδειξη του Θεωρήματος 4.2.** Θα αποδείξουμε τώρα το Θεώρημα 4.2. Ας ξεκινήσουμε με το να αποδείξουμε ότι αν το p είναι βέλτιστη προσέγγιση τότε υπάρχουν τουλάχιστον $N+2$ εναλλασσόμενα σημεία για τη συνάρτηση $\phi = f - p$. Ας γράφουμε

$$E = \|\phi\|_\infty > 0.$$

(Η μόνη περίπτωση να είναι $E = 0$ είναι όταν $f = p$ όταν δηλ. $f \in \mathcal{P}_N$ οπότε και η f είναι η ίδια η βέλτιστη προσέγγιση του εαυτού της και δεν έχουμε τίποτα να αποδείξουμε.)

Αφού η ϕ είναι συνεχής στο $[a,b]$ είναι και ομοιόμορφα συνεχής. Ας είναι λοιπόν $\Delta > 0$ τ.ώ.

$$|x - y| \le \Delta \implies |\phi(x) - \phi(y)| \le E/10. \tag{36}$$

Χωρίζουμε το διάστημα $[a,b]$ σε ίσα διαστήματα μήκους $\le \Delta$ και ας είναι

$$a = t_0 < t_1 < t_2 < \cdots < t_{k-1} < t_k = b$$

τα διαχωριστικά σημεία. Ένα σημείο $x \in [a,b]$ θα το ονομάζουμε $(+)$-σημείο αν $\phi(x) = E$ και $(-)$-σημείο αν $\phi(x) = -E$. Από την (36) έπεται ότι σε ένα διάστημα $[t_j, t_{j+1}]$ δε μπορεί να περιέχεται και ένα $(+)$-σημείο και ένα $(-)$-σημείο, αφού η τιμή της ϕ δε μπορεί να αλλάξει μέσα σ' ένα τέτοιο διάστημα περισσότερο από $E/10$. Μετά από αυτή την παρατήρηση μπορούμε να ονομάσουμε ένα διάστημα $[t_j, t_{j+1}]$ $(+)$-διάστημα αν περιέχει ένα $(+)$-σημείο, $(-)$-διάστημα αν περιέχει ένα $(-)$-σημείο και *ουδέτερο* διάστημα αν δεν περιέχει ούτε από το ένα είδος ούτε από το άλλο.

Γράφουμε S για την ένωση των προσημασμένων διαστημάτων και N για την ένωση των ουδέτερων διαστημάτων. Έχουμε φυσικά $[a,b] = S \cup N$.

Ας απαριθμήσουμε τώρα τα προσημασμένα διαστήματα κινούμενοι από αριστερά προς τα δεξιά, ομαδοποιώντας τα ταυτόχρονα. Κάνουμε την αβλαβή υπόθεση ότι το πρώτο προσημασμένο διάστημα που συναντάμε είναι τύπου $(+)$:

$$\begin{gathered} I_1, I_2, \ldots, I_{k_1}, \text{ είναι } (+)\text{-διαστήματα} \\ I_{k_1+1}, \ldots, I_{k_2}, \text{ είναι } (-)\text{-διαστήματα} \\ \ldots. \\ I_{k_{m-1}+1}, \ldots, I_{k_m} \text{ είναι } (-1)^{m-1}\text{-διαστήματα.} \end{gathered}$$

Για κάθε δείκτη j το διάστημα I_j είναι πριν το I_{j+1}, ισχύει δηλ. $\max I_j \le \min I_{j+1}$.

Θα έχουμε αποδείξει το ζητούμενο αν δείξουμε ότι $m \ge N+2$, μια και τότε μπορούμε να πάρουμε το πρώτο σημείο της εναλλασσόμενης ακολουθίας σημείων από την πρώτη γραμμή διαστημάτων, το δεύτερο από τη δεύτερη γραμμή κ.λ.π. Έστω ότι δεν ισχύει το ζητούμενο, ότι δηλ. έχουμε

$$m \le N+1. \tag{37}$$

Από την (36) έπεται ότι ένα $(+)$-διάστημα δε μπορεί να ακουμπάει ένα $(-)$-διάστημα, αφού δε γίνεται να μεταβληθεί η τιμή της ϕ κατά $2E$ μέσα σε μήκος $\le 2\Delta$. Άρα ανάμεσα στα διαστήματα μιας γραμμής παραπάνω και στα διαστήματα της επόμενης υπάρχει πάντα ένα τουλάχιστον ουδέτερο διάστημα. Επιλέγουμε $m-1$ σημεία $z_1 < z_2 < \cdots < z_{m-1}$ παίρνοντας το z_j να είναι οποιοδήποτε σημείο του συνόλου $[a,b] \setminus S$ ανάμεσα στην j-ομάδα και στην $(j+1)$-ομάδα διαστημάτων. Ορίζουμε έπειτα το πολυώνυμο $q \in \mathcal{P}_{m-1}$

$$q(x) = (z_1 - x)(z_2 - x)\cdots(z_{m-1} - x)$$

το οποίο έχει την ιδιότητα ότι αλλάζει πρόσημο όταν (κινούμενοι από αριστερά προς τα δεξιά) περνάμε κάποιο z_j και σε κάθε διάστημα της πρώτης ομάδας έχει θετικό πρόσημο. Εύκολα βλέπουμε ότι αυτό συνεπάγεται ότι σε κάθε προσημασμένο διάστημα η συνάρτηση $\phi(x) = (f-p)(x)$ και η $q(x)$ έχουν ίδιο πρόσημο.

Θα δείξουμε τώρα ότι μπορούμε να επιλέξουμε ένα κατάλληλα μικρό αλλά θετικό λ τ.ώ. η συνάρτηση

$$p(x) + \lambda q(x) \in \mathcal{P}_N$$

να είναι καλύτερη proσέγγιση της f απ' ότι η συνάρτηση $p(x)$, πράγμα άτοπο, άρα ισχύει τελικά $m \ge N+2$ όπως θα θέλαμε.

Έστω

$$e = \max_{x \in N} |\phi(x)|.$$

Έχουμε $e < E$ μια και σε κάθε ουδέτερο διάστημα έχουμε $|\phi(x)| < E$ για κάθε x στο διάστημα και άρα το supremum της $|\phi(x)|$ σε αυτό το διάστημα είναι κι αυτό $< E$ (αυτό είναι συνέπεια της συνέχειας της $\phi(x)$). Επιλέγουμε τώρα οποιοδήποτε $\lambda > 0$ που να ικανοποιεί της σχέση

$$\lambda < \frac{1}{\|q\|_\infty} \min\{E-e, E/2\}. \tag{38}$$

Θέλουμε τώρα να δείξουμε

$$\|f - (p + \lambda q)\|_\infty < E.$$

Από τη συνέχεια της συνάρτησης $f - (p+\lambda q)$ αρκεί να δείξουμε ότι για κάθε $x \in [a,b]$ ισχύει

$$|(f(x) - p(x)) - \lambda q(x)| < E.$$

Διακρίνουμε τώρα δύο περιπτώσεις για το $x \in [a,b]$:

- $x \in N$
  Από την τριγωνική ανισότητα έχουμε
  $$|(f(x) - p(x)) - \lambda q(x)| \le |f(x) - p(x)| + \lambda |q(x)| \le e + \lambda \|q\|_\infty < E$$
  από την πρώτη ανισότητα της (38).
- $x \in S$
  Έχουμε $|f(x) - p(x)| \ge (9/10)E > \lambda \|q\|_\infty$ (από τη δεύτερη ανισότητα της (38)), και άρα οι αριθμοί $f(x) - p(x)$ και $\lambda q(x)$ έχουν το ίδιο πρόσημο. Για δύο ομόσημους αριθμούς a, b με $|a| \ge |b|$, ισχύει $|a - b| = |a| - |b|$, άρα έχουμε
  $$|(f(x) - p(x)) - \lambda q(x)| = |f(x) - p(x)| - \lambda |q(x)| \le E - \lambda \min_{x \in S} |q(x)| < E.$$
  Η τελευταία ανισότητα οφείλεται στο ότι η συνάρτηση $q(x)$ δε μηδενίζεται πουθενά στο S (αφού οι ρίζες της είναι τα z_j που δεν ανήκουν στο S) και άρα το $\min_{x \in S} |q(x)|$ είναι αυστηρά θετική ποσότητα.

Η απόδειξη είναι πλήρης για το ότι ένα βέλτιστο πολυώνυμο έχει μια εναλλασσόμενη ακολουθία σημείων με τουλάχιστον $N + 2$ σημεία.

Αντίστροφα τώρα, ας υποθέσουμε ότι για το πολυώνυμο $p \in \mathcal{P}_N$ η συνάρτηση $\phi(x) = f(x) - p(x)$ έχει τα εναλλασόμενα σημεία
$$x_0 < x_1 < \cdots < x_{N+1}$$
στο διάστημα $[a, b]$. Θέλουμε να δείξουμε ότι το $p(x)$ είναι μια βέλτιστη προσέγγιση της f στην ∞-νόρμα από το χώρο $\mathcal{P}_N$. Έστω ότι δεν είναι και ότι το πολυώνυμο $q \in \mathcal{P}_N$ είναι καλύτερο, ότι ισχύει δηλ.
$$\|f - q\|_\infty < \|f - p\|_\infty. \tag{39}$$
Για κάθε $j = 0, 1, 2, \ldots, N + 1$ έχουμε
$$|f(x_j) - p(x_j)| = E = \|f - p\|_\infty > \|f - q\|_\infty \ge |f(x_j) - q(x_j)|$$
άρα η συνάρτηση
$$q(x) - p(x) = (f - p)(x) - (f - q)(x)$$
έχει εναλλασσόμενα πρόσημα στα σημεία x_j. Αυτό ισχύει γιατί από την παραπάνω ανισότητα προκύπτει ότι
$$-E < |f(x_j) - q(x_j)| < E$$
για κάθε j, και άρα η διαφορά $(q-p)(x_j)$ είναι θετική όταν $f(x_j)-p(x_j) = E$ και αρνητική όταν $f(x_j)-p(x_j) = -E$. Ανάμεσα σε κάθε εναλλαγή προσήμου μιας συνεχούς συνάρτηση υπάρχει τουλάχιστον ένα σημείο όπου η συνάρτηση μηδενίζεται άρα το πολυώνυμο $q(x)-p(x)$ έχει τουλάχιστον $N+1$ ρίζες. Αφού όμως $\deg(q - p) \le N$ έπεται ότι το $q - p$ είναι το μηδενικό πολυώνυμο, άτοπο λόγω της (39). ´Αρα το πολυώνυμο $p(x)$ είναι όντως μια βέλτιστη προσέγγιση.

Απομένει να δείξουμε ότι η βέλτιστη προσέγγιση είναι μοναδική. Έστω όχι και ας υποθέσουμε ότι το πολυώνυμο $q^*(x)$ είναι επίσης μια βέλτιστη προσέγγιση. Τότε όμως και ο μέσος όρος τους
$$r = \frac{p^* + q^*}{2}$$
είναι επίσης μια βέλτιστη προσέγγιση (αποδείξτε το αυτό; ισχύει για οποιαδήποτε νόρμα, όχι μόνο την ∞-νόρμα). Χρησιμοποιώντας αυτά που έχουμε ήδη αποδείξει, υπάρχουν $N + 2$ σημεία
$$x_0 < x_1 < \cdots < x_{N+1}$$

που είναι εναλλασσόμενα για τη συνάρτηση $f - r$. Αυτό μεταφράζεται στο ότι

$$(f - p^*)(x_j) + (f - q^*)(x_j) = (-1)^j 2E, \quad (j = 0, 1, 2, \ldots, N + 1).$$

Αφού όμως $|(f - p^*)(x_j)| \leq E$ και $|(f - q^*)(x_j)| \leq E$ έπεται ότι για κάθε j

$$(f - p^*)(x_j) = (f - q^*)(x_j) = (-1)^j E,$$

οπότε τα πολυώνυμα $p^*(x)$ και $q^*(x)$ ταυτίζονται σε $N + 2$ σημεία. Αφού ο βαθμός τους είναι $\leq N$ έπεται ότι τα πολυώνυμα ταυτίζονται.

5. Πολυώνυμα Chebyshev και ιδιότητές τους

5.1. **Ορισμός των πολυωνύμων Chebyshev.** Έχοντας αποδείξει το Θεώρημα 4.2 που χαρακτηρίζει τη βέλτιστη προσέγγιση μιας συνάρτησης από ένα χώρο πολυωνύμων $\mathcal{P}_N$ θα το εφαρμόσουμε στο να λύσουμε το εξής πρόβλημα:

> Ποια είναι η βέλτιστη προσέγγιση της συνάρτησης x^n από το χώρο $\mathcal{P}_{n-1}$ (πολυώνυμα βαθμού μέχρι και $n - 1$) και για το διάστημα $[-1, 1]$;

Θέλουμε με άλλα λόγια να λύσουμε το πρόβλημα βελτιστοποίησης

> Για ποιο πολυώνυμο $p(x) \in \mathcal{P}_{n-1}$ ελαχιστοποιείται η νόρμα
> $$\|x^n - p(x)\|_\infty = \max_{-1 \leq x \leq 1} |x^n - p^*(x)|;$$

Η ύπαρξη της λύσης $p(x)$ εξασφαλίζεται από το γεγονός ότι ο χώρος $\mathcal{P}_{n-1}$ έχει πεπερασμένη διάσταση (ίση με n) και η μοναδικότητας από το Θεώρημα 4.2. Το Θεώρημα 4.2 μας εξασφαλίζει επίσης και την ύπαρξη $n + 1 = (n - 1) + 2$ σημείων

$$-1 \leq x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n \leq 1$$

τα οποία είναι εναλασσόμενα για το πολυώνυμο $p(x) = x^n - p^*(x)$, ισχύει δηλ. $|p(x_j)| = M$ για $j = 0, 1, \ldots, n$ και $p(x_j) \cdot p(x_{j+1}) < 0$ για $j = 0, 1, \ldots, n - 1$, όπου

$$M = \|p(x)\|_\infty = \|x^n - p^*(x)\|_\infty.$$

Πρώτα κάνουμε την παρατήρηση ότι σε κάθε ένα από τα εσωτερικά σημεία x_j η παράγωγος $p'(x)$ μηδενίζεται αφού σε κάθε τέτοιο σημείο η $p(x)$ έχει τοπικό μέγιστο (όταν η τιμή της είναι M) ή τοπικό ελάχιστο (όταν η τιμή της είναι $-M$). Αφού $\deg p'(x) \leq n - 1$ συμπεραίνουμε ότι

$$\text{η } p'(x) \text{ μηδενίζεται ακριβώς στα } x_1, x_2, \ldots, x_{n-1} \tag{40}$$

και άρα τα σημεία x_0 και x_n δεν είναι εσωτερικά σημεία του διαστήματος $[-1, 1]$ και συνεπώς

$$x_0 = -1, \quad x_n = 1. \tag{41}$$

Προκύπτει επίσης, από την έλλειψη μηδενικών της $p(x)$ στο διάστημα $[1, +\infty)$ και το γεγονός ότι $\lim_{x \to +\infty} p(x) = +\infty$ ότι η τιμή της $p(x)$ στο 1 είναι

$$p(1) = +M. \tag{42}$$

Το πολυώνυμο $M^2 - p^2(x)$ είναι παντού μη αρνητικό στο διάστημα $[-1, 1]$ άρα κάθε ρίζα του στο $(-1, 1)$ οφείλει να είναι διπλή, τουλάχιστον (αλλιώς η συνάρτηση θα έπαιρνε και θετικές και αρνητικές τιμές σε κάθε γειτονιά μιας ρίζας ρ αφού θα γραφόταν ως $(x - \rho)q(x)$ για κάποιο πολυώνυμο με $q(\rho) \neq 0$).

Η συνάρτηση αυτή μηδενίζεται τουλάχιστον στα σημεία $x_0, x_1, \ldots, x_n$ άρα, λαμβάνοντας υπόψιν ότι $\deg M^2 - p^2(x) \le 2n$,

(43)

όλες οι ρίζες της $M^2 - p^2(x)$ είναι οι x_0 (απλή), x_1 (διπλή), x_2 (διπλή), $\ldots$, x_{n-1} (διπλή), x_n (απλή)

Από τις προτάσεις (40) και (43) προκύπτει ότι τα πολυώνυμα $M^2 - p^2(x)$ και $(1-x^2)(p'(x))^2$ έχουν ακριβώς τις ίδιες ρίζες, άρα είναι το ένα πολλαπλάσιο του άλλου

$$M^2 - p^2(x) = K(1-x^2)(p'(x))^2, \quad \text{(για κάποια σταθερά } K\text{)}.$$

Η σταθερά K προσδιορίζεται κοιτώντας τους μεγιστοβάθμιους συντελεστές των δύο μελών. Αφού

$$p(x) = x^n + \cdots$$

(με «$\cdots$» συμβολίζουμε όρους μικρότερης τάξης) έχουμε $p'(x) = nx^{n-1} + \cdots$ και $(p'(x))^2 = n^2 x^{2(n-1)} + \cdots$ και τέλος

$$(1-x^2)(p'(x))^2 = -n^2 x^{2n} + \cdots$$

Επίσης έχουμε $M^2 - p^2(x) = -x^{2n} + \cdots$ άρα $K = \frac{1}{n^2}$. Παίρνουμε τις θετικές τετραγωνικές ρίζες στη σχέση $M^2 - p^2(x) = \frac{1}{n^2}(1-x^2)(p'(x))^2$ και, κάνοντας την παρατήρηση ότι $p'(x) > 0$ για $x > x_{n-1}$ (αφού δεν έχει μηδενικά πέρα από το x_{n-1} και πρέπει να πάει η $p(x)$ στο $+\infty$) παίρνουμε τη διαφορική εξίσωση

$$\frac{p'(x)}{\sqrt{M^2 - p^2(x)}} = \frac{n}{\sqrt{1-x^2}} \tag{44}$$

που ισχύει τουλάχιστον σε μια γειτονιά του $x = 1$. Λύνουμε τη διαφορική αυτή εξίσωση με τη μέθοδο του χωρισμού μεταβλητών ξαναγράφοντάς την ως

$$\frac{dp}{\sqrt{M^2 - p^2(x)}} = \frac{n\,dx}{\sqrt{1-x^2}}$$

και χρησιμοποιώντας τον τύπο $\int \frac{dx}{\sqrt{1-x^2}} = \arccos x + C$ (αυτό προκύπτει εύκολα από το ότι η $\arccos x$ είναι η αντίστροφη συνάρτηση της $\cos x$). Παίρνουμε έτσι τη λύση

$$p(x) = M \cos(n \arccos x + C) \tag{45}$$

όπου C είναι μια σταθερά που μένει ακόμη να προσδιορισθεί, όπως και η τιμή του M.

Αν $\theta \in [0, \pi]$ και $x = \cos\theta$ (ή αλλιώς $\theta = \arccos x$) βλέπουμε ότι η αντιστοιχία (αλλαγή μεταβλητής) των x και θ είναι μια συνεχής και 1-1 αντιστοιχία των διαστημάτων $x \in [-1, 1]$ και $\theta \in [0, \pi]$. Από δω και πέρα όποτε χρησιμοποιούμε το x ή το θ αυτά πάντα θα συνδέεονται με την παραπάνω σχέση. Έτσι η σχέση (45) μπορεί να γραφεί και ως

$$p(x) = M \cos(n\theta + C). \tag{46}$$

Για $x = 1$ ($\theta = 0$) παίρνουμε (αφού $p(x) = M$) ότι $\cos C = 1$ και άρα το C είναι κάποιο ακέραιο πολλαπλάσιο του 2π, οπότε μπορούμε να το παραλείψουμε τελείως από τη σχέση (45) ή (46) αφού δεν επηρεάζει την τιμή του συνιμητόνου. Για να βρούμε τώρα την τιμή του M στη σχέση

$$p(x) = M \cos(n \arccos x) = M \cos(n\theta) \tag{47}$$

χρησιμοποιούμε τον υπολογισμό που κάναμε στην απόδειξη της (34) όπου βρήκαμε ότι

$$\cos n\theta = \sum_{\substack{j=0 \\ j\,\text{άρτιο}}}^{n} \binom{n}{j} (\cos\theta)^{n-j} (-1)^{j/2} (1-\cos^2\theta)^{j/2}. \tag{48}$$

Αφού ο μεγιστοβάθμιος συντελεστής του x στο αριστερό μέλος της (47) είναι 1 το ίδιο οφείλει να είναι και ο μεγιστοβάθμιος συντελεστής του $\cos\theta$ στο δεξί μέλος της (47). Από κάθε προσθετέο του αθροίσματος στο δεξί μέλος της (48) προκύπτει ακριβώς ένα πολλαπλάσιο του $(\cos x)^n$ με συντελεστή

$$\binom{n}{j}(-1)^{j/2}(-1)^{j/2} = \binom{n}{j}(-1)^j = \binom{n}{j} \quad (\text{αφού } j \text{ άρτιο}).$$

Ο συντελεστής του $(\cos\theta)^n$ δηλ. στο δεξί μέλος της (47) είναι

$$M \sum_{\substack{j=0 \\ j\,\text{άρτιο}}}^{n} \binom{n}{j} = M2^{n-1}. \tag{49}$$

Από αυτό προκύπτει ότι $M = 2^{1-n}$ και η τελική μορφή του πολυωνύμου

$$p(x) = 2^{1-n}\cos(n \arccos x) = 2^{1-n}\cos(n\theta) \tag{50}$$

Παρατήρηση 5.1. Η (49) προκύπτει πολύ εύκολα από το διωνυμικό θεώρημα

$$(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j}$$

παίρνοντας $a = -1$ και $b = 1$ και χρησιμοποιώντας την ταυτότητα

$$2^n = \sum_{j=0}^{n} \binom{n}{j},$$

η οποία επίσης προκύπτει από το διωνυμικό θεώρημα θέτοντας $a = b = 1$.

