

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΑΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΕΦΑΡΜΟΣΜΕΝΕΣ ΜΑΘΗΜΑΤΙΚΕΣ ΕΠΙΣΤΗΜΕΣ»

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
**ΤΟΠΙΚΕΣ ΤΡΙΓΩΝΟΜΕΤΡΙΚΕΣ ΒΑΣΕΙΣ ΚΑΙ Η ΕΦΑΡΜΟΓΗ
ΤΟΥΣ ΣΤΗΝ ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΦΩΝΗΣ**

ΧΡΗΣΤΟΣ Σ. ΜΑΛΛΙΟΠΟΥΛΟΣ

ΑΘΗΝΑ, ΦΕΒΡΟΥΑΡΙΟΣ 2012

ΠΕΡΙΕΧΟΜΕΝΑ

Προλογος	1
ΚΕΦΑΛΑΙΟ 1 Μετασχηματισμοί Gabor	5
1.1 Συνθήκες Απροσδιοριστίας Χρόνου-Συχνότητας	5
1.2 Στοιχεία Χρόνου-Συχνότητας και Μετασχηματισμός Gabor.....	11
1.3 Διασπάσεις σε Διακριτά Πλέγματα Χρόνου-Συχνότητας	14
ΚΕΦΑΛΑΙΟ 2 Τοπικές Τριγωνομετρικές Βάσεις	18
2.1 Τοπικές Ορθοκανονικές Προβολές	19
2.2 Τοπικές Τριγωνομετρικές Βάσεις	28
2.2.1 Βάσεις Malvar.....	29
2.2.2 Lemarie-Meyer Wavelets.....	32
2.3 Παράρτημα	35
ΚΕΦΑΛΑΙΟ 3 Βέλτιστες Τοπικές Βάσεις και Τμηματοποίηση Φωνής.....	37
3.1 Τοπικές Τριγωνομετρικές Βάσεις Διακριτού Χρόνου	38
3.2 Διακριτος Μετασχηματισμος Malvar.....	40
3.3 Βέλτιστες Τοπικές Βάσεις.....	42
3.3.1 Λεξικά Βάσεων και Κοίλα Συναρτησιακά Κόστους	44
3.3.2 Προσαρμοστική Αναζήτηση σε Διαδικά Λεξικά.....	49
3.4 Τμηματοποίηση Φωνής	54
3.4.1 Το Σήμα Φωνής.....	54
3.4.2 Ακρότατα Κλίσης.....	57
3.4.3 Ιεραρχική Ομαδοποίηση (Hierarchical Clustering)	57
3.4.4 Μέγιστα μετασχηματισμών Wavelet (Wavelet Modulus Maxima).....	57
3.4.5 Τοπική Διάσπαση (Temporal Decomposition)	59
3.4.6 Entropy-based Segmentation.....	61
3.5 Παράρτημα	65
Αναφορές	68

ΠΡΟΛΟΓΟΣ

Στην εργασία μελετούμε σήματα πεπερασμένης ενέργειας, δηλαδή τετραγωνικά ολοκληρώσιμες ή αθροίσιμες χρονικές συναρτήσεις και εξετάζουμε κατά πόσον ειδικές ορθοκανονικές αναπαραστάσεις χρόνου-συχνότητας μπορούν να εντοπίσουν τα όρια μεταξύ των φωνημάτων ενός σήματος φωνής.

Αναλυτικότερα, στο πρώτο κεφάλαιο εξετάζουμε τις ορθοκανονικές αναπαραστάσεις Gabor

$$f = \sum_{m,n \in \mathbb{Z}} f_{m,n} \delta_{m,n} \quad (0.1)$$

όπου $g_{m,n}(t) = e^{i2\pi\gamma mt} g(t - nt_0)$, και δείχνουμε ότι η απαίτηση για ορθοκανονικές βάσεις οδηγεί σε δύο βασικούς περιορισμούς:

1. $t_0 \cdot \gamma_0 = 1$ και
2. $\int t^2 |g(t)|^2 dt \cdot \int \gamma^2 |\hat{g}(\gamma)|^2 d\gamma = +\infty$ (θεώρημα Balian-Low).

Ο πρώτος περιορισμός είναι συνέπεια της ορθοκανονικότητας του συστήματος $(g_{m,n})_{m,n \in \mathbb{Z}}$: το πλέγμα $t_0\mathbb{Z} \times \gamma_0\mathbb{Z}$ δεν μπορεί να είναι αυθαίρετα πυκνό. Πυκνά πλέγματα οδηγούν σε πλεονασμούς. Ο δεύτερος περιορίζει τις επιλογές μας ως προς την τοπικότητα των στοιχείων $(g_{m,n})_{m,n \in \mathbb{Z}}$ στον χρόνο και την συχνότητα.

Μπορούμε ωστόσο να κατασκευάσουμε ορθοκανονικές βάσεις όπου τα στοιχεία $g_{m,n}$ είναι καλά εντοπισμένα στον χρόνο και την συχνότητα, εάν αντικαταστήσουμε τα μιγαδικά εκθετικά από τριγωνομετρικές συναρτήσεις (ημίτονα ή συνημίτονα) και επιπλέον επιβάλλουμε κάποιες συνθήκες συμμετρίας στην συνάρτηση g .

Οι παραπάνω βάσεις εντάσσονται στο θεωρητικό πλαίσιο των τοπικών ορθοκανονικών προβολών [HW96] που αναπτύσσεται διεξοδικά στο δεύτερο κεφάλαιο.

Στο τρίτο κεφάλαιο εξετάζουμε αναπαραστάσεις της μορφής (0.1) σε σήματα διακριτού χρόνου και δείχνουμε ότι όλη η θεωρία του L^2 μπορεί να μεταφερθεί χωρίς αλλαγές στον

χώρο των τετραγωνικά αθροισίμων ακολουθιών ℓ^2 ¹. Περιγράφουμε επίσης έναν γρήγορο αλγόριθμο για τον υπολογισμό των προβολών ενός σήματος μήκους N στα στοιχεία $g_{m,n}$ [MALLAT2, §8.4.4].

Όλες οι σύγχρονες βάσεις προφορικού λόγου (speech corpora), ενσωματώνουν πληροφορίες για τα όρια των φωνητικών και γλωσσολογικών τους μονάδων. Οι πληροφορίες αυτές είναι ιδιαίτερα χρήσιμες στην τεκμηρίωση γλωσσολογικών μελετών και την εκπαίδευση συστημάτων αναγνώρισης φωνής. Εξάλλου, η ίδια η διαδικασία αναγνώρισης φωνής βελτιώνεται σημαντικά εάν τα φωνητικά όρια του σήματος εντοπιστούν σε κάποιο αρχικό στάδιο επεξεργασίας.

Οι τρέχουσες μέθοδοι τμηματοποίησης φωνής ξεκινούν από μία πλαισιωμένη φασματική ανάλυση του σήματος φωνής. Στα διανύσματα (πρότυπα) που προκύπτουν – ή σε μετασχηματισμένα παράγωγά τους (π.χ. Mel Frequency Cepstral Coefficients ή partial correlation coefficients μίας AR ανάλυσης του πλαισίου [RS78]) – εφαρμόζεται ακολούθως κάποια μέθοδος κατηγοριοποίησης (clustering). Σε αντίθεση με την αναγνώριση φωνής, όπου κλάσεις είναι τα φωνήματα (ή άλλα ακουστικά συγκροτήματα της γλώσσας), στην τμηματοποίηση ένα πρότυπο ανήκει σε δύο μόνον κλάσεις: είτε βρίσκεται στο σύνορο μεταξύ διαφορετικών φωνημάτων ή στο εσωτερικό ενός φωνήματος. Οι unsupervised classifiers αξιοποιούν την απόσταση μεταξύ των προτύπων όπως αυτή περιγράφεται από κάποιο εκ των προτέρων επιλεγμένο συναρτησιακό στον \mathbb{C}^N και είναι γνωστές με τον όρο “blind speech segmentation”. Οι supervised τεχνικές, προϋποθέτουν κάποια γνώση του γλωσσολογικού περιεχομένου του σήματος – των φωνημάτων δηλαδή από τα οποία αποτελείται.

Άλλες έρευνες χρησιμοποιούν κριτήρια κατηγοριοποίησης, βασισμένα σε μία ή περισσότερες κλίμακες μίας multiresolution analysis του σήματος φωνής καθώς κάποια χαρακτηριστικά – όπως τα όρια εμφώνων και άφωνων ήχων – είναι ευδιάκριτα σε κάποιες κλίμακες ανάλυσης.

Στην εργασία, το πρόβλημα κατηγοριοποίησης αντιμετωπίζεται ως πρόβλημα επιλογής μίας ορθοκανονικής βάσης του \mathbb{C}^N που “αντιπροσωπεύει” το πρότυπο, υπό την εξής έννοια:

Διαφορετικές οικογένειες στοιχείων $(g_{m,n})_{m,n \in \mathbb{Z}}$ παράγουν διαφορετικές ορθοκανονικές βάσεις. Ξεκινώντας από ένα λεξικό βάσεων $\{\mathcal{B}^\gamma, \gamma \in \mathbb{I}\}$ όπου \mathbb{I} ένα σύνολο δεικτών, ενδιαφερόμαστε για την βάση που είναι καλύτερα προσαρμοσμένη στο πρότυπο. Η προσαρμογή μίας βάσης \mathcal{B}^γ σε δεδομένο $f \in \mathbb{C}^N$ είναι καλή, εάν το f περιγράφεται από έναν μικρό αριθμό στοιχείων της \mathcal{B}^γ . Παρουσιάζουμε συναρτήσεις C του \mathbb{I} – συναρτησιακά κό-

¹ Εννοούμε τους χώρους $L^2(\mathbb{R})$ και $\ell^2(\mathbb{Z})$ αντιστοίχως.

στον \mathbb{C}^N – που ελαχιστοποιούνται όταν ο αριθμός των προβολών του f σε μία βάση \mathcal{B}^γ είναι μικρός.

Η παραπάνω ιδιότητα ενός συναρτησιακού κόστους C το καθιστά ιδιαίτερα χρήσιμο στην περίπτωση blind speech segmentation, όταν τα στοιχεία του λεξικού $\{\mathcal{B}^\gamma, \gamma \in \mathbb{I}\}$ αντιπροσωπεύουν τα φασματικά χαρακτηριστικά ενός προτύπου. Στο σύνορο μεταξύ φωνημάτων το C θα πρέπει να προκρίνει βάσεις καλά εντοπισμένες στον χρόνο ενώ στο εσωτερικό ενός φωνήματος, βάσεις καλά εντοπισμένες στην συχνότητα αφού τα χαρακτηριστικά του σήματος διατηρούνται σταθερά.

Αναλυτικότερα, ας υποθέσουμε ότι $\mathbb{I}_{\nu,N}, \mathbb{I}_{\nu+N,N}$ είναι αποσυνδεδετά διαδοχικά διαστήματα του \mathbb{Z} με μήκος N και αριστερά άκρα ν και $\nu + N$ αντιστοίχως και, $\mathbb{I}_{\nu,2N} = \mathbb{I}_{\nu,N} \cup \mathbb{I}_{\nu+N,N}$.

Έστω ακόμη $\mathcal{B}'_N, \mathcal{B}'_{N+N}$ και \mathcal{B}'_{2N} ορθοκανονικές βάσεις εντοπισμένες στα $\mathbb{I}_{\nu,N}, \mathbb{I}_{\nu+N,N}$ και $\mathbb{I}_{\nu,2N}$, έτσι ώστε $\langle x, y \rangle = 0, \forall x \in \mathcal{B}'_N, y \in \mathcal{B}'_{N+N}$ και $\langle \mathcal{B}'_N \rangle \cup \langle \mathcal{B}'_{N+N} \rangle = \langle \mathcal{B}'_{2N} \rangle$ – συμβολίζουμε με $\langle \cdot \rangle$ το εσωτερικό γινόμενο και την γραμμική θήκη ενός διανυσματικού συνόλου – και C ένα συναρτησιακό κόστους. Εάν $f_{\nu,N}, f_{\nu+N,N}$ και $f_{\nu,2N}$ είναι οι διανυσματικές προβολές ενός σήματος f στις $\mathcal{B}'_N, \mathcal{B}'_{N+N}$ και \mathcal{B}'_{2N} αντιστοίχως, τότε τα $\mathbb{I}_{\nu,N}$ και $\mathbb{I}_{\nu+N,N}$ είναι συνοριακά (boundary) τμήματα της f εάν $C(f_{\nu,N}, f_{\nu+N,N}) \leq C(f_{\nu,2N})$.

Στο τελευταίο μέρος του τρίτου κεφαλαίου, δίνουμε ένα παράδειγμα της επιτυχίας της παραπάνω προσέγγισης σε ηχογράφηση αντρικής φωνής.

Η εργασία περιέχει αρκετά πρωτότυπα στοιχεία:

Το πρώτο κεφάλαιο, αποδεικνύει την συνθήκη πλέγματος $t_0 \cdot \gamma_0 = 1$ του μετσηματισμού Gabor, χωρίς καμία αναφορά σε frames. Στην βιβλιογραφία – που έχουμε τουλάχιστον εμείς υπ' όψιν μας [DAU92, CHRIS, GROH] – η συνθήκη πλέγματος αποδεικνύεται στο γενικότερο πλαίσιο των frames².

Το δεύτερο κεφάλαιο βασίζεται στα [MALLAT2 κεφ. 8, HW96 κεφ. 1]. Οι προτάσεις όμως έχουν πρακτικά αναδιατυπωθεί και ο φορμαλισμός είναι διαφορετικός. Η απόδειξη του παραρτήματος είναι δική μας.

Το βασικό πρωτότυπο στοιχείο του τρίτου κεφαλαίου είναι το “Entropy-based Segmentation”. Από έλεγχο που κάναμε στο Google scholar δεν έχουμε υπ' όψιν μας σχετικές δημοσιεύσεις.

² Προφανώς δεν πρόκειται για μειονέκτημα των συγγραμμάτων. Αντιθέτως θα ήταν μάλλον παράδοξο οι συγγραφείς ν' ακολουθήσουν την δική μας προσέγγιση. Ο δικός μας στόχος ήταν απλώς να μην παραπέμψουμε τον αναγνώστη σε επιπλέον θεωρία.

Το πρώτο κεφάλαιο είναι θεωρητικό και οι διατυπώσεις όλων των προτάσεων και οι αποδείξεις τους, γίνονται με όρους αρμονικής ανάλυσης μεταπτυχιακού επιπέδου. Η κατανόησή τους ωστόσο προϋποθέτει στην ουσία μόνον γνώση των ιδιοτήτων του μετασχηματισμού Fourier (συμπεριλαμβανομένου του θεωρήματος Plancherel) και του θεωρήματος Fubini. Σε ένα μόνον σημείο (μετασχηματισμός Gabor) γίνεται επιπλέον χρήση ενός επιχειρήματος πυκνότητας του $L^1 \cap L^2$ στον L^2 . Τα υπόλοιπα κεφάλαια προϋποθέτουν μόνον καλή γνώση γραμμικής άλγεβρας, ειδικότερα των ιδιοτήτων των ορθοκανονικών προβολών. Τέλος, οι συμβολισμοί που χρησιμοποιούνται είναι τυπικοί των σύγχρονων συγγραμμάτων αρμονικής ανάλυσης.

Τα θεωρήματα, προτάσεις, λήμματα και πορίσματα έχουν ενιαία αρίθμηση ανά κεφάλαιο. Έτσι το λήμμα π.χ. 2.3 μπορεί να έπεται του θεωρήματος 2.2 και αυτό της πρότασης 2.1. Επίσης οι αριθμήσεις των σχημάτων και υποσημειώσεων είναι ξεχωριστές και ανεξάρτητες κεφαλαίου.

Βασική επιρροή στην εργασία είχε το βιβλίο του Stephane Mallat “A Wavelet Tour on Signal Processing” [MALLAT2]. Στους αριθμητικούς υπολογισμούς χρησιμοποιήθηκε εκτεταμένα το WaveLab850 (<http://www.waveLab.org>) μία βιβλιοθήκη ρουτινών Matlab από τους Mallat, Donoho, Coifman, Wickerhauser και άλλους. Ειδικότερα, τα σχήματα των παραγράφων 3.3.2 και 3.4.6 έχουν παραχθεί από συναρτήσεις της βιβλιοθήκης αυτής.

Η ιδέα για τις τοπικές τριγωνομετρικές βάσεις – το κεντρικό δηλαδή θέμα της εργασίας – ήταν του επιβλέποντος καθηγητή μου κ. Γεράσιμου (Μάκη) Αθανασούλη στον οποίο είμαι υπόχρεος και ευχαριστώ βαθύτατα. Επίσης θα ήθελα να ευχαριστήσω το Ινστιτούτο Επεξεργασίας του Λόγου και την εταιρία Innoetics για την παραχώρηση των σημάτων φωνής που χρησιμοποιήθηκαν στο τρίτο κεφάλαιο.

Θέλω τέλος να ευχαριστήσω την κόρη μου Ελένη και την σύζυγό μου Τζένη, για την υπομονή τους κατά την συγγραφή της εργασίας αυτής.

////

ΚΕΦΑΛΑΙΟ 1

ΜΕΤΑΣΧΗΜΑΤΙΜΟΙ GABOR

Στο κεφάλαιο αυτό εξετάζουμε τον μετασχηματισμό Fourier στον L^2 , τον χώρο των τετραγωνικά ολοκληρώσιμων συναρτήσεων με πεδίο ορισμού το \mathbb{R} και τιμές εν γένει μιγαδικές. Ειδικότερα ενδιαφερόμαστε για προϋποθέσεις υπό τις οποίες τέτοιες συναρτήσεις μπορούν να προσεγγιστούν και κυρίως, ν' αναπαραχθούν από τις μέσες χρονικές και φασματικές τιμές τους. Θεωρούμε χρονικές συναρτήσεις (σήματα), χωρίς αυτό να περιορίζει σε τίποτε την γενικότητα των αποτελεσμάτων.

Αποδεικνύουμε την συνθήκη (αρχή) απροσδιοριστίας Heisenberg (κατά άλλους Weyl-Heisenberg [COH]), που δείχνει ότι οι διασπορές των χρονικών και φασματικών τιμών των συναρτήσεων αυτών δεν μπορούν να καθοριστούν ανεξάρτητα η μία από την άλλη. Εξετάζουμε τον μετασχηματισμό Gabor και αποδεικνύουμε μία συνθήκη ώστε τα στοιχεία χρόνου-συχνότητας του μετασχηματισμού ν' αποτελούν ορθοκανονική βάση. Τέλος, χρησιμοποιούμε την συνθήκη αυτήν για να δείξουμε ότι εάν κάποια στοιχεία χρόνου-συχνότητας Gabor είναι ορθοκανονική βάση, τότε κανένα τους δεν μπορεί να είναι καλά εντοπισμένο ταυτόχρονα στον χρόνο και την συχνότητα.

Στόχος μας στο κεφάλαιο αυτό είναι να δείξουμε ότι η περιγραφικές ικανότητες ορθοκανονικών βάσεων Gabor είναι περιορισμένες, αιτιολογώντας έτσι την ανάλυση των τοπικών ορθοκανονικών βάσεων που ακολουθεί στο επόμενο κεφάλαιο.

1.1 ΣΥΝΘΗΚΕΣ ΑΠΡΟΣΔΙΟΡΙΣΤΙΑΣ ΧΡΟΝΟΥ-ΣΥΧΝΟΤΗΤΑΣ

Ξεκινούμε με μία απλή αλλά ιδιαίτερα σημαντική παρατήρηση σχετικά με τον φορέα του μετασχηματισμού Fourier (Fourier Transform – FT) συναρτήσεων πεπερασμένης διάρκειας.

Πρόταση 1.1: Εάν η f έχει πεπερασμένη διάρκεια, ο φορέας της \hat{f} δεν μπορεί να μηδενίζεται σε ολόκληρο διάστημα του \mathbb{R} . Το ίδιο ισχύει και εάν η f αντικατασταθεί από την \hat{f} .

Απόδειξη

Αποδεικνύουμε την πρώτη περίπτωση μόνον. Η δεύτερη περίπτωση αποδεικνύεται με ανάλογο τρόπο. Έστω ότι $\hat{f} \neq 0$ και $\hat{f}(\gamma) = 0 \forall \gamma \in \mathbb{I}$ όπου \mathbb{I} ένα διάστημα του \mathbb{R} .

Εάν $\text{supp} f \subset [-\alpha, \alpha]$, τότε $\hat{f}(\gamma) = \int_{-\alpha}^{\alpha} f(t)e^{-i2\pi\gamma t} dt$ και αν γ_0 είναι ένα εσωτερικό σημείο του \mathbb{I} τότε

$$D^n \hat{f}(\gamma_0) = (-i2\pi)^n \int_{-\alpha}^{\alpha} t^n f(t) e^{-i2\pi\gamma_0 t} dt = 0 \quad \forall n \in \mathbb{N} \quad (1.1)$$

Επιπλέον βάσει της (1.1),

$$\begin{aligned} \hat{f}(\gamma) &= \int_{-\alpha}^{\alpha} e^{-i2\pi(\gamma-\gamma_0)t} f(t) e^{-i2\pi\gamma_0 t} dt \\ &= \sum_n \frac{[-i2\pi(\gamma-\gamma_0)]^n}{n!} \int_{-\alpha}^{\alpha} t^n f(t) e^{-i2\pi\gamma_0 t} dt \\ &= 0 \quad \forall \gamma \in \mathbb{R} \end{aligned} \quad (1.2)$$

πράγμα που αντιβαίνει στην αρχική υπόθεση υπόθεση $\hat{f} \neq 0$. Με παρόμοιο τρόπο αποδεικνύεται το ίδιο για τον φορέα της f όταν η \hat{f} έχει συμπαγή φορέα.

QED

Παρατήρηση

Η \hat{f} έχει παραγώγους κάθε τάξης. Ακριβέστερα, $\hat{f} \in L^2 \cap C_0^\infty$. Αυτό είναι αναγκαστική συνέπεια της L^2 ολοκληρωσιμότητας της f και της υπόθεσης ότι η f έχει συμπαγή φορέα. Πράγματι, εάν $\text{supp} f \subset [-\alpha, \alpha]$ τότε

$$\int |(-i2\pi t)^n f(t)|^2 dt \leq (2\alpha)^{2n} \int_{-\alpha}^{\alpha} |f(t)|^2 d\gamma < +\infty$$

και η $D^n \hat{f}$ ορίζεται από την σχέση

$$D^n \hat{f} = \left[(-i2\pi t)^n f(t) \right]^\wedge.$$

και είναι συνεχής λόγω της L^1 ολοκληρωσιμότητας της $(-i2\pi t)^n f(t)$.

Μπορούμε να χαλαρώσουμε την απαίτησή μας για συμπαγή χρονικό φορέα, αναζητώντας συναρτήσεις με μικρή διασπορά τιμών στον χρόνο και την συχνότητα. Θέλουμε δηλαδή να γνωρίζουμε εάν υπάρχουν συναρτήσεις $f \in L^2$ τέτοιες ώστε

$$\left(\int (t-\alpha)^2 |f(t)|^2 dt \right)^{1/2} \left(\int (\gamma-\alpha)^2 |\hat{f}(\gamma)|^2 d\gamma \right)^{1/2} \leq \epsilon \quad \forall \epsilon > 0,$$

και κάποια πραγματικά α, β . Στο επόμενο θεώρημα αποδεικνύεται ότι αυτό δεν είναι δυνατόν.

Ορίζουμε στο L^2 τους τελεστές χρονικής μετατόπισης $\tau_x f(t) = f(t-x)$ και διαμόρφωσης $M_\gamma f(t) = e^{i2\pi\gamma t} f(t)$. Προφανώς πρόκειται για ισομετρικούς ομοιομορφισμούς $L^2 \rightarrow L^2$.

Οι τ και M ικανοποιούν τις ακόλουθες σημαντικές σχέσεις:

$$\widehat{\tau_\alpha f} = M_{-\alpha} \hat{f} \quad \text{και}$$

$$\widehat{M_\alpha f} = \tau_\alpha \hat{f}$$

Θεωρούμε επίσης τους γραμμικούς τελεστές $X, P: L^2 \rightarrow L^2$ με τιμές $(Xf)(t) = tf(t)$ και $(Pf)(t) = \frac{1}{2\pi i} (Df)(t)$ αντιστοίχως. Ο X ορίζεται για κάθε $f \in L^2$ ενώ ο P για κάθε συνάρτηση παραγωγίσιμη σχεδόν παντού και κάθε συνάρτηση $f \in L^2$ με $X\hat{f} \in L^2$ ³. Προφανώς οι X και P δεν είναι φραγμένοι. Η σχέση μεταξύ P και X είναι

$$\widehat{Pf} = D\hat{f}$$

Θεώρημα 1.2: Για κάθε $\alpha, \beta \in \mathbb{R}$ και $f \in L^2$, ισχύει

$$\|(X-\alpha)f\| \|(X-\beta)\hat{f}\| = \left(\int (t-\alpha)^2 |f(t)|^2 dt \right)^{1/2} \left(\int (\gamma-\beta)^2 |\hat{f}(\gamma)|^2 d\gamma \right)^{1/2} \geq \frac{1}{4\pi} \|f\|^2 \quad (1.3)$$

³ Στην δεύτερη περίπτωση η παράγωγος ορίζεται μεσοτετραγωνικά, δηλαδή ως το L^2 όριο των συναρτήσεων $\frac{1}{s}(f - \tau_s f)$ όταν $s \rightarrow 0$. Είναι δηλαδή δυνατόν να δείξουμε (σχετικά εύκολα μάλιστα) ότι αν θέσουμε $\hat{g} = \gamma \hat{f}$ τότε $\int \left| g(t) - \frac{f(t+s) - f(t)}{s} \right|^2 \xrightarrow{s \rightarrow 0} 0$. Σημειώνουμε ότι αυτό δεν είναι γενικά ταυτόσημο με την σχεδόν-παντού (σημειακή) σύγκλιση των παραπάνω λόγων. Ειδικότερα, είναι δυνατόν να μην υπάρχει πουθενά σημειακή σύγκλιση (όμως κάθε ακολουθία έχει υπακολουθία που συγκλίνει σημειακά σχεδόν παντού. Βλ. και [RUDINRCA κεφ. 3]).

Το παραπάνω γινόμενο γίνεται ίσο προς $\frac{1}{4\pi}\|f\|^2$, αν και μόνον εάν η f είναι για κάποιο θετικό s , πολλαπλάσιο της συνάρτησης

$$\tau_\alpha M_\beta \varphi_s(t) = e^{i2\pi\beta(t-\alpha)} e^{-\pi(t-\alpha)^2/s}$$

όπου $\varphi_s(t) = e^{-\pi t^2/s}$ είναι η συνάρτηση Gauss με διασπορά s .

Εάν για κάποια συνάρτηση $f \in L^2$ το αριστερό μέλος της (1.3) είναι άπειρο, η σχέση ικανοποιείται προφανώς τετριμμένα.