Ορισμός 5.1. Τα πολυώνυμα Chebyshev πρώτου είδους $T_n(x)$ είναι τα πολυώνυμα για τα οποία ισχύει για $x \in [-1, 1]$

$$T_n(x) = \cos(n \arccos x), \quad n = 0, 1, 2, \ldots. \tag{51}$$

Παρατήρηση 5.2. Έχουμε ήδη δει ότι το δεξί μέλος της (51) είναι ένα πολυώνυμο του x. Ισοδύναμα αυτό είναι ένα πολυώνυμο του $\cos\theta$. Τα πολυώνυμα Chebyshev ορίζονται φυσικά για κάθε $x \in \mathbb{R}$ αν και το δεξί μέλος της (51) δεν έχει νόημα για $|x| > 1$ αφού δεν ορίζεται η συνάρτηση $\arccos x$ τότε.

5.2. **Ιδιότητες των πολυωνύμων Chebyshev.**

Θεώρημα 5.1. *Για $n \geq 2$ ισχύει*

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x). \tag{52}$$

Από την τριγωνομετρική ταυτότητα $\cos a + \cos b = 2\cos\frac{a-b}{2}\cos\frac{a+b}{2}$ προκύπτει πολύ εύκολα η σχέση

$$\cos n\theta = 2\cos\theta\cos{(n-1)\theta} - \cos{(n-2)\theta}$$

από την οποία προκύπτει το Θεώρημα 5.1 με την αντικατάσταση $x = \cos\theta$.

Η αναδρομική σχέση (52) είναι σημαντική για πολλούς λόγους, ένας από τους οποίους είναι ότι μας επιτρέπει να υπολογίσουμε τα πολυώνυμα $T_n(x)$ πολύ εύκολα, βρίσκοντας από τα $T_0(x) = 1$ και

$T_1(x) = x$ πρώτα το $T_2(x)$, μετά το $T_3(x)$, κ.λ.π. Παραθέτουμε παρακάτω μερικά από τα $T_n(x)$:

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \\ T_5(x) &= 16x^5 - 20x^3 + 5x. \end{aligned}$$

Πρόβλημα 5.1. *Αποδείξτε ότι για n άρτιο τα πολυώνυμα $T_n(x)$ περιέχουν μόνο άρτιες δυνάμεις του x ενώ για n περιττό μόνο περιττές.*

Χρησιμοποιείστε την (52).

Πρόβλημα 5.2. *Αποδείξτε ότι για $x \in \mathbb{R}$ ισχύει*

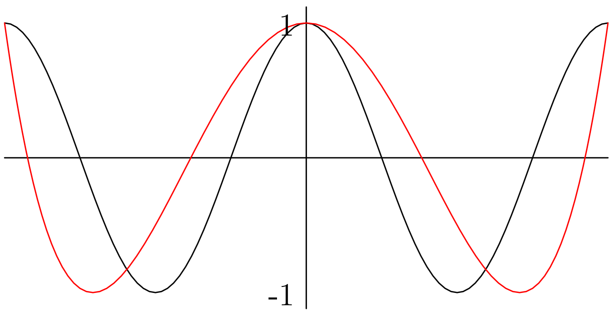
(53) $$T_n(x) = \frac{1}{2}\left((x + \sqrt{x^2 - 1})^n + (x - \sqrt{x^2 - 1})^n\right).$$

Δείξτε το πρώτα για $x \in [-1, 1]$ κάνοντας την αντικατάσταση $x = \cos\theta$. (Προσέξτε ότι για $|x| < 1$ ο αριθμός $\sqrt{x^2 - 1}$ είναι φανταστικός.) Δείξτε έπειτα ότι το δεξί μέλος της (53) *είναι πολυώνυμο του x (ότι δεν υπάρχουν δηλ. τετραγωνικές ρίζες όταν αναπτύξετε τις n-οστές δυνάμεις) και άρα η ισότητα των δύο μελών ισχύει για όλα τα $x \in \mathbb{C}$ μια και τα δύο πολυώνυμα ταυτίζονται σε ένα ολόκληρο διάστημα.*

Από τον τύπο

$$T_n(x) = \cos n\theta, \quad (\text{με } x = \cos\theta,\ \theta \in [0, \pi],\ x \in [-1, 1])$$

προκύπτει ότι οι συναρτήσεις $T_n(\cdot)$ δεν είναι παρά οι συναρτήσεις $\cos(n\cdot)$ όπου όμως ο άξονας των x έχει υποστεί μια αναπαραμέτριση, μια «αλλαγή μεταβλητής» όπως λέμε συνήθως.



Σχήμα 6: Οι συναρτήσεις $T_n(x)$ (κόκκινη) και $\cos(n\theta)$ σχεδιασμένες η μια πάνω στην άλλη. Μπορείτε να δείτε ότι η αλλαγή μεταβλητής $x = \cos\theta$ έχει διατηρήσει τα χαρακτηριστικά της καμπύλης που έχουν να κάνουν μόνο με τον άξονα των y, π.χ. το πόσες φορές τέμνει η καμπύλη τον x-άξονα ή το πόσες φορές αλλάζει η μονοτονία της συνάρτησης. Δεν έχει όμως διατηρήσει κατ' ανάγκη τα χαρακτηριστικά της καμπύλης που εξαρτώνται και από τον άξονα τον x. Ένα χαρακτηριστικό παράδειγμα είναι η κλίση της καμπύλης (παράγωγος). Στα δύο άκρα του διαστήματος ($[-1, 1]$ στη μια περίπτωση και $[0, \pi]$ στην άλλη, τα οποία τα έχουμε σχεδιάσει ως ένα εδώ) φαίνεται καθαρά ότι η μαύρη καμπύλη έχει παράγωγο 0 ενώ η κόκκινη όχι.

Η αλλαγή μεταβλητής αυτή είναι πολύ χρήσιμη για να αποδεικνύουμε διάφορες ιδιότητες των πολυωνύμων Chebyshev ανάγοντας τα αντίστοιχα ερωτήματα στις συναρτήσεις $\cos n\theta$. Για παράδειγμα, αν θέλουμε να βρούμε τα μηδενικά των $T_n(x)$ (για $x \in [-1, 1]$, μια και έχουμε ήδη αποδείξει ότι δεν υπάρχουν άλλα) αρκεί να βρούμε τα μηδενικά της συνάρτησης $\cos n\theta$ για $\theta \in [0, \pi]$ και να πάρουμε τα συνημίτονά τους.

Πρόβλημα 5.3. *Δείξτε ότι τα μηδενικά της συνάρτησης $T_n(x)$ είναι οι αριθμοί*

$$\cos\left(\frac{(2k-1)\pi}{2n}\right), \quad k = 1, 2, \ldots, n.$$

Βρείτε πρώτα τα μηδενικά της $\cos n\theta$ για $\theta \in [0, \pi]$.

Η αλλαγή μεταβλητής αυτή είναι επίσης χρήσιμη στον υπολογισμό διαφόρων ολοκληρωμάτων. Αφού

$$d\theta = \frac{dx}{-\sqrt{1-x^2}}, \quad dx = -\sin\theta\, d\theta$$

τα ορισμένα ολοκληρώματα μετατρέπονται σύμφωνα με τους κανόνες

$$\int_0^\pi \cdots\, d\theta = \int_{-1}^1 \cdots \frac{dx}{\sqrt{1-x^2}},$$

και

$$\int_{-1}^1 \cdots\, dx = \int_0^\pi \cdots \sin\theta\, d\theta.$$

Από το Πρόβλημα 3.7 για παράδειγμα γνωρίζουμε ότι

$$\int_{-\pi}^{\pi} \cos mx \cos nx\, dx = 0 \quad \text{αν } m \neq n$$

ενώ

$$\int_{-\pi}^{\pi} \cos^2 nx\, dx = \pi \quad \text{αν } n \geq 1.$$

Αυτές οι ιδιότητες *ορθογωνιότητας* των συναρτήσεων $\cos n\theta$ μεταφράζονται άμεσα σε ιδιότητες *ορθογωνιότητας* των πολυωνύμων Chebyshev.

Πρόβλημα 5.4. *Αποδείξτε ότι αν $m \neq n$ είναι δύο φυσικοί αριθμοί τότε*

$$\int_{-1}^1 T_m(x) T_n(x) \frac{dx}{\sqrt{1-x^2}} = 0. \tag{54}$$

Για την περίπτωση $m = n$ έχουμε

$$\int_{-1}^1 T_n^2(x) \frac{dx}{\sqrt{1-x^2}} = \begin{cases} \pi/2 & n \geq 1 \\ \pi & n = 0 \end{cases}.$$

Χρησιμοποιείστε την αρτιότητα της συνάρτησης $\cos x$ για να μετατρέψετε τα ολοκληρώματα $\int_{-\pi}^{\pi}$ σε $\int_0^\pi$ και μετά κάντε την αλλαγή μεταβλητής $x = \cos\theta$ όπως περιγράφεται παραπάνω.

Η ιδιότητα (54) ονομάζεται ορθογωνιότητα ως προς το εσωτερικό γινόμενο με *βάρος*

$$w(x) = \frac{1}{\sqrt{1-x^2}}$$

στο διάστημα $[-1, 1]$. Για κάθε αυστηρά θετική συνάρτηση $w(x) > 0$ πάνω σ' ένα διάστημα $[a, b]$ ορίζεται το εσωτερικό γινόμενο

$$\langle f, g\rangle_w = \int_a^b f(x)\overline{g(x)}w(x)\, dx \tag{55}$$

και η αντίστοιχη 2-νόρμα

$$\|f\|_{2,w} = \left(\int_a^b |f(x)|^2 w(x)\, dx\right)^{1/2} \tag{56}$$

που συνδέονται και πάλι με τη σχέση

$$\|f\|_{2,w}^2 = \langle f, f\rangle_w.$$

Ως προς αυτό το εσωτερικό γινόμενο, για $w(x) = (1-x^2)^{-1/2}$, οι συναρτήσεις $T_0(x), T_1(x), \ldots, T_N(x)$ αποτελούν μια ορθογώνια βάση του γραμμικού χώρου $\mathcal{P}_N$.

Πρόβλημα 5.5. *Αποδείξτε ότι ισχύει η ανισότητα Cauchy-Schwarz με βάρος* $w(x) > 0$*:*

$$|\langle f, g\rangle_W| \le \|f\|_{2,w}\|g\|_{2,w}.$$

$f(x)\overline{g(x)}w(x) = f(x)\sqrt{w(x)} \cdot \overline{g(x)\sqrt{w(x)}}$. *Χρησιμοποιείστε τη συνηθισμένη ανισότητα Cauchy-Schwarz, για* $w(x) = 1$ *δηλαδή.*

Το να έχει κανείς μια ορθογώνια βάση είναι πολύτιμο γιατί κάνει πολύ εύκολη τη διαδικασία εύρεσης του αναπτύγματος ενός πολυωνύμου $f(x) \in \mathcal{P}_N$ ως γραμμικού συνδυασμού στοιχείων αυτής της βάσης. Αν

$$f(x) = c_0T_0(x) + c_1T_1(x) + \cdots + c_NT_N(x)$$

για κάποια $c_j \in \mathbb{C}$ ο τρόπος να βρούμε τα c_j είναι να πάρουμε το εσωτερικό γινόμενο και των δύο μελών αυτής της ισότητας με το T_j. Από την ορθογωνιότητα θα «επιζήσει» μόνο ένας όρος δεξιά και παίρνουμε έτσι την ισότητα

$$c_j = \frac{\langle f, T_j\rangle}{\langle T_j, T_j\rangle}.$$

6. Παρεμβολή τιμών σε σημεία

6.1. **Παρεμβολή συναρτήσεων στο $\mathbb{R}$ από πολυώνυμα.** Αν $x_1, x_2, \ldots, x_N \in \mathbb{R}$ είναι κάποια σημεία (διαφορετικά μεταξύ τους) και $d_1, \ldots, d_N \in \mathbb{C}$ είναι κάποιες τιμές, λέμε ότι μια συνάρτηση p (της οποίας το πεδίο ορισμού περιέχει τα x_j) *παρεμβάλλει* τις τιμές d_j στα σημεία x_j αν $p(x_j) = d_j$ για $j = 1, 2, \ldots, N$.

Το πρόβλημα που θα μας απασχολήσει εδώ είναι το αν (και το πώς) μπορούμε να βρούμε τέτοιες συναρτήσεις p οι οποίες επιλέγονται μέσα από κάποια ευρεία κλάση συναρτήσεων, π.χ. ζητάμε οι p να είναι πολυώνυμα. Αν περιορίσουμε τις παρεμβάλλουσες συναρτήσεις p να είναι κάποια πολυώνυμα βαθμού $\le N$ η γενική μορφή του $p(x)$ είναι

$$p(x) = p_0 + p_1x + p_2x^2 + \cdots + p_Nx^N,$$

εξαρτάται δηλ. το πολυώνυμο $p(x)$ από $N+1$ μιγαδικές μεταβλητές (ή «βαθμούς ελευθερίας» όπως λέμε καμιά φορά). Είναι λογικό λοιπόν με τέτοια πολυώνυμα να προσπαθούμε να λύσουμε ένα πρόβλημα παρεμβολής με μέχρι $N+1$ σημεία, αφού κάθε σημείο παρεμβολής δημιουργεί μια επιπλέον εξίσωση (την $p(x_j) = d_j$) που πρέπει να ικανοποιεί το $p(x)$, και δεν περιμένουμε να μπορούμε γενικά να λύσουμε περισσότερες από $N+1$ ταυτόχρονες εξισώσεις με $N+1$ αγνώστους (τα p_j). Για $N+1$ σημεία όμως το πρόβλημα έχει πάντα λύση, όπως λέει το ακόλουθο Θεώρημα.

Θεώρημα 6.1. *Για κάθε $N+1$ διαφορετικά σημεία $x_0, x_1, \ldots, x_N \in \mathbb{R}$ και οποιαδήποτε δεδομένα $d_0, d_1, \ldots, d_N \in \mathbb{C}$ υπάρχει ακριβώς ένα μιγαδικό πολυώνυμο $p(x)$ βαθμού $\le N$ τέτοιο ώστε $p(x_j) = d_j$ για $j = 0, 1, \ldots, N$.*

Επίσης, αν όλες οι τιμές d_j είναι πραγματικές τότε οι συντελεστές του πολυωνύμου παρεμβολής είναι και αυτοί πραγματικοί αριθμοί.

Αν γράψουμε $p(x) = p_0 + p_1 x + p_2 x^2 + \cdots + p_N x^N$ τότε το γραμμικό σύστημα εξισώσεων $p(x_j) = d_j$ που θέλουμε να λύσουμε μπορεί να γραφεί σε μορφή πινάκων ως εξής

$$VP = D \tag{57}$$

όπου $P = (p_0, p_1, \ldots, p_N)^\top$ και $D = (d_0, d_1, \ldots, d_N)^\top$ είναι διανύσματα στήλες μήκους $N+1$ (P είναι το διάνυσμα των αγνώστων) και ο $(N+1) \times (N+1)$ πίνακας Vandermonde V είναι ο

$$V = V(x_0, x_1, \ldots, x_N) = \begin{bmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^N \\ 1 & x_1 & x_1^2 & \ldots & x_1^N \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_N & x_N^2 & \ldots & x_N^N \end{bmatrix}. \tag{58}$$

Το να έχει αυτό το σύστημα λύση P για κάθε δεξί μέλος D είναι ισοδύναμο με το να έχει πάντα μοναδική λύση και ισοδύναμο με το να είναι ο πίνακας V αντιστρέψιμος. Αρκεί λοιπόν να δείξουμε ότι ο πίνακας V πάντα (για κάθε επιλογή σημείων x_j δηλ.) είναι αντιστρέψιμος. Θα δείξουμε ότι η ορίζουσά του δεν είναι 0. Πιο συγκεκριμένα θα δείξουμε ότι

$$\det V(x_0, x_1, \ldots, x_N) = \prod_{0 \le j < i \le N} (x_i - x_j), \tag{59}$$

η οποία ποσότητα δεν είναι μηδέν αφού τα σημεία x_j είναι όλα διαφορετικά. Στο γινόμενο που εμφανίζεται στο δεξί μέλος της (59) έχουμε από ένα παράγοντα για κάθε ζεύγος διαφορετικών σημείων όπου το σημείο με το μεγαλύτερο δείκτη έχει πρόσημο $+$ και το άλλο έχει πρόσημο $-$.

Ας χρησιμοποιήσουμε επαγωγή ως πρός . Για $N = 1$ έχουμε

$$\det \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \end{bmatrix} = x_1 - x_0$$

άρα ο τύπος ισχύει σε αυτή την περίπτωση. Για το επαγωγικό βήμα (από $N-1$ σημεία σε N σημεία) αρκεί να δείξουμε τον τύπο

$$\det V(x_0, x_1, \ldots, x_N) = \det V(x_0, x_1, \ldots, x_{N-1})(x_N - x_0)(x_N - x_1) \cdots (x_N - x_{N-1}).$$

Βολεύει εδώ να θεωρήσουμε το x_N ως μια μεταβλητή γράφοντας στη θέση του x, οπότε η σχέση που έχουμε να δείξουμε γράφεται ως μια ισότητα πολυωνύμων

$$\det V(x_0, x_1, \ldots, x_{N-1}, x) = \det V(x_0, x_1, \ldots, x_{N-1})(x - x_0)(x - x_1) \cdots (x - x_{N-1}).$$

(Το ότι η ποσότητα $\det V(x_0, x_1, \ldots, x_{N-1}, x)$ είναι ένα πολυώνυμο του x βαθμού $\le N$ γίνεται φανερό αν αναπτύξουμε την ορίζουσα ως προς την τελευταία της γραμμή.) Οι ρίζες του πολυωνύμου του δεξιού μέλους είναι προφανώς τα σημεία $x_0, x_1, \ldots, x_{N-1}$. Όμως αυτές είναι και οι ρίζες του αριστερού μέλους αφού η ορίζουσα $\det V(x_0, x_1, \ldots, x_{N-1}, x)$ μηδενίζεται όποτε $x = x_j$ ($1 \le j \le N-1$) αφού αποκτά δύο ίδιες γραμμές. Όταν δύο πολυώνυμα βαθμού έχουν τις ίδιες ρίζες τότε είναι το ένα πολλαπλάσιο του άλλου (επί ένα μιγαδικό αριθμό). Όμως εύκολα βλέπει κανείς ότι το αριστερό και το δεξί μέλος έχουν τον ίδιο μεγιστοβάθμιο συντελεστή, ο οποίος είναι ο αριθμός $\det V(x_0, x_1, \ldots, x_{N-1})$. Αυτό είναι φανερό για το δεξί μέλος και στο αριστερό μέλος είναι επίσης φανερό αν παρατηρήσουμε ότι, στο ανάπτυγμα της $\det V(x_0, x_1, \ldots, x_{N-1}, x)$ ως προς την τελευταία γραμμή (με τις δυνάμεις του x), το x^N πολλαπλασιάζεται με μια υποορίζουσα που είναι ακριβώς η $\det V(x_0, x_1, \ldots, x_{N-1})$.

Συνεπώς τα δύο πολυώνυμα είναι ακριβώς ίδια και η απόδειξη έχει τελειώσει.

Το Θεώρημα 6.1 μπορεί να αποδειχθεί και αλλιώς, τουλάχιστον το κομμάτι που αφορά την ύπαρξη του πολυωνύμου παρεμβολής, με τη λεγόμενη μέθοδο παρεμβολής του Lagrange.

Ορίζουμε τα βοηθητικά πολυώνυμα βαθμού N

$$\ell_i(x) = \prod_{\substack{j=0,1,\ldots,N \\ j \ne i}} \frac{x - x_j}{x_i - x_j}, \quad (i = 0, 1, \ldots, N).$$

Εύκολα βλέπουμε ότι

$$\ell_i(x_j) = \begin{cases} 1 & (i = j) \\ 0 & (i \ne j) \end{cases}.$$

Το πολυώνυμο $p(x)$, βαθμού $\le N$, που παρεμβάλλει τις τιμές d_j στα x_j ($j = 0, 1, \ldots, N$) γράφεται λοιπόν

$$p(x) = d_0 \ell_0(x) + d_1 \ell_1(x) + \cdots + d_N \ell_N(x). \tag{60}$$

Από την (60) γίνεται φανερό ότι αν οι αριθμοί d_j είναι όλοι πραγματικοί τότε το πολυώνυμο παρεμβολής $p(x)$ έχει πραγματικούς συντελεστές.

Πρόβλημα 6.1. *Με τη μέθοδο παρεμβολής Lagrange (60) έχουμε δείξει ότι οποιαδήποτε δεδομένα d_j μπορούν να παρεμβληθούν με πολυώνυμο βαθμού μέχρι N σε οποιαδήποτε $N+1$ σημεία. Γιατί κάθε τέτοιο πολυώνυμο παρεμβολής είναι μοναδικό;*

Εδώ υπάρχουν τουλάχιστον δύο τρόποι για να απαντήσουμε: ένας κοιτώντας το πρόβλημα ως πρόβλημα γραμμικής άλγεβρας (δείτε την (57)) και ένας άλλος που κοιτάει το πλήθος των ριζών πολυωνύμου βαθμού $\le N$.

Η μέθοδος παρεμβολής του Newton εκφράζει το πολυώνυμο παρεμβολής στη μορφή

$$p(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \cdots a_N(x - x_0)(x - x_1)\cdots(x - x_{N-1}). \tag{61}$$

Εκφράζουμε δηλ. το πολυώνυμο παρεμβολής ως γραμμικό συνδυασμό των στοιχείων μιας βάσης του $\mathcal{P}_N$ διαφορετικής από τη συνηθισμένη $\{1, x, x^2, x^3, \ldots, x^{N-1}, x^N\}$, δηλ. της βάσης

$$1, x - x_0, (x - x_0)(x - x_1), \ldots, (x - x_0)(x - x_1)(x - x_2)\cdots(x - x_{N-1}). \tag{62}$$

Πρόβλημα 6.2. *Γιατί είναι οι συναρτήσεις (62) μια βάση του χώρου $\mathcal{P}_N$;*

Πώς μπορεί κανείς να βρει τους συντελεστές a_j της (61) για δεδομένα σημεία x_j και δεδομένα d_j; Η απάντηση είναι πολύ απλή και μεταφράζεται σε ένα αποτελεσματικό αλγόριθμο για τον υπολογισμό των a_j.

Θέτουμε κατ' αρχήν $x = x_0$ στην (61) και παίρνουμε έτσι την ισότητα

$$d_0 = p(x_0) = a_0$$

αφού όλοι οι όροι δεξιά εκτός από τον πρώτο μηδενίζονται. Έτσι βρίσκουμε $a_0 = d_0$. Έχοντας βρει το a_0 θέτουμε στην (61) $x = x_1$ και παρατηρούμε ότι μηδενίζονται όλοι οι όροι δεξια εκτός τους δύο πρώτους. Παίρνουμε έτσι

$$d_1 = p(x_1) = a_0 + a_1(x_1 - x_0)$$

στην οποία όλα είναι γνωστά εκτός από το a_1 για το οποίο λύνουμε και βρίσκουμε

$$a_1 = \frac{d_1 - a_0}{x_1 - x_0}.$$

Συνεχίζουμε θέτοντας $x = x_2$, $x = x_3$, κλπ, και βρίσκουμε με αυτό τον τρόπο διαδοχικά όλους του συντελεστές a_j.

Ποιος είναι όμως ένας λόγος για τον οποίο θα επιθυμούσαμε να χρησιμοποιήσουμε τη μέθοδο Newton έναντι της μεθόδου Lagrange για τον υπολογισμό της παρεμβολής; (Να τονίσουμε εδώ ότι το πολυώνυμο που υπολογίζει η κάθε μέθοδος είναι φυσικά το ίδιο, αλλά είναι γραμμένο με διαφορετικό τρόπο.) Μια απάντηση είναι ότι με τη μέθοδο Newton είναι πολύ ευκολότερο να προσθέσουμε ένα ακόμη σημείο παραμεβολής, χωρίς να χρειαστεί να ξαναυπολογίσουμε τα πάντα (όπως συμβαίνει αν έχουμε το πολυώνυμό μας στη μορφή (60)). Πράγματι, αν υποθέσουμε ότι έχουμε ήδη βρει τους συντελεστές a_j στην (61) και θέλουμε τώρα προσθέσουμε ένα σημείο x_{N+1} και δεδομένο παρεμβολής d_{N+1}, τότε δεν έχουμε παρά να προσθέσουμε ένα ακόμη όρο στο πολυώνυμό μας, οπότε γίνεται

$$\begin{aligned} p(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \cdots a_N(x - x_0)(x - x_1)\cdots(x - x_{N-1}) + \\ a_{N+1}(x - x_0)(x - x_1)\cdots(x - x_N). \end{aligned}$$

Θέτουμε τώρα $x = x_{N+1}$ και παίρνουμε την εξίσωση

$$\begin{aligned} d_{N+1} &= p(x_{N+1}) \\ &= a_0 + a_1(x_{N+1} - x_0) + a_2(x_{N+1} - x_0)(x_{N+1} - x_1) + \cdots \\ &\quad a_N(x_{N+1} - x_0)(x_{N+1} - x_1)\cdots(x_{N+1} - x_{N-1}) + \\ &\quad a_{N+1}(x_{N+1} - x_0)(x_{N+1} - x_1)\cdots(x_{N+1} - x_N) \end{aligned}$$

όπου όλα είναι γνωστά εκτός από το a_{N+1}, το οποίο και υπολογίζουμε λύνοντας αυτή την πρωτοβάθμια εξίσωση.

Δε χρειάστηκε να υπολογίσουμε ξανά τα $a_0, a_1, \ldots, a_N$ αφού στον υπολογισμό αυτών δεν μπαίνει με κανένα τρόπο το x_{N+1} ή το d_{N+1}.

Γενικά, το πολυώνυμο $p(x)$ βαθμού $\le N$ που παρεμβάλλει μια συνεχή συνάρτηση $f \in C([a, b])$ σε κάποια σημεία $x_0, x_1, \ldots, x_N$ δεν έχει καλές ιδιότητες προσέγγισης της f. Αυτό είναι το περιεχόμενο της παρακάτω πρότασης, την οποία δε θα αποδείξουμε.

Θεώρημα 6.2. *Για κάθε διάστημα $[a, b]$ και για κάθε ακολουθία συνόλων κόμβων παρεμβολής*

$$\begin{aligned} & a \le x_0^1 < x_1^1 \le b \\ & a \le x_0^2 < x_1^2 < x_2^2 \le b \\ & \ldots \\ & a \le x_0^N < x_1^N < \cdots < x_N^N \le b \\ & \ldots \end{aligned}$$

(έχουμε δηλαδή για κάθε φυσικό αριθμό $N \ge 1$ ένα σύνολο από $N+1$ κόμβους παρεμβολής στο $[a,b]$) υπάρχει $f \in C([a,b])$ τ.ώ. η ποσότητα $\|f - L_N(f)\|_\infty$ δεν είναι φραγμένη. Εδώ $L_N(f)$ είναι το μοναδικό πολυώνυμο βαθμού $\le N$ που παρεμβάλλει την f στα σημεία $x_0^N, x_1^N, \ldots, x_N^N$.

Δε μπορούμε συνεπώς να περιμένουμε ότι η ποσότητα $\|L_N(f) - f\|_\infty$ θα είναι μικρή αν δεν έχουμε κάποια παραπάνω πληροφορία για την f από τη συνέχειά της στο $[a,b]$. Το επόμενο θεώρημα είναι προς αυτή την κατεύθυνση.

Ορισμός 6.1. Ο χώρος συναρτήσεων $f \in C^n([a,b])$ αποτελείται από τις συναρτήσεις που είναι n φορές παραγωγίσιμες με όλες τις παραγώγους τους συνεχείς. (Στα άκρα εννοούμε τις πλευρικές παραγώγους.)

Θεώρημα 6.3. *Αν $f \in C^{N+1}([a,b])$ και $a \le x_0 < x_1 < \cdots < x_{N+1} \le b$ τότε για κάθε $x \in [a,b]$ ισχύει το φράγμα*

$$|f(x) - L_N(f)(x)| \le \frac{1}{(N+1)!}\left\|f^{(N+1)}\right\|_\infty |(x-x_0)\cdots(x-x_N)|. \tag{63}$$

Έστω $W(x) = (x-x_0)(x-x_1)\cdots(x-x_N)$. Αρκεί να δείξουμε ότι υπάρχει $\xi \in (a,b)$ τ.ώ.

$$f(x) - L_N(f)(x) = \frac{1}{(N+1)!} f^{(N+1)}(\xi) W(x). \tag{64}$$

Ο τύπος (64) ονομάζεται και «τύπος του Lagrange με υπόλοιπο» (παρατηρείστε την ομοιότητα με τον τύπο του Taylor με υπόλοιπο που συνδέει το πολυώνυμο Taylor μιας συνάρτησης με τη συνάρτηση).

Είναι φανερό ότι η (64) ισχύει για $x \in \{x_0, x_1, \ldots, x_N\}$ για οποιοδήποτε ξ μια και μηδενίζονται και τα δύο μέλη. Μπορούμε συνεπώς να υποθέσουμε ότι το x δεν είναι κανένα από τα x_j και άρα $W(x) \ne 0$ και μπορούμε να ορίσουμε

$$\lambda = \frac{f(x) - L_N(f)(x)}{W(x)}, \quad \phi(t) = f(t) - L_N(f)(t) - \lambda W(t).$$

Η συνάρτηση $\phi(t)$ έχει $N+2$ διαφορετικά μηδενικά στο $[a,b]$, τα $x_0, x_1, \ldots, x_N$ και το x. Από το θεώρημα του Rolle έπεται ότι ανάμεσα σε δύο διαδοχικά τέτοια μηδενικά υπάρχει ένα μηδενικό της ϕ', και άρα η ϕ' έχει $N+1$ διαφορετικά μηδενικά στο $[a,b]$. Και πάλι από το θεώρημα του Rolle ανάμεσα σε δύο μηδενικά της ϕ' υπάρχει ένα μηδενικό της ϕ'', οπότε η ϕ'' έχει N διαφορετικά μηδενικά στο $[a,b]$. Συνεχίζοντας κατ' αυτό τον τρόπο βλέπουμε ότι υπάρχει $\xi \in (a,b)$ τ.ώ. $\phi^{(N+1)}(\xi) = 0$.

Όμως $\phi^{(N+1)} = f^{(N+1)} - L_N(f)^{(N+1)} - \lambda W^{(N+1)}$. Αφού $\deg L_N(f) \le N$ έπεται ότι $L_N(f)^{(N+1)} = 0$ αφού $W(x) = x^{N+1} +$ (όροι μικρότερου βαθμού) έχουμε $W^{(N+1)}(\xi) = (N+1)!$. Παίρνουμε έτσι τη σχέση

$$0 = f^{(N+1)}(\xi) - \frac{f(x) - L_N(f)(x)}{W(x)}(N+1)!$$

που είναι ισοδύναμη με την (64).

Πρόβλημα 6.3. *Αν $f(x) = e^x$ δείξτε ότι για κάθε διάστημα $[a,b]$ και για οποιοδήποτε σύστημα κόμβων παρεμβολής x_j^N, $N = 1, 2, \ldots,$, $j = 0, 1, \ldots, N$, (όπως στο Θεώρημα 6.2) η ακολουθία πολυωνύμων $L_N(f)$ συγκλίνει ομοιόμορφα στην f στο $[a,b]$.*

Χρησιμοποιείστε το Θεώρημα 6.3.

6.2. **Παρεμβολή σε παραπάνω διαστάσεις.** Έστω ότι θέλουμε να παρεμβάλλουμε τις τιμές $d_0, d_1, \ldots, d_N$ στα σημεία $x_0, x_1, \ldots, x_N$ με ένα πολυώνυμο βαθμού $\le N$

$$p(x) = p_0 + p_1 x + p_2 x^2 + \ldots + p_N x^N.$$

Έχουμε δει ότι αυτό το πρόβλημα έχει πάντα μοναδική λύση p_j. Αν δούμε το πρόβλημα ως ένα γραμμικό σύστημα, όπως κάναμε στην εξίσωση (57)

$$VP = D$$

όπου $V = V(x_0, x_1, \ldots, x_N)$ είναι ο πίνακας Vandermonde και $D = [d_0, \ldots, d_N]^\top$ είναι το διάνυσμα στήλη των τιμών που θελουμε να παρεμβάλλουμε, τότε η ύπαρξη και μοναδικότητα της λύσης αυτού του συστήματος είναι ισοδύναμες με το να είναι ο V αντιστρέψιμος. Με άλλα λόγια το πρόβλημα έχει μοναδική λύση ακριβώς επειδή

$$\det V(x_0, \ldots, x_N) \neq 0$$

οποτεδήποτε τα σημεία x_j είναι διαφορετικά.

Ένας διαφορετικός τρόπος να εκφράσουμε το πρόβλημα της παρεμβολής είναι να πούμε ότι αναζητούμε ένα γραμμικό συνδυασμό των συναρτήσεων

$$f_0(x) = 1,\ f_1(x) = x,\ f_2(x) = x^2,\ \ldots,\ f_N(x) = x^N \tag{65}$$

ο οποίος παίρνει τιμές d_j στα x_j. Με αυτή τη γλώσσα το πρόβλημα μπορεί να γενικευτεί αλλάζοντας τις συναρτήσεις $f_j(x)$. Για παράδειγμα, μπορούμε να πάρουμε ως συναρτήσεις παρεμβολής τις

$$\begin{aligned} g_0(x) &= 1 \\ g_1(x) &= x - x_0 \\ g_2(x) &= (x - x_0)(x - x_1) \\ &\ldots \\ g_N(x) &= (x - x_0)(x - x_1)\cdots(x - x_{N-1}) \end{aligned} \tag{66}$$

οι οποίες αντιστοιχούν στην παρεμβολή Newton που έχουμε ήδη δεί ότι έχει επίσης μοναδική λύση. Αυτό συνεπάγεται ότι για κάθε $N + 1$ διαφορετικά σημεία $x_0, x_1, \ldots, x_N$ η ορίζουσα του πίνακα

$$g_i(x_j), \quad i, j = 0, 1, \ldots, N,$$

είναι $\neq 0$.

Παρατήρηση 6.1. Η επίλυση του προβλήματος παρεμβολής με γραμμικούς συνδυασμούς των f_j ή των g_j δίνει φυσικά το ίδιο πολυώνυμο (αφού αυτό έχουμε δείξει ότι είναι μοναδικό). Όμως στην πρώτη περίπτωση αυτό είναι γραμμένο σα γραμμικός συνδυασμός των f_j ενώ στη δεύτερη περίπτωση σα γραμμικός συνδυασμός των g_j.

Δεν είναι απαραίτητο οι συναρτήσεις παρεμβολής να είναι πολυώνυμα; αν είναι τότε ο κάθε γραμμικός συνδυασμός τους είναι φυσικά κι αυτός πολυώνυμο. Συνήθως απαιτούμε όμως να είναι συνεχείς συναρτήσεις.

Πρόβλημα 6.4. *Δείξτε ότι για κάθε* $N \ge 1$ *οι συναρτήσεις*

$$e_0(x) = 1, e_1(x) = e^{ix}, e_2(x) = e^{2ix}, \ldots, e_N(x) = e^{Nix}$$

μπορούν να παρεμβάλλουν οποιαδήποτε δεδομένα d_j *σε οποιαδήποτε διαφορετικά σημεία* $x_0, x_1, \ldots, x_N \in [0, 2\pi)$ *με μοναδικό τρόπο.*

Παρά το ότι στη μία διάσταση μπορούμε να έχουμε μοναδική παρεμβολή με πλειάδα συναρτήσεων παρεμβολής (δείτε παραδείγματα f_j, g_j και e_j παραπάνω) σε μεγαλύτερη διάσταση αυτό δεν είναι δυνατό.

Θεώρημα 6.4. *Ας είναι $\Omega \subseteq \mathbb{R}^2$ ένα ανοιχτό σύνολο και $N \geq 1$. Δεν υπάρχουν συνεχείς συναρτήσεις $f_j : \Omega \to \mathbb{R}$, $j = 0, 1, \ldots, N$, τ.ώ. για κάθε διαφορετικά $x_0, x_1, \ldots, x_N \in \Omega$ και για κάθε $d_0, d_1, \ldots, d_N \in \mathbb{C}$ να υπάρχει μοναδικός $\mathbb{C}$-γραμμικός συνδυασμός των f_j που να παίρνει τη τιμή d_j στο x_j για $j = 0, 1, \ldots, N$.*

Παρατήρηση 6.2. Έχει σημασία στην απόδειξη παρακάτω ότι οι συναρτήσεις $f_j(x)$ παίρνουν πραγματικές και όχι μιγαδικές τιμές.

Το φαινόμενο αυτό οφείλεται σε τοπολογικούς λόγους όπως θα γίνει φανερό από την απόδειξη που ακολουθεί.

Ας υποθέσουμε ότι το Θεώρημα 6.4 δεν ισχύει και ότι υπάρχουν τέτοιες συναρτήσεις f_j. Συνέπεια αυτού είναι ότι η συνάρτηση

$$\phi(x_0, x_1, \ldots, x_N) = \det[f_i(x_j)]_{i,j=0,1,\ldots,N}$$

είναι $\neq 0$ οποτεδήποτε όλα τα σημεία x_j είναι διαφορετικά.

Η συνάρτηση $\phi(x_0, x_1, \ldots, x_N)$ είναι (α) πραγματική (αφού όλα τα στοιχεία του πίνακα είναι πραγματικά) και (β) συνεχής συνάρτηση των x_j (αφού οι f_i είναι συνεχείς και η ορίζουσα είναι ένα άθροισμα από όρους κάθε ένας από τους οποίους είναι γινόμενο κάποιων $f_i(x_j)$).

Επιλέγουμε τώρα τα σημεία $x_0, x_1, \ldots, x_N \in \Omega$ με τα x_0 και x_1 να είναι στην ίδια συνεκτική συνιστώσα του Ω. Από την υπόθεσή μας έχουμε

$$0 \neq D = \det[f_i(x_j)]_{i,j=0,1,\ldots,N}.$$

Σταθεροποιούμε τα σημεία $x_2, \ldots, x_N$ και μετακινούμε με συνεχή τρόπο τα σημεία x_0 και x_1 έτσι ώστε

(1) Σε κάθε χρονική στιγμή t τα σημεία $x_0(t), x_1(t), x_2, x_3, \ldots, x_N$ είναι όλα διαφορετικά μεταξύ τους.
(2) Στην αρχική χρονική στιγμή $t = 0$ έχουμε $x_0(0) = x_0$ και $x_1(0) = x_1$.
(3) Στην τελική χρονική στιγμή $t = 1$ έχουμε $x_0(1) = x_1$ και $x_1(1) = x_0$.

Δηλαδή με συνεχή τρόπο ανταλλάσσουν θέση τα x_0 και x_1. Το ότι αυτό είναι δυνατό στο επίπεδο είναι διαισθητικά φανερό και δεν το αποδεικνύουμε αυστηρά μια και θα μας έπαιρνε πολύ χώρο και χρόνο σε κατεύθυνση που δεν είναι βασική γι' αυτό το μάθημα. Είναι επίσης φανερό ότι αυτό δε γίνεται στη μια διάσταση: δύο σημεία δε μπορούν, συνεχώς κινούμενα, να ανταλλάξουν τις θέσεις τους χωρίς να συγκρουστούν.

Η συνάρτηση

$$\psi(t) = \phi(x_0(t), x_1(t), x_2, x_3, \ldots, x_N), \quad 0 \leq t \leq 1,$$

είναι μια συνεχής *πραγματική* συνάρτηση του t. Ισχύει επίσης $\psi(0) = -\psi(1)$ αφού οι δύο αυτές τιμές ισούνται με τις ορίζουσες δύο πινάκων που έχουν τις δύο πρώτες στήλες τους εναλλαγμένες και τις υπόλοιπες στήλες τους ίδιες. Άρα για κάποιο $t \in [0, 1]$ ισχύει $\psi(t) = 0$ πράγμα που αντιφάσκει στο μη μηδενισμό της ορίζουσας $\det[f_i(x_j)]_{i,j=0,1,\ldots,N}$ για κάθε επιλογή διαφορετικών σημείων x_j.

Αυτή η αντίφαση συμπληρώνει την απόδειξη του Θεωρήματος 6.4.

7. Σειρές Fourier

7.1. **Συντελεστές και σειρές Fourier.** Συμβολίζουμε με $C^{2\pi}$ το σύνολο όλων των συνεχών συναρτήσεων επί του $\mathbb{R}$ που είναι επίσης 2π-περιοδικές. Ισοδύναμα, $C^{2\pi}$ είναι το σύνολο όλων των συνεχών συναρτήσεων $f : [0, 2\pi] \to \mathbb{C}$ με $f(0) = f(2\pi)$.

Στο χώρο $C^{2\pi}$ ορίζουμε το εσωτερικό γινόμενο

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(x)\overline{g(x)}\,dx$$

καθώς και τις L^p νόρμες ($1 \le p < \infty$)

$$\|f\|_p = \left(\frac{1}{2\pi}\int_0^{2\pi} |f(x)|^p\,dx\right)^{1/p}.$$

Η 2-νόρμα συνδέεται, ως συνήθως, με το εσωτερικό γινόμενο με τη σχέση

$$\|f\|_2^2 = \langle f, f\rangle.$$

Οι βασικές ιδιότητες του εσωτερικού γινομένου έχουν καλυφθεί στην §3.

Ορισμός 7.1. Αν $f \in C^{2\pi}$ και $n \in \mathbb{Z}$ τότε ο n-οστός συντελεστής Fourier της f ορίζεται ως η ποσότητα

$$\widehat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\,dx = \langle f, e^{inx}\rangle. \tag{67}$$

Παρατήρηση 7.1. Ο παραπάνω ορισμός είναι πολύ περιοριστικός. Το δεξί μέλος της (67) ορίζεται όχι μόνο όταν η f είναι συνεχής αλλά οποτεδήποτε είναι Riemann ολοκληρώσιμη. Για παράδειγμα ορίζεται οποτεδήποτε η f είναι κατά τμήματα συνεχής και δε χρειάζεται καν να έχουμε $f(0) = f(2\pi)$. Για τους σκοπούς του μαθήματος αυτού όμως θα αρκέσει να μιλάμε για συντελεστές Fourier συνεχών και 2π-περιοδικών συναρτήσεων.