Αποδεικνύουμε πρώτα δύο σύντομα λήμματα για τους τελεστές P και X που ορίστηκαν προηγουμένως:

Λήμμα 1.3: Εάν $f \in L^2$ και $(P - \beta)f, (X - \alpha)f \in L^2$, τότε για κάθε $\alpha, \beta \in \mathbb{R}$ ισχύει

$$\|(X - \alpha)f\| \cdot \|(P - \beta)f\| \geq \frac{1}{2} \left| \langle (P - \beta)f, (X - \alpha)f \rangle - \langle (X - \alpha)f, (P - \beta)f \rangle \right|$$

Ισότητα ισχύει αν και μόνον αν $(X - \alpha)f + is(P - \beta)f = 0$ για κάποιο $s \in \mathbb{R}$.

Απόδειξη

$$\begin{aligned} & \frac{1}{2} \left| \langle (P - \beta)f, (X - \alpha)f \rangle - \langle (X - \alpha)f, (P - \beta)f \rangle \right| \\ &= \left| \operatorname{Im} \langle (P - \beta)f, (X - \alpha)f \rangle \right| \\ &\leq \left| \langle (P - \beta)f, (X - \alpha)f \rangle \right| \\ &\leq \|(X - \alpha)f\| \cdot \|(P - \beta)f\| \end{aligned} \tag{1.4}$$

Η (1.4) είναι προφανής λόγω της συζυγούς συμμετρίας του εσωτερικού γινομένου και της ανισότητας Cauchy-Schwartz.

Η δεύτερη από τις παραπάνω ανισότητες γίνεται ισότητα όταν $(P - \alpha)f = \lambda(X - \beta)f$ για κάποιο μιγαδικό λ και η πρώτη εάν $\lambda = is$ για κάποιο πραγματικό s .

QED

Λήμμα 1.4: Εάν $\|Xf\| \|X\hat{f}\| \geq \frac{1}{4\pi} \|f\|^2 \quad \forall f \in L^2$ τότε και $\|(X - \alpha)f\| \|(X - \beta)\hat{f}\| \geq \frac{1}{4\pi} \|f\|^2$ για κάθε $\alpha, \beta \in \mathbb{R}$.

Απόδειξη

Επειδή οι τελεστές τ και M είναι ισομετρίες στον L^2 , έχουμε

$$\|(X - \alpha)f\| = \|X\tau_{-\alpha}f\| = \|M_{\beta}X\tau_{-\alpha}f\| = \|X(M_{\beta}\tau_{-\alpha}f)\|$$

και

$$\|(X - \beta)\hat{f}\| = \|X\tau_{-\beta}\hat{f}\| = \|M_{-\alpha}X\tau_{-\beta}\hat{f}\| = \|X(M_{-\alpha}\tau_{-\beta}\hat{f})\| = \|X(\tau_{-\alpha}M_{\beta}\hat{f})\|$$

Τελικά,

$$\begin{aligned} & \|(X - \alpha)f\| \cdot \|(X - \beta)\hat{f}\| \\ &= \|X(M_{\beta}\tau_{-\alpha}f)\| \cdot \|X(M_{\beta}\tau_{-\alpha}\hat{f})\| \geq \frac{1}{4\pi} \|M_{\beta}\tau_{-\alpha}f\|^2 \\ &= \frac{1}{4\pi} \|f\|^2 \end{aligned} \tag{1.5}$$

QED

Με βάση τα παραπάνω λήμματα, η απόδειξη του θεωρήματος 1.2 είναι απλή:

Απόδειξη (θεωρήματος 1.2)

Βάσει του λήμματος 1.4, αρκεί ν' αποδείξουμε το θεώρημα για $\alpha = \beta = 0$.

Επίσης, από το λήμμα 1.3 έχουμε:

$$\begin{aligned} & \|Xf\| \|X\hat{f}\| = \|Xf\| \|Pf\| \\ & \geq \frac{1}{2} |\langle Pf, Xf \rangle - \langle Xf, Pf \rangle| \\ & = \frac{1}{4\pi} \left| \int XDf(t)\bar{f}(t)dt + \int Xf(t)D\bar{f}dt \right| \end{aligned} \tag{1.6}$$

Αν $\int Xf(t)D\bar{f}(t)dt = -\int (DX)f(t)\bar{f}(t)dt$ (πράγμα που συμβαίνει π.χ. εάν $DXf \in L^2$ ή όταν η f έχει συμπαγή φορέα), η (1.6) γίνεται

$$\|Xf\| \|X\hat{f}\| \geq \frac{1}{4\pi} \left| \int (XD - DX)f(t)\bar{f}(t)dt \right| = \frac{1}{4\pi} \left| -\int f(t)\bar{f}(t)dt \right| = \frac{1}{4\pi} \|f\|^2 \tag{1.7}$$

Από το λήμμα 1.3, ισότητα στην (1.7) έχουμε αν και μόνον αν

$$Df - 2\pi i\beta f = -2\pi s(t - \alpha)f \tag{1.8}$$

Οι λύσεις της παραπάνω πρωτοβάθμιας συνήθους διαφορικής, είναι τα πολλαπλάσια της συνάρτησης $\tau_\alpha M_\beta \varphi_s$ (και μόνον αυτά)⁴. Επίσης πρέπει $s > 0$ ώστε $\varphi_s \in L^2$.

QED

Εάν \mathcal{H} είναι ο δ.χ. των συναρτήσεων f για τις οποίες Pf, Xf, XPf και $PXf \in L^2$ μπορούμε να δείξουμε ότι η $\|f\|_{\mathcal{H}} = \left(\|Xf\|_2^2 + \|Pf\|_2^2 \right)^{1/2}$ είναι μία νόρμα στον \mathcal{H} . Με την νόρμα αυτήν, ο \mathcal{H} εμφυτεύεται συνεχώς στον L^2 :

Πόρισμα 1.5: Εάν $f \in L^2$, τότε

$$\|f\|_2^2 \leq 2\pi \left(\|Xf\|_2^2 + \|Pf\|_2^2 \right) \quad (1.9)$$

και ισότητα ισχύει αν και μόνον αν $f(x) = Ce^{-\pi x^2}$.

Απόδειξη

Χρησιμοποιούμε το θεώρημα 1.2 και την ταυτότητα $2|xy| \leq x^2 + y^2$ με $x = \|Xf\|_2$, $y = \|Pf\|_2$ και $\alpha = \beta = 0$.

Για ισότητα στην (1.9) χρειαζόμαστε ισότητα στην (1.3) και $x = y$, δηλαδή $Pf = -iXf$. Οι μόνες λύσεις της διαφορικής είναι οι Gaussians $Ce^{-\pi x^2}$, $C \in \mathbb{C}$.

QED

Παρατηρήσεις

1. Για να γράψουμε την (1.3) στην συνήθη μορφή της, θέτουμε

$$\Delta_f t^2 = \min_{\alpha \in \mathbb{R}} \|f\|^{-2} \int (t - \alpha)^2 |f(t)|^2 dt \quad \text{και} \quad \Delta_f \xi^2 = \min_{\beta \in \mathbb{R}} \|\hat{f}\|^{-2} \int (\xi - \beta)^2 |\hat{f}(\xi)|^2 d\xi$$

Οι παραπάνω ποσότητες γίνονται ελάχιστες όταν

$$\alpha = \bar{t}_f = \|f\|^{-2} \int t |f(t)|^2 dt \quad \text{και} \quad \beta = \bar{\gamma}_f = \|\hat{f}\|^{-2} \int \gamma |\hat{f}(\gamma)|^2 d\gamma$$

⁴ Αντικαθιστώντας την f με την συνάρτηση $e^{i2\pi\beta(t-\alpha)} e^{-\pi(t-\alpha)^2/s}$ η συνήθης διαφορική (1.8) επαληθεύεται. Επιπλέον κάθε διαφορική πρώτου βαθμού (με συνοριακή συνθήκη $f(t_0) = f_0$) έχει μοναδική λύση σε ένα διάστημα εάν η λύση είναι ομαλή στο διάστημα αυτό. Π.χ. βλ. R. Agarwal "An Introduction to ODE" Lecture 10.

όταν δηλαδή τα α και β ισούνται με τις μέσες τιμές των f και \hat{f} αντιστοίχως:

Πράγματι, έστω $g = \|f\|^{-2} |f|^2$ και $\bar{g}(t) = g(t + \bar{t}_f)$. Είναι εύκολο να δειχθεί ότι

$$\bar{t}_{\bar{g}} = \int t \bar{g} dt = 0 \quad (1.10)$$

Εξάλλου,

$$\begin{aligned} \Delta_f t^2 &= \int (t - \alpha)^2 g(t) dt \\ &= \int \left(t - (\alpha - \bar{t}_f) \right)^2 \bar{g}(t) dt \\ &= \int t^2 \bar{g}(t) dt + (\alpha - \bar{t}_f)^2 \int \bar{g}(t) dt - 2(\alpha - \bar{t}_f) \int t \bar{g}(t) dt \end{aligned} \quad (1.11)$$

που λόγω της (1.10) ελαχιστοποιείται ακριβώς όταν $\alpha = \bar{t}_f = \|f\|^{-2} \int t |f(t)|^2 dt$.

Η (1.3) γράφεται τότε,

$$\Delta_f t \Delta_f \gamma \geq 1/4\pi$$

2. Η (1.3) γενικεύεται στον L^2 , αλλά όχι με προφανή τρόπο. Στην βιβλιογραφία που εμείς έχουμε υπ' όψιν μας, η απόδειξη απαιτεί την χρήση επιχειρημάτων πυκνότητας ειδικών υποχώρων του L^2 [BENFRA].
3. Το θεώρημα 1.2 είναι γνωστό σαν «αρχή απροσδιοριστίας Heisenberg» επειδή αρχικά διατυπώθηκε στην κβαντομηχανική ως φυσικός νόμος, μία συνθήκη ανάμεσα στην θέση και την ορμή ενός ελεύθερου κβαντικού σωματιδίου.

1.2 ΣΤΟΙΧΕΙΑ ΧΡΟΝΟΥ-ΣΥΧΝΟΤΗΤΑΣ ΚΑΙ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΣ GABOR

Ο μετασχηματισμός Gabor (Gabor Transform – GT), επίσης γνωστός και ως «πλαισιωμένος μετασχηματισμός Fourier» (windowed Fourier transform), συσχετίζει ένα σήμα με μία οικογένεια συναρτήσεων (στοιχείων) χρόνου-συχνότητας,

$$V_g(x, \gamma) = \int f(t) \bar{g}(t-x) e^{-i2\pi\gamma t} dt$$

όπου $g_{x,\gamma}(t) = g(t-x) e^{i2\pi\gamma t}$, $(x, \gamma \in \mathbb{R})$. Εάν $g \in L^2$, $g_{x,\gamma} \in L^2$, $\forall x, \gamma \in \mathbb{R}$ και ο μετασχηματισμός είναι καλά ορισμένος για κάθε x, γ αφού $|V_g(x, \gamma)| = |\langle f, g_{x,\gamma} \rangle| \leq \|f\| \|g\|$. Επιπλέον η $V_g(x, \gamma)$ είναι συνεχής συνάρτηση στο \mathbb{R}^2 :

Πρόταση 1.6: Για κάθε $x, x', \gamma, \gamma' \in \mathbb{R}$, $V_g(x', \gamma') \xrightarrow[\gamma' \rightarrow \gamma]{x' \rightarrow x} V_g(x, \gamma)$

Απόδειξη

Έστω $f, g \in L^2$. Με βάση τους τελεστές χρονικής μετατόπισης $\tau_x g(t) = g(t-x)$ και διαμόρφωσης $M_\gamma g(t) = e^{i2\pi\gamma t} g(t)$, έχουμε $g_{x,\gamma} = M_\gamma \tau_x g$. Επειδή η g είναι ομοιόμορφα συνεχής σε κάθε κλειστό διάστημα του γ , $M_{\gamma'} g(t) \xrightarrow{\gamma' \rightarrow \gamma} M_\gamma g(t) \forall t$. Επίσης λόγω της ισομετρίας του $M_\gamma \forall \gamma$, $\|M_\gamma g\|^2 = \|g\|^2 < +\infty$. Εφαρμογή του θεωρήματος κυριαρχημένης σύγκλισης μας δίνει αμέσως $\|M_{\gamma'} g - M_\gamma g\|^2 \xrightarrow{\gamma' \rightarrow \gamma} 0$.

Εξάλλου

$$\|\tau_{x'} g - \tau_x g\|^2 = \|M_{-x'} \hat{g} - M_{-x} \hat{g}\|^2 \xrightarrow{x' \rightarrow x} 0$$

λόγω της ισομετρίας του FT στον L^2 . Επομένως,

$$\|M_{\gamma'} \tau_{x'} g - M_\gamma \tau_x g\| = \|M_{\gamma'} \tau_{x'} g - M_\gamma \tau_x g \pm M_{\gamma'} \tau_x g\| \leq \|\tau_{x'} g - \tau_x g\| + \|M_{\gamma'} \tau_x g - M_\gamma \tau_x g\| \xrightarrow[\gamma' \rightarrow \gamma]{x' \rightarrow x} 0.$$

Τελικά,

$$|V_g f(x', \gamma') - V_g f(x, \gamma)| = \left| \langle f, M_{\gamma'} \tau_{x'} g - M_\gamma \tau_x g \rangle \right| \leq \|f\| \cdot \|M_{\gamma'} \tau_{x'} g - M_\gamma \tau_x g\| \xrightarrow[\gamma' \rightarrow \gamma]{x' \rightarrow x} 0$$

QED

Θεώρημα 1.7: Οι τιμές του GT είναι στοιχεία του $L^2(\mathbb{R}^2)$. Ειδικότερα, ο GT είναι ισομετρία. Ως εκ τούτου ο V_g είναι ισομετρικός ομοιομορφισμός στο πεδίο τιμών του. Επιπλέον κάθε συνάρτηση f του L^2 μπορεί ν' ανακτηθεί από τον GT της και ισχύει:

$$f(t) = \iint V_g f(x, \gamma) g(t-x) e^{i2\pi\gamma t} d\gamma dx \quad (1.12)$$

Απόδειξη

Η απόδειξη αποτελείται από δύο μέρη. Αποδεικνύουμε πρώτα την ισομετρία του GT.

Η σχέση ορισμού του GT γράφεται ισοδύναμα

$$V_g f(x, \gamma) = e^{-i2\pi\gamma x} \int f(t) \bar{g}(t-x) e^{-i2\pi\gamma(x-t)} dt = e^{-i2\pi\gamma x} \left(f * \overline{jM_\gamma g}(x) \right) \quad (1.13)$$

όπου ο τελεστής M_γ έχει οριστεί στην απόδειξη της πρότασης 1.6, * είναι το σύμβολο συνέλιξης δύο συναρτήσεων και j η ανάκλαση $jg(t) = g(-t)$. Χρησιμοποιούμε τις γνωστές σχέσεις

$$\widehat{jh} = \overline{\widehat{h}} \text{ και } \widehat{M_\gamma h} = \tau_\gamma \widehat{h}, \quad (h \in L^2(\mathbb{R})),$$

οπότε ο FT της (1.13) γράφεται

$$\widehat{V_g f}(\xi, \gamma) = \left[M_{-\gamma} (f * \overline{jM_\gamma g}) \right]^\wedge = \tau_{-\gamma} (\hat{f} \cdot \tau_\gamma \hat{g}) = \hat{f}(\xi + \gamma) \overline{\hat{g}(\xi)} \quad (1.14)$$

Εάν η $f \in L^1$, η \hat{f} είναι φραγμένη και

$$V_g f(\cdot, \gamma) \in L^2 \quad \forall \gamma \quad (1.15)$$

Επομένως, εφαρμόζοντας διαδοχικά τα θεωρήματα Fubini (για θετικές συναρτήσεις) και Plancherel και, χρησιμοποιώντας την (1.14), τελικά έχουμε

$$\|V_g f\|^2 = \iint |V_g f(x, \gamma)|^2 dx d\gamma = \int |\hat{g}(\xi)|^2 \int |\hat{f}(\xi + \gamma)|^2 d\gamma d\xi = \|f\|^2 \quad (1.16)$$

Δηλαδή ο GT είναι ισομετρία για φραγμένες και τετραγωνικά ολοκληρώσιμες συναρτήσεις. Δεδομένου ότι ο $L^2 \cap L^1$ είναι πυκνός στον L^2 , η (1.16) είναι ισομετρία σε όλον τον L^2 .

Απόδειξη της (1.12):

Εστω $t \in \mathbb{R}$. Εάν $f \in L^2 \cap L^1$ από την (1.15) και το θεώρημα Plancherel έχουμε

$$\int V_g f(x, \gamma) g(t-x) e^{i2\pi\gamma t} dx = \int \hat{f}(\xi + \gamma) |\hat{g}(\xi)|^2 e^{i2\pi(\gamma + \xi)t} d\xi \quad (1.17)$$

Επιπλέον, $\iint |\hat{f}(\xi + \gamma)| |\hat{g}(\xi)|^2 d\xi d\gamma \leq \|f\|_\infty \|g\|_2^2$ οπότε τ' ακόλουθα είναι αποτέλεσμα εφαρμογής του θεωρήματος Fubini:

$$\begin{aligned} \iint V_g f(x, \gamma) g(t-x) e^{i2\pi\gamma t} dx d\gamma &= \int \left(\int V_g f(x, \gamma) g(t-x) e^{i2\pi\gamma t} dx \right) d\gamma \\ &= \int \left(\int \hat{f}(\xi + \gamma) |\hat{g}(\xi)|^2 e^{i2\pi(\xi + \gamma)t} d\xi \right) d\gamma = \int |\hat{g}(\xi)|^2 \left(\int \hat{f}(\xi + \gamma) e^{i2\pi(\xi + \gamma)t} d\gamma \right) d\xi \end{aligned} \quad (1.18)$$

Η (1.12) προκύπτει επειδή $\int |\hat{g}(\xi)|^2 \left(\int \hat{f}(\xi + \gamma) e^{i2\pi(\xi + \gamma)t} d\gamma \right) d\xi = f(t)$ t -σχεδόν παντού.

Όπως και στο πρώτο σκέλος, η (1.12) ισχύει τελικά σε όλον τον L^2 λόγω της πυκνότητας του $L^2 \cap L^1$ στον L^2 .

QED

Παρατηρήσεις

1. Στην πρώτη ισότητα της σχέσης (1.18) έχουμε σιωπηρά χρησιμοποιήσει το θεώρημα Fubini. Αυτό δικαιολογείται από την ανισότητα Hölder

$$\iint |V_g f(x, \gamma) g(t-x) e^{i2\pi\gamma t}| dx \leq \|V_g f(x, \gamma)\| \|g(t-x) e^{i2\pi\gamma t}\|$$

2. Η (1.17) ισχύει και χωρίς την υπόθεση $f \in L^1$. Ωστόσο, η ολοκληρωσιμότητα της f χρειάζεται ούτως ή άλλως στην συνέχεια και γι' αυτό την χρησιμοποιήσαμε εξ' αρχής στο δεύτερο σκέλος της απόδειξης. Στην γενική περίπτωση, η (1.17) προκύπτει ως εξής:

$$\int \left(\int |V_g f(x, \gamma)|^2 dx \right) d\gamma = \iint |V_g f(x, \gamma)|^2 dx d\gamma < \infty \text{ επομένως } \int |V_g f(x, \gamma)|^2 dx < \infty$$

γ -σχεδόν παντού.

3. Η ισομετρία του GT συνεπάγεται ότι οι (πεπερασμένοι) γραμμικοί συνδυασμοί των στοιχείων είναι πυκνοί στον L^2 αφού

$$\langle f, M_\gamma \tau_x g \rangle = 0 \rightarrow V_g f(x, \gamma) = 0 \rightarrow \|f\| = \|V_g f\| = 0 \rightarrow f(t) = 0 \text{ σ.π.} - t$$

1.3 ΔΙΑΣΠΑΣΕΙΣ ΣΕ ΔΙΑΚΡΙΤΑ ΠΛΕΓΜΑΤΑ ΧΡΟΝΟΥ-ΣΥΧΝΟΤΗΤΑΣ

Η αναπαράσταση (1.12) μας λέει ότι κάθε τετραγωνικά ολοκληρώσιμη συνάρτηση, περιγράφεται από τις μέσες τιμές της γύρω από (όλα) τα σημεία (x, γ) του επιπέδου χρόνου – συχνότητας. Επιπλέον η αναπαράσταση είναι ευσταθής, υπό την έννοια ότι δύο συναρτήσεις είναι παρόμοιες αν και μόνον εάν ο GT τους είναι επίσης παρόμοιες.

Ο GT είναι αριθμητικά χρήσιμος, εάν το συνεχές του επιπέδου (x, γ) μπορεί ν' αντικατασταθεί από ένα **διακριτό** πλέγμα σημείων (x, γ) . Με τον τρόπο αυτόν η σχέση ανακατασκευής (1.12) γράφεται

$$f(t) = \sum_{x, \gamma} v_{x, \gamma} g_{x, \gamma}(t), \quad (g_{x, \gamma}(t) = g(t-x) e^{i2\pi\gamma t}) \quad (1.19)$$

Επιπλέον, εάν τα στοιχεία $g_{x, \gamma}(t) = g(t-x) e^{i2\pi\gamma t}$ είναι ορθοκανονική βάση του L^2 , η διαδικασία (1.19) απλοποιείται υπολογιστικά επειδή

$$f(t) = \sum_{x, \gamma} \langle f, g_{x, \gamma} \rangle g_{x, \gamma}(t) \quad (1.20)$$

Στην ενότητα αυτήν, αναζητούμε συνθήκες υπό τις οποίες ομοιόμορφα πλέγματα της μορφής $\{(x, \gamma) = (nx_0, m\gamma_0) \mid n, m \in \mathbb{Z}\}$ μπορούν να παράγουν ορθοκανονικές βάσεις $g_{n,m} = g_{nx_0, m\gamma_0}$ του L^2 .

Αναλυτικότερα, αποδεικνύουμε ότι για να συμβεί το παραπάνω, θα πρέπει $x_0\gamma_0 = 1$. Στην περίπτωση μάλιστα αυτήν, η σχέση μεταξύ του χρονικού και φασματικού φορέα του στοιχείου g είναι «κακή», δηλαδή ισχύει $\Delta_g t \Delta_g \gamma = +\infty$.

Πρόταση 1.8: Εάν το σύστημα $(g_{n,m}(t) = g(t - nx_0)e^{i2\pi m\gamma_0 t}, n, m \in \mathbb{Z})$ είναι ορθοκανονική βάση του L^2 τότε $x_0\gamma_0 = 1$.

Απόδειξη

Αποδεικνύουμε πρώτα ότι αν το σύστημα $(g_{n,m}, n, m \in \mathbb{Z})$ είναι ορθοκανονική βάση, τότε

$$\sum_n |g_n(t)|^2 = \gamma_0, \quad (g_n(t) = g(t - nt_0)) \quad (1.21)$$

Η παραπάνω σχέση οφείλεται στο γεγονός ότι το σύστημα $(e_m(t) = \sqrt{\gamma_0} e^{i2\pi\gamma_0 mt}, m \in \mathbb{Z})$ είναι ορθοκανονική βάση για κάθε συνάρτηση f με φορέα που περιέχεται σε ένα διάστημα μήκους $1/\gamma_0$. Τότε, για κάθε $n \in \mathbb{Z}$ έχουμε:

$$\sum_m \left| \left\langle f, g_{n,m} \right\rangle \right|^2 = \sum_m \left| \left\langle f \bar{g}_n, \frac{1}{\sqrt{\gamma_0}} e_m \right\rangle \right|^2 = \frac{1}{\gamma_0} \int |f(t)|^2 |g(t - nt_0)|^2 dt$$

επομένως,

$$\|f\|^2 = \sum_{n,m} \left| \left\langle f, g_{n,m} \right\rangle \right|^2 = \frac{1}{\gamma_0} \sum_n \int |f(t)|^2 |g(t - nt_0)|^2 dt = \frac{1}{\gamma_0} \int |f(t)|^2 \sum_n |g(t - nt_0)|^2 dt \quad (1.22)$$

Εάν $\sum_n |g(t - nt_0)|^2 > \gamma_0$ σε κάποιο σύνολο μη μηδενικού μέτρου Δ , τότε για την $f\chi_\Delta$ έχουμε βάσει της (1.22) $\|f\chi_\Delta\|^2 > \int |f\chi_\Delta(t)|^2 dt$ πράγμα άτοπο. Για τον ίδιο λόγο δεν μπορεί να είναι $\sum_n |g(t - nt_0)|^2 < \gamma_0$ σε κανένα σύνολο μη μηδενικού μέτρου.

Η πρόταση τώρα αποδεικνύεται εύκολα από την ορθοκανονικότητα των $(g_{n,m})$:

$$1 = \|g\|^2 = \int |g(t)|^2 dt = \int_0^{t_0} \sum_n |g(t - nt_0)|^2 dt = t_0 \gamma_0$$

QED

Θεώρημα 1.9 (Balian—Low [HW96]): Εάν το σύστημα

$$(g_{n,m}(t) = g(t - nt_0)e^{i2\pi m\gamma_0 t}, n, m \in \mathbb{Z})$$

είναι ορθοκανονική βάση του L^2 τότε $\Delta_g t \Delta_g \gamma = +\infty$

Απόδειξη

Για τον λόγο που αναφέρεται στην απόδειξη του θεωρήματος 1.2, θεωρούμε ότι η g έχει μηδενικές μέσες τιμές σε χρόνο και συχνότητα. Επιπλέον, βάσει της πρότασης 1.8 εάν το σύστημα $g_{n,m}$ είναι ορθοκανονική βάση τότε $x_0 \gamma_0 = 1$.

Χωρίς βλάβη της γενικότητας υποθέτουμε τέλος ότι $x_0 = \gamma_0 = 1$ (βλ. παρατήρηση), οπότε

μένει να δειχθεί ότι εάν το σύστημα $(g_{n,m}(t) = g(t - n)e^{i2\pi m t}, n, m \in \mathbb{Z})$ είναι ορθοκανονική

βάση του L^2 , τότε

$$\Delta_g t \Delta_g \gamma = \int t^2 |g(t)|^2 dt \cdot \int \gamma^2 |\hat{g}(\gamma)|^2 d\gamma = +\infty \quad (1.23)$$

Αποδεικνύουμε την (1.23) με εις άτοπον απαγωγή και ορίζουμε όπως και στην ενότητα 1.1 τον τελεστή X από την σχέση:

$Xg(t) = tg(t)$. Η σχέση των X και D με την (1.23) είναι

$$\int |Xg(t)|^2 dt \int |Dg(t)|^2 dt = \frac{1}{4\pi^2} \Delta_g t \Delta_g \gamma \quad (1.24)$$

Ισχυριζόμαστε ότι ισχύουν τ' ακόλουθα:

$$\langle Xg, Dg \rangle = \sum_{m,n} \langle Xg, g_{m,n} \rangle \langle g_{m,n}, Dg \rangle \quad (1.25)$$

$$\langle Xg, g_{m,n} \rangle = \langle g_{-m,-n}, Xg \rangle \quad (1.26)$$

και

$$\langle Dg, g_{m,n} \rangle = \langle g_{-m,-n}, Dg \rangle \quad (1.27)$$

Οι σχέσεις (1.25), (1.26) και (1.27) συνεπάγονται ότι

$$\langle Xg, Dg \rangle = -\langle Dg, Xg \rangle \quad (1.28)$$

Αυτό όμως δεν μπορεί να ισχύει γιατί ολοκληρώνοντας την (1.28) κατά μέρη έχουμε

$$\langle Xg, Dg \rangle = \int tg(t)\bar{g}'(t)dt = -\int (g'(t)t + g(t))\bar{g}(t)dt = -1 - \langle Dg, Xg \rangle = \langle Xg, Dg \rangle - 1$$

που είναι άτοπο⁵.

⁵ Όπως και στην απόδειξη της συνθήκης απροσδιοριστίας, θεωρούμε ότι τα Dg , Xg και DXg είναι στοιχεία του L^2 . Μία επίσης ικανή συνθήκη με βάση το παράρτημα είναι η απόλυτη συνέχεια της g και η τετραγωνική ολοκληρωσιμότητα της Dg .

Για την (1.25) έχουμε:

$$\langle \chi_g, Dg \rangle = \left\langle \sum_{m,n} \langle \chi_g, g_{m,n} \rangle g_{m,n}, Dg \right\rangle = \sum_{m,n} \langle \chi_g, g_{m,n} \rangle \langle g_{m,n}, Dg \rangle$$

Για την (1.26) χρησιμοποιούμε τις ταυτότητες, $e^{-i2\pi mn} = 1$ και $-n \langle g, g_{m,n} \rangle = 0$. Η ισχύς της δεύτερης για $n \neq 0$ προκύπτει από την ορθογωνιότητα των $g_{m,n}$.

$$\langle \chi_g, g_{m,n} \rangle = \int t g(t) \bar{g}(t-n) e^{i2\pi mt} dt = \int (y+n) g(y+n) \bar{g}(y) e^{i2\pi m(y+n)} dy = \langle g_{-m,-n}, \chi_g \rangle$$

Για την (1.27) χρησιμοποιούμε ολοκλήρωση κατά μέρη, οπότε

$$\begin{aligned} \langle Dg, g_{m,n} \rangle &= - \int g(t) (\bar{g}'(t-n) - i2\pi m \bar{g}(t-n)) e^{-i2\pi mt} dt \\ &= - \int g(t) \bar{g}'(t-n) e^{-i2\pi mt} dt \\ &= - \int g(y+n) \bar{g}'(y) e^{-i2\pi mt} dy \\ &= - \langle g_{-m,-n}, Dg \rangle \end{aligned}$$

QED

Παρατήρηση

Υπό την προϋπόθεση ότι $x_0 \gamma_0 = 1$, η παραδοχή $x_0 = \gamma_0 = 1$ δεν περιορίζει την γενικότητα επειδή ο τελεστής $\rho_x g(t) = \sqrt{x} g(xt)$ είναι ορθομοναδιαίος (unitary) και επιπλέον

$$\rho_{\gamma_0} M_m \tau_n g = M_{m\gamma_0} \tau_{nt_0} \rho_{\gamma_0} g$$

όπου M και τ είναι οι ορθομοναδιαίοι τελεστές διαμόρφωσης και μετατόπισης που έχουν οριστεί στην ενότητα 1.2. Αναλυτικότερα, οι ορθομοναδιαίοι τελεστές διατηρούν τα εσωτερικά γινόμενα, επομένως

$$\langle M_{m'\gamma_0} \tau_{n't_0} \rho_{\gamma_0} g, M_{m\gamma_0} \tau_{nt_0} \rho_{\gamma_0} g \rangle = \langle \rho_{\gamma_0} M_{m'} \tau_{n'} g, \rho_{\gamma_0} M_m \tau_n g \rangle = \delta_{m'm} \delta_{n'n}.$$

Επιπλέον για τις διασπορές των g και $\rho_{\gamma_0} g$ πολύ εύκολα αποδεικνύεται ότι ισχύει

$$\Delta_{\rho_{\gamma_0} g} t = \frac{1}{\gamma_0} \Delta_g t \text{ και } \Delta_{\rho_{\gamma_0} g} \gamma = \gamma_0 \Delta_g \gamma \text{ οπότε } \Delta_{\rho_{\gamma_0} g} t \Delta_{\rho_{\gamma_0} g} \gamma < +\infty \leftrightarrow \Delta_g t \Delta_g \gamma < +\infty.$$

////

ΚΕΦΑΛΑΙΟ 2

ΤΟΠΙΚΕΣ ΤΡΙΓΩΝΟΜΕΤΡΙΚΕΣ ΒΑΣΕΙΣ

“ The greatest challenge to any thinker is stating the problem in a way that will allow a solution „
Bertrand Russell

Βάσει του θεωρήματος Balian-Low οι χρονικές και φασματικές τιμές ορθοκανονικών συστημάτων Gabor $(g_{n,m}(t) = g(t - nt_0)e^{i2\pi m\gamma_0 t}, n, m \in \mathbb{Z})$, έχουν κακές διασπορές υπό την έννοια $\Delta_g t \Delta_g \gamma = +\infty$. Οι τοπικές όμως ιδιότητες ενός τέτοιου συστήματος μπορούν να βελτιωθούν – στο πλαίσιο πάντοτε της συνθήκης απροσδιοριστίας $(\Delta_g t \Delta_g \gamma \geq \frac{1}{4\pi})$ – εάν τα μιγαδικά εκθετικά αντικατασταθούν από ειδικές τριγωνομετρικές συναρτήσεις σε συνδυασμό με την επιβολή κάποιων συνθηκών συμμετρίας στην g .

Σε αυτό το κεφάλαιο κατασκευάζουμε ορθοκανονικές βάσεις της μορφής

$$g_{x,\gamma}(t) = g(t-x)\cos(\pi\gamma(t-x)) \quad \text{και} \quad g_{x,\gamma}(t) = g(t-x)\sin(\pi\gamma(t-x))$$

όπου η g είναι λεία (πολλές φορές συνεχώς παραγωγίσιμη) με συμπαγή φορέα και, τα x και γ ανήκουν σε ένα διακριτό πλέγμα του \mathbb{R}^2 . Λόγω της παραγωγισιμότητας της g , τα στοιχεία $g_{\gamma,x}$ είναι καλά εντοπισμένα στην συχνότητα αφού τότε

$$\int |\gamma^j \hat{g}_{x,\gamma}(\xi)|^2 d\xi < +\infty, \forall j \in \mathbb{N}.$$

Τα παραπάνω συστήματα είναι γνωστά ως “τοπικές ορθοκανονικές βάσεις” [HW96] ή “βάσεις Malvar” [MEYER93] ή “lapped orthogonal transforms” [MALLAT2].

Δείχνουμε επιπλέον ότι οι βάσεις Malvar μπορούν να χρησιμοποιηθούν για την κατασκευή των wavelets των Lemarie και Meyer, πράγμα εντυπωσιακό αφού οι βάσεις Malvar παράγονται από τις τριγωνομετρικές διαμορφώσεις των στοιχείων $\tau_x g$ ενώ, τα wavelets προκύπτουν από διαστολές (dilations) των παραπάνω στοιχείων.

2.1 ΤΟΠΙΚΕΣ ΟΡΘΟΚΑΝΟΝΙΚΕΣ ΠΡΟΒΟΛΕΣ

Στην ενότητα αυτήν, δείχνουμε ότι κάθε τετραγωνικά ολοκληρώσιμη συνάρτηση μπορεί να γραφεί ως ορθοκανονικό άθροισμα τοπικών προβολών της. Αναλυτικότερα αποδεικνύουμε ότι σε κάθε διαμέριση $\mathbb{R} = \cup_{j \in \mathbb{Z}} \mathbb{I}_j$ του \mathbb{R} σε διαστήματα \mathbb{I}_j μπορούμε ν' αντιστοιχίσουμε μία διάσπαση $L^2 = \oplus_{j \in \mathbb{Z}} V_j$ του L^2 σε ορθογώνιους υποχώρους του, ώστε κάθε V_j να περιλαμβάνει μόνον συναρτήσεις εντοπισμένες στον διάστημα \mathbb{I}_j . Με τον τρόπο αυτόν μπορούμε να κατασκευάσουμε τοπικές ορθοκανονικές βάσεις στον L^2 , ενώνοντας ορθοκανονικές βάσεις των V_j .

Το ουσιώδες στοιχείο, που καθιστά την παραπάνω κατασκευή μη-τετριμμένη, είναι ότι τα στοιχεία κάθε V_j αν και εντοπισμένα στο διάστημα \mathbb{I}_j έχουν γενικά φορείς που υπερβαίνουν τα όρια του \mathbb{I}_j και επομένως η καθετότητα μεταξύ των προβολών f_j μίας συνάρτησης $f \in L^2$ στους V_j , δεν είναι καθόλου προφανής. Ξεκινούμε ορίζοντας ορθοκανονικές προβολές στο $[0, +\infty)$.

Πρόταση 2.1: Έστω $\delta > 0$ και $g \in L^2(\mathbb{R})$ με

$$\text{supp } g = [-\delta, +\infty), \quad |g(t)|^2 + |jg(t)|^2 = 1 \quad (\forall t \in \mathbb{R}) \quad (2.1)$$

όπου $jg(t) = g(-t)$ (βλ. κεφ. 1). Τότε ο τελεστής

$$S_0 f = \bar{g}(gf + j(gf)), \quad (f \in L^2) \quad (2.2)$$

είναι ορθοκανονική προβολή. Θα λέμε ειδικότερα ότι ο S_0 είναι **ορθοκανονική προβολή στο διάστημα $[0, +\infty)$** , εννοώντας ότι οι φορείς των εικόνων του βρίσκονται εντός του $[-\delta, +\infty)$.

Απόδειξη

Αποδεικνύουμε πρώτα ότι ο S_0 είναι αυτοσυζυγής⁶:

⁶ Ένας γραμμικός τελεστής $T: \mathcal{H} \rightarrow \mathcal{H}$ είναι ορθοκανονική προβολή στον \mathcal{H} αν ο T είναι ταυτοτικός στα στοιχεία του πεδίου τιμών του $T(\mathcal{H})$ και, επιπλέον ο πυρήνας του $\mathcal{N}(T)$ είναι κάθετος στο $T(\mathcal{H})$. Προφανώς, η πρώτη συνθήκη γράφεται ισοδύναμα $T^2 x = Tx$ ($x \in \mathcal{H}$). Εξάλλου, εάν ο T είναι αυτοσυζυγής έχουμε $\langle T^2 x, x \rangle = \langle Tx, Tx \rangle = \langle Tx, x \rangle \Rightarrow \langle Tx, x - Tx \rangle = 0$ οπότε $\langle T(\mathcal{H}), \mathcal{N}(T) \rangle = 0$.

$$\begin{aligned}
 & \langle S_0 f, h \rangle \\
 &= \langle \bar{g} g f, h \rangle + \langle \bar{g} j(gf), h \rangle \\
 &= \langle f, \bar{g} g h \rangle + \langle g f, j(g h) \rangle \\
 &= \langle f, \bar{g} (g h + j(g h)) \rangle \\
 &= \langle f, S_0 h \rangle
 \end{aligned} \tag{2.3}$$

Οι μετασχηματισμοί στην τρίτη από τις παραπάνω ισότητες δικαιολογούνται εύκολα, γράφοντας τα εσωτερικά γινόμενα ως ολοκληρώματα και αλλάζοντας την μεταβλητή ολοκλήρωσης $-t \rightarrow t$.

Σκοπεύουμε να δείξουμε ότι ο S_0 είναι idempotent, δηλ. $(S_0)^2 = S_0$. Αυτό είναι άμεση απόρροια των ακόλουθων προτάσεων:

$$(2.4) \quad \text{Εάν η } h \text{ είναι άρτια, τότε } S_0(fh) = hS_0f \text{ και}$$

$$(2.5) \quad S_0\bar{g} = \bar{g}$$

Πράγματι, εάν η h είναι άρτια τότε $jh = h$ και η (2.4) προκύπτει αμέσως αντικαθιστώντας την f με την hf στον ορισμό (2.2).

Η πρόταση (2.5) είναι απόρροια της συνθήκης $|g(t)|^2 + |jg(t)|^2 = 1$ και προκύπτει αντικαθιστώντας την f με την \bar{g} στην σχέση ορισμού (2.2).

Επειδή η $(gf + j(gf))$ είναι άρτια έχουμε

$$(S_0)^2 f = S_0[\bar{g}(gf + j(gf))] = (gf + j(gf))S_0\bar{g} = \bar{g}(gf + j(gf)) = S_0 f.$$

Τελικά $\forall f \in L^2$,

$$\langle S_0 f, f - S_0 f \rangle = \langle S_0 f, f \rangle - \langle (S_0)^2 f, f \rangle = 0 \tag{2.6}$$

δηλαδή ο S_0 είναι ορθοκανονική προβολή.

QED

Παρατηρήσεις

1. Ο S_0 έχει δύο επιπλέον αναπαραστάσεις που θα φανούν ιδιαίτερα χρήσιμες στην συνέχεια:

Από τον ορισμό (2.1) προκύπτει $|g(t)| = 1$ αν $t \in [\delta, +\infty)$ και $|g(t)| = 0$ αν $t \in (-\infty, -\delta]$ επομένως,

$$S_0 f = S_0(\chi_{[-\delta, \delta]} f) + \chi_{(\delta, +\infty)} f = \chi_{[-\delta, \delta]} S_0 f + \chi_{(\delta, +\infty)} f \tag{2.7}$$

όπου, $S_0(\chi_{[-\delta, \delta]} f) = \chi_{[-\delta, \delta]} S_0 f$ βάσει της (2.4) επειδή η $\chi_{[-\delta, \delta]}$ είναι άρτια.

Εξάλλου,

$$S_0 f = \bar{g}(gf + j(gf)) = \bar{g}f_0 \quad (2.8)$$

όπου f_0 είναι μία άρτια συνάρτηση στο $[-\delta, \delta]$ και ίση με την f στο $[\delta, +\infty)$.

2. Ο S_0 γενικεύεται με πολλούς τρόπους:

- i.* Γράφοντας $S_\alpha = \tau_\alpha S_0 \tau_{-\alpha}$ προκύπτουν ορθοκανονικές προβολές στο διάστημα $[\alpha, +\infty)$, $\alpha \in \mathbb{R}$. Οι τελεστές $S_\alpha = \tau_\alpha S_0 \tau_{-\alpha}$ είναι ορθοκανονικές προβολές, επειδή ο τελεστής μετατόπισης τ_α είναι ορθομοναδιαίος (unitary)⁷.

Μετά τις απαραίτητες αντικαταστάσεις στον ορισμό του S_α , προκύπτει

$$S_\alpha f(t) = \bar{g}(t-\alpha)[g(t-\alpha)f(t) + g(\alpha-t)f(2\alpha-t)] \quad (2.9)$$

Θέτοντας $\bar{g}(t-\alpha) = \bar{g}_\alpha$ η (2.9) γράφεται ισοδύναμα

$$S_\alpha f(t) = \bar{g}_\alpha(t)[g_\alpha(t)f(t) + g_\alpha(2\alpha-t)f(2\alpha-t)] \quad (2.10)$$

Παρατηρώντας ότι η $f(2\alpha-t)$ είναι αντικατοπτρισμός της f ως προς το α , η (2.9) γράφεται ισοδύναμα

$$S_\alpha f(t) = \bar{g}(t-\alpha)f_\alpha = \bar{g}_\alpha f_\alpha$$

όπου $f_\alpha(\alpha+t) = f_\alpha(\alpha-t)$ στο $[-\delta, \delta]$ και $f_\alpha(t) = f(t)$ στο $[\alpha+\delta, +\infty)$.

- ii.* Επιλέγοντας την συνάρτηση πλαισίωσης g στον ορισμό (2.1) έτσι ώστε

$$\text{supp } g = (-\infty, \delta], \quad |g(t)|^2 + |jg(t)|^2 = 1 \quad (\forall t \in \mathbb{R})$$

προκύπτουν ορθοκανονικές προβολές στο $(-\infty, 0]$

$$S^0 f = \bar{g}(gf + j(gf)), \quad (f \in L^2, \text{ supp } g \in (-\infty, \delta]) \quad (2.11)$$

και βάσει της *i*, ορθοκανονικές προβολές $S^\beta = \tau_\beta S^0 \tau_{-\beta}$ στο διάστημα $(-\infty, \beta] \forall \beta \in \mathbb{R}$.

- iii.* Εύκολα επίσης αποδεικνύεται ότι αν αντί της (2.2) ορίσουμε

$$S_0 f = \bar{g}(gf - j(gf)), \quad (f \in L^2) \quad (2.12)$$

τότε ο S_0 είναι επίσης ορθοκανονική προβολή. Η μόνη διαφορά σε σχέση με την απόδειξη της πρότασης 2.1 είναι η (2.4) που τώρα γράφεται:

⁷ $S_\alpha^2 = \tau_\alpha S_0 \tau_{-\alpha} \tau_\alpha S_0 \tau_{-\alpha} = \tau_\alpha S_0 \tau_{-\alpha} = S_\alpha$ και ο S_α είναι αυτοσυζηγής επειδή $\tau_\alpha^* = \tau_\alpha^{-1} = \tau_{-\alpha}$.

$$(2.4)' \quad \text{Εάν η } h \text{ είναι περιττή, τότε } \bar{g}(gh - j(gh)) = h[\bar{g}(gf + j(gf))].$$

Η (2.12) διαφέρει από την (2.2) μόνον στο πρόσημο εντός της παρένθεσης. Η αναπαράσταση (2.7) παραμένει η ίδια και, η $S_0 f$ είναι τώρα το γινόμενο της \bar{g} με μία **περιττή** συνάρτηση στο $[-\delta, \delta]$.

Λόγω της άρτιας συμμετρίας της $gf + j(gf)$ στο $[-\delta, \delta]$, αναφέρουμε τις προβολές $S_0 f = \bar{g}(gf + j(gf))$, ($f \in L^2$) ως **συμμετρικές** ή **ισοδύναμα** λέμε ότι έχουν **άρτια συμμετρία**. Αντιστοίχως οι προβολές (2.12) λέγονται **αντισυμμετρικές** ή ότι έχουν **περιττή συμμετρία**.

Είναι προφανές ότι και οι τρεις παραπάνω γενικεύσεις μπορούν να εφαρμοστούν στην αρχική σχέση (2.2) ανεξάρτητα η μία από την άλλη. Έτσι, για οποιοδήποτε $\alpha > 0$ μπορούμε να κατασκευάσουμε συμμετρικές ή αντισυμμετρικές ορθοκανονικές προβολές στο $(-\infty, \alpha]$ ή το $[\alpha, +\infty)$, για οποιοδήποτε θετικό αριθμό δ .

Η συνάρτηση πλαισίωσης g στην (2.1) επιλέγεται συνήθως πραγματική και θετική, επομένως αύξουσα. Επιπλέον απαιτούμε η g να παραγωγίζεται αρκετές φορές ώστε η διασπορά των φασματικών τιμών της να είναι μικρή.

Στο Σχ. 1 που ακολουθεί, εικονίζονται μία συνάρτηση f , η g και η προβολή $S_0 f$ που ορίζεται στην (2.2).

Η f δίνεται από την σχέση:

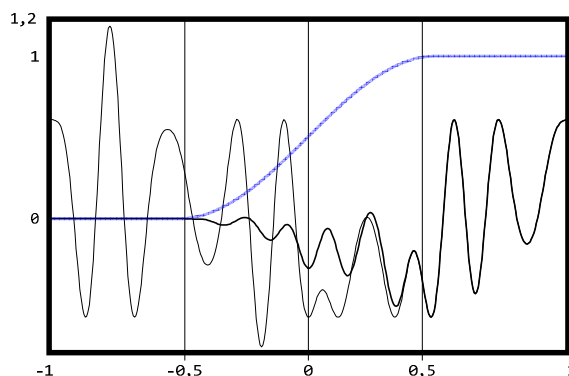
$$f(t) = \left(\frac{1}{3} \cdot y^3(t) - y(t) \right) / \left(1 + \frac{1}{10} t^2 \right), \quad y(t) = \sin(2\pi t) + \sin\left(5\pi\left(t - \frac{1}{3}\right)\right).$$

Ως συνάρτηση πλαισίωσης έχει επιλεγεί η $g(t) = \sin\left(\frac{\pi}{4}(1 + \sin \pi t)\right) \chi_{[-\frac{1}{2}, \frac{1}{2}]}(t) + \chi_{(\frac{1}{2}, +\infty)}(t)$ που ικανοποιεί τις συνθήκες (2.1) στο διάστημα $[-\frac{1}{2}, \frac{1}{2}]$.

Πράγματι,

$$g(t) = \sin\left(\frac{\pi}{4}(1 + y(t))\right), \quad y(t) = \sin \pi t$$

$$\text{και } \sin^2\left(\frac{\pi}{4}(1 - x)\right) = \cos^2\left(\frac{\pi}{2} - \frac{\pi}{4}(1 - x)\right) = \cos^2\left(\frac{\pi}{4}(1 + x)\right) \text{ επομένως } g^2(x) + g^2(-x) = 1.$$



Σχ. 1: Συμμετρική προβολή της συνάρτησης f στο $[0, +\infty)$, $S_0 f = \bar{g}(gf + j(gf))$. Με την λεπτή γραμμή εικονίζεται η αρχική συνάρτηση f και με την έντονη η $S_0 f$.

Επιπλέον η y είναι περιττή στο $[-\frac{1}{2}, \frac{1}{2}]$ οπότε $y(-t) = -y(t)$ και το ζητούμενο προκύπτει με την αντικατάσταση $x \rightarrow y(t)$.

Η πρώτη παράγωγος της g μηδενίζεται στα σημεία $\pm \frac{1}{2}$ αλλά όχι και οι παράγωγοι ανώτερης τάξης. Είναι ωστόσο δυνατόν να κατασκευάσουμε αύξουσες θετικές συναρτήσεις που διαθέτουν συνεχείς παραγώγους κάθε τάξης και επιπλέον ικανοποιούν την συνθήκη (2.1)).

Στο [HW96, §1.1.3] περιγράφεται ένας τρόπος κατασκευής συναρτήσεων $g \in C^\infty$ που ικανοποιούν την συνθήκη $|g(t)|^2 + |g(-t)|^2 = 1$ ($\forall t \in \mathbb{R}$) και στο [MALLAT2] υπονοείται ένας αναδρομικός τρόπος κατασκευής τέτοιων συναρτήσεων, χωρίς απόδειξη. Στο Παράρτημα στο τέλος του κεφαλαίου, παρουσιάζουμε αναλυτικά την αναδρομική κατασκευή του [MALLAT2].

Πρόταση 2.2: Εάν $\alpha < \beta$ και $\delta, \delta' > 0$ με $\delta + \delta' < \beta - \alpha$, ο τελεστής $S_\alpha^\beta = S_\alpha S^\beta$ είναι ορθοκανονική προβολή στο διάστημα $[\alpha, \beta]$. Όπως και στην πρόταση 1.6, εννοούμε ότι ο φορέας των εικόνων του S_α^β βρίσκεται εντός του $[\alpha - \delta, \beta + \delta']$.

Απόδειξη

Αρκεί να δείξουμε ότι οι S^α και S_β αντιμετατίθενται αφού τότε

$$(S_\alpha S^\beta)^2 = S_\alpha S^\beta S_\alpha S^\beta = (S_\alpha)^2 (S^\beta)^2 = S_\alpha S^\beta$$

οπότε ο S_α^β είναι idempotent. Επιπλέον ο S_α^β είναι αυτοσυζυγής κατά προφανή τρόπο, ως σύνθεση αυτοσυζυγών τελεστών.

Έστω

$\Delta_1 = (-\infty, \alpha - \delta]$, $\Delta_2 = [\alpha - \delta, \alpha + \delta]$, $\Delta_3 = [\alpha + \delta, \beta - \delta']$, $\Delta_4 = [\beta - \delta', \beta + \delta']$, $\Delta_5 = [\beta + \delta', +\infty)$
και $\chi_{\Delta_j}, j=1,2,3,4,5$ οι χαρακτηριστικές συναρτήσεις τους.

Από την (2.7) έπεται $S_\alpha f = 0$ στο Δ_1 και $S_\alpha f = f$ στα $\Delta_3, \Delta_4, \Delta_5$. Ομοίως, $S^\beta f = f$ στα $\Delta_1, \Delta_2, \Delta_3$ και $S^\beta f = 0$ στο Δ_5 . Τέλος, λόγω της συμμετρίας των χ_{Δ_2} και χ_{Δ_4} γύρω από τα α και β αντιστοίχως, έχουμε $S_\alpha(\chi_{\Delta_2} f) = \chi_{\Delta_2} S_\alpha f$, $S^\beta(\chi_{\Delta_4} f) = \chi_{\Delta_4} S^\beta f$. Μετά από πράξεις προκύπτει,

$$S_\alpha^\beta f = S_\alpha S^\beta f = S^\beta S_\alpha f = \chi_{\Delta_2} S_\alpha f + \chi_{\Delta_3} f + \chi_{\Delta_4} S^\beta f \quad (2.13)$$

QED

Η αναπαράσταση (2.13) μας δίνει την δυνατότητα να περιγράψουμε την προβολή $S_\alpha^\beta f$ ως γινόμενο μίας συνάρτησης και ενός πλαισίου με φορέα $[\alpha - \delta, \beta + \delta']$:

Έστω $j_\alpha h(t) = h(2\alpha - t)$ ο τελεστής ανάκλασης της συνάρτησης h γύρω από το α . Βάσει της (2.10), $S_\alpha f = \bar{g}_\alpha(g_\alpha f + j_\alpha(g_\alpha f))$, $S^\beta f = \bar{g}^\beta(g^\beta f - j_\beta(g^\beta f))$, $f_\alpha = g_\alpha f + j_\alpha(g_\alpha f)$, $f^\beta = g^\beta f - j_\beta(g^\beta f)$ και $g_\alpha^\beta = g_\alpha g^\beta$ (η επιλογή αντισυμμετρικής S^β είναι τυχαία).

Είναι προφανές ότι $g_\alpha^\beta = g_\alpha$ στο Δ_2 , $g_\alpha^\beta = g^\beta$ στο Δ_4 , $g_\alpha^\beta = g_\alpha = g^\beta = 1$ στο Δ_3 και $g_\alpha^\beta = 0$ στο $\Delta_1 \cup \Delta_5$. Από την (2.13) προκύπτει

$$S_\alpha^\beta f = \bar{g}_\alpha^\beta(\chi_{\Delta_2} f_\alpha + \chi_{\Delta_3} f + \chi_{\Delta_4} f^\beta) = \bar{g}_\alpha^\beta f_\alpha^\beta \quad (2.14)$$

όπου f_α^β είναι μία συνάρτηση συμμετρική ή αντισυμμετρική στα α και β , ανάλογα με την συμμετρία των S_α και S^β αντιστοίχως.

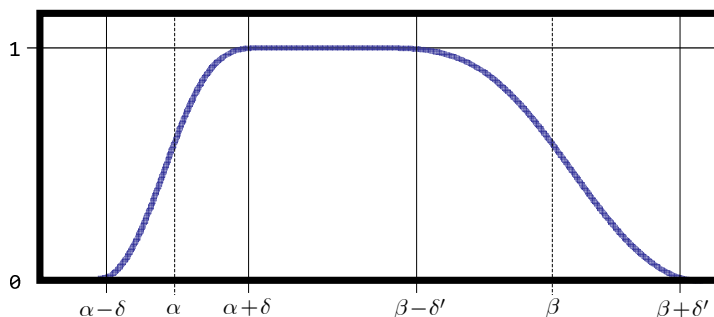
Αντιστρόφως, εάν h είναι μία συνάρτηση συμμετρική στο α και αντισυμμετρική στο β , είναι να εύκολο να δειχθεί με βάση την (2.14) ότι

$$S_\alpha^\beta(\bar{g}_\alpha^\beta h) = h S_\alpha^\beta \bar{g}_\alpha^\beta = \bar{g}_\alpha^\beta h \quad (2.15)$$

Έχουμε δείξει το εξής

Πόρισμα 2.3: Μία συνάρτηση f είναι εικόνα του S_α^β αν και μόνον εάν $f = \bar{g}_\alpha^\beta h$ και η h έχει στα α και β τις συμμετρίες του S_α^β .