Σε μια συνεχή συνάρτηση αντιστοιχούμε τη «σειρά Fourier» της

$$\sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{inx}$$

καθώς και τα *συμμετρικά μερικά της αθροίσματα*

$$S_N(f)(x) = \sum_{n=-N}^{N} \widehat{f}(n)e^{inx}.$$

Δεν κάνουμε εξ αρχής κάποιο ισχυρισμό για τη σύγκλιση της σειράς Fourier της f, που ορίζεται να σημαίνει τη σύγκλιση των $S_N(f)(x)$ σε κάποιο μιγαδικό αριθμό. Η «ιδανική» κατάσταση είναι η σειρά αυτή να συγκλίνει στην f. Αυτό όμως δεν ισχύει γενικά (υπάρχουν συνεχείς συναρτήσεις των οποίων η σειρά Fourier δε συγκλίνει πουθενά) αν και υπάρχουν φυσιολογικές συνθήκες που μας το εξασφαλίζουν αυτό (π.χ. η f να έχει συνεχή δεύτερη παράγωγο).

Πρόβλημα 7.1. *Αν $p(x) = \sum_{k=-N}^{N} p_k e^{ikx}$ είναι ένα τριγωνομετρικό πολυώνυμο βρείτε τις ποσότητες $\widehat{p}(n)$ σα συνάρτηση των p_k. Δείξτε ότι η σειρά Fourier της $p(x)$ συγκλίνει στην $p(x)$.*

Πάρτε το εσωτερικό γινόμενο της $p(x)$ με την e^{inx} και χρησιμοποιείστε την ορθογωνιότητα των εκθετικών συναρτήσεων.

Πρόβλημα 7.2. *Δείξτε ότι* $\left|\widehat{f}(n)\right| \le \|f\|_1$.

Πρόβλημα 7.3. *Υπολογίστε τους συντελεστές Fourier των συναρτήσεων $f(x) = 1$, $g(x) = \chi_{[a,b]}(x)$ όπου $[a, b] \subseteq [0, 2\pi]$. (Η συνάρτηση $g(x)$ δεν είναι βέβαια συνεχής αλλά και γι' αυτήν οι συντελεστές Fourier ορίζονται με τον ίδιο τύπο.)*

Θεώρημα 7.1 (Λήμμα Riemann-Lebesgue)**.** *Αν $f \in C^{2\pi}$ τότε $\widehat{f}(n) \to 0$ για $|n| \to \infty$.*

Το Θεώρημα 7.1 ισχύει κατά προφανή τρόπο αν η f είναι τριγωνομετρικό πολυώνυμο αφού η ακολουθία των συντελεστών Fourier της όχι απλά πάει στο 0 αλλά είναι τελικά μηδενική (αν το $|n|$ ξεπεράσει το βαθμό του πολυωνύμου).

Για να αποδείξουμε το Θεώρημα 7.1 για μια γενική $f \in C^{2\pi}$ χρησιμοποιούμε το Θεώρημα του Weierstrass για προσέγγιση από τριγωνομετρικά πολυώνυμα (Θεώρημα 3.3).

Έστω $\epsilon > 0$. Υπάρχει τότε τριγωνομετρικο πολυώνυμο $p(x)$ τ.ώ. $\|f - p\|_\infty < \epsilon$, και άρα ισχύει και $\|f - p\|_1 < \epsilon$ η οποία συνεπάγεται (Πρόβλημα 7.2)

$$\left|\widehat{f}(n) - \widehat{p}(n)\right| < \epsilon.$$

Αφού για $|n|$ αρκετά μεγάλο το $\widehat{p}(n)$ μηδενίζεται έπεται ότι για $|n|$ αρκετά μεγάλο $\left|\widehat{f}(n)\right| < \epsilon$, που σημαίνει ακριβώς ότι $\lim_{|n|\to\infty} \left|\widehat{f}(n)\right| = 0$.

Θεώρημα 7.2. *Έστω $f \in C^{2\pi}$. Τότε για κάθε $n \in \mathbb{Z}$ ισχύουν τα παρακάτω.*

(1) *Αν $g(x) = f(x - \alpha)$ τότε $\widehat{g}(n) = \widehat{f}(n)e^{-i\alpha n}$.*
(2) *Αν $g(x) = e^{ikx}f(x)$ τότε $\widehat{g}(n) = \widehat{f}(n - k)$.*
(3) *Αν επίσης $f' \in C^{2\pi}$ τότε $\widehat{f'}(n) = in\widehat{f}(n)$.*

Η απόδειξη του Θεωρήματος 7.2 αφήνεται ως άσκηση για τον αναγνώστη ο οποίος μπορεί να αποδείξει το ζητούμενο χρησιμοποιώντας τον ορισμό των συντελεστών Fourier αντικαταστάσεις μεταβλητών και ολοκλήρωση κατά μέρη.

Ορισμός 7.2 (Συνέλιξη συναρτήσεων)**.** Αν $f, g \in C^{2\pi}$ η *συνέλιξη* των f και g ορίζεται να είναι η συνάρτηση $f * g \in C^{2\pi}$ που δίνεται από τον τύπο

$$f * g(x) = \frac{1}{2\pi}\int_0^{2\pi} f(t)g(x - t)\, dt.$$

Πρόβλημα 7.4. *Αποδείξτε ότι η $f * g$ είναι συνεχής και 2π-περιοδική και ότι $f * g = g * f$.*

Για τη συνέχεια της $f * g$ βρείτε ένα άνω φράγμα για την ποσότητα $f * g(x + h) - f * g(x)$ χρησιμοποιώντας των ομοιόμορφη συνέχεια των f ή g.

Πρόβλημα 7.5. (Συντελεστές Fourier συνέλιξης)
Αν $f, g \in C^{2\pi}$ δείξτε ότι $\widehat{f * g}(n) = \widehat{f}(n)\widehat{g}(n)$.

Χρησιμοποιείστε τον ορισμό των συντελεστών Fourier και αλλάξτε τη σειρά ολοκλήρωσης.

Ορισμός 7.3 (Συνέλιξη ακολουθιών). Αν $a_n, b_n \in \mathbb{C}, n \in \mathbb{Z}$, είναι δύο (διπλής κατεύθυνσης) ακολουθίες τέτοιες ώστε η a_n είναι φραγμένη και η b_n ικανοποιεί την $\sum_{n=-\infty}^{\infty} |b_n| < \infty$ (ή το ανάποδο), τότε η συνέλιξή τους ορίζεται ως η ακολουθία $a * b$ των οποίων οι όροι δίνονται από τον τύπο

$$(a * b)_n = \sum_{k=-\infty}^{\infty} a_k b_{n-k}. \tag{68}$$

Πρόβλημα 7.6. *Δείτε ότι οι υποθέσεις του ορισμού της συνέλιξης ακολουθιών εγγυώνται τη σύγκλιση της σειράς (68) για κάθε $n \in \mathbb{Z}$. Δείτε επίσης ότι $(a * b)_n = (b * a)_n$ για $n \in \mathbb{Z}$.*

Απόλυτη σύγκλιση μιας σειράς συνεπάγεται σύγκλιση.

Πρόβλημα 7.7. (Συντελεστες Fourier γινομένου)
Αν $f \in C^{2\pi}$ και $p(x)$ τριγωνομετρικό πολυώνυμο δείξτε ότι

$$\widehat{f \cdot p}(n) = (\widehat{f} * \widehat{p})_n.$$

Οι $\widehat{f}(n)$ είναι φραγμένοι από $\|f\|_1$ και η ακολουθία $\widehat{p}(n)$ είναι τελικά μηδενική, άρα ορίζεται η συνέλιξη των δύο ακολουθιών. Δείξτε το ζητούμενο πρώτα για $p(x) = e^{ikx}$.

7.2. **Ο πυρήνας του Dirichlet.** Κεντρικό αντικείμενο για τη μελέτη της κατά σημείο σύγκλισης

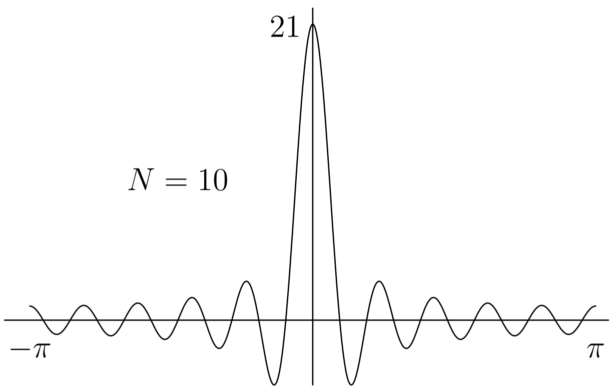
$$S_N(f)(x) \to f(x)$$

είναι ο λεγόμενος πυρήνας του Dirichlet τάξης N, το τριγ. πολυώνυμο δηλ. που ορίζεται ως

$$D_N(x) = \sum_{k=-N}^{N} e^{ikx}. \tag{69}$$

Δεν είναι δύσκολο να βρει κανείς ένα κλειστό τύπο για το $D_N(x)$:

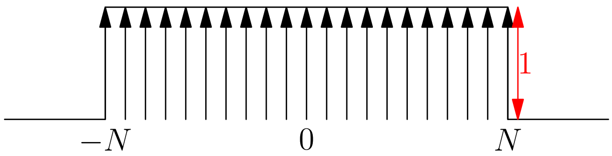
$$D_N(x) = \frac{\sin{(N + \frac{1}{2})x}}{\sin\frac{x}{2}}. \tag{70}$$



Σχήμα 7: Ο πυρήνας του Dirichlet

Για να δείξουμε την (70) χρησιμοποιούμε τον τύπο για το άθροισμα της πεπερασμένης γεωμετρικής σειράς

$$1 + z + z^2 + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z}, \quad (z \neq 1), \tag{71}$$



Σχήμα 8: Οι συντελεστές Fourier του πυρήνα του Dirichlet $D_N(x)$ για $N = 10$

(με e^{ix} στη θέση του z) και τον τύπο για τη διαφορά συνημιτόνων

$$\cos A - \cos B = 2\sin\frac{A+B}{2}\sin\frac{B-A}{2}. \tag{72}$$

Πρόβλημα 7.8. *Αποδείξτε τις (71) και (72).*

Πρόβλημα 7.9. *Κάντε τις πράξεις μόνοι σας για εξάσκηση και αποδείξτε την (70). Θυμηθείτε ότι μια εν γένει καλή στρατηγική όταν έχετε ένα κλάσμα με μιγαδικό παρανομαστή είναι να πολλαπλασιάζετε αριθμητή και παρανομαστή με το συζυγή του παρανομαστή ώστε να γίνεται πραγματικός ο παρανομαστής.*

Πρόβλημα 7.10. *Αποδείξτε ότι αν $f \in C^{2\pi}$ τότε για κάθε $N \ge 0$ ισχύει*

$$S_N(f)(x) = f * D_N(x).$$

Συνέπεια του Πυθαγορείου Θεωρήματος (Πρόβλημα 3.6) είναι η ταυτότητα του Parseval για τριγωνομετρικά πολυώνυμα. Για την απόδειξη απλά χρησιμοποιούμε την ορθοκανονικότητα των εκθετικών συναρτήσεων. Θυμόμαστε επίσης (δείτε το Πρόβλημα 7.1) ότι οι συντελεστές Fourier ενός τριγωνομετρικού πολυωνύμου είναι οι ίδιοι οι συντελεστές του πολυωνύμου ($\widehat{p}(k) = p_k$).

Θεώρημα 7.3 (Ταυτότητα Parseval για τριγ. πολυώνυμα)**.**
Αν $p(x) = \sum_{k=-N}^{N} p_k e^{ikx}$ τότε

$$\|p(x)\|_2^2 = \sum_{k=-N}^{N} |p_k|^2 = \sum_{k=-N}^{N} |\widehat{p}(k)|^2. \tag{73}$$

Ένα βασικό ερώτημα της ανάλυσης Fourier είναι το κατά πόσον μπορούμε να ανακατασκευάσουμε μια συνάρτηση από τους συντελεστές Fourier της. Ένας φυσιολογικός τρόπος να εξειδικεύσει κανείς αυτό το ερώτημα είναι να ρωτήσει αν τα μερικά αθροίσματα της σειράς Fourier συγκλίνουν στην ίδια τη συνάρτηση

$$S_N(f) \to f;$$

Για να είναι πλήρως καθορισμένο το πρόβλημα πρέπει κανείς να εξηγήσει τι είδους σύγκλιση εννοεί. Αν για παράδειγμα η σύγκλιση εννοείται κατά σημείο ($\forall x : S_N(f)(x) \to f(x)$) ή ομοιόμορφα ($\|S_N(f) - f\|_\infty \to 0$) τότε η απάντηση γενικά είναι αρνητική. Αυτό σημαίνει ότι υπάρχει συνεχής συνάρτηση f για την οποία δεν ισχύει η κατά σημείο σύγκλιση παντού (ούτε φυσικά και η ομοιόμορφη σύγκλιση). Αυτό δε θα το αποδείξουμε σε αυτό το μάθημα. Αν όμως η σύγκλιση εννοείται στην L^2 νόρμα τότε η απάντηση είναι καταφατική (Θεώρημα 7.5) και αυτό έχει τεράστια σημασία για τις εφαρμογές της ανάλυσης Fourier.

Πολλές φορές στην Ανάλυση μια ιδιότητα σύγκλισης είναι συνέπεια μιας ανισότητας. Έτσι κι εδώ, για να αποδείξουμε το Θεώρημα 7.5 θα χρειαστούμε ένα φράγμα για το γραμμικό τελεστή $f \to S_N(f)$. Πριν διατυπώσουμε αυτό το φράγμα (ανισότητα του Bessel) ορίζουμε την έννοια της ορθογώνιας προβολής σε ένα υπόχωρο.

Ορισμός 7.4. Αν $e_1, \ldots, e_k$ είναι ένα ορθοκανονικό σύστημα συναρτήσεων στο $C^{2\pi}$

$$\langle e_j, e_j \rangle = 1 \text{ και } \langle e_i, e_j \rangle = 0 (\text{για } i \neq j),$$

και $V = \mathrm{span}\{e_1, \ldots, e_k\}$ είναι ο γραμμικός χώρος που παράγουν η *ορθογώνια προβολή* μιας συνάρτησης $f \in C^{2\pi}$ στο V είναι η συνάρτηση

$$P_V(f) = \sum_{j=1}^{k} \langle f, e_j \rangle e_j.$$

Παρατήρηση 7.2. Είναι φανερό ότι $P_V(f) \in V$. Έχουμε επίσης ότι το διάνυσμα $f - P_V(f)$ είναι κάθετο στο χώρο V, είναι δηλ. κάθετο σε κάθε διάνυσμα του V. Αν $v \in V$ πρέπει να δείξουμε ότι $\langle f - P_V(f), v \rangle = 0$. Αρκεί να το κάνουμε για τα διανύσματα $e_1, \ldots, e_k$ στη θέση του v μια και είναι μια γραμμική σχέση ως προς v. Αυτό επαληθεύεται εύκολα (κάντε το).

Κάτι άλλο που οφείλουμε να πούμε εδώ είναι ότι κάθε γραμμικός χώρος (πάνω στον οποίο έχουμε ορίσει κάποιο εσωτερικό γινόμενο (ώστε να έχει νόημα να μιλάμε για ορθογωνιότητα) έχει κάποια ορθοκανονική βάση $e_1, \ldots, e_k$ όπου $k = \dim V$. Αυτό είναι εύκολο να δειχτεί επαγωγικά ως προς τη διάσταση k και μπορεί επίσης να αποδειχτεί με χρήση της λεγόμενης ορθοκανονικοποίησης Gram–Schmidt την οποία θα δούμε αργότερα.

Πρόβλημα 7.11. *Δείξτε ότι*

$$\|P_V(f)\|_2^2 = \sum_{j=1}^{k} |\langle f, e_j \rangle|^2. \tag{74}$$

Εφαρμόστε το Πυθαγόρειο Θεώρημα (Πρόβλημα 3.6).

Πρόβλημα 7.12. *Αποδείξτε ότι το διάνυσμα $P_V(f)$ είναι το μοναδικό διάνυσμα v του V τ.ώ. $\langle f - v, w \rangle = 0$ για κάθε $w \in V$.*

Αν υπάρχει και άλλο τέτοιο διάνυσμα v' τότε το τρίγωνο fvv' έχει ορθή γωνία και στην κορυφή v και στην κορυφή v'. Δείξτε ότι αυτό είναι αδύνατο εφαρμόζοντας δύο φορές το Πυθαγόρειο θεώρημα. Με άλλα λόγια η υποτείνουσα είναι πάντα η αυστηρά μεγαλύτερη πλευρά σε ένα ορθογώνιο τρίγωνο και άρα δε μπορεί να υπάρχουν δύο υποτείνουσες.

Πρόβλημα 7.13. *Αν $V = \mathrm{span}\{e_1, \ldots, e_k\}$ και τα e_j είναι ανά δύο ορθογώνια αλλά όχι κατ' ανάγκη μοναδιαία, από ποιον τύπο δίνεται η προβολή $P_V(f)$;*

Κανονικοποιείστε τα e_j.

Πρόβλημα 7.14. *Αποδείξτε ότι το διάνυσμα (συνάρτηση) $P_V(f)$ δεν εξαρτάται από τα $e_1, e_2, \ldots, e_k$ αλλά μόνο από το χώρο V. Αν δηλ. $e'_1, \ldots, e'_k$ είναι ένα άλλο ορθοκανονικό σύστημα στο χώρο V (οπότε αυτόματα $V = \mathrm{span}\{e'_1, \ldots, e'_k\}$) τότε ισχύει και πάλι*

$$P_V(f) = \sum_{j=1}^{k} \langle f, e'_j \rangle e'_j.$$

$P_V(f) = \sum_{j=1}^{k} \lambda_j e'_j$ *για κάποια $\lambda_j \in \mathbb{C}$ αφού τα e'_j παράγουν το V. Παίρνοντας εσωτερικό γινόμενο αυτής της ισότητας με τα e'_i παίρνετε το ζητούμενο.*

Πρόβλημα 7.15. *Δείξτε ότι τα συμμετρικά μερικά αθροίσματα $S_N(f)$ της σειράς Fourier μιας $f \in C^{2\pi}$ είναι ακριβώς η προβολή της f στο γραμμικό χώρο που παράγουν οι συναρτήσεις*

$$e^{-iNx}, e^{-i(N-1)x}, \ldots, 1, e^{ix}, e^{i2x}, \ldots, e^{iNx}.$$

Το επόμενο αποτέλεσμα είναι πολύ σημαντικό για τη Θεωρία Προσέγγισης και την ανάλυση γενικότερα

Θεώρημα 7.4 (Η προβολή είναι η βέλτιστη προσέγγιση)**.**
Αν V είναι ένας υπόχωρος του $C^{2\pi}$ πεπερασμένης διάστασης και $f \in C^{2\pi}$ δείξτε ότι το διάνυσμα $P_V(f)$ είναι το μοναδικό διάνυσμα $v \in V$ που ελαχιστοποιεί την απόσταση $\|f - v\|_2$.

Αρκεί να δείξουμε ότι αν $v \in V$ είναι διαφορετικό από το $P_V(f)$ τότε

$$\|f - P_V(f)\|_2 < \|f - v\|_2.$$

Όμως το τρίγωνο $f P_V(f) v$ είναι ορθογώνιο με ορθή γωνία στην κορυφή $P_V(f)$, άρα η υποτείνουσα fv έχει μεγαλύτερο μήκος από την κάθετη πλευρά $f P_V(f)$.

Συνέπεια (γιατί;) του Θεωρήματος 7.4 είναι ότι

$$\|P_V(f)\|_2 \le \|f\|_2,$$

το οποίο αναδιατυπώνουμε ως πόρισμα.

Πόρισμα 7.1 (Ανισότητα Bessel)**.**
Αν $f_1, \ldots, f_k$ είναι ένα ορθοκανονικό σύστημα συναρτήσεων στο $C^{2\pi}$

$$\langle f_j, f_j \rangle = 1 \text{ και } \langle f_i, f_j \rangle = 0 \text{ (για } i \ne j),$$

και $f \in C^{2\pi}$ τότε

$$\|P_V(f)\|_2^2 = \sum_{j=1}^{k} |\langle f, f_j \rangle|^2 \le \|f\|_2^2$$

όπου $V = \mathrm{span}\{f_1, \ldots, f_k\}$.