Στο Σχ. 2 εικονίζεται το πλαίσιο $g_\alpha^\beta(t) = g_0(t-\alpha)g_0(\frac{\beta-t}{2})$ για την g του Σχ. 1 δηλαδή, $g_0(t) = \sin(\frac{\pi}{4}(1 + \sin \pi t))$.



Σχ. 2: Γράφημα του πλαισίου $g_\alpha^\beta(t) = g_0(t-\alpha) + g_0(\frac{\beta-t}{2})$

Μελετούμε τώρα το πρώτο βασικό θεώρημα του κεφαλαίου, που μας επιτρέπει να συνθέσουμε τις ορθοκανονικές προβολές S_λ^μ και S_ν^μ δύο διαδοχικών διαστημάτων $\Delta_1 = [\lambda, \mu]$ και $\Delta_2 = [\mu, \nu]$ σε μία ορθοκανονική προβολή στο διάστημα $\Delta_1 \cup \Delta_2 = [\lambda, \nu]$.

Λήμμα 2.4: Έστω $\delta > 0$, $\text{supp } g_0 \in [-\delta, +\infty)$ και $g^0 = \overline{jg_0}$. Εάν $S_0 f = g_0(g_0 f + j(g_0 f))$ και $S^0 f = g^0(g^0 f - j(g^0 f))$ τότε $(S_0 + S^0)f = f \forall f \in L^2$. Το ίδιο ισχύει εάν οι συμμετρίες των S_0 και S^0 (άρτια και περιττή αντιστοίχως) αντιστραφούν.

Απόδειξη

Έστω $\Delta_1 = (-\infty, -\delta)$, $\Delta_2 = [-\delta, \delta]$, $\Delta_3 = (\delta, +\infty)$. Στο Δ_1 $g_0 = 0$ και $|g^0| = 1$ και στο Δ_3 $g^0 = 0$ και $|g_0| = 1$ επομένως το λήμμα ισχύει στο $\Delta_1 \cup \Delta_3$. Στο Δ_2 έχουμε

$$\begin{aligned} & \bar{g}_0(g_0 f + j(g_0 f)) + jg_0(\overline{jg_0} f + j(\bar{g}_0 f)) \\ &= f + (\bar{g}_0 j g_0) f - ((jg_0) \bar{g}_0) f = f \end{aligned}$$

QED

Είναι προφανές ότι το παραπάνω λήμμα ισχύει και για κάθε $S_\alpha, \alpha \in \mathbb{R}$ αφού

$$(S_\alpha + S^\alpha) = \tau_\alpha (S_0 + S^0) \tau_{-\alpha} = \tau_\alpha \tau_{-\alpha} = I.$$

Ορισμός 2.5: Έστω $\Delta_1 = [\lambda, \mu]$, $\Delta_2 = [\mu, \nu]$, $\lambda < \mu < \nu$. Θα λέμε ότι τα Δ_1 και Δ_2 έχουν **συμβατά πλαίσια** $g_\lambda^\mu = g_\lambda g^\mu$, $g_\mu^\nu = g_\mu g^\nu$ εάν

i. $\text{supp} g_\lambda^\mu = [\lambda - \delta, \mu + \delta']$ και $\text{supp} g_\mu^\nu = [\mu - \delta', \nu + \delta'']$ για κάποια $0 < \delta < \delta' < \delta''$

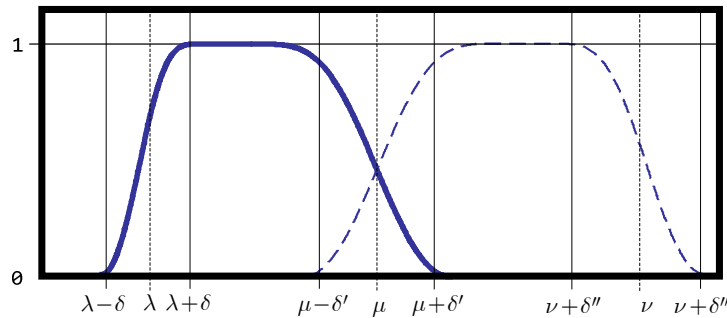
ii. $\lambda + \delta < \mu - \delta' < \mu + \delta' < \nu - \delta''$ και επιπλέον

iii. ισχύει για τις g^μ και g_μ η προϋπόθεση του λήμματος 2.4, δηλαδή $g^\mu = \overline{J_\mu g_\mu}$.

Εάν τα πλαίσια g_λ^μ , g_μ^ν είναι συμβατά, οι προβολές $S_\lambda^\mu f = S_\lambda S^\mu f = g_\lambda^\mu f_\lambda^\mu$ και

$S_\mu^\nu f = S_\mu S^\nu f = g_\mu^\nu f_\mu^\nu$ θα ονομάζονται **συμβατές**, εάν οι S^μ και S_μ έχουν αντίθετες συμμετρίες στο μ .

Στο Σχ. 3 φαίνονται τα συμβατά πλαίσια g_λ^μ , g_μ^ν του ορισμού 1.2.



Σχ. 3: Συμβατά πλαίσια του ορισμού 1.2: g_λ^μ (συνεχής έντονη γραμμή) και g_μ^ν (διακεκομμένη).

Θεώρημα 2.6: Εάν $S_\lambda^\mu f = S_\lambda S^\mu f = \overline{g_\lambda^\mu} f_\lambda^\mu$ και $S_\mu^\nu f = S_\mu S^\nu f = \overline{g_\mu^\nu} f_\mu^\nu$ είναι συμβατές προβολές στα διαστήματα $\Delta_1 = [\lambda, \mu]$, $\Delta_2 = [\mu, \nu]$, $\lambda < \mu < \nu$ αντιστοίχως, τότε

$$S_\lambda^\mu + S_\mu^\nu = S_\lambda^\nu \tag{2.16}$$

και

$$S_\lambda^\mu S_\mu^\nu = 0 \tag{2.17}$$

Απόδειξη

Βάσει του λήμματος 2.4 η απόδειξη της (2.16) είναι σχεδόν προφανής:

Έστω

$$\Delta_1 = [\lambda - \delta, \lambda + \delta), \Delta_2 = [\lambda + \delta, \nu - \delta''], \Delta_3 = [\nu - \delta'', \nu + \delta''] \text{ και}$$

$$\Delta'_1 = [\lambda + \delta, \mu - \delta'], \Delta'_2 = [\mu - \delta', \mu + \delta'], \Delta'_3 = [\mu + \delta', \nu - \delta''].$$

Είναι

$$S'_\lambda f = \chi_{\Delta'_1} \bar{g}_\lambda f_\lambda + \chi_{\Delta'_2} f + \chi_{\Delta'_3} \bar{g}'' f'' \quad (2.18)$$

και

$$\begin{aligned} (S^\mu_\lambda + S^\nu_\mu) f = & \\ & \chi_{\Delta'_1} \bar{g}_\lambda f_\lambda + \chi_{\Delta'_3} \bar{g}'' f'' + \\ & \chi_{\Delta'_1} f + \chi_{\Delta'_3} f + \\ & \chi_{\Delta'_2} \bar{g}_\mu f_\mu + \chi_{\Delta'_2} \bar{g}'' f'' \end{aligned} \quad (2.19)$$

Από το λήμμα 2.4, $\chi_{\Delta'_2} \bar{g}_\mu f_\mu + \chi_{\Delta'_2} \bar{g}'' f'' = \chi_{\Delta'_2} f$ επομένως $S^\mu_\lambda + S^\nu_\mu = S^\nu_\lambda$

Η (2.17) είναι συνέπεια μίας γενικής ιδιότητας των ορθοκανονικών προβολών: εάν οι P, Q και $P + Q$ είναι ορθοκανονικές προβολές τότε $PQ = QP = 0$.

Πράγματι, από το "idempotency" των P, Q και $P + Q$

$$\begin{aligned} (P + Q)^2 &= P + Q \\ \Rightarrow P^2 + Q^2 + PQ + QP &= P + Q \\ \Rightarrow PQ + QP &= 0 \end{aligned} \quad (2.20)$$

Τότε $PQ = P^2Q = -P(QP) = QP^2 = QP$ και λόγω της (2.20) $PQ = QP = 0$.

QED

Παρατηρήσεις

Το θεώρημα 2.6 επιτρέπει την διάσπαση του L^2 σε ένα ευθύ άθροισμα υποχώρων του, με τον τρόπο που αναφέρθηκε στην εισαγωγή του κεφαλαίου. Αναλυτικότερα, εάν $(\alpha_j)_{j \in \mathbb{Z}}$ μία ακολουθία με $\alpha_j \xrightarrow{j \rightarrow \pm \infty} \pm \infty$ και $(\delta_j)_{j \in \mathbb{Z}}$ τέτοια ώστε $\alpha_j + \delta_j < \alpha_{j+1} - \delta_{j+1}, \forall j$, μπορούμε να θέσουμε $V_j = S^{\alpha_{j+1}}_{\alpha_j}(L^2)$ αρκεί οι τελεστές $S^{\alpha_{j+1}}_{\alpha_j}$ να είναι συμβατοί μεταξύ τους, υπό την έννοια του ορισμού 2.5.

Λιγότερο προφανές είναι το γεγονός ότι η ίδια μέθοδος μπορεί να εφαρμοστεί για ορθοκανονικές διασπάσεις υποχώρων του L^2 της μορφής $L^2(\mathbb{I})$ όπου \mathbb{I} είναι ένα οποιοδήποτε του \mathbb{R} .

Π.χ. εάν $\mathbb{I} = [0, +\infty)$ (μία περίπτωση που θα μας απασχολήσει παρακάτω, στην κατασκευή των wavelets των Lemarie και Meyer), τότε $L^2(\mathbb{I}) = \bigoplus_{j \in \mathbb{Z}} S^{\alpha_{j+1}}_{\alpha_j}(L^2)$ αρκεί

$$\alpha_j \xrightarrow{j \rightarrow -\infty} 0.$$

2.2 ΤΟΠΙΚΕΣ ΤΡΙΓΩΝΟΜΕΤΡΙΚΕΣ ΒΑΣΕΙΣ

Πρόταση 2.7: Έστω $S_\alpha^\beta f = \bar{g}_\alpha^\beta f_\alpha^\beta$ μία τοπική ορθοκανονική προβολή στο $[\alpha, \beta]$ με

$$\text{supp} S_\alpha^\beta f \subset [\alpha - \delta, \beta + \delta'] \quad \forall f \in L^2 \quad \text{και} \quad \Delta_1 = [\alpha - \delta, \alpha], \quad \Delta_2 = [\alpha, \beta], \quad \Delta_3 = [\beta, \beta + \delta'].$$

Έστω ακόμη $\{e_m, m \in \mathbb{Z}\}$ μία ορθοκανονική βάση του $L^2(\Delta_2)$, και v_m οι επεκτάσεις των e_m στο $[\alpha - \delta, \beta + \delta']$ με βάση τις συμμετρίες του S_α^β στα α και β ⁸.

Τότε:

1. το $\{\bar{g}_\alpha^\beta v_m, m \in \mathbb{Z}\}$ είναι ορθοκανονική βάση του $S_\alpha^\beta(L^2)$ και,
2. $\langle S_\alpha^\beta f, \bar{g}_\alpha^\beta v_m \rangle = \langle f_\alpha^\beta, e_m \rangle \quad \forall f \in L^2.$

Απόδειξη

Θεωρούμε χωρίς βλάβη της γενικότητας ότι $\alpha = 0, \beta = 1$ και ότι ο S_α^β είναι αντισυμμετρικός στο 0 και συμμετρικός στο 1.

Αποδεικνύουμε πρώτα την ορθοκανονικότητα του $(\bar{g}_0^1 v_m)_{m \in \mathbb{Z}}$ στο $[-\delta, 1 + \delta']$:

Αυτό είναι συνέπεια της συμμετρίας των στοιχείων $v_m, m \in \mathbb{Z}$ και της συνθήκης (2.1) που για την g_0^1 στο 1 γράφεται

$$\left| g_0^1(t) \right|^2 + \left| g_0^1(2-t) \right|^2 = 1 \quad (2.21)$$

Είναι,

$$\begin{aligned} & \int_{-\delta}^{1+\delta'} v_m(t) \bar{v}_n(t) \left| g_0^1(t) \right|^2 dt \\ &= \left\{ \int_{-\delta}^0 + \int_0^\delta + \int_\delta^{1-\delta'} + \int_{1-\delta'}^1 + \int_1^{1+\delta'} \right\} \left\{ v_m(t) \bar{v}_n(t) \left| g_0^1(t) \right|^2 \right\} dt \end{aligned} \quad (2.22)$$

Όμως,

$$\begin{aligned} & \int_1^{1+\delta'} v_m(t) \bar{v}_n(t) \left| g_0^1(t) \right|^2 dt = \int_{1-\delta'}^1 v_m(2-u) \bar{v}_n(2-u) \left| g_0^1(2-u) \right|^2 du \\ &= \int_{1-\delta'}^1 v_m(u) \bar{v}_n(u) \left| g_0^1(2-u) \right|^2 du \end{aligned}$$

επομένως λόγω της (2.21)

$$\left\{ \int_{1-\delta'}^1 + \int_1^{1+\delta'} \right\} \left\{ v_m(t) \bar{v}_n(t) \left| g_0^1(t) \right|^2 \right\} dt = \int_{1-\delta'}^1 v_m(t) \bar{v}_n(t) dt$$

⁸ Π.χ. $e_m(t) = e_m(2\alpha - t), \forall t \in [\alpha, \beta]$ εάν η S_α^β είναι συμμετρική στο α και

$e_m(t) = -e_m(2\beta - t), \forall t \in [\alpha, \beta]$ εάν η S_α^β είναι αντισυμμετρική στο β .

Με τον ίδιο τρόπο αποδεικνύεται ότι

$$\left\{ \int_{-\delta}^0 + \int_0^{\delta} \right\} \left| v_m(t) \bar{v}_n(t) |g_0^1(t)|^2 \right| dt = \int_0^{\delta} v_m(t) \bar{v}_n(t) dt$$

οπότε

$$\int_{-\delta}^{1+\delta'} v_m(t) \bar{v}_n(t) |g_0^1(t)|^2 dt = \int_0^1 v_m(t) \bar{v}_n(t) dt$$

Για τα υπόλοιπα σκέλη της πρότασης, έχουμε τα εξής:

Εάν $t \in \Delta_2$, εξ' υποθέσεως $f_0^1(t) = \sum_m c_m e_m(t) = \sum_m c_m v_m(t)$ όπου $c_m = \int_0^1 f_0^1 e_m$.

Εάν $t \in \Delta_1$

$$f_0^1(t) = -f_0^1(-t) = -\sum_m c_m e_m(-t) = \sum_m c_m v_m(t)$$

Εάν $t \in \Delta_3$

$$f_0^1(t) = f_0^1(2-t) = \sum_m c_m e_m(2-t) = \sum_m c_m v_m(t).$$

Επομένως $\bar{g}_0^1 f_0^1(t) = \sum_j c_j (\bar{g}_0^1 v_m)(t)$ στο $\Delta_1 \cup \Delta_2 \cup \Delta_3 = [-\delta, 1 + \delta']$ οπότε

$S_0^1(L^2) \subset cl \langle \bar{g}_0^1 v_m, m \in \mathbb{Z} \rangle$ ⁹ και $\bar{g}_0^1 v_m \in S_0^1(L^2) \forall m$, λόγω του πορίσματος 2.3.

Επιπλέον με βάση τα παραπάνω, $c_m = \langle f_0^1, e_m \rangle = \langle \bar{g}_0^1 f_0^1, \bar{g}_0^1 v_m \rangle = \langle S_0^1 f, \bar{g}_0^1 v_m \rangle$.

QED

Παρατήρηση

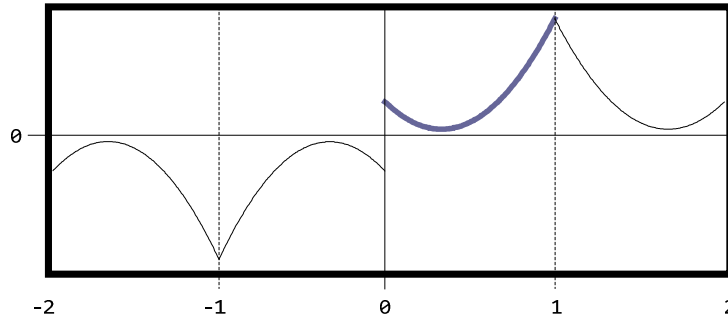
Για συντομία, στην απόδειξη αναλύθηκαν μόνον τα ολοκληρώματα γύρω από το 1 όπου τα v_m έχουν άρτια συμμετρία. Εάν η συμμετρία είναι περιττή δεν αλλάζει τίποτε στην απόδειξη επειδή το γινόμενο $v_m \bar{v}_m$ εξακολουθεί να είναι άρτιο.

2.2.1 Βάσεις Malvar

Κατασκευάζουμε τώρα μία πραγματική ορθοκανονική τριγωνομετρική βάση του $L^2[0,1]$ με τις συμμετρίες της πρότασης 2.7, δηλαδή περιττή στο 0 και άρτια στο 1.

⁹ Συμβολίζουμε με $cl \langle \cdot \rangle$ την κλειστότητα της γραμμικής θήκης ενός συνόλου. Η παραπάνω σχέση δηλώνει ότι ο $S_0^1(L^2)$ είναι ένας υπόχωρος του L^2 . Στην πραγματικότητα, $S_0^1(L^2) = cl \langle \bar{g}_0^1 v_m, m \in \mathbb{Z} \rangle$. Ο $S_0^1(L^2)$ είναι **κλειστός** επειδή ο S_0^1 είναι ορθοκανονική προβολή. Εξάλλου $\bar{g}_0^1 v_m \in S_0^1(L^2), \forall m \in \mathbb{Z}$ επομένως ο $S_0^1(L^2)$ είναι και **πυκνός** στον L^2 .

Θεωρούμε μία συνάρτηση f στο $[0,1]$ και της επέκτασή της στο $[-2,2]$ με $f(t) = f(2-t)$ αν $t \in [1, 2]$ και $f(t) = -f(-t)$ αν $t \in [-2,0]$. Μία τέτοια συνάρτηση φαίνεται στο Σχ. 4.



Σχ. 4

Θεωρούμε επίσης την τυπική τριγωνομετρική βάση στο $L^2[-2,2]$,

$$\left\{ \frac{1}{2}, \frac{1}{\sqrt{2}} \sin \frac{\pi kt}{2}, \frac{1}{\sqrt{2}} \cos \frac{\pi lt}{2}, (k, l = 1, 2, \dots) \right\}$$

Επειδή η f είναι περιττή στο $[-2,2]$, η σταθερά και οι συνημιτονικοί όροι εξαιρούνται από την αναπαράσταση της f . Το ίδιο συμβαίνει με τα στοιχεία $\sin \pi kt$ που είναι περιττά στο 1 και το -1 . Τελικά, στην αναπαράσταση της f συμμετέχουν μόνον οι όροι $\sin\left(\frac{2k+1}{2}\pi t\right)$ και η f γράφεται

$$f(t) = \sum_k c_k \sin\left(\frac{2k+1}{2}\pi t\right), \quad c_k = \frac{1}{2} \int_{-2}^2 f(t) \sin\left(\frac{2k+1}{2}\pi t\right) dt \quad (2.23)$$

όπου η σειρά υγκλίνει μεσοτετραγωνικά στο $[-2,2]$.

Λόγω των συμμετριών των $\sin\left(\frac{2k+1}{2}\pi t\right)$ στα $-1, 0$ και 1 είναι εύκολο ν' αποδείξουμε ότι το σύνολο $\left\{ \sqrt{2} \sin\left(\frac{2k+1}{2}\pi t\right), k = 1, 2, \dots \right\}$ είναι ορθοκανονικό στο $[0,1]$:

Εάν $e_j(t) = \sin\left(\frac{2j+1}{2}\pi t\right)$ και $e_k(t) = \sin\left(\frac{2k+1}{2}\pi t\right)$ με $j \neq k$,

$$\int_0^1 e_j(t) e_k(t) dt = \frac{1}{4} \left[\int_{-2}^{-1} + \int_{-1}^0 + \int_0^1 + \int_1^2 \right] (e_j(t) e_k(t)) dt = \frac{1}{4} \int_{-2}^2 e_j(t) e_k(t) dt = 0$$

Εάν $j = k$,

$$1 = \frac{1}{2} \int_{-2}^2 e_j^2(t) dt = \frac{1}{2} \left[\int_{-2}^{-1} + \int_{-1}^0 + \int_0^1 + \int_1^2 \right] (e_j(t) e_k(t)) dt = \frac{4}{2} \int_0^1 e_j^2(t) dt$$

Επομένως, βάσει της πρότασης 2.7 το $\left\{ \sqrt{2} g_0^1(t) \sin\left(\frac{2k+1}{2}\pi t\right), k = 1, 2, \dots \right\}$ είναι ορθοκανονική βάση του $S_0^1(L^2)$.

Επειδή η παραπάνω κατασκευή γενικεύεται για οποιεσδήποτε συμμετρίες στα 0 και 1 προκύπτει το ακόλουθο

Θεώρημα 2.8: Κάθε ένα από τα παρακάτω συστήματα είναι ορθοκανονική βάση του χώρου στα δεξιά του. Τα σύμβολα + και - αντιστοιχούν σε συμμετρικές και αντισυμμετρικές προβολές στα 0 και 1:

1	$\left\{ \sqrt{2}g_0^1(t)\sin\left(\frac{2k+1}{2}\pi t\right), k=0,1,2,\dots \right\}$	$RangeS_{0-}^{1+}$
2	$\left\{ \sqrt{2}g_0^1(t)\sin(k\pi t), k=1,2,\dots \right\}$	$RangeS_{0-}^{1-}$
3	$\left\{ \sqrt{2}g_0^1(t)\cos\left(\frac{2k+1}{2}\pi t\right), k=0,1,2,\dots \right\}$	$RangeS_{0+}^{1-}$
4	$\left\{ 1, \sqrt{2}g_0^1(t)\cos(k\pi t), k=1,2,\dots \right\}$	$RangeS_{0+}^{1+}$

///

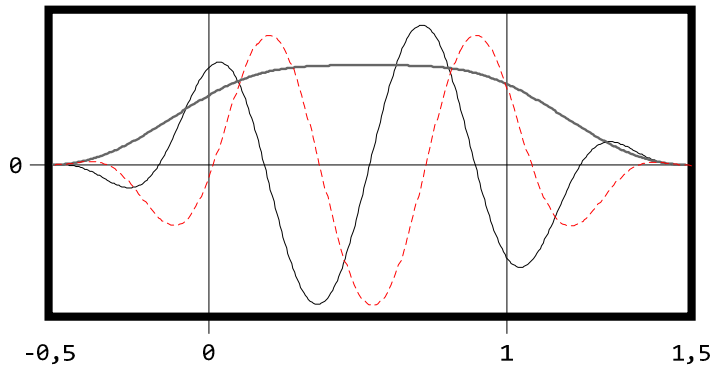
Οι ορθοκανονικές βάσεις για ένα γενικό διάστημα $[\alpha, \beta]$ προκύπτουν με μετάθεση και διαστολή των βάσεων του διαστήματος $[0,1]$.

Αναλυτικότερα, εάν $\beta - \alpha = \ell$ ο παραπάνω πίνακας γίνεται:

1	$\left\{ \sqrt{\frac{2}{\ell}}g_\alpha^\beta(t)\sin\left(\frac{(2k+1)\pi}{2\ell}(t-\alpha)\right), k=0,1,2,\dots \right\}$	$RangeS_{\alpha-}^{\beta+}$	(2.24)
2	$\left\{ \sqrt{\frac{2}{\ell}}g_\alpha^\beta(t)\sin\left(k\frac{\pi}{\ell}(t-\alpha)\right), k=1,2,\dots \right\}$	$RangeS_{\alpha-}^{\beta-}$	
3	$\left\{ \sqrt{\frac{2}{\ell}}g_\alpha^\beta(t)\cos\left(\frac{(2k+1)\pi}{2\ell}(t-\alpha)\right), k=0,1,2,\dots \right\}$	$RangeS_{\alpha+}^{\beta-}$	
4	$\left\{ \sqrt{\frac{1}{\ell}}, \sqrt{\frac{2}{\ell}}g_\alpha^\beta(t)\cos\left(k\frac{\pi}{\ell}(t-\alpha)\right), k=1,2,\dots \right\}$	$RangeS_{\alpha+}^{\beta+}$	

Οι (2.24) είναι γνωστές ως βάσεις Malvar. Στο Σχ. 5 εικονίζεται η συνάρτηση

$$\sqrt{2}g_0^1(t)\cos(3\pi t) \text{ με } g_0^1(t) = g_0(t)g_0(1-t) \text{ και } g_0(t) = \sin\left(\frac{\pi}{4}(1 + \sin \pi t)\right).$$



Σχ. 5: Συναρτήσεις Malvar $\sqrt{2}g_0^1(t)\cos(3\pi t)$ (συνεχής γραμμή), $\sqrt{2}g_0^1(t)\sin(3\pi t)$ (διακεκομμένη γραμμή) και $g_0^1(t) = g_0(t)g_0(1-t)$ (έντονη γραμμή)

2.2.2 Lemarie-Meyer Wavelets

Οι βάσεις Malvar (2.24) αποτελούνται σε κάθε περίπτωση από πλαίσια g_α^β , διαμορφωμένα (modulated) από ημιτονικούς ή συνημιτονικούς παράγοντες. Υπό την έννοια αυτήν, μπορούμε να ισχυριστούμε ότι πρόκειται για “στοιχεία τύπου Gabor”.

Τα wavelets είναι οικογένειες συναρτήσεων της μορφής

$$2^{n/2}\psi(2^n t - m), \quad m, n \in \mathbb{Z}$$

προκύπτουν δηλαδή από διαστολές και όχι διαμορφώσεις μίας βασικής συνάρτησης ψ .

Στην ενότητα αυτήν, κατασκευάζουμε ένα σημαντικό wavelet, το wavelet των Lemarie και Meyer, ως ειδική περίπτωση βάσης Malvar, δείχνοντας έτσι τις περιγραφικές δυνατότητες της θεωρίας των τοπικών ορθοκανονικών προβολών.