Πόρισμα 7.2. *Αν $f \in C^{2\pi}$ τότε*

$$\|S_N(f)\|_2^2 = \sum_{k=-N}^{N} \left|\widehat{f}(k)\right|^2 \le \|f\|_2^2. \tag{75}$$

Θεώρημα 7.5 (Σύγκλιση της σειράς Fourier κατά L^2)**.**
Αν $f \in C^{2\pi}$ τότε $S_N(f) \to f$ στην L^2 νόρμα. Δηλαδή

$$\lim_{N\to\infty} \|S_N(f) - f\|_2 = 0. \tag{76}$$

Το Θεώρημα 7.5 είναι προφανές αν η f είναι τριγωνομετρικό πολυώνυμο, αφού σε αυτή την περίπτωση μόλις το $N \ge \deg f$ η συνάρτηση $S_N(f)$ δεν ξαναλλάζει και μένει σταθερή και ίση με f. Για τη γενική περίπτωση χρησιμοποιούμε το Θεώρημα του Weierstrass για τριγωνομετρικά πολυώνυμα (Θεώρημα 3.3). Αν $\epsilon > 0$ τότε υπάρχει τριγωνομετρικό πολυώνυμο $p(x)$ τέτοιο ώστε $|p(x) - f(x)| \le \epsilon$ για κάθε $x \in \mathbb{R}$. Αυτή η ανισότητα συνεπάγεται φυσικά την ανισότητα

$$\|f - p\|_2 \le \epsilon. \tag{77}$$

Έχουμε λοιπόν

$$\begin{aligned}
\|f - S_N(f)\|_2 &= \|f - p + p - S_N(p) + S_N(p) - S_N(f)\|_2 \\
&\le \|f - p\|_2 + \|p - S_N(p)\|_2 + \|S_N(p) - S_N(f)\|_2 && \text{(Τριγωνική ανισότητα)} \\
&\le \epsilon + \|p - S_N(p)\|_2 + \|S_N(f - p)\| && \text{(από την (77))} \\
&\le \epsilon + \|p - S_N(p)\|_2 + \|f - p\|_2 && \text{(ανισότητα Bessel)} \\
&\le 2\epsilon + \|p - S_N(p)\|_2 && \text{(από την (77)).}
\end{aligned}$$

Αν τώρα επιλέξουμε το N να είναι τουλάχιστον $\deg p$ τότε $S_N(p) = p$ οπότε έχουμε από την παραπάνω ανισότητα ότι $\|f - S_N(f)\|_2 < 2\epsilon$, πράγμα που σημαίνει ακριβώς ότι $\|f - S_N(f)\|_2 \to 0$.

Έχοντας αποδείξει τη σύγκλιση $S_N(f) \xrightarrow{L^2} f$ μπορούμε τώρα εύκολα να επεκτείνουμε την ταυτότητα του Parseval σε κάθε συνάρτηση $f \in C^{2\pi}$. Το παρακάτω πρόβλημα το κάνει πιο εύκολο.

Πρόβλημα 7.16. *Αν $f_n, g \in C^{2\pi}$ και $\|f_n - g\|_2 \to 0$ τότε $\|f_n\|_2 \to \|g\|_2$. (Ισχύει και για οποιαδήποτε άλλη νόρμα.)*

Γράψτε $\|f_n\|_2 = \|f_n - g + g\|_2 \le \|f_n - g\|_2 + \|g\|_2$ και ομοίως $\|g\|_2 = \|g - f_n + f_n\|_2 \le \|g - f_n\|_2 + \|f_n\|_2$ για να αποδείξετε ότι $\limsup \|f_n\|_2 \le \|g\|_2 \le \liminf \|f_n\|_2$.

Θεώρημα 7.6 (Ταυτότητα Parseval)**.**
Αν $f \in C^{2\pi}$ τότε

$$\|f\|_2^2 = \sum_{k=-\infty}^{\infty} \left|\widehat{f}(k)\right|^2. \tag{78}$$

Η απόδειξη του Θεωρήματος 7.6 έγκειται στο να παρατηρήσουμε ότι το δεξί μέλος της (78) είναι το όριο των $\|S_N(f)\|_2^2$ και να χρησιμοποιήσουμε τη σύγκλιση $S_N(f) \xrightarrow{L^2} f$ μαζί με το Πρόβλημα 7.16.

Μπορεί μια συνεχής συνάρτηση να «ανακατασκευαστεί» από την ακολουθία των συντελεστών Fourier της; Η απάντηση είναι καταφατική και αυτό είναι πάρα πολύ σημαντικό γιατί μας εξασφαλίζει ότι αν αποφασίσουμε να δούμε μια συνάρτηση μόνο στο «πεδίο Fourier» τότε δεν έχουμε χάσει πληροφορία.

Θεώρημα 7.7 (Θ. Μοναδικότητας)**.** *Αν οι $f, g \in C^{2\pi}$ έχουν $\widehat{f}(n) = \widehat{g}(n)$ για κάθε $n \in \mathbb{Z}$ τότε $f(x) = g(x)$ για κάθε $x \in \mathbb{R}$.*

Έστω $h = f - g$. Τότε $h \in C^{2\pi}$ και $\widehat{h}(n) = 0$ για κάθε $n \in \mathbb{Z}$. Πρέπει να δείξουμε ότι υπό αυτή την υπόθεση $h \equiv 0$. Αλλά από την ταυτότητα του Parseval (Θεώρημα 7.6) έχουμε

$$\|h\|_2^2 = \sum_{k=-\infty}^{\infty} \left|\widehat{h}(k)\right|^2 = 0,$$

άρα $\int_0^{2\pi} |h(x)|^2 \, dx = 0$. Από τη συνέχεια της h προκύπτει ότι η h είναι παντού 0.

Κλείνουμε το Κεφάλαιο των Σειρών Fourier με μια περίπτωση στην οποία συμβαίνει το «ιδανικό», η σειρά δηλ. Fourier μιας συνάρτησης συγκλίνει κατά σημείο (και μάλιστα ομοιόμορφα) στη συνάρτηση.

Θεώρημα 7.8. *Αν $f'' \in C^{2\pi}$ (η 2π-περιοδική δηλ. f έχει συνεχή δεύτερη παράγωγο παντού) τότε*

$$S_N(f)(x) \to f(x)$$

ομοιόμορφα για $x \in \mathbb{R}$.

Το κλειδί στην απόδειξη είναι ότι το μέγεθος των συντελεστών Fourier μιας τέτοιας συνάρτησης είναι μικρό, και αυτό έχει σα συνέπεια τη σύγκλιση της σειράς. Πράγματι χρησιμοποιώντας δύο φορές τον κανόνα (δείτε Θεώρημα 7.2)

$$\widehat{f'}(n) = in\widehat{f}(n)$$

παίρνουμε $\widehat{f''}(n) = -n^2\widehat{f}(n)$ και άρα, για $n \neq 0$,

$$\widehat{f}(n) = \frac{-1}{n^2}\widehat{f''}(n).$$

Χρησιμοποιώντας το φράγμα $\left|\widehat{f''}(n)\right| \leq \|f''\|_1$ παίρνουμε

$$\left|\widehat{f}(n)\right| \leq \frac{\|f''\|_1}{n^2} \quad (n \neq 0).$$

Από αυτό έπεται ότι η σειρά Fourier της f

$$\sum_{k=-\infty}^{\infty} \widehat{f}(k)e^{ikt}$$

συγκλίνει σε κάποια 2π-περιοδική συνάρτηση $g(x)$ επειδή συγκλίνει απόλυτα. Αυτό το τελευταίο είναι συνέπεια του ότι

$$\sum_{k=-\infty}^{\infty} \frac{\|f''\|_1}{k^2} < \infty. \tag{79}$$

Η σύγκλιση όμως είναι και ομοιόμορφη αφού η «ουρά» της σειράς Fourier

$$\sum_{|k|>N} \widehat{f}(k)e^{ikt}$$

φράσσεται (τριγωνική ανισότητα) από την ουρά της συγκλίνουσας σειράς (79) η οποία (επειδή η σειρά είναι συγκλίνουσα) πάει στο 0 με το N.

Επειδή τα μερικά αθροίσματα της σειράς Fourier είναι τριγωνομετρικά πολυώνυμα και άρα συνεχείς συναρτήσεις έπεται ότι και το όριο $g(x)$ είναι συνεχής συνάρτηση. Απομένει να δείξουμε ότι $g(x) = f(x)$ για κάθε $x \in \mathbb{R}$.

Θα δείξουμε ότι $\widehat{f}(n) = \widehat{g}(n)$ για κάθε $n \in \mathbb{Z}$ και άρα, από το Θεώρημα Μοναδικότητας (Θεώρημα 7.7) θα έχουμε $f \equiv g$. Αφού $S_N(f) \to g$ ομοιόμορφα έχουμε (γιατί;) ότι για κάθε n ότι

$$\widehat{S_N(f)}(n) \xrightarrow{N\to\infty} \widehat{g}(n).$$

Όμως η ακολουθία $\widehat{S_N(f)}(n)$ είναι σταθερή και ίση με $\widehat{f}(n)$ για $N \geq n$, άρα έχουμε δείξει το ζητούμενο.

8. Ορθογώνια πολυώνυμα

8.1. **Ορθογωνιοποίηση Gram–Schmidt.** Όπως είδαμε και στο Θεώρημα 7.4 το να μπορεί κανείς να υπολογίσει την ορθογώνια προβολή ενός διανύσματος f σε ένα γραμμικό χώρο V ισοδυναμεί με το να βρεί το πλησιέστερο διάνυσμα από το χώρο V στο διάνυσμα f. Το αποδείξαμε αυτό στο Θεώρημα 7.4. Εκεί η απόδειξη έγινε για ένα συγκεκριμένο διανυσματικό χώρο και εσωτερικό γινόμενο αλλά ισχύει σε οποιαδήποτε περίπτωση. Και το να υπολογίσουμε την ορθογώνια προβολή του f αν διαθέτουμε ήδη μια ορθοκανονική βάση $e_1, \ldots, e_k$ του V είναι πολύ εύκολο και δίνεται από τον Ορισμό 7.4.

Οι περισσότεροι γραμμικοί χώροι που μας απασχολούν εδώ είναι φυσικά χώροι συναρτήσεων και τα διανύσματα είναι συναρτήσεις, αλλά δε χάνουμε τίποτε με το να την κρύψουμε αυτή την πληροφορία σε αυτή τη φάση, μια και το είδος διανυσμάτων για το οποίο μιλάμε δεν ενδιαφέρει (ακόμη) αλλά μόνο το ότι μπορούμε αυτά να τα προσθέτουμε μεταξύ τους και να τα πολλαπλασιάζουμε με αριθμούς παραμένοντας στον ίδιο χώρο. Αυτό που χρειαζόμαστε τώρα, μέχρι να αρχίσουμε να μιλάμε για χώρους πολυωνύμων, είναι ακριβώς αυτή η γραμμική δομή και το εσωτερικό γινόμενο που θεωρούμε ότι υπάρχει ορισμένο στο γραμμικό μας χώρο.

Είναι λοιπόν πολύτιμο το να έχουμε μια ορθοκανονική βάση του V. Αυτό το επιτυγχάνει κανείς με μια αλγοριθμική διαδικασία, τη λεγόμενη ορθοκανονικοποίηση Gram–Schmidt.

Η διαδικασία αυτή παίρνει ως είσοδο μια ακολουθία $f_1, f_2, \ldots$ από γραμμικώς ανεξάρτητα διανύσματα σε κάποιο γραμμικό χώρο V (ο χώρος V μπορεί να είναι και απειροδιάστατος και η ακολουθία f_n μπορεί και να είναι μια άπειρη ακολουθία διανυσμάτων). Η διαδικασία παράγει μια άλλη ορθοκανονική ακολουθία $e_1, e_2, \ldots$.

Θεώρημα 8.1 (Η διαδικασία Gram–Schmidt)**.** *Έστω V γραμμικός χώρος με εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$ και $f_1, f_2, \ldots \in V$ μια γραμμικώς ανεξάρτητη ακολουθία διανυσμάτων. Τα διανύσματα $e_1, e_2, \ldots \in V$ (και η βοηθητική ακολουθία $v_2, v_3, \ldots$) ορίζονται ως εξής:*

$$\begin{aligned} e_1 &= \frac{1}{\|f_1\|_2} f_1 \\ v_k &= f_k - \langle f_k, e_1 \rangle e_1 + \cdots \langle f_k, e_{k-1} \rangle e_{k-1} && \text{(για } k \geq 2\text{)} \\ e_k &= \frac{1}{\|v_k\|_2} v_k && \text{(για } k \geq 2\text{).} \end{aligned}$$

Τότε τα e_j είναι ανά δύο ορθογώνια και έχουν $\|e_j\|_2 = 1$ και επίσης παράγουν τους ίδιους γραμμικούς χώρους με τα f_j, δηλ. για κάθε $k \geq 1$

$$\text{span}\{f_1, \ldots, f_k\} = \text{span}\{e_1, \ldots, e_k\}.$$

Το ότι η νόρμα των e_j είναι 1 είναι άμεσο από τον ορισμό.

Αποδεικνύουμε με επαγωγή ως προς k ότι τα διανύσματα $e_1, \ldots, e_k$ είναι ορθοκανονικά και παράγουν τον ίδιο χώρο με τα $f_1, \ldots, f_k$. Αυτό είναι προφανές για $k = 1$ αφού το e_1 είναι πολλαπλάσιο του f_1. Αν υποθέσουμε ότι ισχύει η πρόταση για το $k - 1$ αποδεικνύουμε κατ' αρχήν ότι το e_k είναι κάθετο προς τα $e_1, \ldots, e_{k-1}$. Αυτό είναι φανερό μια και το v_k ισούται με το f_k μείον την προβολή του στο χώρο $\text{span}\{e_1, \ldots, e_{k-1}\}$ και άρα (Πρόβλημα 7.12) είναι κάθετο σε ολόκληρο το χώρο $\text{span}\{e_1, \ldots, e_{k-1}\}$ και άρα και στα ίδια τα $e_1, \ldots, e_k$. Το διάνυσμα e_k είναι απλά η κανονικοποίηση του v_k και άρα είναι κι αυτό ορθογώνιο στα $e_1, \ldots, e_k$. Τέλος, αφού

$$f_k - v_k \in \text{span}\{e_1, \ldots, e_{k-1}\} = \text{span}\{f_1, \ldots, f_{k-1}\}$$

προκύπτει ότι

$$\text{span}\{e_1, \ldots, e_{k-1}, v_k\} = \text{span}\{f_1, \ldots, f_{k-1}, f_k\}$$

και άρα, αφού το v_k είναι πολλαπλάσιο του e_k, έχουμε και το επιθυμητό

$$\text{span}\{e_1, \ldots, e_{k-1}, e_k\} = \text{span}\{f_1, \ldots, f_{k-1}, f_k\}.$$

και η επαγωγική απόδειξη είναι πλήρης.

Κατά κάποιο τρόπο η διαδικασία Gram–Schmidt εξετάζει τα στοιχεία f_k ένα προς ένα και κρατάει από κάθε f_k το «κομμάτι» του που είναι ορθογώνιο με e_j που έχουν υπολογιστεί μέχρι εκείνη τη στιγμή, δηλ. τα $e_1, \ldots, e_{k-1}$. Είναι σημαντικό να τονίσουμε ότι στον ορισμό του e_k μέσω του βοηθητικού διανύσματος v_k (που είναι ουσιαστικά το e_k πριν κανονικοποιηθεί) όλα τα στοιχεία που εμφανίζονται στο δεξί μέλος έχουν ήδη υπολογιστεί στα προηγούμενα στάδια της διαδικασίας και άρα γνωρίζουμε ό,τι χρειάζεται για τον υπολογισμό.

Πρόβλημα 8.1. *Δείξτε ότι κάθε χώρος πεπερασμένης διάστασης με εσωτερικό γινόμενο έχει ορθοκανονική βάση.*

Πρόβλημα 8.2. *Εφαρμόστε τη διαδικασία ορθοκανονικοποίησης Gram–Schmidt στα διανύσματα*

$$f_1(x) = 1,\ f_2(x) = 1 - x^2,\ f_3(x) = 1 - x$$

στο γραμμικό χώρο $C([0,1])$ με εσωτερικό γινόμενο $\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\,dx$.

Πρόβλημα 8.3. *Στο χώρο $C([a,b])$ των συνεχών συναρτήσεων στο διάστημα $[a,b]$ βρείτε ένα τύπο για το εσωτερικό γινόμενο $\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx$ αν*

$$f(x) = \sum_{j=0}^{m} f_j x^j,\ \ g(x) = \sum_{j=0}^{n} g_j x^j$$

είναι δύο πολυώνυμα, μέσω των συντελεστών f_j, g_j.

Πρόβλημα 8.4. *Ας είναι V ο γραμμικός χώρος των τμηματικά συνεχών συναρτήσεων στο $[0,1]$ με το συνηθισμένο εσωτερικό γινόμενο $\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\,dx$. Ας είναι επίσης W ο υπόχωρος του V που παράγεται από τα διανύσματα*

$$\chi_{[0,1]}(x),\ \chi_{[\frac{1}{2},1]}(x).$$

Αν $f(t) = t^2 + 1$ είναι ένα στοιχείο του V βρείτε την ορθογώνια προβολή του στο χώρο W.

Πρόβλημα 8.5. *Στο χώρο $C([0,1])$ με εσωτερικό γινόμενο $\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\,dx$ βρείτε την ορθογώνια προβολή της $f(x) = e^x$ στο χώρο που παράγεται από τα διανύσματα*

$$f_1(x) = 1,\ f_2(x) = x.$$

Πρόβλημα 8.6. *Σε ένα χώρο V με εσωτερικό γινόμενο τα διανύσματα $f_1, f_2, \cdots$ είναι γραμμικώς ανεξάρτητα και οι (πεπερασμένοι) γραμμικοί συνδυασμοί τους είναι πυκνοί στο V. Ας είναι $e_1, e_2, \cdots$ η ορθοκανονική ακολουθία που παράγεται από τη διαδικασία Gram–Schmidt. Δείξτε ότι οι (πεπερασμένοι) γραμμικοί συνδυασμοί των e_j είναι επίσης πυκνοί στο V (ένα σύνολο διανυσμάτων λέγεται* πυκνό *σε ένα άλλο σύνολο αν κάθε στοιχείο του άλλου συνόλου μπορεί να προσεγγιστεί όσο καλά θέλουμε από στοιχεία του πρώτου συνόλου).*

8.2. **Η ακολουθία ορθογωνίων πολυωνύμων ως προς μια συνάρτηση βάρους σε ένα διάστημα.** Με δεδομένη την τεράστια σημασία που έχουν οι χώροι πολυωνύμων $\mathcal{P}_n$ στη θεωρία προσέγγισης καταλαβαίνει εύκολα κανείς πόσο σημαντικό είναι το ακόλουθο αποτέλεσμα που μας δίνει ένα απλό ρόπο να βρούμε μια ορθογώνια ακολουθία από μονικά πολυώνυμα όλων των βαθμών.

Υποθέτουμε στα παρακάτω ότι έχουμε σταθεροποιήσει ένα κλειστό και φραγμένο διάστημα $[a,b] \subseteq \mathbb{R}$ και μια θετική συνάρτηση βάρους $w(x)$ πάνω στο διάστημα αυτό, μέσω της οποίας ορίζεται ένα εσωτερικό γινόμενο

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}w(x)\,dx$$

και η αντίστοιχη 2-νόρμα $\|f\|_2^2 = \langle f, f\rangle$.

Θεώρημα 8.2 (Η κατασκευή των ορθογωνίων πολυωνύμων). *Έστω η ακολουθία πολυωνύμων $Q_n(x)$ που ορίζεται ως εξής:*

$$\begin{aligned} Q_0(x) &= 1 \\ Q_1(x) &= (x-a_0)Q_0(x) = x - a_0 \\ Q_{n+1}(x) &= (x-a_n)Q_n(x) - b_nQ_{n-1}(x) \qquad \text{(για } n \geq 1\text{)} \end{aligned}$$

όπου

$$a_n = \frac{\langle xQ_n(x), Q_n(x)\rangle}{\langle Q_n(x), Q_n(x)\rangle}, \quad b_n = \frac{\langle xQ_n(x), Q_{n-1}(x)\rangle}{\langle Q_{n-1}(x), Q_{n-1}(x)\rangle}. \tag{80}$$

Τότε $\deg Q_n = n$, *το $Q_n(x)$ είναι μονικό (δηλ. $Q_n(x) = x^n + \cdots$) και τα πολυώνυμα $Q_n(x)$ είναι ανά δύο ορθογώνια. Επίσης τα πολυώνυμα $Q_n(x)$ είναι πραγματικά πολυώνυμα.*

Πρόβλημα 8.7. *Αφού πρώτα υπολογίσετε και το $Q_2(x)$ αποδείξτε ότι τα πολυώνυμα Q_0, Q_1, Q_2 είναι ανά δύο ορθογώνια.*

Πρόβλημα 8.8. *Αποδείξτε με επαγωγή ως προς n ότι το $Q_n(x)$ είναι μονικό, πραγματικό πολυώνυμο βαθμού n.*

Αποδεικνύουμε την ορθογωνιότητα των $Q_0, \ldots, Q_n$ με επαγωγή ως προς n. Για $n = 0, 1, 2$ αυτό είναι το αντικείμενο του Προβλήματος 8.7. Αν υποθέσουμε ότι τα $Q_0, Q_1, \ldots, Q_n$ είναι ανά δύο ορθογώνια πρέπει, για να ολοκληρώσουμε την επαγωγική απόδειξη, να δείξουμε ότι το Q_{n+1} είναι ορθογώνιο προς τα $Q_0, Q_1, \ldots, Q_n$. Δείχνουμε λοιπόν ότι $\langle Q_{n+1}, Q_k\rangle = 0$ για $k \leq n$ διαχωρίζοντας 3 περιπτώσεις για το k.

Περίπτωση $k = n$:

Χρησιμοποιώντας τον ορισμό του a_n και το γεγονός (επαγωγική ως προς n υπόθεση) ότι $\langle Q_n, Q_{n-1}\rangle = 0$ έχουμε

$$\begin{aligned} \langle Q_{n+1}, Q_n\rangle &= \langle (x-a_n)Q_n - b_nQ_{n-1}, Q_n\rangle \\ &= \langle xQ_n, Q_n\rangle - a_n\langle Q_n, Q_n\rangle - b_n\cancel{\langle Q_{n-1}, Q_n\rangle} \\ &= \langle xQ_n, Q_n\rangle - \frac{\langle xQ_n, Q_n\rangle}{\langle Q_n, Q_n\rangle}\langle Q_n, Q_n\rangle \\ &= \langle xQ_n, Q_n\rangle - \langle xQ_n, Q_n\rangle \\ &= 0. \end{aligned}$$

Περίπτωση $k = n-1$:

$$\begin{aligned} \langle Q_{n+1}, Q_{n-1}\rangle &= \langle xQ_n, Q_{n-1}\rangle - a_n\cancel{\langle Q_n, Q_{n-1}\rangle} - b_n\langle Q_{n-1}, Q_{n-1}\rangle \\ &= \langle xQ_n, Q_{n-1}\rangle - \frac{\langle xQ_n, Q_{n-1}\rangle}{\langle Q_{n-1}, Q_{n-1}\rangle}\langle Q_{n-1}, Q_{n-1}\rangle \\ &= 0. \end{aligned}$$

Περίπτωση $k < n-1$:

$$
\begin{aligned}
\langle Q_{n+1}, Q_k\rangle &= \langle xQ_n, Q_k\rangle - a_n\cancel{\langle Q_n, Q_k\rangle} - b_n\cancel{\langle Q_{n-1}, Q_k\rangle}\\
&= \langle xQ_n, Q_k\rangle\\
&= \langle Q_n, xQ_k\rangle\\
&= 0.
\end{aligned}
$$

Στην προτελευταία ισότητα χρησιμοποιήθηκε η συγκεκριμένη μορφή του εσωτερικού γινομένου η οποία συνεπάγεται την ταυτότητα

$$\langle f(x)g(x), h(x)\rangle = \langle f(x), \overline{g(x)}h(x)\rangle$$

για οποιεσδήποτε συναρτήσεις f, g, h (το χρησιμοποιήσαμε για τη συνάρτηση $g(x) = x$). Τέλος, $\deg(xQ_k) < n$ και άρα $\langle Q_n, xQ_k\rangle = 0$ αφού το Q_n είναι ορθογώνιο (από την επαγωγική μας υπόθεση) προς όλα τα $Q_0, Q_1, \ldots, Q_{n-1}$ άρα και προς όλους τους γραμμικούς τους συνδυασμούς που είναι όλος ο χώρος $\mathcal{P}_{n-1}$.

Η απόδειξη του Θεωρηματος 8.2 είναι πλήρης.

Πρόβλημα 8.9. *Αποδείξτε ότι $b_n > 0$ στο Θεώρημα 8.2.*

Πρόβλημα 8.10. *Αν $p \in \mathcal{P}_n$ ποιοι είναι οι συντελεστές του p ως προς την ορθογώνια βάση $Q_0, Q_1, \ldots, Q_n$ του $\mathcal{P}_n$;*

Πρόβλημα 8.11. *Αν $F_0(x) = 1, F_1(x) = x + C, \cdots$ είναι μια ακολουθία μονικών ορθογωνίων πολυωνύμων με $\deg F_k = k$ τότε δείξτε ότι το $F_n(x)$ είναι το μονικό πολυώνυμο βαθμού n με την ελάχιστη L^2 νόρμα, και είναι επίσης το μοναδικό τέτοιο πολυώνυμο, και άρα η ακολουθία μονικών ορθογωνίων πολυωνύμων ως προς ένα εσωτερικό γινόμενο είναι μοναδική.*

Πρόβλημα 8.12. *Αν T_n είναι τα πολυώνυμα Chebyshev στο $[-1, 1]$ δείξτε ότι τα πολυώνυμα $Q_n = 2^{1-n}T_n$ είναι τα ορθογώνια πολυώνυμα για το διάστημα $[-1, 1]$ και το βάρος*

$$w(x) = \frac{1}{\sqrt{1-x^2}}.$$

Δείξτε επίσης ότι η αναδρομική σχέση (52) για τα πολυώνυμα Chebyshev είναι ίδια με την αναδρομική σχέση που περιγράφεται στο Θεώρημα 8.2.

Θα δείξουμε το ακόλουθο σημαντικό αποτέλεσμα το οποίο αργότερα θα εφαρμόσουμε σε μεθόδους αριθμητικής ολοκλήρωσης.

Θεώρημα 8.3 (Ρίζες ορθογωνίων πολυωνύμων).
Έστω $[a, b]$ ένα κλειστό φραγμένο διάστημα στο $\mathbb{R}$ και $w(x) > 0$ μια συνεχής συνάρτηση βάρους στο $[a, b]$. Ας είναι $Q_0(x), Q_1(x), Q_2(x), \ldots$ η ακολουθία μονικών ορθογωνίων πολυωνύμων για το εσωτερικό γινόμενο

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}w(x)\,dx.$$

Τότε για κάθε $n > 0$ το πολυώνυμο $Q_n(x)$ έχει όλες του τις ρίζες απλές και στο διάστημα (a, b).

Το Θεώρημα 8.3 είναι άμεση συνέπεια του παρακάτω Λήμματος.

Λήμμα 8.1. *Με τους ορισμούς του Θεωρήματος 8.3 αν μια συνάρτηση $f \in C([a, b])$ είναι ορθογώνια προς όλα τα πολυώνυμα $p \in \mathcal{P}_{n-1}$ τότε η $f(x)$ έχει τουλάχιστον n διαφορετικές ρίζες στο (a, b).*

Πράγματι, αφού το Q_n είναι ορθογώνιο προς τα $Q_0, Q_1, \ldots, Q_{n-1}$ είναι και ορθογώνιο προς κάθε γραμμικό συνδυασμό τους, δηλ. προς κάθε $p \in \mathcal{P}_{n-1}$ και, σύμφωνα με το προηγούμενο Λήμμα, έχει n διαφορετικές ρίζες στο (a, b). Αυτές είναι όλες οι ρίζες του Q_n αφού $\deg Q_n = n$.

Για να αποδείξουμε το Λήμμα 8.1 θα χρειαστούμε την παρακάτω πρόταση.

Λήμμα 8.2. *Η $f \in C([a, b])$ είναι ορθογώνια προς το $\mathcal{P}_{n-1}$ αν και μόνο αν υπάρχει συνάρτηση u στο $[a, b]$, n φορές παραγωγίσιμη στο $[a, b]$ και τ.ώ. $u^{(n)} = fw$ και $u^{(k)}(a) = u^{(k)}(b) = 0$ για $k = 0, 1, \ldots, n-1$.*

Είναι πολύ εύκολο να βρει κανείς μια συνάρτηση u της οποίας η n-οστή παράγωγος να είναι η fw. Για παράδειγμα η συνάρτηση $v(x) = \int_a^x f(t)w(t)\,dt$ ικανοποιεί $v'(x) = f(x)w(x)$ και μπορούμε να επαναλάβουμε αυτή τη διαδικασία n φορές ώστε να βρούμε μια τέτοια u. Προσθέτοντας ένα οποιοδήποτε πολυώνυμο $p \in \mathcal{P}_{n-1}$ σε αυτή τη u δεν πρόκειται να αλλάξει τη n-οστή της παράγωγο (αφού $p^{(n)} \equiv 0$) άρα έχουμε επιπλέον n βαθμούς ελευθερίας με τους οποίους εύκολα μπορούμε να ικανοποιήσουμε τις n συνοριακές συνθήκες $u^{(k)}(a) = 0$. Το σημαντικό είναι ότι μπορούμε ταυτόχρονα να ικανοποιήσουμε και τις συνοριακές συνθήκες και στο άλλο άκρο του διαστήματος, πράγμα που δε φαίνεται κατ' αρχήν δυνατό με τους βαθμούς ελευθερίας που έχουμε στη διάθεσή μας, αλλά τελικά μπορούμε να το κάνουμε λόγω της υπόθεσης της ορθογωνιότητας της f προς όλα τα στοιχεία του $\mathcal{P}_{n-1}$.

Έστω λοιπόν u μια συνάρτηση τ.ώ. $u^{(n)} = fw$ και $u^{(k)}(a) = 0$ για $k = 0, 1, 2, \ldots, n-1$. Θα δείξουμε ότι ικανοποιεί και τις άλλες συνοριακές συνθήκες $u^{(k)}(b) = 0$ για $k = 0, 1, 2, \ldots, n-1$.

Θα χρειαστούμε τον παρακάτω τύπο που γενικεύει τον τύπο ολοκλήρωσης κατά μέρη:

$$\int_a^b u^{(n)}(x)v(x)\,dx = \sum_{k=1}^{n}(-1)^{k-1}\, u^{(n-k)}(x)v^{(k-1)}(x)\Big|_{x=a}^{x=b} + (-1)^n \int_a^b u(x)v^{(n)}(x)\,dx. \tag{81}$$

Πρόβλημα 8.13. *Αποδείξτε τον τύπο (81) με επαγωγή ως προς n. Για $n = 1$ έχουμε το συνηθισμένο τύπο ολοκλήρωσης κατά μέρη.*

Αν τώρα $p \in \mathcal{P}_{n-1}$ τότε χρησιμοποιώντας τον τύπο (81) και το ότι $p^{(n)} \equiv 0$ έχουμε

$$\int_a^b f\overline{p}w = \sum_{k=1}^{n}(-1)^{k-1}u^{(n-k)}(b)p^{(k-1)}(b). \tag{82}$$

Όμως οι αριθμοί $p(b), p'(b), p^{(2)}(b), \ldots, p^{(k-1)}(b)$ είναι τελείως στη διαθεσή μας όπως λέει το επόμενο Πρόβλημα.

Πρόβλημα 8.14. *Αν $a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}$ και $b \in \mathbb{R}$ τότε υπάρχει $p \in \mathcal{P}_{n-1}$ τ.ώ. $p^{(k)}(b) = a_k$, για $k = 0, 1, \ldots, n-1$.*

Για να είναι λοιπόν το αριστερό μέλος της (82) ίσο με 0 για κάθε $p \in \mathcal{P}_{n-1}$ ο μόνος τρόπος είναι να είναι όλοι οι συντελεστές $u^{(n-k)}(b) = 0$ για $k = 1, \ldots, n$. Με άλλα λόγια πρέπει και αρκεί $u^{(k)}(b) = 0$ για $k = 0, 1, \ldots, n-1$, και η απόδειξη του Λήμματος 8.2 είναι πλήρης.

Επανερχόμαστε τώρα στην απόδειξη του Λήμματος 8.1. Ας είναι $f \in C([a, b])$ ορθογώνια προς το $\mathcal{P}_{n-1}$ (δηλ. ορθογώνιο προς όλες τα στοιχεία του $\mathcal{P}_{n-1}$). Από το Λήμμα 8.2 έχουμε ότι υπάρχει συνάρτηση $u \in C^n([a, b])$ τ.ώ. $fw = u^{(n)}$ και όλες οι παράγωγοι της u τάξης μικρότερης του n μηδενίζονται στα άκρα του διαστήματος. Αφού $u(a) = u(b)$ από το θεώρημα του Rolle έχουμε ότι η u' έχει κάποια ρίζα στο διάστημα (a, b). Αφού η u' μηδενίζεται στα δύο άκρα και σε ένα ενδιάμεσο σημείο προκύπτει, και πάλι από το θεώρημα του Rolle ότι η $u^{(2)}$ έχει δύο διαφορετικές ρίζες στο (a, b). Συνεχίζοντας κατ' αυτόν τον τρόπο, εφαρμόζοντας δηλ. συνεχώς το θεώρημα του Rolle ώστε να «κερδίζουμε» από μια επιπλέον ρίζα κάθε φορά που ανεβάζουμε την τάξη της παραγώγου της u, καταλήγουμε τελικά ότι η $u^{(n)}$ έχει n διαφορετικές ρίζες στο (a, b). Η απόδειξη του Λήμματος 8.1 είναι πλήρης και άρα το Θεώρημα 8.3 έχει επίσης αποδειχτεί πλήρως.

9. Αριθμητική ολοκλήρωση

9.1. **Απλοί και σύνθετοι κανόνες αριθμητικής ολοκλήρωσης.** Πολύ συχνά δεν είναι δυνατό να υπολογίσουμε ακριβώς ένα ολοκλήρωμα. Π.χ. το ολοκλήρωμα

$$\int_A^B e^{-x^2}\,dx$$

δεν είναι δυνατό να υπολογιστεί ακριβώς (εκτός από ειδικές τιμές των A και B) αφού έχει αποδειχτεί ότι το αόριστο ολοκλήρωμα της e^{-x^2} δε μπορεί να γραφεί σε κλειστό τύπο, χρησιμοποιώντας δηλ. τις συνηθισμένες συναρτήσεις (τριγωνομετρικές, λογαρίθμους, εκθετικές, κλπ) και αλγεβρικές πράξεις. Αναγκαστικά λοιπόν το προσεγγίζουμε με αριθμητικές μεθόδους. Πολλές φορές μάλιστα, όταν το μόνο που μας ενδιαφέρει είναι η τιμή ενός ολοκληρώματος για συγκεκριμένο διάστημα ολοκλήρωσης, είναι προτιμότερο να υπολογίζουμε ένα ολοκλήρωμα αριθμητικά (προσεγγιστικά) πολύ απλά γιατί είναι κατά πολύ ευκολότερο.

Σε αδρές γραμμές μια τέτοια διαδικασία γίνεται σε δυο στάδια. Πρώτα χωρίζουμε το διάστημα ολοκλήρωσης $[A, B]$ σε μικρότερα διαστήματα, για απλότητα ας πούμε ότι χωρίζουμε σε N ίσα διαστήματα, και σε κάθε ένα από τα μικρά αυτά διαστήματα χρησιμοποιούμε ένα «απλό κανόνα αριθμητικής ολοκλήρωσης» για να προσεγγίσουμε το ολοκλήρωμα εκεί μέσα. Η προσέγγιση του όλου ολοκληρώματος ονομάζεται «σύνθετος κανόνας αριθμητικής ολοκλήρωσης».

Με αυτές τις παραδοχές ας δούμε μερικά παραδείγματα απλών κανόνων αριθμητικής ολοκλήρωσης. Όλοι οι κανόνες που θα εξετάσουμε είναι γραμμικοί, μιμούμενοι το πραγματικό (ακριβές) ολοκλήρωμα που είναι επίσης μια γραμμική διαδικασία. Σταθεροποιούμε λοιπόν ένα διάστημα $[a, b]$ και συμβολίζουμε κατ' αρχήν με $I(f)$ το ακριβές ολοκλήρωμα της f στο $[a, b]$

$$I(f) = \int_a^b f(x)\,dx.$$

Η γραμμικότητα του ολοκληρώματος στην οποία αναφερθήκαμε σημαίνει απλά ότι για κάθε δύο συναρτήσεις f και g (ας πούμε συνεχείς στο $[a, b]$) και κάθε δύο σταθερές $\lambda, \mu \in \mathbb{C}$ ισχύει

$$I(\lambda f + \mu g) = \lambda I(f) + \mu I(g).$$

Το ίδιο ισχύει και για τους απλούς κανόνες αριθμητικής ολοκλήρωσης που περιγράφουμε παρακάτω.

Κανόνας αριστερού σημείου

(83) $$I_L(f) = (b - a) f(a).$$

Εδώ η ποσότητα που προσεγγίζει το ολοκλήρωμα $I(f)$ είναι το ολοκλήρωμα που θα είχε η συνάρτηση αν ήταν σταθερή στο $[a, b]$ με τιμή ίση με $f(a)$. Ομοίως ορίζεται φυσικά ο κανόνας δεξιού σημείου.

Κανόνας μέσου σημείου

$$I_M(f) = (b-a) f\left(\frac{a+b}{2}\right). \tag{84}$$

Κανόνας τραπεζίου

$$I_T(f) = (b-a)\frac{f(a)+f(b)}{2}. \tag{85}$$

Κανόνας του Simpson

$$I_S(f) = (b-a)\left(\frac{1}{6}f(a) + \frac{4}{6}f\left(\frac{a+b}{2}\right) + \frac{1}{6}f(b)\right). \tag{86}$$

Παρατηρείστε ότι ολοι αυτοί οι κανόνες είναι της μορφής

$$\widetilde{I}(f) = \sum_{j=1}^{n} A_j f(t_j), \tag{87}$$

όπου $t_1, t_2, \ldots, t_n$ είναι σημεία στο $[a,b]$ και $A_1, A_2, \ldots, A_N$ είναι σταθερές.

Ορισμός 9.1 (Τάξη του κανόνα ολοκλήρωσης)**.**
Τάξη ενός κανόνα αριθμητικής ολοκλήρωσης είναι ο μέγιστος ακέραιος k τ.ώ. ο κανόνας υπολογίζει σωστά το ολοκλήρωμα κάθε πολυωνύμου βαθμού το πολύ k.

Για να βρούμε ποια είναι η τάξη ενός κανόνα $\widetilde{I}$ απλά βρίσκουμε ποιο είναι το μεγαλύτερο k τ.ώ.

$$I(x^r) = \widetilde{I}(x^r), \quad \text{για κάθε } r \in \{0, 1, \ldots, k\}.$$

Πράγματι, λόγω της γραμμικότητας, είναι φανερό ότι αν ο κανόνας μας δουλεύει σωστά πάνω στις συναρτήσεις $1, x, x^2, \ldots, x^k$ τότε δουλεύει σωστά και πάνω σε κάθε γραμμικό συνδυασμό αυτών, δηλ. πάνω σε κάθε πολυώνυμο βαθμού το πολύ k.

Πρόβλημα 9.1. *Αποδείξτε ότι ο κανόνας αριστερού σημείου έχει τάξη 0, ο κανόνας του μέσου σημείου και ο κανόνας του τραπεζίου έχουν τάξη 1 και ο κανόνας του Simpson έχει τάξη 3, σε κάθε διάστημα* $[a,b]$.

Η σημασία που έχει η τάξη ενός κανόνα αριθμητικής ολοκλήρωσης φαίνεται αν εξετάσουμε την απόδοση ενός σύνθετου κανόνα ολοκλήρωσης σε σχέση με την παράμετρο N που δείχνει σε πόσα μικρά διαστήματα χωρίζουμε το διάστημα ολοκλήρωσης.

Θεώρημα 9.1 (Απόδοση ενός σύνθετου κανόνα ολοκλήρωσης τάξης r)**.**
Έστω $[A,B]$ ένα κλειστό και φραγμένο διάστημα και $f \in C^{r+1}([A,B])$ μια συνάρτηση με $r+1$ συνεχείς παραγώγους (r είναι κάποιος ακέραιος). Αν $\widetilde{I}$ είναι ένας απλός κανόνας αριθμητικής ολοκλήρωσης τάξης r τότε αν $\widetilde{I}_N$ είναι ο σύνθετος κανόνας αριθμητικής ολοκλήρωσης που προκύπτει από τον $\widetilde{I}$ (με χωρισμό του διαστήματος σε N ίσα διαστήματα) τότε

$$\left|\int_A^B f - \widetilde{I}_N(f)\right| \le K' \frac{1}{N^{r+1}}, \tag{88}$$

όπου η σταθερά K' δεν εξαρτάται από το N παρά μόνο από την f και τον απλό κανόνα ολοκλήρωσης $\widetilde{I}$.

Το διάστημα $[A, B]$ υποδιαιρείται σε N ίσα διαστήματα μέσω των σημείων

$$x_j = A + j\frac{B-A}{N}, \quad j = 0, 1, \ldots, N.$$

Αν συμβολίσουμε με $\widetilde{I}(f, x_j, x_{j+1})$ το αποτέλεσμα του απλού κανόνα ολοκλήρωσης εφαρμοσμένου στη συνάρτηση f στο διάστημα $[x_j, x_{j+1}]$ τότε έχουμε από την τριγωνική ανισότητα

$$\left|\int_A^B f - \widetilde{I}_N(f)\right| \le \sum_{j=0}^{N-1} \left|\int_{x_j}^{x_{j+1}} f - \widetilde{I}(f, x_j, x_{j+1})\right| \tag{89}$$

οπότε θα σταθεροποιήσουμε ένα διάστημα $[a, b] = [x_j, x_{j+1}]$ και θα εκτιμήσουμε το σφάλμα που προκύπτει μέσα σε αυτό το διάστημα από τη χρήση του απλού κανόνα ολοκλήρωσης. Κατόπιν θα αθροίσουμε τα επιμέρους σφάλματα και θα προκύψει ένα φράγμα για το συνολικό σφάλμα μέσω της (89).