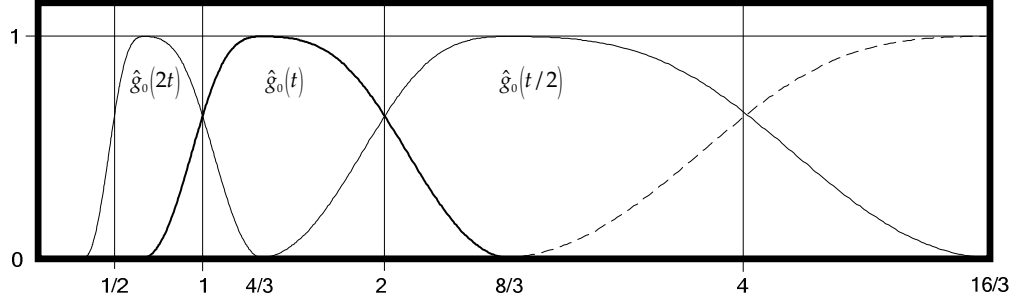
Στις παρατηρήσεις του θεωρήματος 2.6 αναφέρθηκε οι τοπικές προβολές μπορούν να χρησιμοποιηθούν για την ορθοκανονική διάσπαση του $L^2[0, \infty)$ επιλέγοντας κατάλληλες ακολουθίες $(\alpha_n)_{n \in \mathbb{Z}}$, $(\delta_n)_{n \in \mathbb{Z}}$.

Αντί για χρόνο, εργαζόμαστε στον άξονα των συχνοτήτων και θεωρούμε το πλαίσιο \hat{g}_0 με $\text{supp} \hat{g}_0 = [2^0 - 2^0 \frac{1}{3}, 2^0 + 2^0 \frac{1}{3}] = [\frac{2}{3}, \frac{8}{3}]$.

Θέτουμε $\alpha_n = 2^n$ και $\delta_n = \frac{1}{3}(\alpha_{n+1} - \alpha_n) = 2^n/3$ οπότε προκύπτει η διαμέριση του $[0, +\infty)$ στα διαστήματα $\Delta_n = [\alpha_n, \alpha_{n+1}] = [2^n, 2^{n+1}]$, $n \in \mathbb{Z}$ και, τα πλαίσια $\hat{g}_n(t) = \hat{g}_0(2^{-n}t)$ με

$$\text{supp} \hat{g}_n = [\alpha_n - \delta_n, \alpha_{n+1} + \delta_{n+1}] = [2^{n+1}/3, 2^{n+2}/3].$$

Τα \hat{g}_n καλύπτουν όλον τον θετικό ημιάξονα. Στο Σχ. εικονίζονται τρία πλαίσια $\hat{g}_{-1}, \hat{g}_0, \hat{g}_1$.



Σχ. 6

Έστω $v_{n,m}^+$ και $u_{n,m}^+$, $n \in \mathbb{Z}, m = 0, 1, 2, \dots$ οι βάσεις Malvar της (2.24), τύπου 1 και 3 αντιστοίχως. Επειδή $\hat{g}_n(t) = \hat{g}(2^{-n}t)$ έχουμε

$$u_{n,m}(t) = 2^{-n/2} \sqrt{2} \hat{g}_0(2^{-n}t) \sin \left[\pi \left(m + \frac{1}{2} \right) 2^{-n} (t - 2^n) \right] \quad (2.25)$$

και

$$v_{n,m}(t) = 2^{-n/2} \sqrt{2} \hat{g}_0(2^{-n}t) \cos \left[\pi \left(m + \frac{1}{2} \right) 2^{-n} (t - 2^n) \right] \quad (2.26)$$

Χρησιμοποιώντας τις ταυτότητες:

$$\sin \left(x - m\pi - \frac{\pi}{2} \right) = (-1)^{m+1} \cos(x) \text{ και}$$

$$\cos \left(x - m\pi - \frac{\pi}{2} \right) = (-1)^{m+1} \sin(x)$$

οι σχέσεις (2.25) και (2.26) γράφονται

$$u_{n,m}^+(t) = 2^{-n/2} \sqrt{2} \hat{g}_0(2^{-n}t) \cos \left[\pi \left(m + \frac{1}{2} \right) 2^{-n} t \right] \quad (2.27)$$

και

$$v_{n,m}^+(t) = 2^{-n/2} \sqrt{2} \hat{g}_0(2^{-n}t) \sin \left[\pi \left(m + \frac{1}{2} \right) 2^{-n} t \right] \quad (2.28)$$

Οι όροι $(-1)^{m+1}$ έχουν παραληφθεί, επειδή δεν επηρεάζουν την ορθοκανονικότητα των $u_{n,m}^+$ και $v_{n,m}^+$.

Επεκτείνουμε άρτια τις \hat{g}_n σε όλο το \mathbb{R} , και θέτουμε $\hat{g} = \hat{g}_0 + j\hat{g}_0$. Η επέκταση αυτή παράγει άρτιες και περιττές επεκτάσεις των (2.27) και (2.28) αντιστοίχως,

$$u_{n,m} = u_{n,m}^+ + ju_{n,m}^+ \text{ και } v_{n,m} = v_{n,m}^+ - jv_{n,m}^+.$$

Το σύνολο $\left\{ \frac{1}{\sqrt{2}} u_{n,m}, \frac{1}{\sqrt{2}} v_{n,m}, n \in \mathbb{Z}, m = 0, 1, 2, \dots \right\}$ είναι ορθοκανονική βάση του $L^2(\mathbb{R})$.

Πράγματι, είναι εύκολο να δειχθεί ότι

$$\langle u_{n,m}, u_{n',m'} \rangle = \langle v_{n,m}, v_{n',m'} \rangle = \langle u_{n,m}, v_{n',m'} \rangle = 2\delta_{n-n', m-m'}$$

επειδή κάθε μία από τις u_n^+ και v_n^+ έχει αποσυνδετούς φορείς με τις ju_n^+ και jv_n^+ .

Εξάλλου εάν f^e και f^o είναι το άρτιο και περιττό αντιστοίχως μέρος της f ,

$$f = f^e + f^o = \sum_{n,m} \langle f^e, u_{n,m} \rangle u_{n,m} + \sum_{n,m} \langle f^o, v_{n,m} \rangle v_{n,m}.$$

Τότε και το σύνολο

$$\left\{ \frac{1}{2}(u_{n,m} + iv_{n,m}), \frac{1}{2}(u_{n,m} - iv_{n,m}), n \in \mathbb{Z}, m = 0, 1, 2, \dots \right\}$$

είναι εύκολο να δειχθεί ότι είναι επίσης ορθοκανονικό.

Παρατηρούμε ότι

$$\frac{1}{2}(u_{n,m} + iv_{n,m}) = 2^{-(n+1)/2} \hat{g}_0(2^{-n}\gamma) e^{i\pi(m+1/2)2^{-n}\gamma}, n \in \mathbb{Z}, m = 0, 1, 2, \dots$$

και

$$\frac{1}{2}(u_{n,m} - iv_{n,m}) = 2^{-(n+1)/2} \hat{g}_0(2^{-n}\gamma) e^{i\pi(m+1/2)2^{-n}\gamma}, n \in \mathbb{Z}, m = -1, -2, \dots$$

Το συμπέρασμα είναι ότι το σύνολο

$$2^{-(n+1)/2} \hat{g}_0(2^{-n}\gamma) e^{i\pi(m+1/2)2^{-n}\gamma}, n \in \mathbb{Z}, m \in \mathbb{Z}$$

είναι ορθοκανονική βάση του $L^2(\mathbb{R})$.

Θέτουμε $\psi(t) = \frac{1}{2}g(\frac{t}{2} + \frac{1}{4})$ οπότε $\hat{\psi}(\gamma) = \hat{g}_0(2\gamma)e^{i\pi\gamma}$ και ο FT του $2^{(n+1)/2}\psi(2^{n+1}t+m)$ γίνεται

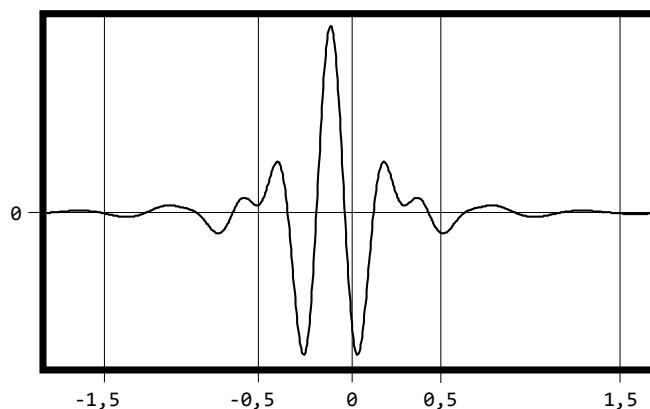
$$2^{-(n+1)/2} \hat{\psi}(2^{-(n+1)}\gamma) e^{i\pi m 2^{-n}\gamma} = 2^{-(n+1)/2} \hat{g}_0(2^{-n}\gamma) e^{i\pi(m+1/2)2^{-n}\gamma}$$

QED

Στο σχήμα που ακολουθεί, εικονίζεται η συνάρτηση $\psi(t) = \frac{1}{2}g_0(\frac{t}{2} + \frac{1}{4})$ για

$$\hat{g}_0(t) = \hat{g}_0^+(t) + \hat{g}_0^+(-t) \text{ και}$$

$$\hat{g}_0^+(t) = \chi_{[2/3, 4/3]}(t) \sin \frac{\pi}{4} \left(1 + \sin \frac{2\pi}{3}(t-1) \right) + \chi_{[4/3, 8/3]}(t) \sin \frac{\pi}{4} \left(1 + \sin \frac{\pi}{3}(2-t) \right).$$



Σχ. 7 Χρονική αναπαράσταση του Le Marie-Meyer wavelet

2.3 ΠΑΡΑΡΤΗΜΑ

Πρόταση 2.9: Έστω $f, \varphi \in C^\infty$ και $f_{n+1} = f_n \circ \varphi$. Εάν $\varphi(t_0) = t_0$ και η πρώτη παράγωγος $D\varphi$ της φ μηδενίζεται στο t_0 , τότε η f_{n+1} έχει $2^{n+1} - 1$ μηδενικές παραγώγους στο t_0 .

Επιπλέον, εάν η φ είναι περιττή και $|f(t)|^2 + |jf(t)|^2 = 1$ ($\forall t \in \mathbb{R}$) τότε και

$$|f_n(t)|^2 + |jf_n(t)|^2 = 1 \quad (\forall t \in \mathbb{R}) \quad (2.29)$$

Αποδεικνύουμε πρώτα το ακόλουθο

Λήμμα 2.10: Έστω $\varphi \in C^\infty$, $\varphi(t_0) = t_0$,

$$\varphi_{n+1} = \varphi_n \circ \varphi \quad (2.30)$$

και $M > 0$. Εάν $D^M \varphi_n(t_0) = 0$, τότε $D^{2M+1} \varphi_{n+1}(t_0) = 0$.

Απόδειξη

Η πρώτη παράγωγος της (2.30) γράφεται $D\varphi_{n+1} = D(\varphi_n \circ \varphi) = (D\varphi_n \circ \varphi)D\varphi$. Γενικότερα, από j διαδοχικές παραγωγίσεις του $D\varphi_n \circ \varphi$ λαμβάνουμε $(D^j \varphi_n \circ \varphi)(D\varphi)^j$ οπότε η $(M+1)$ -στη παράγωγος της (2.30) αποτελείται από όρους της μορφής

$$C_{j,\nu,\lambda} (D^j \varphi_n \circ \varphi)(D\varphi)^j \prod_{\nu,\lambda \leq M+1} (D^\nu \varphi)^\lambda, \quad j \leq M+1 \quad (2.31)$$

όπου $C_{j,\nu,\lambda}$ είναι κατάλληλες σταθερές.

Αφού $\varphi(t_0) = t_0$, $(D^j \varphi_n \circ \varphi)(t_0) = D^j \varphi_n(t_0)$ και οι όροι στην (2.31) μηδενίζονται για $j = 1, 2, \dots, M$ εξ' υποθέσεως επειδή $D^M \varphi_n(t_0) = 0$ ενώ

$$C_{M+1,\nu,\lambda} (D^{M+1} \varphi_n \circ \varphi)(D\varphi)^{M+1} \prod_{\nu,\lambda \leq M+1} (D^\nu \varphi)^\lambda = 0 \quad \text{στο } t_0$$

επειδή $D\varphi(t_0) = 0$.

Μπορούμε επομένως να παραγωγίσουμε άλλες M φορές, διατηρώντας όλους τους όρους μηδενικούς οπότε η (2.30) έχει συνολικά $2M+1$ μηδενικές παραγώγους στο t_0 .

QED

Αναδρομικά, βρίσκουμε εύκολα ότι η φ_n έχει συνολικά $2^n - 1$ μηδενικές παραγώγους (η φ_1 $2^1 - 1$, η φ_2 $2 \cdot 1 + 1 = 2^2 - 1$ κ.ο.κ.).

Απόδειξη (πρότασης 2.9):

$f_{n+1} = f_n \circ \varphi = f_{n-1} \circ \varphi_2 = \dots = f \circ \varphi_{n+1}$ και $Df_{n+1} = (Df \circ \varphi_{n+1})D\varphi_{n+1}$. Από το λήμμα 2.10 η φ_{n+1} έχει $2^{n+1} - 1$ μηδενικές παραγώγους επομένως το ίδιο ισχύει και για την f_{n+1} .

Εάν η φ είναι περιττή τότε $\varphi_n(-t) = -\varphi_n(t) \forall n$ επομένως

$$|f \circ \varphi_n(t)|^2 + |j(f \circ \varphi_n)(t)|^2 = 1 \quad (\forall t \in \mathbb{R})$$

QED

ΚΕΦΑΛΑΙΟ 3

ΒΕΛΤΙΣΤΕΣ ΤΟΠΙΚΕΣ ΒΑΣΕΙΣ ΚΑΙ ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΦΩΝΗΣ

Η θεωρία των τοπικών ορθοκανονικών προβολών του προηγούμενου κεφαλαίου ισχύει χωρίς ιδιαίτερες αλλαγές στον $l^2(\mathbb{Z})$, τον διανυσματικό χώρο των τετραγωνικά αθροισμένων ακολουθιών $\{f(n), n \in \mathbb{Z}\}$ με $\sum_n |f(n)|^2 < \infty$. Επίσης, όπως αποδεικνύεται στην ενότητα 3.1, το ίδιο ισχύει και για τις βάσεις Malvar της παραγράφου 2.2.1 εάν, οι τριγωνομετρικές συναρτήσεις (2.24) αντικατασταθούν από τις αντίστοιχες συναρτήσεις διακριτού χρόνου. Στο πρώτο μέρος του κεφαλαίου, αξιοποιώντας τις συμμετρίες των τοπικών ορθοκανονικών προβολών στα άκρα ενός διαστήματος, υπολογίζουμε την αναπαράσταση του διανύσματος f στο $\{-\delta, \dots, N-1+\delta'\}$ χρησιμοποιώντας τον διακριτό μετασχηματισμό συνημιτόνου του f στο $\{0, \dots, N-1\}$. Ο υπολογισμός γίνεται σε $O(N \log N)$ βήματα.

Οι γραμμικές προσεγγίσεις, προβάλλουν ένα σήμα μήκους N , σε $M \leq N$ διανύσματα μίας ορθοκανονικής βάσης του \mathbb{C}^N που έχει επιλεγεί εκ των προτέρων. Η προσέγγιση βελτιώνεται σημαντικά, εάν η βάση επιλέγεται ανάλογα με το σήμα. Στην §3.3 θεωρούμε οικογένειες (λεξικά) βάσεων του \mathbb{C}^N και αναπτύσσουμε μία μέθοδο για την επιλογή της βάσης που αναπαριστά ένα σήμα καλύτερα από τα υπόλοιπα μέλη της οικογένειάς της. Οι βάσεις Malvar είναι ένα καλό παράδειγμα: διαφορετικές διαμερίσεις του \mathbb{Z} παράγουν διαφορετικές τοπικές τριγωνομετρικές βάσεις. Εάν οι διαμερίσεις επιλεγούν κατάλληλα, οι βάσεις μπορούν να συγκριθούν μεταξύ τους και επομένως μπορούμε να εφαρμόσουμε την παραπάνω μέθοδο για την επιλογή της βάσης Malvar που αναπαριστά καλύτερα ένα δεδομένο σήμα. Οι Coifman και Wickerhauser [COIF92] χρησιμοποίησαν πρώτοι οικογένειες βάσεων Malvar στην συμπίεση σήματος, με ιδιαίτερη επιτυχία.

Εναλλακτικά, μπορούμε να δούμε το πρόβλημα επιλογής σε λεξικά βάσεων, ως πρόβλημα αναγνώρισης. Η βασική ιδέα είναι ότι, αν κάθε βάση ενός λεξικού αντιπροσωπεύει με κάποιον τρόπο μία κλάση, τότε η επιλογή της καλύτερης βάσης ισοδυναμεί με την κατη-

γοριοποίηση ενός σήματος στην κλάση που η βάση αυτή αντιπροσωπεύει. Έτσι, στο τελευταίο μέρος του κεφαλαίου, χρησιμοποιούμε βάσεις Malvar για την αυτόματη τμηματοποίηση φωνής, τον εντοπισμό δηλαδή των ακουστικών ορίων μεταξύ διαδοχικών φωνημάτων στον συνεχή λόγο.

3.1 ΤΟΠΙΚΕΣ ΤΡΙΓΩΝΟΜΕΤΡΙΚΕΣ ΒΑΣΕΙΣ ΔΙΑΚΡΙΤΟΥ ΧΡΟΝΟΥ

Όλη η θεωρία των τοπικών ορθοκανονικών προβολών στον $L^2(\mathbb{R})$ συμπυκνώνεται στις επόμενες προτάσεις

1. Οι τελεστές $S_0 f = \bar{g}(gf \pm j(gf))$, ($f \in L^2(\mathbb{R})$) είναι ορθοκανονικές προβολές.
2. Οι τελεστές $S_\lambda^\mu = S_\lambda S^\mu = S^\mu S_\lambda$ είναι ορθοκανονικές προβολές για όλα τα μ, λ με $\lambda < \mu$.
3. Εάν $\Delta_1 = [\lambda, \mu]$, $\Delta_2 = [\mu, \nu]$, $\lambda < \mu < \nu$ και οι προβολές S_λ^μ και S_μ^ν είναι συμβατές υπό την έννοια του ορισμού 2.5 τότε $S_\lambda^\mu + S_\mu^\nu = S_\lambda^\nu$.

Εάν $\lambda, \mu, \nu \in \mathbb{Z}$, η απόδειξη των παραπάνω προτάσεων παραμένει η ίδια στην περίπτωση $f, g \in \ell^2(\mathbb{Z})$ – τόσο σε επίπεδο ουσίας όσο και συμβολισμού – και το μόνο σημείο που χρειάζεται αλλαγή είναι ο ορισμός του εσωτερικού γινομένου στον $\ell^2(\mathbb{Z})$:

$$\langle f, h \rangle = \sum_{n \in \mathbb{Z}} f(n) \bar{h}(n)$$

Για τον λόγο που περιγράφεται παρακάτω (πρόταση 3.2), οι προηγούμενες προτάσεις πρέπει να ισχύουν και όταν $\lambda, \mu, \nu \in \mathbb{Z} - \frac{1}{2}$.

Πρόταση 3.1: Έστω $g_\alpha \in \ell^2(\mathbb{Z})$ με $|g_\alpha(n)|^2 + |g_\alpha(2\alpha - n)|^2 = 1$ και j_α ο τελεστής ανάκλασης γύρω από το α με $j_\alpha f(n) = f(2\alpha - n)$. Ο τελεστής S_α με $S_\alpha f = \bar{g}_\alpha (g_\alpha f + j_\alpha(g_\alpha f))$ είναι ορθοκανονική προβολή στον $\ell^2(\mathbb{Z})$ $\forall \alpha \in \mathbb{Z} - \frac{1}{2}$.

Απόδειξη

Έστω $g \in \ell^2(\mathbb{Z})$ με $|g(n)|^2 + |g(-1-n)|^2 = 1$. Ο τελεστής $S_{-1/2} f = \bar{g}(gf + j_{-1/2}(gf))$ είναι ορθοκανονική προβολή. Η απόδειξη είναι η ίδια με την σχέση (2.3). Η αλλαγή μεταβλητής στην τρίτη ισότητα είναι $-1-n \rightarrow n$.

Για $\alpha = m - \frac{1}{2}$ ($m \in \mathbb{Z}$) γράφουμε $S_\alpha = S_{m-1/2} = \tau_m S_{-1/2} \tau_{-m}$, και ο S_α είναι ορθοκανονική προβολή επειδή ο τ_m είναι ορθομοναδιαίος στον $\ell^2(\mathbb{Z})$ για κάθε m .

QED

Οι τοπικές τριγωνομετρικές βάσεις στον $\ell^2(\mathbb{Z})$ παράγονται ξεκινώντας από τα στοιχεία διακριτών βάσεων Fourier, ακριβώς όπως συμβαίνει με τις βάσεις Malvar στον L^2 .

Ειδικότερα, έστω $f \in \mathbb{C}^N$ και $n \in \{0, \dots, N-1\}$. Σε αντιστοιχία με όσα κάναμε στην παράγραφο 2.2.1, επεκτείνουμε το f συμμετρικά γύρω από το $N - \frac{1}{2}$ και αντισυμμετρικά γύρω από το $-\frac{1}{2}$ έτσι ώστε $f(n) = f(2N-1-n) = -f(-1-n)$, ($n \in \{-2N, \dots, 2N-1\}$).

Στο $\{-2N, \dots, 2N-1\}$ θεωρούμε το τριγωνομετρικό σύνολο

$$\left\{ \frac{1}{2\sqrt{N}}, \sqrt{\frac{1}{2N}} \sin \frac{\pi}{2N} k \left(n + \frac{1}{2} \right), \sqrt{\frac{1}{2N}} \cos \frac{\pi}{2N} m \left(n + \frac{1}{2} \right), k = 1, \dots, 2N-1, m = 1, \dots, 2N \right\} \quad (3.1)$$

Πρόταση 3.2: Το (3.1) είναι ορθοκανονική βάση στον \mathbb{C}^{4N} , για $n \in \{-2N, \dots, 2N-1\}$. Επιπλέον οι όροι $\frac{1}{\sqrt{2N}} \cos \frac{\pi}{2N} m \left(n + \frac{1}{2} \right)$ είναι συμμετρικοί γύρω από το $-\frac{1}{2}$ για κάθε m και οι όροι $\frac{1}{\sqrt{2N}} \sin \frac{\pi}{2N} k \left(n + \frac{1}{2} \right)$ αντισυμμετρικοί γύρω από το $N - \frac{1}{2}$ για $k = 2, 4, \dots, 2N-2$.

Απόδειξη

Οι συμμετρίες στα $-\frac{1}{2}$ και $N - \frac{1}{2}$ προκύπτουν αμέσως κάνοντας τις αντικαταστάσεις $n \rightarrow -1-n$ στους συνημιτονικούς όρους και $n \rightarrow 2N-1-n$ στους ημιτονικούς και χρησιμοποιώντας βασικές τριγωνομετρικές ταυτότητες.

Για την απόδειξη της ορθοκανονικότητας του (3.1) γράφουμε

$$\cos \frac{\pi}{2N} m \left(n + \frac{1}{2} \right) = \frac{1}{2} e^{i \frac{\pi}{4N} m} \left(e^{i \frac{\pi}{2N} m n} + e^{-i \frac{\pi}{2N} m n} \right) \quad (3.2)$$

και

$$\sin \frac{\pi}{2N} m' \left(n + \frac{1}{2} \right) = \frac{1}{2i} e^{i \frac{\pi}{4N} m'} \left(e^{i \frac{\pi}{2N} m' n} - e^{-i \frac{\pi}{2N} m' n} \right) \quad (3.3)$$

Η ορθοκανονικότητα των (3.2) και (3.3) για $m = 1, \dots, 2N$ και $m' = 1, \dots, 2N-1$ αντιστοίχως, προκύπτει από την ορθοκανονικότητα του συνόλου $\left\{ \frac{1}{2\sqrt{N}} e^{i \frac{\pi}{2N} m n}, m = 0, \dots, 4N-1 \right\}$ στο $\{-2N, \dots, 2N-1\}$.

QED

Λόγω των συμμετριών των στοιχείων της (3.1) στα $-\frac{1}{2}$ και $N - \frac{1}{2}$, είναι εύκολο ν' αποδειχθεί (με τα ίδια ακριβώς βήματα που ακολουθήθηκαν στην 2.2.1) ότι τα συστήματα του παρακάτω πίνακα, είναι ορθοκανονικές βάσεις του \mathbb{C}^N για $n = 0, 1, \dots, N-1$. Επιπλέον,

κάθε σύστημα έχει τις συμμετρίες που φαίνονται στην δεύτερη στήλη του πίνακα (συμβολίζουμε με + τις άρτιες και με − τις περιττές συμμετρίες):

1	$\left\{ \sqrt{\frac{2}{N}} \sin \frac{\pi}{N} \left(m + \frac{1}{2}\right) \left(n + \frac{1}{2}\right), n, m = 0, 1, 2, \dots, N-1 \right\}$	$(-1/2)^-, (N-1/2)^+$
2	$\left\{ \sqrt{\frac{2}{N}} \sin \frac{\pi}{N} m \left(n + \frac{1}{2}\right), n, m = 0, 1, 2, \dots, N-1 \right\}$	$(-1/2)^-, (N-1/2)^-$
3	$\left\{ \sqrt{\frac{2}{N}} \cos \frac{\pi}{N} \left(m + \frac{1}{2}\right) \left(n + \frac{1}{2}\right), n, m = 0, 1, 2, \dots, N-1 \right\}$	$(-1/2)^+, (N-1/2)^-$
4	$\left\{ \sqrt{\frac{1}{N}}, \sqrt{\frac{2}{N}} \cos \frac{\pi}{N} m \left(n + \frac{1}{2}\right), m = 1, 2, \dots \right\}$	$(-1/2)^+, (N-1/2)^+$

(3.4)

Ας συμβολίσουμε με $v_{-1/2, m}$, ($m = 0, \dots, N-1$), τα στοιχεία μίας (οποιασδήποτε) από τις παραπάνω βάσεις 1–4 και έστω $\alpha \in \mathbb{Z} - 1/2$. Επειδή ο τελεστής μετατόπισης $\tau_{\alpha+1/2}$ είναι ορθομοναδιαίος στον $\ell^2(\mathbb{Z})$, είναι σαφές ότι το σύνολο

$$v_{\alpha, m} = \tau_{\alpha+1/2} v_{-1/2, m}, \quad (m = 0, \dots, N-1)$$

είναι επίσης ορθοκανονική βάση του \mathbb{C}^N για $n = \alpha + 1/2, \dots, \alpha + N - 1/2$.

Βάσει της πρότασης 2.7, το σύνολο $\bar{g}_{\alpha}^{\alpha+N} v_{\alpha, m}$, ($m = 0, \dots, N-1$) είναι ορθοκανονική βάση του $S_{\alpha}^{\alpha+N}(\ell^2)$. Επιπλέον, από τις προτάσεις 3.1 και 2.2 και, το θεώρημα 2.6 έπεται ότι οι διανυσματικοί χώροι $S_{\alpha}^{\alpha+N}(\ell^2)$ συνιστούν μία ορθοκανονική διάσπαση του $\ell^2(\mathbb{Z})$ όταν τα α είναι στοιχεία μίας ακολουθίας α με $\alpha \xrightarrow{\pm\infty} \pm\infty$.