Σταθεροποιούμε λοιπόν ένα απλό κανόνα ολοκλήρωσης τον οποίο τον περιγράφουμε στο διάστημα $[0, 1]$

$$\widetilde{I}(f, 0, 1) = \sum_{j=1}^{n} A_j f(t_j)$$

όπου A_j κάποιες σταθερές και t_j κάποια σημεία στο $[0, 1]$. Ο κανόνας αυτός, εφαρμοζόμενος σε ένα τυχαίο διάστημα $[a, b]$, παίρνει τη μορφή

$$\widetilde{I}(f, a, b) = \sum_{j=1}^{n} (b-a) A_j f(a + t_j(b-a)). \tag{90}$$

Το σημαντικό στην απόδειξη είναι να γράψουμε την f στο διάστημα αυτό ως ένα πολυώνυμο βαθμού $\le r$ συν κάποιο σφάλμα. Χρησιμοποιούμε γι' αυτό το θεώρημα του Taylor με κέντρο το a που μας λέει ότι

$$f(x) = p(x) + R(x)$$

όπου $p(x)$ είναι το πολυώνυμο Taylor της f βαθμού r

$$p(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(r)}(a)}{r!}(x-a)^r,$$

και $R(x)$ είναι το σφάλμα για το οποίο ισχύει ότι για κάθε $x \in [a, b]$ υπάρχει ένα $\xi \in (a, b)$ τ.ώ. να ισχύει

$$R(x) = \frac{f^{(r+1)}(\xi)}{(r+1)!}(x-a)^{r+1}. \tag{91}$$

Αφού ο κανόνας ολοκλήρωσης που χρησιμοποιούμε έχει τάξη r το πολυώνυμο p ολοκληρώνεται ακριβώς άρα το σφάλμα στην ολοκλήρωση της f οφείλεται αποκλειστικά στην αριθμητική ολοκλήρωση της συνάρτησης $R(x)$, κι έτσι έχουμε

$$\begin{aligned}
\left|\int_a^b f - \widetilde{I}(f,a,b)\right| &\le \left|\int_a^b R - \widetilde{I}(R,a,b)\right| \\
&\le \left|\int_a^b R\right| + \left|\widetilde{I}(R,a,b)\right| && \text{(τριγωνική ανισότητα)} \\
&\le (b-a)\|R\|_\infty + \sum_{j=1}^{n}(b-a)|A_j|\|R\|_\infty && \text{(από την (90))} \\
&\le (b-a)K\|R\|_\infty
\end{aligned}$$

όπου $K = 1 + \sum_{j=1}^n |A_j|$ είναι μια θετική σταθερά που δεν εξαρτάται από το διάστημα αλλά μόνο από τον κανόνα ολοκλήρωσης.

Εφαρμόζοντας τώρα την (91) παίρνουμε

$$\|R\|_\infty \le \frac{\|f^{(r+1)}\|_\infty}{(r+1)!}(b-a)^{r+1} \le \frac{\|f^{(r+1)}\|_\infty}{(r+1)!}\frac{(B-A)^{r+1}}{N^{r+1}}.$$

το οποίο μας δίνει

$$\left|\int_a^b f - \widetilde{I}(f,a,b)\right| \le K\frac{(B-A)^{r+2}}{N^{r+2}}$$

για το σφάλμα της ολοκλήρωσης στο μικρό διάστημα $[a,b]$. Αθροίζοντας για τα N διαφορετικά μικρά διαστήματα $[a,b] = [x_j, x_{j+1}]$ παίρνουμε

$$\left|\int_A^B f - \widetilde{I}_N(f)\right| \le \frac{K'}{N^{r+1}} \tag{92}$$

όπου η σταθερά $K' = K(B-A)^{r+2}$ δεν εξαρτάται από το N, και η απόδειξη του Θεωρήματος 9.1 είναι πλήρης.

Η σημασία λοιπόν της τάξης ενός απλού κανόνα ολοκλήρωσης φαίνεται όταν χρησιμοποιήσουμε τον κανόνα αυτό σε ένα σύνθετο κανόνα ολοκλήρωσης και παίρνουμε ολοένα και μεγαλύτερες τιμές του N. Το άνω φράγμα για το σφάλμα ολοκλήρωσης φθίνει σαν την $r+1$ δύναμη του $1/N$. Για παράδειγμα αν χρησιμοποιούμε ένα κανόνα ολοκλήρωσης τάξης 0 (π.χ. τον κανόνα I_L του αριστερού σημείου) τότε αν θέλουμε να υποδεκαπλασιάσουμε το εγγυημένο άνω φράγμα για το σφάλμα μας τότε πρέπει να πολλαπλασιάσουμε το N που χρησιμοποιήσαμε με το 10. Αν χρησιμοποιήσουμε ένα κανόνα τάξης 3, όπως ο κανόνας I_S του Simpson, τότε αρκεί να πολλαπλασιάσουμε το N μόνο με τον αριθμό $10^{1/4} = 1.77827941$.

9.2. **Κανόνες ολοκλήρωσης τύπου Gauss.** Στην παράγραφο αυτή θα δούμε τις μεθόδους ολοκλήρωσης τύπου Gauss, των οποίων η υψηλή τάξη οφείλεται σε κατάλληλη επιλογή των σημείων στα οποία υπολογίζουμε τη συνάρτηση f που προσπαθούμε να ολοκληρώσουμε αριθμητικά. Κοιτάμε το πρόβλημα λίγο γενικότερα από πριν και, δεδομένου ενός κλειστού και φραγμένου διαστήματος $[a, b]$ και μιας θετικής (και συνεχούς) συνάρτησης βάρους $w(x) > 0$ ορισμένης πάνω στο $[a, b]$ προσπαθούμε να υπολογίσουμε αριθμητικά το ολοκλήρωμα της f ως προς το βάρος w, δηλ. την ποσότητα

$$I(f) = \int_a^b f(x)w(x)\,dx.$$

Οι απλοί κανόνες αριθμητικής ολοκλήρωσης που χρησιμοποιούμε είναι της μορφής

$$\widetilde{I}(f) = \sum_{j=1}^{n} A_j f(x_j) \tag{93}$$

όπου $x_1, \ldots, x_n \in [a, b]$ και A_j είναι σταθερές.

Το ερώτημα που απαντάμε πρώτα είναι «αν κάποιος έχει επιλέξει τα n σημεία $x_1, \ldots, x_n$ για μας ποια πρέπει να είναι η επιλογή των A_j;»

Θεώρημα 9.2. *Οι συντελεστές A_j στον κανόνα (93) μπορούν πάντα να επιλεγούν ώστε ο κανόνας να έχει τάξη τουλάχιστον $n - 1$.*

Για να πετύχουμε να ολοκληρώνουμε ακριβώς κάθε πολυώνυμο βαθμού $n - 1$ (το οποίο καθορίζεται me n αριθμούς) αρκεί να έχουμε στη διάθεσή μας n βαθμούς ελευθερίας. Αυτό είναι ουσιαστικά το περιεχόμενο τυ Θεωρήματος 9.2.

Ο πιο εύκολος τρόπος απόδειξης είναι χρησιμοποιώντας τα πολυώνυμα παρεμβολής Lagrange που αντιστοιχούν στα σημεία $x_1, \ldots, x_n$:

$$\ell_i(x) = \prod_{\substack{j=1,2,\ldots,n\\ j\neq i}} \frac{x - x_j}{x_i - x_j}, \quad i = 1, 2, \ldots, n, \tag{94}$$

τα οποία έχουν την ιδιότητα

$$\ell_i(x_j) = \delta_{i,j} := \begin{cases} 1 & (\text{για } i = j) \\ 0 & (\text{για } i \neq j) \end{cases}.$$

Αν συμβολίσουμε με $L(x)$ το πολυώνυμο, βαθμού $n - 1$, που παρεμβάλλει τις τιμές της f στα σημεία $x_1, \ldots, x_n$

$$L(x) = \sum_{i=1}^{n} f(x_i)\ell_i(x)$$

μπορούμε να ορίσουμε τον επιθυμητό κανόνα ολοκλήρωσης από τον τύπο

$$\widetilde{I}(f) = I(L).$$

Με άλλα λόγια, για να προσεγγίσουμε αριθμητικά το ολοκλήρωμα της f αυτό που κάνουμε είναι ότι πρώτα βρίσκουμε το πολυώνυμο που έχει τις ίδιες τιμές με την f στα σημεία x_j και έπειτα υπολογίζουμε το *ακριβές* ολοκλήρωμα αυτού του πολυωνύμου. Αυτή η τιμή αποτελεί και την προσέγγισή μας.

Ως στρατηγική ακούγεται λογική, αλλά είναι αυτός ο τρόπος αριθμητικής ολοκλήρωσης της μορφής (93); Είναι εύκολο να δούμε πως ναι αφού (η ολοκλήρωση είναι μια γραμμική πράξη)

$$I(L) = \sum_{i=1}^{n} I(\ell_i) f(x_i)$$

και άρα μπορούμε να πάρουμε

$$A_i = I(\ell_i) = \int_a^b \ell_i(x) w(x)\, dx$$

για τους συντελεστές μας.

Τέλος, ο κανόνας μας έχει τάξη $n-1$ μια και αν η f είναι πολυώνυμο βαθμού $\le n-1$ τότε (μοναδικότητα του πολυωνύμου παρεμβολής, Πρόβλημα 6.1) $f = L$ και άρα ο κανόνας μας ολοκληρώνει την f ακριβώς.

Το Θεώρημα 9.2 αποτελεί απλά ένα κάτω φράγμα για την δυνατή τάξη ενός κανόνα με n σημεία. Έχουμε ήδη δει παραδείγματα (ο κανόνας του Simpson, με 3 σημεία και τάξη 3, παραπάνω από το 2 που προβλέπεται από το Θεώρημα 9.2) όπου η τάξη μπορεί να είναι καλύτερη από $n-1$. Βεβαίως υπάρχουν και παραδείγματα (π.χ. ο κανόνας του τραπεζίου, με 2 σημεία και τάξη 1) όπου η τάξη είναι ακριβώς $n-1$. Θα δούμε στο υπόλοιπο της παραγράφου ότι επιλέγοντας κατάλληλα τα $x_1, \ldots, x_n$ μπορούμε να πάρουμε κανόνες ολοκλήρωσης τάξης $2n-1$, μια σημαντική βελτίωση.

Θεώρημα 9.3 (Κανόνες ολοκλήρωσης Gauss)**.**
Ας είναι $[a,b]$ ένα κλειστό και φραγμένο διάστημα και $w(x) \in C([a,b])$ μια αυστηρά θετική, συνεχής συνάρτηση. Ας είναι ακόμη $Q_0(x) = 1, Q_1(x), Q_2(x), \ldots$ η ακολουθία μονικών ορθογωνίων πολυωνύμων ως προς το βάρος $w(x)$ στο διάστημα $[a,b]$, και

$$x_1^{(n)}, x_2^{(n)}, \ldots, x_n^{(n)} \in (a,b)$$

τα μηδενικά του Q_n (είναι όλα διαφορετικά και στο (a,b) από το Θεώρημα 8.3). Τότε ένας απλός κανόνας αριθμητικής ολοκλήρωσης της μορφής

$$\widetilde{I}(f) = \sum_{j=1}^{n} A_j f(x_j^{(n)}) \quad (A_j \text{ σταθερές}) \tag{95}$$

ο οποίος είναι τάξης τουλάχιστον $n-1$ είναι αναγκαστικά τάξης τουλάχιστον $2n-1$.

Απόδειξη. Έστω $p \in \mathcal{P}_{2n-1}$. Διαιρούμε το πολυώνυμο p με το πολυώνυμο Q_n και έστω q το πηλίκο και r το υπόλοιπο (αυτό σημαίνει ότι $\deg r < \deg Q_n = n$)

$$p = Q_n q + r.$$

Αφού $\deg p < 2n$ και $\deg Q_n = n$ έπεται ότι $\deg p < n$, και άρα $p, r \in \mathcal{P}_{n-1}$. Κάνουμε την παρατήρηση ότι το Q_n είναι ορθογώνιο προς όλα τα πολυώνυμα του $\mathcal{P}_{n-1}$ αφού είναι ορθογώνιο προς κάθε στοιχείο μιας βάσης του $\mathcal{P}_{n-1}$, αυτής που απαρτίζεται από τα πολυώνυμα $Q_0(x), Q_1(x), \ldots, Q_{n-1}(x)$. Άρα

$$\begin{aligned}
\int_a^b p(x)w(x)\,dx &= \int_a^b Q_n(x)q(x)w(x)\,dx + \int_a^b r(x)w(x)\,dx \\
&= \langle Q_n, q\rangle + \int_a^b r(x)w(x)\,dx \\
&= \int_a^b r(x)w(x)\,dx && (\text{αφού } Q_n \perp \mathcal{P}_{n-1}) \\
&= \widetilde{I}(r) && (\text{αφού ο } \widetilde{I} \text{ είναι ακριβής στο } \mathcal{P}_{n-1}) \\
&= \widetilde{I}(qQ_n) + \widetilde{I}(r) \\
&= \widetilde{I}(p).
\end{aligned}$$

Η προτελευταία ισότητα παραπάνω προκύπτει επειδή το πολυώνυμο Q_n μηδενίζεται στα σημεία όπου ο κανόνας $\widetilde{I}$ υπολογίζει την προς ολοκλήρωση συνάρτηση. Έχουμε λοιπόν δείξει ότι $\widetilde{I}(p) = \int_a^b p(x)w(x)\,dx$ για κάθε πολυώνυμο p βαθμού $\leq 2n-1$ και η απόδειξη είναι πλήρης. □

Πόρισμα 9.1. *Για κάθε φραγμένο και κλειστό διάστημα $[a,b]$ και συνάρτηση βάρους $w(x) \in C([a,b])$, $w(x) > 0$, και κάθε ακέραιο $n \geq 1$ υπάρχει απλός κανόνας ολοκλήρωσης της μορφής (95) που έχει τάξη τουλάχιστον $2n-1$.*

Απόδειξη. Επιλέγουμε ως $x_1, \ldots, x_n$ να είναι οι ρίζες του ορθογωνίου πολυωνύμου Q_n που αντιστοιχεί στο διάστημα $[a,b]$ και στο βάρος $w(x)$ και επιλέγουμε τους συντελεστές A_j, $j = 1, 2, \ldots, n$, χρησιμοποιώντας το Θεώρημα 9.2. Ως συνέπεια του Θεωρήματος 9.3 ο κανόνας που φτιάξαμε έχει τάξη τουλάχιστον $2n-1$. □

Παραδείγματα. Αν ακολουθήσουμε τη διαδικασία που περιγράφεται στο Πόρισμα 9.1 για το διάστημα $[-1,1]$ και το βάρος $w(x) \equiv 1$ παίρνουμε, αφού πρώτα υπολογίσουμε τα αντίστοιχα ορθογώνια πολυώνυμα Q_0, Q_1, Q_2, Q_3 με βάση τον αλγόριθμο που περιγράφεται στο Θεώρημα 8.2 (εδώ φαίνεται η αξία της αναδρομικής σχέσης που περιγράφεται σε αυτό το Θεώρημα) και βρούμε τις ρίζες των Q_2 και Q_3 τους κανόνες Gauss με $n = 2$ και $n = 3$ σημεία αντίστοιχα και με τάξη ≥ 3 και ≥ 5 αντίστοιχα

$$\widetilde{I}_2(f) = 2 \cdot \left(\frac{1}{2} f\left(\frac{-1}{\sqrt{3}}\right) + \frac{1}{2} f\left(\frac{1}{\sqrt{3}}\right)\right) \tag{96}$$

και

$$\widetilde{I}_3(f) = 2 \cdot (0.5555 f(-0.7745) + 0.8888 f(0) + 0.5555 f(0.7745)). \tag{97}$$

Για την περίπτωση $n = 2$ οι ρίζες του πολυωνύμου Q_2 υπολογίζονται πολύ εύκολα σε κλειστή μορφή. Αυτό είναι εφικτό και για το πολυώνυμο Q_3 (που είναι 3ου βαθμού) αλλά ήδη η εύρεση των ριζών σε κλειστή μορφή γίνεται αρκετά πολύπλοκη και αμφιβόλου χρησιμότητας, αφού και για να υπολογίσουμε μια απλή ποσότητα όπως το $\sqrt{3}$ πάλι σε αριθμητικές (προσεγγιστικές) μεθόδους θα καταλήξουμε. Η εύρεση των ριζών στις περισσότερες περιπτώσεις γίνεται με αριθμητικές μεθόδους

Τα ορθογώνια πολυώνυμα για το $[-1,1]$ και βάρος $w \equiv 1$ ονομάζονται πολυώνυμα Legendre. Τα τρία πρώτα είναι:

$$\begin{aligned} Q_0(x) &= 1 \\ Q_1(x) &= x \\ Q_2(x) &= \frac{3}{2}x^2 - \frac{1}{2} \\ Q_3(x) &= \frac{5}{2}x^3 - \frac{3}{2}x. \end{aligned}$$

Πρόβλημα 9.2. *Ο κανόνας (96) και ο κανόνας Simpson (86) είναι της ίδιας τάξης 3. Ο κανόνας (96) χρησιμοποιεί δύο σημεία (δύο υπολογισμούς της f στο διάστημα ολοκλήρωσης) ενώ ο κανόνας Simpson χρησιμοποιεί τρία σημεία. Φαίνεται λοιπόν ότι ο κανόνας (96) είναι προτιμότερος από τον κανόνα του Simpson μια και υπερτερεί αυτού σε ταχύτητα (πλήθος υπολογισμών της f) για το ίδιο αποτέλεσμα. Λαμβάνοντας υπόψιν ότι τελικά χρησιμοποιούμε σύνθετους κανόνες ολοκλήρωσης (υποδιαιρούμε το διάστημα ολοκλήρωσης σε N ίσα διαστήματα και σε καθένα από αυτά χρησιμοποιούμε τον απλό κανόνα ολοκλήρωσης) δείξτε ότι αυτη η διαφορά στην ταχύτητα εξανεμίζεται όταν το N είναι μεγάλο.*

10. Το θεώρημα Stone–Weierstrass

10.1. **Άλγεβρες συναρτήσεων.** Θα δούμε σε αυτό το Κεφάλαιο μια πολύ γενική επέκταση του θεωρήματος προσέγγισης του Weierstrass που λέει ότι οι συνεχείς συναρτήσεις σε ένα κλειστό φραγμένο διάστημα μπορούν να προσεγγιστούν ομοιόμορφα από πολυώνυμα. Για απλότητα θα υποθέσουμε γνωστό το θεώρημα προσέγγισης του Weierstrass στην απόδειξή μας αυτού του γενικότερου θεωρήματος, του θεωρήματος Stone–Weierstrass, αν και χωρίς πολύ κόπο θα μπορούσαμε να παρακάμψουμε τη χρήση του θεωρήματος Weierstrass στην απόδειξη του Θεωρήματος 10.1 παρακάτω και να έχουμε έτσι μια εναλλακτική απόδειξη του θεωρήματος Weierstrass ως απλό πόρισμα του θεωρήματος Stone–Weierstrass.

Μια *άλγεβρα* A είναι ένας διανυσματικός χώρος στον οποίο υπάρχει επιπλέον ορισμένη μια πράξη πολλαπλασιασμού των στοιχείων του διανυσματικού χώρου μεταξύ τους (πέρα από τον πολλαπλασιασμό στοιχείων του διανυσματικού χώρου με αριθμούς που είναι ορισμένος σε κάθε διανυσματικό χώρο). Πιο συγκεκριμένα η πράξη αυτή στέλνει κάθε δύο στοιχεία f, g του διανυσματικού χώρου σε ένα άλλο στοιχείο που το συμβολίζουμε με fg ή $f \cdot g$ και η πράξη αυτή ικανοποιεί τα παρακάτω αξιώματα σε σχέση με τις άλλες συνηθισμένες πράξεις του διανυσματικού χώρου (πρόσθεση στοιχείων του και πολλαπλασιασμό με τα στοιχεία του σώματος, συνήθως $\mathbb{R}$ ή $\mathbb{C}$):

(1) $(fg)h = f(gh)$ για κάθε $f, g, h \in A$,
(2) $(f+g)h = fh + gh$ για κάθε $f, g, h \in A$,
(3) $\lambda(fg) - (\lambda f)g = f(\lambda g)$ για όλους τους αριθμούς λ και $f, g \in A$.

Με άλλα λόγια το σύνολο A, πέρα από τη δομή του διανυσματικού χώρου πάνω από ένα σώμα K (το $\mathbb{R}$ ή το $\mathbb{C}$ συνήθως) έχει και δομή δακτυλίου όπου η πράξη του πολλαπλασιασμού είναι συμβατή με τις πράξεις του διανυσματικού χώρου.

Αν ικανοποιείται επιπλέον και η ιδιότητα $fg = gf$ για κάθε $f, g \in A$ τότε μιλάμε για *αντιμεταθετική άλγεβρα*. Πρόκειται για *άλγεβρα με μοναδιαίο στοιχείο* αν υπάρχει $e \in A$ (το μοναδιαίο στοιχείο) τέτοιο ώστε $fe = ef = f$ για κάθε $f \in A$.

Αν ο διανυσματικός χώρος υπόβαθρο της άλγεβρας A είναι εφοδιασμένος με νόρμα $\|\cdot\|$ τότε απαιτούμε να ισχύει

$$\|fg\| \leq \|f\|\|g\|$$

και μιλάμε τότε για μια *άλγεβρα με νόρμα*. (Αν ο χώρος με αυτή τη νόρμα είναι πλήρης τότε μιλάμε για μια *Banach άλγεβρα*). Αν $B \subseteq A$ είναι επίσης μια άλγεβρα (με τις πράξεις της A) τότε λέμε ότι η B είναι *υποάλγεβρα* της A (με άλλα λόγια B είναι διανυσματικός υπόχωρος που είναι κλειστός και ως προς την πράξη πολλαπλασιασμού των στοιχείων του).