3.2 ΔΙΑΚΡΙΤΟΣ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΣ MALVAR

Έστω $\alpha, \beta \in \mathbb{Z} - 1/2$, $\delta, \delta' \in \mathbb{N}$, $\alpha + \delta < \beta - \delta'$, g_{α}^{β} ένα πλαίσιο συμμετρικό στο α και αντισυμμετρικό στο β και $f \in \ell^2(\mathbb{Z})$. Λόγω των συμμετριών του g_{α}^{β} στα α και β , θεωρούμε την ορθοκανονική βάση $\{\bar{g}_{\alpha}^{\beta} v_{\alpha, m}^{\beta}, (m = 0, \dots, \beta - \alpha - 1)\}$ του $S_{\alpha}^{\beta}(\ell^2)$ με,

$$v_{\alpha, m}^{\beta}(n) = \sqrt{\frac{2}{\beta - \alpha}} \cos \frac{\pi}{N} \left(m + \frac{1}{2}\right) \left(\frac{n - \alpha}{\beta - \alpha}\right), \quad (m = 0, 1, \dots, N-1, n \in \mathbb{Z})$$

Τα διανύσματα $v_{\alpha, m}^{\beta}$ είναι τα στοιχεία της βάσης 3 της (3.4), μετά την αντικατάσταση $n \rightarrow n - \alpha - 1/2$.

Ο διακριτός μετασχηματισμός Malvar της f στο $\{\alpha + 1/2, \dots, \beta - 1/2\}$ ορίζεται ως

$$M_\alpha^\beta f(m) = \langle f, \bar{g}_\alpha^\beta v_{\alpha,m}^\beta \rangle = \sum_{n=-\infty}^{+\infty} f(n) \bar{g}_\alpha^\beta(n) v_{\alpha,m}^\beta(n), \quad m = 0, \dots, \beta - \alpha - 1 \quad (3.5)$$

Επειδή ο S_α^β είναι ορθοκανονική προβολή,

$$\langle f, \bar{g}_\alpha^\beta v_{\alpha,m}^\beta \rangle = \langle S_\alpha^\beta f, \bar{g}_\alpha^\beta v_{\alpha,m}^\beta \rangle, \quad m = 0, \dots, \beta - \alpha - 1$$

Χρησιμοποιώντας το δεύτερο συμπέρασμα της πρότασης 2.7 έχουμε

$$\langle S_\alpha^\beta f, \bar{g}_\alpha^\beta v_{\alpha,m}^\beta \rangle = \langle f_\alpha^\beta, v_{\alpha,m}^\beta \rangle$$

όπου το εσωτερικό γινόμενο στο πρώτο μέλος της παραπάνω ισότητας υπολογίζεται στο $\{\alpha + 1/2 - \delta, \dots, \beta - 1/2 + \delta'\}$, ενώ το δεύτερο μέλος στο $\{\alpha + 1/2, \dots, \beta - 1/2\}$.

Επομένως,

$$M_\alpha^\beta f(m) = \langle f_\alpha^\beta, v_{\alpha,m}^\beta \rangle = \sum_{n=-\alpha+1/2}^{\beta-1/2} f_\alpha^\beta(n) v_{\alpha,m}^\beta(n), \quad m = 0, \dots, \beta - \alpha - 1 \quad (3.6)$$

Κατά τα γνωστά, θέτουμε

$$\Delta_1 = \{\alpha + 1/2, \dots, \alpha + 1/2 + \delta - 1\},$$

$$\Delta_2 = \{\alpha + 1/2 + \delta, \dots, \beta + 1/2 - \delta'\},$$

$$\Delta_3 = \{\beta + 1/2 - \delta', \dots, \beta - 1/2\}$$

Η f_α^β δίνεται από την σχέση

$$\begin{aligned} f_\alpha^\beta(n) = & \chi_{\Delta_1}(n) [\bar{g}_\alpha^\beta(n) f_\alpha^\beta(n) + \bar{g}_\alpha^\beta(2\alpha - n) f_\alpha^\beta(2\alpha - n)] + \\ & \chi_{\Delta_2}(n) f(n) + \\ & \chi_{\Delta_3}(n) [\bar{g}_\alpha^\beta(n) f_\alpha^\beta(n) - \bar{g}_\alpha^\beta(2\beta - n) f_\alpha^\beta(2\beta - n)] \end{aligned} \quad (3.7)$$

Τα εσωτερικά γινόμενα στην (3.6) είναι ο **διακριτός μετασχηματισμός συνημιτόνου** τύπου IV (DCT-IV) του f_α^β και, όταν το f_α^β είναι πραγματικό σήμα, μπορεί να υπολογιστεί μέσω ενός FFT $(\beta - \alpha)/2$ σημείων, σε $O((\beta - \alpha) \log(\beta - \alpha))$ (βλ. παράρτημα). Συμπεριλαμβανομένων των $O(\beta - \alpha)$ πράξεων που απαιτούνται για τους υπολογισμούς (3.7), τελικά η πολυπλοκότητα του μετασχηματισμού Malvar είναι $O((\beta - \alpha) \log(\beta - \alpha))$.

Ο αντίστροφος μετασχηματισμός Malvar ανακατασκευάζει το διάνυσμα $f_\lambda^\mu \in \mathbb{C}^{\mu - \lambda}$ από τις τιμές $M_\lambda^\mu f(m)$, $m = 0, \dots, \mu - \lambda - 1$. Για ν' αναπαράγουμε το f στο $\{\lambda + 1/2, \dots, \mu - 1/2\}$ συν-

δυσάζουμε τα διανύσματα f_κ^λ , f_λ^μ και f_μ^ν του πλαισίου $\{\lambda + \frac{1}{2}, \dots, \mu - \frac{1}{2}\}$ και των γειτονικών του $\{\kappa + \frac{1}{2}, \dots, \lambda - \frac{1}{2}\}$ και $\{\mu + \frac{1}{2}, \dots, \nu - \frac{1}{2}\}$.

Στο $\Delta_2 = \{\lambda + \frac{1}{2} + \delta, \dots, \mu + \frac{1}{2} - \delta'\}$:

$$f(n) = f_\lambda^\mu(n)$$

και στα $\Delta_1 = \{\alpha + \frac{1}{2}, \dots, \alpha + \frac{1}{2} + \delta - 1\}$ και $\Delta_3 = \{\beta + \frac{1}{2} - \delta', \dots, \beta - \frac{1}{2}\}$

$$f(n) = \bar{g}_\kappa^\lambda(n) f_\kappa^\lambda(n) + \bar{g}_\lambda^\mu(n) f_\lambda^\mu(n) \text{ και}$$

$$f(n) = \bar{g}_\lambda^\mu(n) f_\lambda^\mu(n) + \bar{g}_\mu^\nu(n) f_\mu^\nu(n)$$

αντιστοίχως, βάσει του θεωρήματος 2.6.

Εάν οι g_κ^λ , g_λ^μ και g_μ^ν έχουν προκύψει από μεταθέσεις της ίδιας συνάρτησης g , ισχύει

$g_\kappa^\lambda(n) = \bar{g}_\kappa^\mu(2\lambda - n)$ και $g_\mu^\nu(n) = \bar{g}_\lambda^\mu(2\mu - n)$ (λήμμα 2.4). Επομένως οι προηγούμενες σχέσεις γράφονται:

$$f(n) = \bar{g}_\lambda^\mu(2\lambda - n) f_\kappa^\lambda(n) + g_\lambda^\mu(n) f_\lambda^\mu(n), \quad n \in \{\lambda + \frac{1}{2}, \dots, \lambda + \frac{1}{2} + \delta - 1\}$$

$$f(n) = f_\lambda^\mu(n), \quad n \in \{\lambda + \frac{1}{2} + \delta, \dots, \mu + \frac{1}{2} - \delta'\} \text{ και}$$

$$f(n) = \bar{g}_\lambda^\mu(2\mu - n) f_\mu^\nu(n) + g_\lambda^\mu(n) f_\lambda^\mu(n), \quad n \in \{\mu + \frac{1}{2} - \delta' + 1, \dots, \mu - \frac{1}{2}\}$$

3.3 ΒΕΛΤΙΣΤΕΣ ΤΟΠΙΚΕΣ ΒΑΣΕΙΣ

Η σπουδαιότητα των ορθοκανονικών βάσεων στην προσέγγιση ενός σήματος, οφείλεται σε δύο κυρίως ιδιότητές τους:

- i* Λόγω της γραμμικής ανεξαρτησίας μεταξύ των στοιχείων τους, η προσέγγιση γίνεται με τον ελάχιστο δυνατό αριθμό όρων.
- ii* Η ανακατασκευή του σήματος γίνεται απ' ευθείας από τις ίδιες τις προβολές του στα στοιχεία της βάσης.

Οι βάσεις των χώρων $L^2(K)$ όπου K ένα διάστημα του \mathbb{R} , είναι απειροδιάστατες και η ιδιότητα *i* δεν έχει ιδιαίτερη σημασία. Το ενδιαφέρον εστιάζεται σε βάσεις $\{g_\gamma, \gamma \in \mathbb{I}\}$ με καλή συμπεριφορά στο άπειρο, σε βάσεις δηλαδή όπου οι προβολές μίας συνάρτησης $f \in L^2(K)$, $\langle f, g_\gamma \rangle$, έχουν μεγάλη απομείωση (decay rate).

Ειδικότερα όταν το K είναι φραγμένο, η απομείωση των όρων του μετασχηματισμού Fourier βραχέως χρόνου (Short-time Fourier Transform – STFT) της f στο άπειρο, εξαρτάται από την παραγωγισιμότητα της περιοδικής επέκτασης f_K της συνάρτησης $\chi_K f$ με,

$f_K(t) = \sum_n \chi_K(t+n)f(t+n)$: Εάν η $f_K \in L^2(K)$ έχει $m \geq 0$ συνεχείς παραγώγους, τότε [KATZ3] ¹⁰

$$\sum_n n^{2m} |\hat{f}_K(n)|^2 < \infty \quad (3.8)$$

Το πρόβλημα με την αναπαράσταση STFT είναι ότι ο πολλαπλασιασμός με την χ_K , ενδέχεται να εισάγει ασυνέχειες στην f_K ακόμη και αν η f είναι πολλές φορές συνεχώς παραγωγίσιμη. Αναλυτικότερα, έστω $K = [0,1]$. Αν $f(0) \neq f(1)$ η f_K δεν είναι συνεχής και η (3.8) γίνεται ¹¹

$$\sum_n |\hat{f}_K(n)| = \infty$$

Στον \mathbb{C}^K ($K \in \mathbb{N}$) το πρόβλημα προσέγγισης ενός διανύσματος x , ανάγεται στην επιλογή ορθοκανονικών βάσεων που προσεγγίζουν το x καλύτερα από άλλες βάσεις του \mathbb{C}^K . Μία τέτοια βάση, εξαρτάται από το x ή τις ιδιότητες μίας κλάσης διανυσμάτων της οποίας το x είναι μέλος. Οι ορθοκανονικές βάσεις που προκύπτουν από την διάσπαση ενός πίνακα σε ιδιάζουσες τιμές (Singular Value Decomposition – SVD), είναι χαρακτηριστική περίπτωση:

Αν $x_0, x_1, \dots, x_{N-1} \in \mathbb{C}^K$ και $K' \leq K \leq N$, είναι γνωστό ότι το ορθοκανονικό σύνολο K' διανυσμάτων που προσεγγίζει καλύτερα στα x_0, x_1, \dots, x_{N-1} είναι τα ιδιοδιανύσματα των K' μεγαλύτερων ιδιοτιμών του πίνακα αυτοσυσχέτισης X^*X των δεδομένων x_0, x_1, \dots, x_{N-1} , όπου $X = [x_1 \dots x_{N-1}]$. Η προσαρμογή είναι βέλτιστη υπό την έννοια ότι το σφάλμα προσέγγισης κάθε x_j , $j = 0, \dots, N-1$, είναι μικρότερο σε σχέση με οποιαδήποτε άλλη ορθοκανονική προσέγγιση του $\{x_0, x_1, \dots, x_{N-1}\}$ από K' διανύσματα [GOL3]. Στην περίπτωση που τα x_0, x_1, \dots, x_{N-1} είναι τιμές του ίδιου τυχαίου διανύσματος, η παραπάνω προσέγγιση είναι γνωστή ως “βάση” ή “μετασχηματισμός Karhunen-Loeve” ή “Ανάλυση Πρωτευουσών Συνιστωσών” (Principle Component Analysis) [MALLAT2].

¹⁰ Υπάρχει ένα γενικό θεώρημα που συνδέει το σφάλμα προσέγγισης της f από τους N πρώτους όρους κάθε ορθοκανονικής βάσης (όχι απαραίτητα Fourier), με την απομοίωση των προβολών της f στα στοιχεία της βάσης [MALLAT2, §9.1].

¹¹ Για την ακρίβεια, είναι εύκολο ν' αποδειχθεί ότι αν $\sum_n |\hat{f}_K(n)| < \infty$ τότε η σύγκλιση της $\sum_n \hat{f}_K(n)e^{i2\pi nt}$ είναι ομοιόμορφη. Επειδή τα μερικά αθροίσματα είναι συνεχή, το ίδιο ισχύει και για το ομοιόμορφο όριό τους $f_K(t) = \sum_n \hat{f}_K(n)e^{i2\pi nt}$ (π.χ. [RUDPMA], κεφ. 7). Αντιστρέφοντας τον συλλογισμό, εάν η f δεν είναι συνεχής, η σύγκλιση δεν είναι ομοιόμορφη και επομένως η $\sum_n |\hat{f}_K(n)|$ αποκλίνει.

Εκτός του μεγάλου υπολογιστικού κόστους της SVD, τα ιδιοδιανύσματα της δεν σχετίζονται γενικά με τις “φυσικές” ιδιότητες των x_0, x_1, \dots, x_{N-1} και, ως εκ τούτου, δεν μπορούν να χρησιμοποιηθούν για τον χαρακτηρισμό των τελευταίων. Για παράδειγμα, εάν τα x_0, x_1, \dots, x_{N-1} είναι τμήματα ενός ακουστικού σήματος, γνωρίζουμε ότι τα διανύσματα $e^{i2\pi mn/K}$ του DFT αντιπροσωπεύουν το φασματικό περιεχόμενο κάθε x_j ενώ δεν ισχύει το ίδιο για τα ιδιοδιανύσματα του X^*X .

3.3.1 Λεξικά Βάσεων και Κοίλα Συναρτησιακά Κόστους

Θεωρούμε ένα λεξικό (σύνολο) \mathcal{D} , ορθοκανονικών βάσεων σε έναν χώρο σήματος πεπερασμένης διάστασης N :

$$\mathcal{D} = \bigcup_{\lambda \in \mathbb{I}} \mathcal{B}^\lambda$$

Το \mathbb{I} είναι ένα αυθαίρετο σύνολο δεικτών και οι \mathcal{B}^λ ορθοκανονικές βάσεις N διανυσμάτων

$$\mathcal{B}^\lambda = \{g_m^\lambda, m = 0, \dots, N-1\}$$

Π.χ. το \mathcal{D} μπορεί να περιέχει τις τοπικές τριγωνομετρικές βάσεις για όλες τις δυνατές διαμερίσεις του $\{0, \dots, N-1\}$ ¹².

Στόχος μας είναι η επιλογή της βάσης του \mathcal{D} που προσεγγίζει ένα δεδομένο διάνυσμα $f \in \mathbb{C}^N$ με το μικρότερο δυνατό σφάλμα, υπό την εξής έννοια: αν \mathbb{I}_M^λ το υποσύνολο M στοιχείων της \mathcal{B}^λ με τις μεγαλύτερες προβολές $\left| \langle f, g_m^\lambda \rangle \right|$ για το f , το σφάλμα προσέγγισης (M τάξης) της f από την \mathcal{B}^λ είναι

$$\epsilon^\lambda(M) = \sum_{j \notin \mathbb{I}_M^\lambda} \left| \langle f, g_j^\lambda \rangle \right|^2 \quad (3.9)$$

Ορισμός 3.3: Η βάση \mathcal{B}^α προσεγγίζει το f καλύτερα από την \mathcal{B}^λ , αν

$$\epsilon^\alpha(M) \leq \epsilon^\lambda(M) \quad \forall M \leq N \quad (3.10)$$

Το κριτήριο (3.10) είναι μία σχέση μερικής διάταξης στο σύνολο \mathcal{D} . Οι \mathcal{B}^α και \mathcal{B}^λ δεν είναι συγκρίσιμες όταν υπάρχουν $M_1, M_1 \leq N$ τέτοια ώστε

$$\epsilon^\alpha(M_1) < \epsilon^\lambda(M_1) \quad \text{και} \quad \epsilon^\lambda(M_2) < \epsilon^\alpha(M_2) \quad (3.11)$$

Από την (3.9) έπεται ότι, το κριτήριο (3.10) μπορεί να γραφεί ισοδύναμα

¹² Ακριβέστερα, είτε ο χώρος θα είναι ο $S_{-1/2}^{N-1/2}(\ell^2(\mathbb{Z}))$ είτε τα άκρα (αριστερό και δεξί αντίστοιχα) των πλαισίων g στα 0 και $N-1$ θα πρέπει να τροποποιηθούν ώστε $\delta_0 = \delta_{N-1} = 0$.

$$\sum_{j \in \mathbb{P}_M^\alpha} \left| \langle f, g_j^\alpha \rangle \right|^2 \geq \sum_{j \in \mathbb{P}_M^\lambda} \left| \langle f, g_j^\lambda \rangle \right|^2 \quad \forall M \leq N \quad (3.12)$$

Το ακόλουθο θεώρημα, γνωστό ως “Schur Concavity Test” [MALLAT2] παρέχει ένα σημαντικό κριτήριο για τις προσεγγίσεις ενός διανύσματος από δύο βάσεις:

Θεώρημα 3.4: Η βάση \mathcal{B}^α προσεγγίζει το f καλύτερα από την \mathcal{B}^γ αν και μόνον αν για όλες τις κοίλες συναρτήσεις φ ισχύει

$$\sum_{m=1}^N \varphi \left(\frac{\langle f, g_m^\alpha \rangle^2}{\|f\|^2} \right) \leq \sum_{m=1}^N \varphi \left(\frac{\langle f, g_m^\gamma \rangle^2}{\|f\|^2} \right) \quad (3.13)$$

Η απόδειξη του θεωρήματος βασίζεται στο ακόλουθο λήμμα των Hardy, Littlewood & Pólya.

Λήμμα 3.5: Έστω $x_m \geq 0$ και $y_m \geq 0$ δύο ακολουθίες μήκους N με

1. $x_m \geq x_{m+1}$ και $y_m \geq y_{m+1}$ για $1 \leq m \leq N$
2. $\sum_{m=1}^N x_m = \sum_{m=1}^N y_m$

Για κάθε $M \leq N$ οι ακολουθίες αυτές ικανοποιούν την σχέση

$$\sum_{m=1}^M x_m \geq \sum_{m=1}^M y_m \quad (3.14)$$

αν και μόνον εάν για όλες τις κοίλες συναρτήσεις φ ισχύει

$$\sum_{m=1}^N \varphi(x_m) \leq \sum_{m=1}^N \varphi(y_m) \quad (3.15)$$

Απόδειξη

Δείχνουμε πρώτα ότι η (3.14) συνεπάγεται την (3.15):

Έστω φ μία κοίλη συνάρτηση και \mathbb{H} το σύνολο των ακολουθιών της μορφής 1, δηλαδή των ακολουθιών z N στοιχείων για τις οποίες ισχύει

$$z_1 \geq z_2 \geq \dots \geq z_N \geq 0 \quad (3.16)$$

Για κάθε $z \in \mathbb{H}$, θεωρούμε τα μερικά αθροίσματα των M πρώτων όρων του z

$$Z_M = \sum_{m=1}^M z_m, \quad M = 1, \dots, N$$

και την συνάρτηση

$$\begin{aligned} \Theta(Z_1, Z_2, \dots, Z_N) &= \sum_{m=1}^N \varphi(z_m) \\ &= \varphi(z_1) + \sum_{m=1}^N \varphi(Z_m - Z_{m-1}) \end{aligned}$$

Οι ακολουθίες x και y ανήκουν στο \mathbb{H} εξ' ορισμού και από την 2 γνωρίζουμε ότι έχουν το ίδιο άθροισμα, δηλαδή $X_N = Y_N$.

Η συνθήκη (3.14) γράφεται ισοδύναμα $X_M \geq Y_M, \forall 1 \leq M < N$. Για να δείξουμε επομένως την (3.15) μπορούμε ισοδύναμα να δείξουμε ότι η Θ είναι μία φθίνουσα συνάρτηση ως προς κάθε όρισμά της ξεχωριστά, δηλαδή ότι για κάθε $1 \leq k \leq N$

$$\Theta(Z_1, Z_2, \dots, Z_N) \leq \Theta(Z_1, \dots, Z_{k-1}, Z_k + \eta, Z_{k+1}, \dots, Z_N)$$

που, βάσει της σχέσης ορισμού του Θ γράφεται

$$\sum_{m=1}^N \varphi(z_m) \geq \sum_{m=1}^{k-1} \varphi(z_m) + \varphi(z_k + \eta) + \varphi(z_{k+1} - \eta) + \sum_{m=k+2}^N \varphi(z_m) \quad (3.17)$$

Ασφαλώς, προκειμένου να παραμείνουμε στο \mathbb{H} θα πρέπει επίσης το η να ικανοποιεί την συνθήκη

$$z_{k-1} \geq z_k + \eta \geq z_{k+1} - \eta \geq z_{k+2}$$

Η ανισότητα (3.17) αποδεικνύεται αν

$$\varphi(z_k) + \varphi(z_{k+1}) \geq \varphi(z_k + \eta) + \varphi(z_{k+1} - \eta) \quad (3.18)$$

δεδομένου ότι οι υπόλοιποι όροι της (3.17) είναι ίδιοι.

Η (3.18) είναι συνέπεια της κοιλότητας της φ . Εξ' ορισμού, η φ είναι κοίλη, αν $\forall x, y$ και

$$0 \leq \alpha \leq 1$$

$$\varphi(\alpha x + (1-\alpha)y) \geq \alpha \varphi(x) + (1-\alpha)\varphi(y) \quad (3.19)$$

Θέτουμε

$$\alpha = \frac{z_k - z_{k+1} + \eta}{z_k - z_{k+1} + 2\eta}$$

οπότε $0 \leq \alpha \leq 1$ και επιπλέον

$$z_k = \alpha(z_k + \eta) + (1-\alpha)(z_{k+1} - \eta) \quad \text{και}$$

$$z_{k+1} = (1-\alpha)(z_k + \eta) + \alpha(z_{k+1} - \eta)$$

Υπολογίζοντας το άθροισμα $\varphi(z_k) + \varphi(z_{k+1})$ και χρησιμοποιώντας την συνθήκη κοιλότητας (3.19) προκύπτει η (3.18).

Για το αντίστροφο, αποδεικνύουμε ότι η (3.15) συνεπάγεται την (3.14) για μία ειδική οικογένεια κοίλων συναρτήσεων που ορίζονται από τις σχέσεις

$$\varphi_M(u) = \begin{cases} x_M - u & (u \geq x_M) \\ 0 & (u < x_M) \end{cases}, M = 1, \dots, N$$

Εύκολα αποδεικνύεται ότι

$$\sum_{m=1}^N \varphi_M(x_m) = Mx_m - \sum_{m=1}^M x_m \quad (3.20)$$

Εξ' υποθέσεως,

$$\sum_{m=1}^N \varphi_M(x_m) \leq \sum_{m=1}^N \varphi_M(y_m)$$

Επιπλέον $\varphi_M(u) \leq 0$ και $\varphi_M(u) \leq x_M - u$ οπότε,

$$Mx_m - \sum_{m=1}^M x_m \leq \sum_{m=1}^N \varphi_M(y_m) \leq \sum_{m=1}^M \varphi_M(y_m) \leq Mx_m - \sum_{m=1}^M y_m$$

που αποδεικνύει την (3.14) και τελικά, το λήμμα.

Το θεώρημα 3.4 είναι τώρα άμεση συνέπεια του λήμματος:

Απόδειξη (θεωρήματος 3.4)

Για κάθε βάση \mathcal{B}^λ διατάσσουμε σε φθίνουσα σειρά τα εσωτερικά γινόμενα $\left| \langle f, g_m^\lambda \rangle \right|$ και

θέτουμε

$$x_m^\lambda = \frac{\left| \langle f, g_m^\lambda \rangle \right|^2}{\|f\|^2} \geq x_{m+1}^\lambda = \frac{\left| \langle f, g_{m+1}^\lambda \rangle \right|^2}{\|f\|^2}$$

Η διατήρηση της ενέργειας σε ορθοκανονικές βάσεις συνεπάγεται $\sum_{m=1}^N x_m^\lambda = 1 \forall \lambda$. Από την συνθήκη (3.12) προκύπτει επομένως ότι η \mathcal{B}^α είναι καλύτερη από την \mathcal{B}^λ αν και μόνον εάν

$$\sum_{m=1}^M x_m^\alpha \geq \sum_{m=1}^M x_m^\lambda$$

Το λήμμα 3.5 αποδεικνύει ότι αυτό, ισοδυναμεί με την σχέση

$$\sum_{m=1}^N \varphi(x_m^\alpha) \leq \sum_{m=1}^N \varphi(x_m^\lambda)$$

για όλες τις κοίλες συναρτήσεις φ .

QED

Στην πράξη, δύο βάσεις συγκρίνονται χρησιμοποιώντας μία μόνον κοίλη συνάρτηση φ .

Το κόστος προσέγγισης του f σε μία βάση \mathcal{B}^λ υπολογίζεται από το άθροισμα Schur

$$C(f, \lambda) = \sum_{m=1}^N \varphi \left(\frac{\left| \langle f, g_m^\lambda \rangle \right|^2}{\|f\|^2} \right)$$

Το θεώρημα 3.4 αποδεικνύει ότι, αν η \mathcal{B}^α είναι (για την f) καλύτερη από την \mathcal{B}^λ , τότε

$$C(f, \alpha) \leq C(f, \lambda) \tag{3.21}$$

Η συνθήκη αυτή, είναι ασφαλώς αναγκαία αλλά όχι ικανή, ενδέχεται δηλαδή οι και \mathcal{B}^λ να μην είναι συγκρίσιμες, υπό την έννοια της (3.11), αφού ελέγχουμε μία μόνον κοίλη συνάρτηση.

Οι Coifman και Wickerhauser [COIF92] βρίσκουν την καλύτερη βάση σε ένα λεξικό \mathcal{D} , ελαχιστοποιώντας το κόστος της f στο \mathcal{D} :

$$C(f, \alpha) = \min_{\lambda \in \mathbb{I}} C(f, \lambda)$$

Σύμφωνα με τα παραπάνω, δεν υπάρχει καλύτερη βάση στο \mathcal{D} (υπό την έννοια της (3.11)) που προσεγγίζει την f . Επιπλέον, ενδέχεται η βέλτιστη ως προς C βάση στο \mathcal{D} , να μην είναι βέλτιστη για άλλα συναρτησιακά κόστους, δεν μπορεί όμως να είναι χειρότερη.

Ιδανικές και Διαχυτικές Βάσεις

Η ιδανική βάση για ένα δεδομένο διάνυσμα f είναι εκείνη που το αναπαριστά με έναν μόνον όρο, επομένως το f είναι γραμμικώς εξαρτημένο από κάποιο στοιχείο της, π.χ.

$$g_m = \eta f \text{ για κάποιο } \eta \in \mathbb{C}. \text{ Εάν } \varphi(0) = 0, \text{ το κόστος του } f \text{ στην εν λόγω βάση είναι } \varphi(1).$$

Απεναντίας, η χειρότερη βάση αναπαράστασης του f , είναι εκείνη στην οποία οι προβολές του διαχέονται ομοιόμορφα στα στοιχεία της, δηλαδή

$$|\langle f, g_m \rangle|^2 = \frac{\|f\|^2}{N}, \quad m = 1, 2, \dots, N \quad (3.22)$$

Το κόστος της f σε μία διαχυτική βάση είναι

$$C(f) = N\varphi^{-1}\left(\frac{1}{N}\right) \quad (3.23)$$

Πρόταση 3.6: Κάθε βάση \mathcal{B} είναι χειρότερη από την ιδανική και καλύτερη από την διαχυτική βάση

Απόδειξη

Η ιδανική βάση είναι προφανώς καλύτερη από κάθε άλλη βάση – υπό την έννοια του ορισμού 3.3 – αφού το σφάλμα αναπαράστασης της f είναι 0 $\forall M \geq 1$. Το σφάλμα αναπαράστασης του f από M διανύσματα μίας διαχυτικής βάσης είναι

$$\|f\|^2 \frac{N-M}{N}$$

Για να δείξουμε ότι κάθε βάση είναι καλύτερη από μία διαχυτική, παρατηρούμε ότι για κάθε $j \notin \mathbb{I}_M$

$$|\langle f, g_j^\lambda \rangle|^2 \leq \sum_{m \notin \mathbb{I}_M} |\langle f, g_m^\lambda \rangle|^2 \leq \frac{\|f\|^2}{M} \quad (3.24)$$

Επομένως, το σφάλμα προσέγγισης από M διανύσματα γίνεται

$$\epsilon^\lambda(M) = \sum_{m \notin \mathbb{I}_M} |\langle f, g_m^\lambda \rangle|^2 \leq \|f\|^2 \frac{N-M}{M}$$

που δείχνει ότι είναι μικρότερο του σφάλματος M όρων της διαχυτικής, για κάθε βάση \mathcal{B}^λ .

QED

Παραδείγματα Συναρτησιακών Κόστους

Εντροπία: Η συνάρτηση $\varphi(x) = -x \ln x$ είναι κοίλη για κάθε $x \geq 0$. Το συναρτησιακό κόστος που προκύπτει από την φ δίνεται από την σχέση

$$C(f) = -\sum_{m=1}^N \frac{|\langle f, g_m \rangle|^2}{\|f\|^2} \ln \left(\frac{|\langle f, g_m \rangle|^2}{\|f\|^2} \right) \quad (3.25)$$

Η πρόταση 3.6 δείχνει ότι

$$0 \leq C(f) \leq \ln N \quad (3.26)$$

και $C(f) = \ln N$ σε διαχυτικές βάσεις.

Πρέπει να παρατηρήσουμε ότι ο ορισμός (3.25) δεν έχει καμία σχέση με την εντροπία Shannon. Το θεώρημα Shannon [SHA49] μας λέει ότι ένα κάτω όριο του αριθμού των bits που απαιτούνται για την δυαδική κωδικοποίηση των N όρων $\frac{|\langle f, g_m \rangle|^2}{\|f\|^2}$, είναι η εντροπία της κατανομής πιθανότητας των N τιμών $\frac{|\langle f, g_m \rangle|^2}{\|f\|^2}$. Η κατανομή πιθανότητας, μπορεί διαφέρει σημαντικά από τις τιμές καθ' αυτές. Για παράδειγμα, στην περίπτωση μίας διαχυτικής βάσης, το κόστος προσέγγισης δίνεται από την (3.23) ενώ η εντροπία Shannon είναι μηδενική¹³.

Κόστος ℓ^p : Για $p < 2$ η $\varphi(x) = x^{p/2}$ ($x \geq 0$) είναι κοίλη. Το αντίστοιχο συναρτησιακό κόστος για μία βάση \mathcal{B}^λ είναι το

$$C(f, \lambda) = \sum_{m=1}^N \frac{|\langle f, g_m^\lambda \rangle|^p}{\|f\|^p}$$

Βάσει της (3.23)

$$1 \leq C(f) \leq N^{1-p/2}$$

3.3.2 Προσαρμοστική Αναζήτηση σε Δυαδικά Λεξικά

Στην παράγραφο αυτήν, κατασκευάζουμε λεξικά τοπικών ορθοκανονικών βάσεων του $\{0, \dots, N-1\}$ ($N = 2^L$, $L > 0$), που προκύπτουν από τις δυαδικές διαμερίσεις του $[-1/2, N-1/2]$ και δείχνουμε πώς μπορεί να επιλεγεί η βέλτιστη βάση υπό την έννοια του θεωρήματος 3.4.

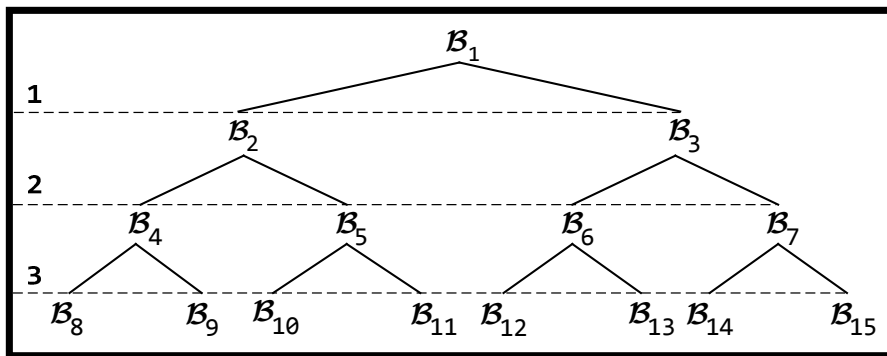
¹³ Εάν x είναι η τυχαία μεταβλητή με τιμές τα $\frac{|\langle f, g_m \rangle|^2}{\|f\|^2}$, $m = 1, \dots, N$, τότε $P\{x = 1/N\} = 1$ και η εντροπία Shannon είναι $-1 \cdot \log_2(1) = 0$.

Εάν $\{\alpha_0, \alpha_1, \dots, \alpha_{M-1}\}$ είναι μία τέτοια διαμέριση ¹⁴, συμβολίζουμε με δ_m το μήκος του διαστήματος επικάλυψης των πλαισίων g_m και g_{m+1} γύρω από το α_m επομένως το g_m είναι για κάθε m , ένα λείο πλαίσιο με μη μηδενικές τιμές στα ακέραια σημεία του διαστήματος $[\alpha_m - \delta_m, \alpha_{m+1} + \delta_{m+1}]$.

Επιπλέον, όπως αναφέρθηκε στην υποσημείωση 12, για την αναπαράσταση ενός διανύσματος $f \in \mathbb{C}^N$ κάνουμε κάποιες υποθέσεις, ή για τις τιμές του f εκτός του $\{0, \dots, N-1\}$ ή για την μορφή των πλαισίων στα 0 και $N-1$.

Τέλος, υποθέτουμε ότι οι επικαλύψεις όλων των πλαισίων όλων των δυαδικών διαμερίσεων έχουν το ίδιο μήκος δ . Για να εξασφαλίσουμε ότι μόνον γειτονικά πλαίσια επικαλύπτονται, θα πρέπει να ισχύει $\alpha_{m+1} - \alpha_m \geq 2\delta$. Εάν $M = 2^{-j}N = 2^{L-j}$, $\alpha_{m+1} - \alpha_m = 2^j$ και $1 \leq \delta \leq 2^{j-1}$.

Για ένα δεδομένο $1 < j \leq L$, υπάρχουν συνολικά 2^{L-j} δυαδικές διαμερίσεις πλάτους 2^j . Οι βάσεις που προκύπτουν από τις διαμερίσεις αυτές, μπορούν να αναπαρασταθούν με τους κόμβους ενός πλήρους δυαδικού δέντρου, βάθους $L-j$. Ένα τέτοιο δέντρο φαίνεται στο Σχ. 8. Η αριθμοδότηση των κόμβων είναι τυπική και έχει γίνει ξεκινώντας από το 1 για την μηδενική διαμέριση και αναθέτοντας στα παιδιά του κόμβου \mathcal{B}_n τους αριθμούς $2n$ για το αριστερό και $2n+1$ για το δεξί. Ο πατέρας του κόμβου \mathcal{B}_n είναι τότε ο $\mathcal{B}_{\lfloor n/2 \rfloor}$, όπου $\lfloor x \rfloor$ είναι το ακέραιο μέρος του x .



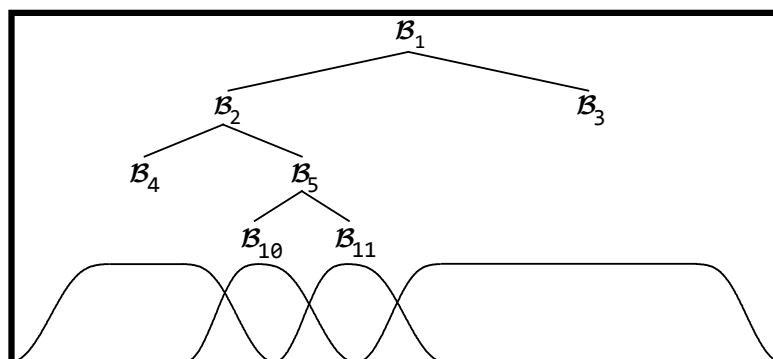
Σχ. 8: Πλήρες δέντρο τοπικών βάσεων Malvar, που αναπαριστά 3 διαδοχικές δυαδικές διαμερίσεις. Κάθε διαμέριση αντιστοιχεί στους κόμβους ενός από τα επίπεδα 1 – 3, και προκύπτει από την διχοτόμηση των διαστημάτων της προηγούμενης.

¹⁴ σημειώνουμε ότι εφ' όσον $N = 2^j$ και οι διαμερίσεις είναι δυαδικές, θα ισχύει $M = 2, 4, 8, \dots$ οπότε αναγκαστικά $\alpha_m \in \mathbb{Z} + \frac{1}{2}$ ($m = 0, \dots, M-1$)

Μία βάση μπορεί να προκύψει από τους κόμβους του ίδιου ή διαφορετικών επιπέδων του δέντρου. Το κριτήριο “καταλληλότητας” είναι το εξής:

Εάν ένας κόμβος δεν συμμετέχει στην βάση, τότε θα πρέπει να συμμετέχουν και τα δύο του παιδιά.

Μία βάση που περιέχει στοιχεία από διαμερίσεις διαφορετικών επιπέδων, φαίνεται στο Σχ. 9.



Σχ. 9: Αναπαράσταση σήματος από βάσεις διαφορετικών επιπέδων σε ένα δυαδικό λεξικό

Με βάση το παραπάνω κριτήριο καταλληλότητας, αποδεικνύουμε ότι:

Πρόταση 3.7: Ο αριθμός N_J των βάσεων που περιέχονται σε ένα πλήρες δυαδικό δέντρο βάθους J , ικανοποιεί την σχέση

$$2^{2^{J-1}} \leq N_J \leq 2^{\frac{5}{4}2^{J-1}} \quad (3.27)$$

Απόδειξη

Με αναδρομή στο J , για $J = 0$ υπάρχει μία μόνον βάση, που αντιστοιχεί στην μηδενική διαμέριση του $\{0, \dots, N-1\}$.

Ο αριθμός βάσεων N_{J+1} σε ένα δέντρο βάθους $J+1$, προκύπτει από τον αριθμό διαμερίσεων των N_J των δύο υποδέντρων του βάθους J και είναι

$$N_{J+1} = 1 + N_J^2 \quad (3.28)$$

Επειδή $N_{J+1} \geq N_J^2$ το αριστερό μέλος της (3.27) προκύπτει ευθέως από την αναδρομική σχέση (3.28).

Για το δεξί μέλος της (3.27) εργαζόμαστε ως εξής:

¹⁵ Για κάθε βάση του αριστερού υποδέντρου έχουμε N_J επιλογές για το δεξί. Επιπλέον έχουμε την επιλογή της μηδενικής διαμέρισης.

$$N_{J+1} = N_J^2(1 + N_J^{-2}) \Rightarrow \lg N_{J+1} \leq 2\lg N_J + \lg(1 + N_J^{-2})$$

όπου με \lg συμβολίζουμε τον λογάριθμο βάσης 2.

Είναι $N_J \geq 2 \forall J \geq 1$ και, $\lg(x) \leq x \forall x > 0$, επομένως η παραπάνω σχέση γίνεται

$$\lg N_{J+1} \leq 2\lg N_J + \frac{1}{4} \quad (3.29)$$

Κάνοντας διαδοχικές αντικαταστάσεις στην (3.29) με θέτοντας $N_1 = 2$, τελικά προκύπτει

$$\lg N_{J+1} \leq 2^J + \frac{1}{4} \sum_{j=0}^{J-1} 2^j \leq 2^J + 2^J/4 \Rightarrow N_{J+1} \leq 2^{\frac{5}{4}2^J}$$

QED

Στην αρχή της παραγράφου δείξαμε ότι το πλάτος επικάλυψης δ των πλαισίων μίας διαμέρισης και το πλάτος 2^j της διαμέρισης συνδέονται με την σχέση $1 \leq \delta \leq 2^{j-1}$. Στην περίπτωση αυτή, το βάθος του δυαδικού δέντρου είναι $J = L - j$. Θέτουμε $\eta = 2^{-L}\delta$ οπότε η προηγούμενη σχέση γράφεται

$$J \leq -\lg(2\eta)$$

Όταν το η είναι δεδομένο, το μέγιστο βάθος του δυαδικού δέντρου είναι επομένως $-\lg(2\eta)$ και από την πρόταση 3.7 προκύπτει ο μέγιστος αριθμός βάσεων του λεξικού

$$2^{1/(4\eta)} \leq N_J \leq 2^{3/(8\eta)}$$

Προσθετικά συναρτησιακά Κόστους και ο Αλγόριθμος Coifman-Wickerhauser

Σε ένα δυαδικό λεξικό βάσεων Malvar, κάθε κόμβος \mathcal{B}_n αντιστοιχεί σε έναν υπόχωρο V_n του $V_1 = \mathbb{C}^N$ που, παράγεται από την βάση \mathcal{B}_1 . Στον V_n αντιστοιχούν δύο βάσεις: η \mathcal{B}_n και η $\mathcal{B}_{2n} \cup \mathcal{B}_{2n+1}$. Η βάση που αναπαριστά καλύτερα τον V_n μπορεί τότε να υπολογιστεί συγκρίνοντας τις τιμές ενός κοίλου συναρτησιακού κόστους C , στις βάσεις \mathcal{B}_n και $\mathcal{B}_{2n} \cup \mathcal{B}_{2n+1}$. Ο υπολογισμός είναι ιδιαίτερα απλός για τα συναρτησιακά της προηγούμενης παραγράφου που είναι προσθετικά δηλαδή,

$$C(\mathcal{B}_{2n} \cup \mathcal{B}_{2n+1}) = C(\mathcal{B}_{2n}) + C(\mathcal{B}_{2n+1}) \quad (3.30)$$

Πρόταση 3.8 [COIF92]: Έστω $\mathbb{I}_1, \mathbb{I}_2$ δύο διαδοχικά διαστήματα του \mathbb{Z} , ίσου μήκους N και $\mathbb{I} = \mathbb{I}_1 \cup \mathbb{I}_2$. Έστω ακόμη $\mathcal{B}_1, \mathcal{B}_2$ οι βέλτιστες βάσεις (των περιορισμών) της f στα \mathbb{I}_1 και \mathbb{I}_2 αντίστοιχως και C ένα κοίλο προσθετικό συναρτησιακό κόστους. Η βέλτιστη βάση \mathcal{B} της f στο \mathbb{I} , δίνεται από την σχέση:

$$\mathcal{B} = \begin{cases} \mathcal{B}_1 \cup \mathcal{B}_2 & \text{αν } C(f, \mathcal{B}_1) + C(f, \mathcal{B}_2) < C(f, \mathcal{B}) \\ \mathcal{B} & \text{αν } C(f, \mathcal{B}_1) + C(f, \mathcal{B}_2) \geq C(f, \mathcal{B}) \end{cases} \quad (3.31)$$

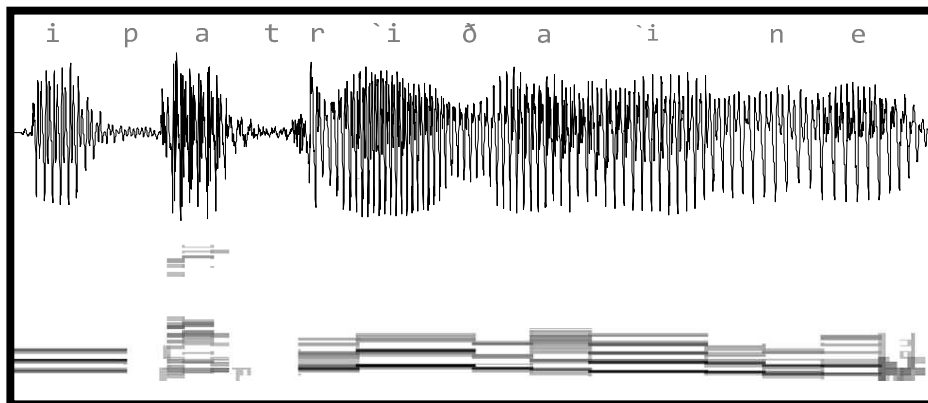
////

Σε ένα δυαδικό λεξικό βάθους J , ξεκινούμε από τα φύλλα του δέντρου που αντιστοιχούν στην βάση

$$\bigcup_{n=0}^{2^J-1} \mathcal{B}_{2^{J-1}+n},$$

και επισκεπτόμαστε όλους τους κόμβους έως την ρίζα του δέντρου. Όταν η τιμή του C στον κόμβο \mathcal{B}_j είναι μικρότερη από το άθροισμα των τιμών του στους \mathcal{B}_{2j} και \mathcal{B}_{2j+1} αντικαθιστούμε την βάση $\mathcal{B}_{2j} \cup \mathcal{B}_{2j+1}$ από την \mathcal{B}_j . Στην ρίζα, όλοι οι δυνατοί συνδυασμοί βάσεων έχουν ελεγχθεί και η βέλτιστη βάση για το f έχει υπολογιστεί.

Στο Σχ. 10 εικονίζεται η βέλτιστη βάση Malvar με βάση τον αλγόριθμο Coifman-Wickerhauser για ένα σήμα φωνής. Επάνω από την χρονική αναπαράσταση του σήματος φαίνεται η μεταγραφή της φράσης "η Πατρίδα είναι" σε σύμβολα του Διεθνούς Φωνητικού Αλφαβήτου (IPA symbols). Κάθε σύμβολο έχει τοποθετηθεί περίπου στο μέσον της περιοχής που αντιπροσωπεύει στο ακουστικό σήμα. Αξίζει να σημειωθεί η εμφανής σχέση του μήκους του βέλτιστου πλαισίου ανάλυσης με την διάρκεια κάθε φωνήματος.



Σχ. 10: Χρονική αναπαράσταση της φράσης "Η πατρίδα είναι" (επάνω μέρος) και φασματογράφημα της βέλτιστης βάσης Malvar.

3.4 ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΦΩΝΗΣ

Όλες οι σύγχρονες βάσεις προφορικού λόγου, ενσωματώνουν πληροφορίες για τα όρια των γλωσσολογικών τους στοιχείων όπως φωνήματα, συλλαβές και λέξεις ¹⁶.

Οι πληροφορίες αυτές είναι ιδιαίτερα χρήσιμες στην τεκμηρίωση γλωσσολογικών μελετών και την εκπαίδευση συστημάτων αναγνώρισης φωνής. Εξάλλου, η ίδια η διαδικασία αναγνώρισης φωνής βελτιώνεται σημαντικά εάν τα φωνητικά όρια του σήματος εντοπιστούν σε κάποιο αρχικό στάδιο της επεξεργασίας [VM90].

Επίσης σημαντική είναι η σχέση της αυτόματης τμηματοποίησης με την κωδικοποίηση του φωνητικού σήματος: εκ των προτέρων γνώση των ακουστικών ορίων των φωνημάτων, βελτιώνει την επίδοση των αλγορίθμων παραμετροποίησης, καθορίζοντας ένα σταθερό περιβάλλον για την εξαγωγή των παραμέτρων.

Στις επόμενες παραγράφους, γίνεται μία επισκόπηση των κυριότερων μεθόδων τμηματοποίησης και περιγράφεται πώς οι βέλτιστες βάσεις της ενότητας 3.3 μπορούν να χρησιμοποιηθούν για το ίδιο πρόβλημα ως ένα εργαλείο χαμηλού υπολογιστικού κόστους.

3.4.1 Το Σήμα Φωνής

Ο λόγος είναι αποτέλεσμα της κίνησης των αρθρωτών (articulators) του φωνητικού μας συστήματος (χείλη, γλώσσα, σιαγώνες, ρινικές κοιλότητες και γλωττίδα) και ως εκ τούτου, τα χαρακτηριστικά του ακουστικού σήματος μεταβάλλονται με τον χρόνο.

Θα μπορούσε κανείς να υποθέσει ότι, κάθε φωνητικός φθόγγος (φώνημα) καθορίζεται – έστω και κατά προσέγγιση – από την ίδια πάντοτε στιγμιαία διάταξη των αρθρωτών, και το γεγονός αυτό αντανακλάται στα χαρακτηριστικά του ακουστικού σήματος. Η υπόθεση δεν ευσταθεί για διάφορους λόγους:

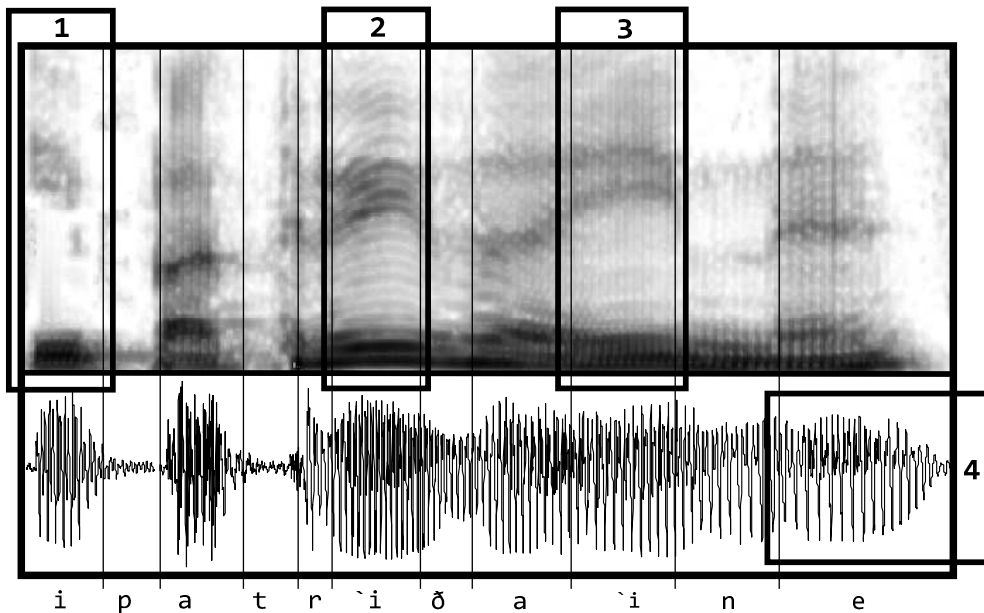
- Οι ταχύτητες κίνησης των αρθρωτών είναι γενικά μικρές επομένως η θέση τους, κατά την εκφορά ενός φθόγγου, επηρεάζεται σε σημαντικό βαθμό από το φωνητικό του περιβάλλον, τα φωνήματα δηλαδή που προηγούνται ή έπονται χρονικά του φθόγγου. Στο Σχ. 11 τα μέρη 2 και 3 αντιστοιχούν στο ίδιο φώνημα /'i/. Τα φασματικά ωστόσο περιεχόμενα των δύο (γραμματικά ταυτόσημων) ήχων διαφέρουν σημαντικά ενώ,

¹⁶ ή άλλα μορφολογικά συγκροτήματα σε γλώσσες που δεν έχουν αυστηρή συλλαβική δομή (π.χ. Αγγλικά).

εμφανής είναι και η διαφορά στις διάρκειές τους. Το ίδιο συμβαίνει και με τα δύο /a/ της φράσης ¹⁷.

- Οι αρθρωτές διαφέρουν από άνθρωπο σε άνθρωπο. Οι διαφορές δεν είναι μόνον ανατομικές αλλά και φυσιολογικές (λειτουργικές) και αυτό αντανακλάται στα ακουστικά χαρακτηριστικά των εκφερόμενων φθόγγων.
- Η γραμματική, συντακτική και σημασιολογική πληροφορία του φωνητικού μηνύματος – που, στον γραπτό λόγο αναπαρίσταται με τον τόνο, την στίξη αλλά και την γραμματική και συντακτική κατηγορία κάθε λέξης – επιφέρει σημαντικές μεταβολές στην εκφορά των φωνημάτων όταν τονίζονται ή βρίσκονται κοντά σε σημεία στίξης ή όταν ανήκουν σε λέξεις με ειδικό γραμματικό ή συντακτικό ρόλο. Π.χ. η φασματική δομή του φωνήματος /i/ στο μέρος 1 του Σχ. 11 είναι πολύ διαφορετική από το φάσμα των 2 και 3 επειδή το φώνημα βρίσκεται στην αρχή πρότασης – ακολουθεί μετά από μία περίοδο παύσης. Εξάλλου η διάρκεια του /e/ στο μέρος 4 του ίδιου σχήματος είναι σχετικά μεγάλη αποδίδοντας με τον τρόπο αυτόν το κόμμα που έπεται στην εκφερόμενη πρόταση.

¹⁷ Η περίπτωση του /i/ στα μέρη 2 και 3 είναι ενδεικτική της επίπτωσης που έχει στον προφορικό λόγο η διακύμανση της συχνότητας κίνησης των φωνητικών μας χορδών. Ο (γραμματικός) τόνος στο μέρος 2 αποδίδεται με υψηλή συχνότητα ενώ, ο τόνος στο 3 με χαμηλή. Επομένως στο 2 οι αρμονικές του σήματος είναι αραιότερα τοποθετημένες και αυτό φαίνεται στο Short-time Fourier φάσμα.



Σχ. 11: Φασματογράφημα και κυματομορφή της φράσης “Η πατρίδα είναι,” τμήματος της πρότασης “Η Πατρίδα είναι, νιώθεται, μας δίνει το ίχνος της καταγωγής και την προβολή μας στο μέλλον”. Ο οριζόντιος άξονας στο φασματογράφημα είναι χρόνος. Ο κατακόρυφος άξονας συχνότητα. Οι αποχρώσεις του γκρι αντιστοιχούν στο log-magnitude του πλαισιωμένου DFT του σήματος από λεία πλαίσια Hann. Τα φωνητικά φαινόμενα 1, 2, 3 και 4 αναλύονται στο κείμενο.

Από τα παραπάνω γίνεται σαφές ότι, ο υπολογισμός των ορίων των φωνημάτων μέσω της ταυτοποίησης κάθε φωνήματος ξεχωριστά, παρουσιάζει ιδιαίτερες δυσκολίες, λόγω της σημαντικής επίδρασης στο ακουστικό σήμα ανατομικών και γλωσσολογικών παραγόντων. Οι αντίστοιχες μέθοδοι τμηματοποίησης, προϋποθέτουν την ύπαρξη εκτεταμένων λεξικών με ακουστικά πρότυπα των φωνημάτων και έχουν αυξημένες απαιτήσεις τόσο σε χώρο όσο και σε επεξεργασία ([PARK], [LI99]).