Πρόβλημα 10.1. *Σε μια άλγεβρα με νόρμα αποδείξτε ότι ο πολλαπλασιασμός στοιχείων της είναι συνεχής: αν $f_n \to f$ και $g_n \to g$ τότε $f_n g_n \to fg$.*

Πρόβλημα 10.2. *Σε μια άλγεβρα A με νόρμα δείξτε ότι η κλειστότητα $\overline{B}$ μιας υποάλγεβρας B είναι επίσης υποάλγεβρα της A.*

Πρόβλημα 10.3. *Δείξτε ότι οι πολυωνυμικές συναρτήσεις στο $[a, b]$ αποτελούν μια υποάλγεβρα της $C([a, b])$. Ομοίως οι πολυωνυμικές συναρτήσεις πάνω σε ένα σύνολο $K \subseteq \mathbb{R}^n$ αποτελούν υποάλγεβρα της $C(K)$.*

Παρακάτω δίνουμε μερικά παραδείγματα αλγεβρών (μας ενδιαφέρουν κυρίως οι άλγεβρες συναρτήσεων αλλά όχι μόνο) και αφήνουμε στον αναγνώστη να επαληθεύσει ότι πληρούνται τα διάφορα αξιώματα της άλγεβρας.

Παραδείγματα.

(1) Ο χώρος $\mathbb{R}^n$ με πολλαπλασιασμό κατά συνιστώσες και νόρμα την $\|\cdot\|_\infty$ είναι αντιμεταθετική Banach άλγεβρα με μοναδιαίο στοιχείο το $(1, 1, \ldots, 1)$.

Πρόβλημα 10.4. *Δείξτε ότι όλες οι υποάλγεβρες του $\mathbb{R}^2$ είναι οι*

$$\{(0,0)\},\ \{(x,x) : x \in \mathbb{R}\},\ \{(x,0) : x \in \mathbb{R}\},\ \{(0,x) : x \in \mathbb{R}\},\ \mathbb{R}^2.$$

Κάθε υποάλγεβρα πρέπει να είναι και γραμμικός υπόχωρος.

Πρόβλημα 10.5. *Στο Παράδειγμα 1 παραπάνω είναι το $\mathbb{R}^n$ άλγεβρα με νόρμα αν χρησιμοποιήσουμε την ℓ^1 ή την ℓ^2 νόρμα αντί για την ℓ^∞;*

(2) Αν X είναι οποιοδήποτε σύνολο και $B(X)$ είναι οι φραγμένες πραγματικές συναρτήσεις $X \to \mathbb{R}$ εφοδιασμένες με την νόρμα $\|\cdot\|_\infty$ και με πολλαπλασιασμό κατά σημείο (ο συνηθισμένος πολλαπλασιασμός συναρτήσεων δηλαδή) τότε η $B(X)$ είναι αντιμεταθετική άλγεβρα με μοναδιαίο τη συνάρτηση που είναι η σταθερά 1. Οι σταθερές συναρτήσεις πάνω στο X σχηματίζουν μια υποάλγεβρα της $B(X)$ που είναι «ισόμορφη» με το $\mathbb{R}$ ως άλγεβρα (δακτύλιο).

(3) Αν X είναι μετρικός χώρος και $C(X)$ είναι οι *συνεχείς* συναρτήσεις $X \to \mathbb{R}$ με τον κατά σημείο πολλαπλασιασμό τότε $C(X)$ είναι αντιμεταθετική άλγεβρα με μοναδιαίο (τη σταθερά 1) αλλά δεν έχει νόρμα. Αν $C_b(X)$ είναι οι φραγμένες συνεχείς συναρτήσεις τότε η άλγεβρα $C_b(X)$ είναι υποάλγεβρα και της $C(X)$ και της $B(X)$ και είναι εφοδιασμένη με την $\|\cdot\|_\infty$ νόρμα. Αν ο χώρος X είναι συμπαγής τότε όλες οι συνεχείς συναρτήσεις είναι φραγμένες και $C_b(X) = C(X)$ οπότε $C(X)$ είναι υποάλγεβρα της $B(X)$ σε αυτή την περίπτωση.

(4) Η άλγεβρα των πολυωνύμων περιορισμένων πάνω σε ένα διάστημα $[a, b]$ (με την $\|\cdot\|_\infty$ νόρμα) είναι μια πυκνή υποάλγεβρα της $C([a, b])$. Ομοίως τα τριγωνομετρικά πολυώνυμα είναι μια πυκνή υποάλγεβρα της άλγεβρας $C^{2\pi}$ των 2π-περιοδικών συνεχών συναρτήσεων. Η πυκνότητα αυτών των υποαλγεβρών είναι το θεώρημα προσέγγισης του Weierstrass για αλγεβρικά και τριγωνομετρικά πολυώνυμα αντίστοιχα.

(5) Ας είναι $\mathbb{R}^{n\times n}$ ο γραμμικός χώρος (με συντελεστές από το $\mathbb{R}$) των $n\times n$ πραγματικών πινάκων. Αν δούμε κάθε πίνακα $T\in\mathbb{R}^{n\times n}$ ως ένα γραμμικό τελεστή $\mathbb{R}^n\to\mathbb{R}^n$, και υποθέσουμε ότι το $\mathbb{R}^n$ το έχουμε εφοδιάσει με μια νόρμα $\|\cdot\|_\alpha$ τότε η *νόρμα τελεστή* του T ορίζεται ως
$$\|T\| = \sup \frac{\|Tx\|_\alpha}{\|x\|_\alpha} :\ x\in\mathbb{R}^n\setminus\{0\}.$$

Πρόβλημα 10.6. *Δείξτε ότι η νόρμα που μόλις ορίσαμε ικανοποιεί όντως τα αξιώματα μιας νόρμας.*

Μπορούμε να δούμε το διανυσματικό χώρο $\mathbb{R}^{n\times n}$ ως μια άλγεβρα με πολλαπλασιασμό το συνηθισμένο πολλαπλασιασμό πινάκων (που αντιστοιχεί βεβαίως στη σύνθεση των αντιστοίχων τελεστών). Τότε αυτή η άλγεβρα είναι μια άλγεβρα με νόρμα τη νόρμα τελεστή που ορίσαμε παραπάνω. Η κρίσιμη ιδιότητα
$$\|ST\| \le \|S\|\cdot\|T\|$$
αποδεικνύεται εύκολα αν παρατηρήσει κανείς ότι προκύπτει από την ανισότητα
$$\|STx\|_\alpha \le \|S\|\cdot\|T\|\cdot\|x\|_\alpha.$$

Πρόβλημα 10.7. *Δείξτε ότι η υποάλγεβρα του $C([a,b])$ που παράγεται από τις συναρτήσεις 1 και x είναι η άλγεβρα των πολυωνύμων. (Η άλγεβρα που παράγεται από κάποια στοιχεία μιας άλγεβρας είναι η ελάχιστη υποάλγεβρα που περιέχει τα στοιχεία. Με άλλα λόγια είναι η τομή όλων των υποαλγεβρών που περιέχουν τα στοιχεία αυτά.)*

Οι παρακάτω δύο ιδιότητες είναι σημαντικές για άλγεβρες συναρτήσεων.

(1) Μια άλγεβρα A πραγματικών συναρτήσεων επί του X *διαχωρίζει σημεία* αν για κάθε διαφορετικά $x,y\in X$ υπάρχει $f\in A$ τ.ώ. $f(x)\neq f(y)$.
(2) Μια άλγεβρα A πραγματικών συναρτήσεων επί του X *δε μηδενίζεται πουθενά στο* X αν δεν υπάρχει $x\in X$ τ.ώ. $f(x)=0$ για κάθε $f\in A$.

Παραδείγματα.

(1) Κάθε υποάλγεβρα της $C([a,b])$ που περιέχει τις δύο συναρτήσεις 1 και x διαχωρίζει σημεία και δε μηδενίζεται πουθενά στο $[a,b]$.
(2) Η υποάλγεβρα της $C([-1,1])$ που αποτελείται από τις άρτιες συναρτήσεις δε μηδενίζεται πουθενά στο $[-1,1]$ αλλά δε διαχωρίζει σημεία.
(3) Η υποάλγεβρα της $C([-1,1])$ που αποτελείται από τις συναρτήσεις που μηδενίζονται στο 0 διαχωρίζει σημεία στο $[-1,1]$ αλλά μηδενίζεται στο 0.

10.2. **Το θεώρημα Stone–Weierstrass για πραγματικές συναρτήσεις.**

Θεώρημα 10.1. *Αν X συμπαγής μετρικός χώρος και $A\subseteq C(X)$ υποάλγεβρα που διαχωρίζει σημεία και δε μηδενίζεται πουθενά τότε η A είναι πυκνή στην $C(X)$. Κάθε συνεχής συνάρτηση δηλ. πάνω στο X μπορεί να προσεγγιστεί ομοιόμορφα από συναρτήσεις στην A.*

Απόδειξη του Θεωρήματος 10.1. Θα χρειαστούμε μερικά λήμματα για την απόδειξη του Θεωρήματος 10.1.

Λήμμα 10.1. *Αν A άλγεβρα πραγματικών συναρτήσεων πάνω στο X που διαχωρίζει σημεία και δε μηδενίζεται πουθενά τότε για κάθε διαφορετικά $x,y\in X$ και κάθε δύο $a,b\in\mathbb{R}$ υπάρχει $f\in A$ τ.ώ.*
$$f(x)=a,\quad f(y)=b.$$

Απόδειξη. Έστω διαφορετικά $x, y \in X$. Το σύνολο

$$\tilde{A} = \{(f(x), f(y)) : f \in A\}$$

είναι υποάλγεβρα του $\mathbb{R}^2$, διαφορετική από τις υποάλγεβρες

$$\{(0,0)\}, \ \{(x,x) : x \in \mathbb{R}\}, \ \{(x,0) : x \in \mathbb{R}\}, \ \{(0,x) : x \in \mathbb{R}\}.$$

Άρα $\tilde{A} = \mathbb{R}^2$ (δείτε Πρόβλημα 10.4). Άρα για κάθε $(a,b) \in \mathbb{R}^2$ υπάρχει $f \in A$ τ.ώ. $(a,b) = (f(x), f(y))$. □

Με $\overline{A}$ συμβολίζουμε την κλειστότητα της άλγεβρας με νόρμα A.

Λήμμα 10.2. *Αν A υποάλγεβρα της $C(X)$ και $f \in A$ τότε και $|f| \in \overline{A}$. Επίσης αν $f, g \in A$ τότε και οι συναρτήσεις*

$$f \vee g = \max\{f,g\} \text{και } f \wedge g = \min\{f,g\}$$

ανήκουν στην $\overline{A}$.

Απόδειξη. Έστω $\epsilon > 0$, $f \in A$ και ας είναι $p(x) = \sum_{k=0}^{n} a_k x^k$ ένα πολυώνυμο τ.ώ. να έχουμε

$$||x| - p(x)| \leq \epsilon \text{ για } x \in J = [-\|f\|_\infty, \|f\|_\infty].$$

Ειδικότερα $|a_0| = |p(0)| \leq \epsilon$. (Το $p(x)$ υπάρχει από το θεώρημα προσέγγισης του Weierstrass.)

Αφού $f(x) \in J$ έχουμε λοιπόν $||f(x)| - p(f(x))| \leq \epsilon$ για κάθε $x \in X$. Είναι φανερό ότι $p(f(x)) \in A$. Γράφουμε

$$g(x) = p(f(x)) - a_0 = a_1 f(x) + a_2 f(x)^2 + \cdots + a_n f(x)^n \in A$$

και έχουμε

$$|f(x) - g(x)| \leq |a_0| + \epsilon \leq 2\epsilon.$$

Αφού το ϵ είναι οποιοαδήποτε έπεται ότι $|f| \in \overline{A}$.

Για να δούμε ότι $f \vee g$ και $f \wedge g$ ανήκουν στην $\overline{A}$ όταν οι *πραγματικές συναρτήσεις* $f, g \in A$ δεν έχουμε παρά να εφαρμόσουμε αυτό που έχουμε αποδείξει μέχρι στιγμής για την απόλυτη τιμή αφού παρατηρήσουμε ότι

$$f \vee g = \frac{1}{2}(f + g + |f-g|) \text{ και } f \wedge g = \frac{1}{2}(f + g - |f-g|).$$

□

Μπορούμε τώρα να υποθέσουμε ότι $A = \overline{A}$, αφού οι ιδιότητες της A που έχουμε υποθέσει ισχύουν αυτόματα και για την $\overline{A}$. Υποθέτουμε λοιπόν ότι η άλγεβρα A είναι κλειστή.

Λήμμα 10.3. *Για κάθε $f \in C(X)$ και $x \in X$ υπάρχει $g_x \in A$ τ.ώ. $g_x(x) = f(x)$ και $g_x(y) > f(y) - \epsilon$ για κάθε $y \in X$.*

Απόδειξη. Από το Λήμμα 10.1 μπορούμε να βρούμε για κάθε $y \in X \setminus \{x\}$ συνάρτηση $h_y \in A$ τ.ώ. $h_y(x) = f(x)$ και $h_y(y) = f(y)$. Από τη συνέχεια της $h_y - f$ έχουμε ότι το σύνολο

$$U_y = \{t \in X : h_y(t) > f(t) - \epsilon\}$$

είναι ανοιχτό και περιέχει τα x, y. Άρα

$$X = \bigcup_{y \neq x} U_y$$

είναι ένα ανοιχτό κάλυμμα του X και, επειδή ο X είναι συμπαγής χώρος, μπορούμε να διαλέξουμε πεπερασμένα στο πλήθος $y_1, y_2, \ldots, y_n$ τ.ώ.

$$X = U_{y_1} \cup \cdots \cup U_{y_n}.$$

Αν ορίσουμε

$$g_x = \max\{h_{y_1}, \ldots, h_{y_n}\}$$

τότε $g_x \in A$ (αφού η άλγεβρα A είναι κλειστή ως προς τα μέγιστα) και $g_x(x) = f(x)$ αφού $h_{y_j}(x) = f(x)$ για όλα τα $j = 1, 2, \ldots, n$.

Έχουμε το ζητούμενο $g_x(y) > f(y) - \epsilon$ για κάθε $y \in X$ γιατί για κάθε y υπάρχει ένα y_j τ.ώ. $y_j \in U_{y_j}$ και $g \geq h_{y_j}$. □

Έστω τώρα $f \in C(X)$ και $\epsilon > 0$. Θα βρούμε $h \in A$ τ.ώ. $\|f - h\|_\infty < \epsilon$. Από το Λήμμα 10.3 μπορούμε να βρούμε για κάθε $x \in X$ μια συνάρτηση $g_x \in A$ τ.ώ. $g_x(x) = f(x)$ και $g_x(y) > f(y) - \epsilon$ για κάθε $y \in X$. Τα σύνολα

$$V_x = \{y \in X : \ g_x(y) < f(y) + \epsilon\}$$

είναι ανοιχτά για κάθε $x \in X$ και $x \in V_x$, οπότε

$$X = \bigcup_{x \in X} V_x$$

είναι ένα ανοιχτό κάλυμμα του X, οπότε υπάρχει ένα πεπερασμένο υποκάλυμμα αυτού

$$X = V_{x_1} \cup \cdots \cup V_{x_m},$$

για κάποια $x_1, \ldots, x_m \in X$. Θέτουμε

$$h = \min\{g_{x_1}, \ldots, g_{x_m}\} \in A.$$

Αν $y \in X$ τότε υπάρχει κάποιο x_i τ.ώ. $y \in V_{x_i}$ που σημαίνει ότι $g_{x_i}(y) < f(y) + \epsilon$ και άρα $h(y) < f(y) + \epsilon$.

Επίσης για κάθε i έχουμε $g_{x_i}(y) > f(y) - \epsilon$ οπότε το ίδιο ισχύει για το ελάχιστο των συναρτήσεων g_{x_i}, δηλ. τη συνάρτηση h, άρα έχουμε και την ανισότητα $h(y) > f(y) - \epsilon$ και έχουμε αποδείξει το ζητούμενο

$$f(y) - \epsilon < h(y) < f(y) + \epsilon \text{ για κάθε } y \in X.$$

□

Πόρισμα 10.1. *Αν X, Y είναι συμπαγείς μετρικοί χώροι τότε οι πεπερασμένοι γραμμικοί συνδυασμοί από συναρτήσεις της μορφής*

$$f(x, y) = g(x)h(y)$$

με $g \in C(X)$, $h \in C(Y)$, είναι πυκνές στο χώρο $C(X \times Y)$.

Πόρισμα 10.2. *Αν $K \subseteq \mathbb{R}^n$ είναι συμπαγές τότε τα πολυώνυμα n μεταβλητών είναι πυκνά στο $C(K)$.*

10.3. **Μιγαδικοί συντελεστές.** Το Θεώρημα10.1 δεν ισχύει αν επιτρέψουμε μιγαδικούς συντελεστές στην άλγεβρα A και μιγαδικές τιμές στο χώρο $C(X)$. Η άλγεβρα των πολυωνύμων $\mathbb{C}[z]$, περιορισμένη σε ένα συμπαγές $X \subseteq \mathbb{C}$ προφανώς διαχωρίζει σημεία και δε μηδενίζεται πουθενά στο X αλλά δεν είναι πυκνή στο $C(X)$ όταν το X έχει εσωτερικό. Ένας τρόπος να το δούμε αυτό είναι να επικαλεστούμε τη θεωρία των αναλυτικών συναρτήσεων η οποία μας λέει ότι η αναλυτικότητα μιας συνάρτησης διατηρείται από τα ομοιόμορφα όρια. Ξεκινώντας λοιπόν από μια συνάρτηση στο X που δεν είναι αναλυτική, π.χ. τη συνάρτηση $f(z) = \overline{z}$, θα παίρναμε αντίφαση αν αυτή μπορούσε να προσεγγιστεί ομοιόμορφα στο X από πολυώνυμα (που είναι αναλυτικές συναρτήσεις).

Ένας εναλλακτικός τρόπος για να αποδείξουμε ότι η συνάρτηση $\overline{z}$ δεν είναι ομοιόμορφο όριο πολυωνύμων σε κάποια συμπαγή υποσύνολα του $\mathbb{C}$ είναι ο παρακάτω. Ας πάρουμε $X = \left\{e^{it} : t \in [0, 2\pi)\right\}$ να είναι ο μοναδιαίος κύκλος στο $\mathbb{C}$. Η συνάρτηση $f(t) = e^{-it}$ πάνω στο X είναι συνεχής. Αν ήταν ομοιόμορφο όριο πολυωνύμων $p_n(z) = p_n(e^{it})$ τότε θα είχαμε

$$\widehat{p_n}(k) \to \widehat{f}(k) \text{ για κάθε } k \in \mathbb{Z}.$$

Αλλά $\widehat{p_n}(-1) = 0$ για κάθε n ενώ $\widehat{f}(-1) = 1$ (οι συντελεστές Fourier τριγωνομετρικών πολυωνύμων είναι οι συντελεστές τους).

Για να επεκτείνουμε το Θεώρημα 10.1 για μιγαδικές συναρτήσεις χρειαζόμαστε μια επιπλέον συνθήκη.

Θεώρημα 10.2. *Αν X συμπαγής μετρικός χώρος και $A \subseteq C(X)$ (μιγαδικές τιμές) υποάλγεβρα (πάνω από το $\mathbb{C}$) που διαχωρίζει σημεία, δε μηδενίζεται πουθενά και $f \in A \Rightarrow \overline{f} \in A$ τότε η A είναι πυκνή στην $C(X)$. Κάθε συνεχής συνάρτηση δηλ. πάνω στο X μπορεί να προσεγγιστεί ομοιόμορφα από συναρτήσεις στην A.*

Απόδειξη. Από την υπόθεση $f \in A \Rightarrow \overline{f} \in A$ παίρνουμε ότι για κάθε $f \in A$ το πραγματικό και φανταστικό μέρος της είναι επίσης στην A:

$$\mathrm{Re}\,[f] = (f + \overline{f})/2, \;\; \mathrm{Im}\,[f] = (f - \overline{f})/(2i).$$

Ας ορίσουμε $A' \subseteq A$ να είναι οι πραγματικές συναρτήσεις της A. Η A' είναι επίσης μια άλγεβρα αλλά με πραγματικούς συντελεστές. Είναι μάλιστα μια υποάλγεβρα της άλγεβρας των πραγματικών συναρτήσεων στην $C(X)$.

Εύκολα βλέπουμε, κοιτώντας τα πραγματικά και φανταστικά μέρη ρων $f \in A$, ότι αφού η A διαχωρίζει σημεία και δε μηδενίζεται πουθενά το ίδιο συμβαίνει και με την A'. Συνεπώς, από το Θεώρημα 10.1 η A' είναι πυκνή στις πραγματικές συναρτήσεις της $C(X)$.

Ας είναι τώρα $f \in C(X)$ και $\epsilon > 0$. Βρίσκουμε κάποιες συναρτήσεις $g, h \in A'$ τ.ώ. $\|\mathrm{Re}\,[f] - g\|_\infty \leq \epsilon/2$ και $\|h - \mathrm{Im}\,[f]\|_\infty \leq \epsilon/2$. Η συνάρτηση $g + ih$ είναι κι αυτή στην A και $\|(g + ih) - f\|_\infty \leq \epsilon$. □

Πόρισμα 10.3. *Τα τριγωνομετρικά πολυώνυμα είναι πυκνά στις συνεχείς 2π-περιοδικές συναρτήσεις.*