Στον αντίποδα βρίσκονται οι μέθοδοι τμηματοποίησης που βασίζονται στον εντοπισμό των περιοχών απότομης μεταβολής των χαρακτηριστικών του ακουστικού σήματος θεωρώντας ότι, διαφορετικά φωνήματα έχουν αναγκαστικά, εμφανώς διαφορετικά ακουστικά χαρακτηριστικά. Στην προσέγγιση αυτή δεν προϋποθέτουμε καμία γνώση του φωνητικού περιεχομένου του σήματος και για τον λόγο αυτόν οι μέθοδοι είναι γνωστές με τους όρους unconstrained ή blind speech segmentation.

Θεωρούμε ότι διαθέτουμε τις παραμετρικές αναπαραστάσεις διαδοχικών πλαισίων ανάλυσης του σήματος φωνής. Οι αναπαραστάσεις αυτές μπορεί να είναι συντελεστές DFT (συνήθως τα λογαριθμικά πλάτη τους), συντελεστές cepstrum (ή οι αντίστοιχοι Mel-frequency warped coefficients) [RS78], ή οι παράμετροι ενός μοντέλου του πλαισίου (π.χ. οι συντελεστές μερικής αυτοσυσχέτισης ενός auto-regressive μοντέλου [BD]). Τα πλαίσια έ-

χουν εν γένει σταθερό πλάτος, αρκετά μικρότερο από την μέση διάρκεια ενός φωνήματος.

3.4.2 Ακρότατα Κλίσης

Το αποτέλεσμα της παραμετρικής επεξεργασίας διαδοχικών πλαισίων του σήματος είναι μία συνάρτηση $f(t) \in \mathbb{C}^P$, $P \geq 1$. Τα σημεία μέγιστης μεταβολής της είναι τ' ακρότατα της Df . Συνήθως χρειάζεται να εξομαλύνουμε την f , αντικαθιστώντας την από τις μέσες τιμές $f * g$

$$f_g(t) = f * g(t) = \int f(x)g(t-x)dx$$

όπου g είναι μία λεία θετική συνάρτηση πλαισίωσης, οπότε

$$Df_g = D(f * g) = f * Dg \quad (3.32)$$

3.4.3 Ιεραρχική Ομαδοποίηση (Hierarchical Clustering)

Η ιεραρχική ομαδοποίηση [ZUE, SVEN] αφορά σε μία οικογένεια αλγορίθμων όπου γειτονικά πλαίσια κατατάσσονται στην ίδια ή διαφορετικές ομάδες και, τα όρια των φωνημάτων προκύπτουν από τα όρια διαδοχικών ομάδων. Κοινό χαρακτηριστικό των μεθόδων αυτών είναι ότι η ομαδοποίηση είναι γενικά πολυεπίπεδη δημιουργώντας τμήματα σε διαφορετικές κλίμακες ανάλυσης.

Στην απλούστερή της μορφή, η ιεραρχική ομαδοποίηση ξεκινά κατατάσσοντας κάθε πλαίσιο ανάλυσης σε διαφορετική ομάδα. Ακολούθως, διαδοχικά πλαίσια με κοντινά (υπό την συνήθη L^2 έννοια) χαρακτηριστικά, κατατάσσονται στην ίδια ομάδα και τα διανύσματα χαρακτηριστικών τους αντικαθίστανται από τον μέσο όρο τους. Η διαδικασία επαναλαμβάνεται έως ότου καταλήξουμε σε ένα μόνον διάνυσμα ή όσο οι αποστάσεις μεταξύ διαδοχικών ομάδων διατηρούνται πάνω από ένα κατώφλι.

Γενικότερα, η κατάταξη πλαισίων σε ομάδες, γίνεται με βάση την τιμή ενός συναρτησιακού του \mathbb{C}^P που μετρά με κάποιον τρόπο την ομοιογένεια των χαρακτηριστικών διαδοχικών πλαισίων ανάλυσης.

3.4.4 Μέγιστα μετασηματισμών Wavelet (Wavelet Modulus Maxima)

Οι μέθοδοι της προηγούμενης παραγράφου δείχνουν ότι, η τμηματοποίηση φωνής μπορεί ν' αναχθεί στον εντοπισμό μεταβολών μεταξύ διαδοχικών πλαισίων σήματος, σε διαφορετικές κλίμακες ανάλυσης.

Ο μετασχηματισμός wavelet (WT) εξ' ορισμού αναπαριστά ένα σήμα από τις μεταβολές αυτές. Αν υποθέσουμε $P = 1$ και ότι το $f(t) \in \mathbb{C}$ έχει προέλθει από την ανάλυση του ακουστικού σήματος σε κάποια κλίμακα, είναι δυνατόν ν' αποδειχθεί ότι, τα μέγιστα της πρώτης παραγώγου του f , αντιστοιχούν στα ακρότατα του WT της f , $Wf(s,t) = \langle f, \psi_{s,t} \rangle$, αρκεί το ψ να έχει μηδενική ροπή πρώτης τάξης¹⁸. Γενικότερα, είναι δυνατόν ν' αποδειχθεί η ακόλουθη πρόταση [MALLAT, §6.1.2]:

Πρόταση 3.9: Ένα wavelet ψ με ισχυρή απομείωση έχει n μηδενικές ροπές αν και μόνον αν υπάρχει συνάρτηση φ ισχυρά απομειούμενη, ώστε

$$\psi(t) = (-1)^n D^n \varphi(t)$$

Κατά συνέπεια, ο WT μίας συνάρτησης f γράφεται:

$$Wf(s,t) = \langle f, \psi_{s,t} \rangle = s^n D^n (f * \varphi_s)(t) \quad (3.33)$$

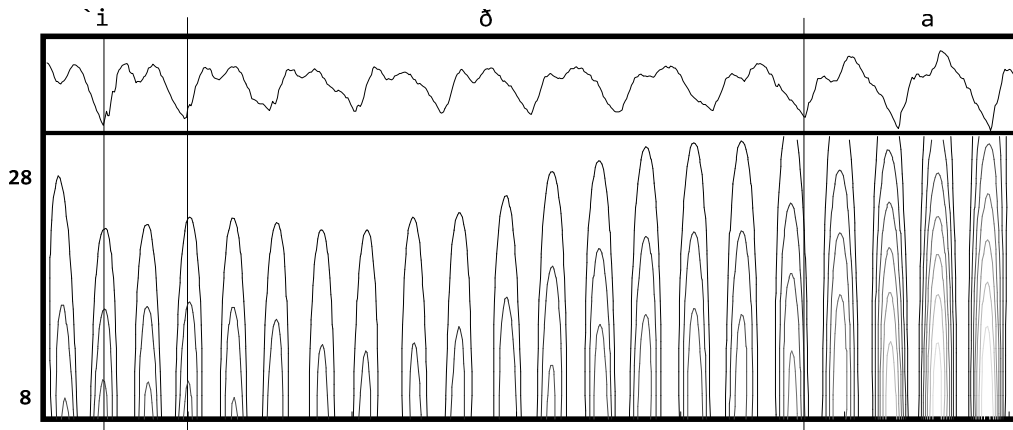
όπου $\varphi_s(x) = s^{-1/2} \varphi(-x/s)$ και (κατά τα γνωστά) $\psi_{s,t}(x) = s^{-1/2} \psi\left(\frac{x-t}{s}\right)$.

////

Η (3.33) είναι ο υπολογισμός (3.32) σε διαφορετικές κλίμακες ανάλυσης.

Σε αντίθεση με την §3.4.3, η παραπάνω σχέση δεν παρέχει έναν τρόπο συσχετισμού των ακροτάτων της §3.4.2 σε διαφορετικές κλίμακες ανάλυσης αφού, ακρότατα του Wf μπορούν να εμφανίζονται ή να εξαφανίζονται με αυθαίρετο τρόπο σε διαφορετικές κλίμακες. Εάν ωστόσο οι γεωμετρικοί τόποι των ακροτάτων είναι συνεχείς καμπύλες του χώρου $t \times s$, ο συσχετισμός αυτός είναι δυνατός: ξεκινώντας από ένα ακρότατο $m(s_0, t_0) = |Wf(s_0, t_0)|$, ακολουθούμε την καμπύλη $m(s, t_s)$ που διέρχεται από το (s_0, t_0) με φορά προς τις μεγαλύτερες κλίμακες. Μία τέτοια περίπτωση φαίνεται στο **Σχ. 12**.

¹⁸ Η ροπή μηδενικής τάξης $\int \psi(t) dt = \hat{\psi}(0)$ πρέπει να είναι μηδενική για να ορίζεται ο WT.



Σχ. 12: Στο επάνω μέρος εικονίζεται το τμήμα /`iδa/ του σήματος φωνής του Σχ. 11 (ολόκληρο το δ και μέρος μόνον των `i και a). Στο κάτω μέρος, τα μέγιστα του WT καθώς η κλίμακα (scale) αυξάνεται από 8 σε 32 (μεγάλες κλίμακες αντιστοιχούν σε μικρές αναλύσεις). Περιοχές μεγάλης μεταβολής, χαρακτηρίζονται από μέγιστα που διατηρούνται σε πολλές κλίμακες.

Η πρόταση που ακολουθεί, παρέχει την θεωρητική βάση για τα παραπάνω [MALLAT, §6.2.1]:

Πρόταση 3.10: Εάν η συνάρτηση φ της 3.9 είναι Gaussian, τότε για κάθε $f \in L^2$ τα μέγιστα του WT της f ανήκουν σε συνδεδεμένες καμπύλες που δεν διακόπτονται καθώς η κλίμακα αυξάνεται (η ανάλυση μειώνεται).

////

3.4.5 Τοπική Διάσπαση (Temporal Decomposition)

Στην μέθοδο "τοπικής διάσπασης" [ATAL83] θεωρούμε ότι καθένα από N διαδοχικά διανύσματα χαρακτηριστικών $f(n) \in \mathbb{C}^P$ ($n = 0, \dots, N-1$) του σήματος φωνής, προέρχονται από $M \leq P$ διανύσματα $s(k) \in \mathbb{C}^P$ ($k = 1, \dots, M$), κάθε ένα από τα οποία αντιπροσωπεύει ένα ολόκληρο φωνητικό τμήμα του υπό ανάλυση σήματος. Ακριβέστερα θεωρείται ότι τα $s(k)$ παράγουν τον δ.χ. των $f(n)$ και επομένως ισχύει η ακόλουθη σχέση μεταξύ των $f(n)$ και $s(n)$:

$$f(n) = \sum_{j=1}^M g_j(n) s(j), \quad (n = 0, \dots, N-1) \quad (3.34)$$

Τα $g_j \in \mathbb{C}^N$ ($j = 1, \dots, M$) αντιπροσωπεύουν την συνεισφορά (συντελεστή βάρους) κάθε $s(j)$ στον υπολογισμό του $f(n)$ και, ως εκ τούτου, ακολούθως θ' αναφέρονται και ως πλαίσια. Θέτοντας

$$F(:, n) = f(n),$$

$$G(n, j) = g_j(n) \text{ και}$$

$$S(:, j) = s(j)$$

έχουμε $F \in \mathbb{C}^{N \times P}$, $G \in \mathbb{C}^{N \times M}$ και $S \in \mathbb{C}^{P \times M}$ αντιστοίχως και η (3.34) γράφεται ισοδύναμα

$$F = SG^* \quad (3.35)$$

όπου το (*) δηλώνει τον συζυγή¹⁹ ενός πίνακα.

Στόχος είναι ο προσδιορισμός των S και G από τον F ώστε οι τιμές $g_j(n)$ κάθε πλαισίου g_j να είναι εντοπισμένες γύρω από j . Εάν ο G είναι πλήρους βαθμού, η λύση της (3.35) είναι μοναδική για κάθε G , $S = F(G^*)^{-1}$. Σε κάθε περίπτωση, η λύση ελαχίστων τετραγώνων της (3.35) είναι η

$$S = F(G^*)^\dagger$$

όπου $(G^*)^\dagger$ είναι ο ψευδοαντίστροφος του G^* .

Κατά τα γνωστά, θεωρούμε έναν πίνακα $X \in \mathbb{C}^{N \times M}$ ως γραμμική απεικόνιση $\mathbb{C}^M \rightarrow \mathbb{C}^N$ και συμβολίζουμε $X(\mathbb{C}^M)$ το πεδίο τιμών του, \mathcal{N}_X τον πυρήνα του και \mathcal{N}_X^\perp την γραμμική θήκη των γραμμών του (*rowspan*). Αναστρέφοντας την (3.35) προκύπτει αμέσως ότι $\mathcal{N}_F^\perp \subset G(\mathbb{C}^P)$ επομένως είναι σκόπιμο να θέσουμε $\mathcal{N}_F^\perp = G(\mathbb{C}^P)$. Ο υπολογίζεται κρατώντας τα ιδιοδιανύσματα των $M \leq P$ μεγαλύτερων ιδιοτιμών του V στην SVD ανάλυση $F = UDV^*$ του F . Ακριβέστερα είναι

$$\mathcal{N}_F^\perp \subset V(\mathbb{C}^P) = \text{colspan}(V) = \text{span}\{v_1, \dots, v_P\}$$

και για κάθε j θα έχουμε

$$g_j = \sum_{k=1}^M b_{jk} v_j \quad (3.36)$$

Οι συντελεστές υπολογίζονται έτσι, ώστε να ελαχιστοποιείται το ακόλουθο τετραγωνικό κριτήριο κόστους

$$\partial(j)^2 = \|g_j\|^{-2} \sum_{n=0}^{N-1} (n-j)^2 g_j(n)^2 \quad (3.37)$$

που λόγω της (3.36) γράφεται ισοδύναμα

$$\min \partial(j)^2 = \min b^* (V^* N_j^2 V) b = \min b_j^* R b_j \quad (3.38)$$

με $\sum_{k=1}^M b_{jk}^2 = 1$.

Ο R είναι αυτοσυζυγής και θετικά ορισμένος, πλήρους βαθμού, επομένως έχει θετικές, μη μηδενικές ιδιοτιμές και ορθοκανονικά ιδιοδιανύσματα (π.χ. [CHR, κεφ. 1]). Το b_j που ελαχιστοποιεί την (3.38) είναι το ιδιοδιάνυσμα της μικρότερης ιδιοτιμής του R .

¹⁹ Δηλαδή τον adjoint ενός πίνακα. Στην ανάλυση που ακολουθεί, οι πίνακες θεωρούνται τελεστές $\mathbb{C}^M \rightarrow \mathbb{C}^N$ και επομένως η ορολογία είναι η αντίστοιχη. Σε κάθε περίπτωση, εννοούμε τον ανάστροφο πίνακα συζυγών στοιχείων.

Η διαδικασία (3.38) πρέπει να επαναληφθεί για όλα τα δυνατά j , που είναι N . Επειδή όμως τα g_j προκύπτουν από τα ίδια M v_j , τα περισσότερα g_j θα είναι τα ίδια. Τα σημεία μεταβολής των g_j είναι τα ζητούμενα όρια των φωνημάτων.

Ο αλγόριθμος αρχικά είχε προταθεί από τον Atal για την συμπίεση σημάτων φωνής, θεωρώντας ότι ο πίνακας F περιέχει τα log-area ratios (partial correlation coefficients) της auto-regressive ανάλυσης διαδοχικών πλαισίων του σήματος. Είναι προφανές ότι η αριθμητική πολυπλοκότητα του αλγορίθμου είναι μεγάλη και οφείλεται στην SVD ανάλυση του F και το πρόβλημα ιδιοτιμών (3.38).

3.4.6 Entropy-based Segmentation

Η τοπική διάσπαση (temporal decomposition) προτάθηκε αρχικά ως μία μέθοδος συμπίεσης ενός σήματος φωνής αλλά κυρίως χρησιμοποιείται στην τμηματοποίηση φωνής (blind segmentation). Κατ' αντιστοιχία, μπορούμε να δούμε την μέθοδο των Coifman-Wickerhauser ως έναν τρόπο εντοπισμού των ορίων διαδοχικών φωνημάτων (ή κατηγοριών φωνημάτων ή φωνητικών συγκροτημάτων). Ήδη από το Σχ. 10 φαίνεται ότι η μέθοδος επιλέγει τοπικές βάσεις αντίστοιχες της διάρκειας κάθε φωνήματος. Μεγάλες μεταβολές στο φασματικό περιεχόμενο του σήματος, οδηγούν σε αναπαραστάσεις από διαφορετικές τοπικές βάσεις.

Είναι αξιοσημείωτο το γεγονός ότι, σε ό,τι τουλάχιστον αφορά τον εντοπισμό των ορίων γειτονικών φωνημάτων, μπορούμε να λάβουμε αντίστοιχα αποτελέσματα περιορίζοντας σημαντικά την αναζήτηση σε ένα δυαδικό λεξικό βάσεων Malvar. Ειδικότερα, μπορούμε να εντοπίσουμε τα όρια φωνημάτων συγκρίνοντας τους κόμβους δύο μόνον επιπέδων του λεξικού (§3.3.2). Το κατώτερο επίπεδο αντιπροσωπεύει την διακριτική ικανότητα της αναπαράστασης (μεταβολές συντομότερες από το μήκος των πλαισίων του επιπέδου αυτού δεν είναι δυνατόν να εντοπιστούν) και κωδικοποιεί αποτελεσματικά απότομες φασματικές μεταβολές. Αντιθέτως, το ανώτερο επίπεδο θα πρέπει να αναπαριστά αποτελεσματικότερα, περιοχές σταθερού φάσματος.

Με βάση την 3.3.2, εάν το μήκος του σήματος είναι 2^N και το μήκος του πλαισίου στα φύλλα του δυαδικού λεξικού 2^P , οι βάσεις θα είναι οι $\mathcal{B}_{2^{L-1}}, \dots, \mathcal{B}_{2^L}$ όπου $L = N - P + 1$.

Ο υπολογισμός του κόστους δύο διαδοχικών βάσεων $\mathcal{B}_{2^{L-1}+2j}$ και $\mathcal{B}_{2^{L-1}+2j+1}$ γίνεται εύκολα λόγω της προσθετικότητας του entropy cost:

$$C(\mathcal{B}_{2^{L-1}+2j} \cup \mathcal{B}_{2^{L-1}+2j+1}) = C(\mathcal{B}_{2^{L-1}+2j}) + C(\mathcal{B}_{2^{L-1}+2j+1}) \quad (3.39)$$

Το κόστος αναπαράστασης στην αντίστοιχη βάση του προηγούμενου επιπέδου είναι

$$C(\mathcal{B}_{2^{L-2}+j}). \text{ Εάν}$$

$$C(\mathcal{B}_{2^{L-2}+j}) > C(\mathcal{B}_{2^{L-1}+2j} \cup \mathcal{B}_{2^{L-1}+2j+1})$$

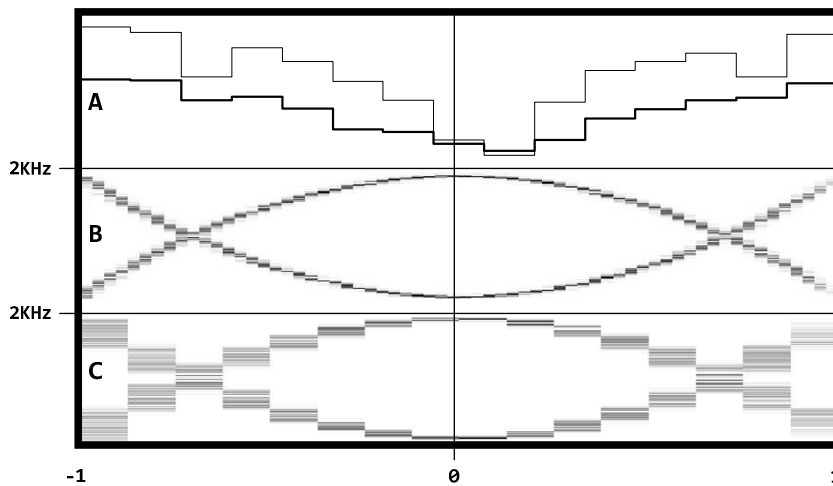
τα υπό εξέταση πλαίσια θεωρούνται οριακά διαφορετικά, εσωτερικά. Παρά την απλότητά της, η μέθοδος παράγει ενθαρρυντικά αποτελέσματα.

Στο Σχ. 13 φαίνεται η επίδοση της μεθόδου σε ένα multi-chirp αποτελούμενο από το άθροισμα ενός κυρτού και ενός κοίλου τετραγωνικού chirp:

$$f(t) = \cos(f_0 t + \alpha f_1 t^3) + \cos(f_1 t - \alpha f_0 t^3)$$

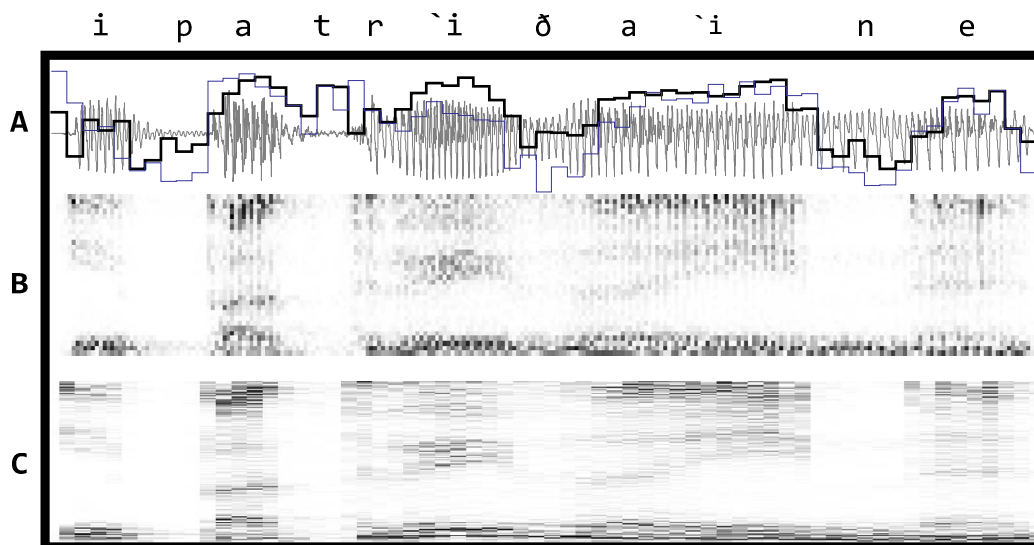
Για την ανάλυση έχουν χρησιμοποιηθεί $8192 = 2^{13}$ δείγματα της f στο διάστημα $[-1, 1]$ που αντιστοιχούν σε δειγματοληψία $\sim 4\text{KHz}$. Τα f_0 και f_1 έχουν τιμές 50 και 1950 Hz αντιστοίχως.

Όπως είναι αναμενόμενο, στις περιοχές γρήγορης μεταβολής των συχνοτήτων (άκρα των φασματογραφήματων **B** και **C** του σχήματος), προκρίνεται το συναρτησιακό κόστος που αντιστοιχεί σε μικρά πλαίσια ανάλυσης.



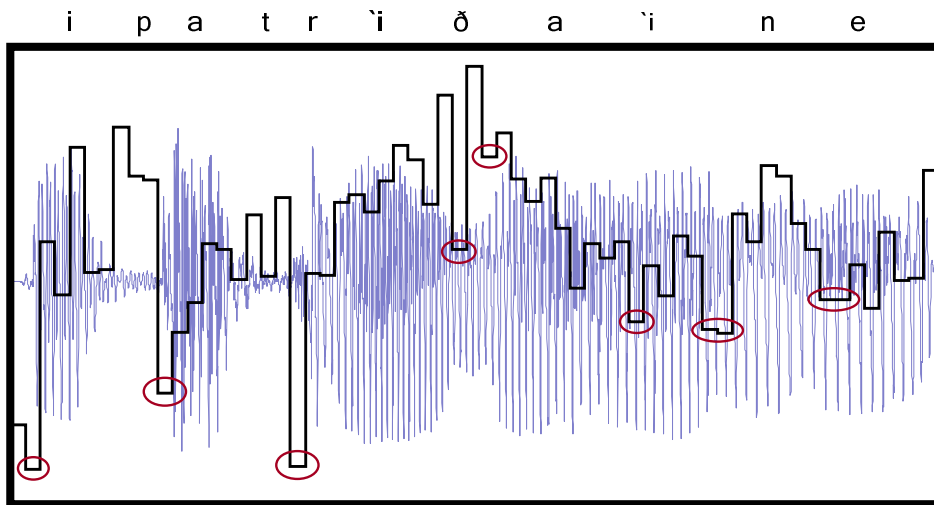
Σχ. 13: Entropy-based Segmentation τετραγωνικού multi-chirp. **A:** με λεπτή γραμμή η εντροπία μεγάλου μήκους πλαισίου (512 δείγματα). Με έντονη γραμμή, η εντροπία μικρού μήκους πλαισίου (128 δείγματα). **B:** φασματογραφήμα των συντελεστών Malvar για τα πλαίσια 128 δειγμάτων. **C:** φασματογραφήμα για τα πλαίσια 512 δειγμάτων.

Παρόμοια συμπεριφορά παρατηρείται σε συνθετότερα σχήματα, όπως η φωνή. Στο Σχ. 14 φαίνεται το αποτέλεσμα εφαρμογής της μεθόδου στο σήμα φωνής των σχημάτων 10 και 11. Έχουμε υπολογίσει την εντροπία δύο επιπέδων του δυαδικού λεξικού, που αντιστοιχούν σε πλαίσια $32 = 2^5$ και $128 = 2^7$ δειγμάτων. Για συχνότητα δειγματοληψίας 8kHz, οι διάρκειες των πλαισίων αυτών είναι περίπου 4 και 16 ms αντιστοίχως. Παρατηρούμε ότι η εντροπία του κατώτερου επιπέδου γίνεται μικρότερη στις περιοχές κοντά στις άκρες των φωνημάτων. Αξιοσημείωτη είναι η επίδοση της μεθόδου στα κρουστικά **p** και **t**.



Σχ. 14 A: Χρονική αναπαράσταση του σήματος φωνής "Η Πατρίδα είναι" και τιμές των συναρτησιακών κόστους για δύο επίπεδα βάσεων Malvar. Με έντονη γραμμή φαίνεται το κόστος μεγάλης ανάλυσης (μικρό μήκος πλαισίου). **B, C:** φασματογραφική απεικόνιση των συντελεστών Malvar για μικρό και μεγάλο πλαίσιο ανάλυσης αντιστοίχως. Οι ενισχυμένες υψηλές συχνότητες στα φασματογραφήματα οφείλονται στην προέμφαση του σήματος με συντελεστή 0.85.

Στο Σχ. 15 εικονίζεται το ίδιο φωνητικό σήμα. Αντί όμως των εντροπιών C_s και C_l μικρού και μεγάλου πλαισίου αντιστοίχως, εικονίζεται η διαφορά τους $C_s - C_l$. Σε κάποια σημεία του σήματος, φαίνεται ότι μεταβολή της σχετικής εντροπίας μπορεί να βελτιώσει την επίδοση της μεθόδου.



Σχ. 15: Μεταβολή της διαφοράς των εντροπιών σε σχέση με το σήμα "Η Πατρίδα είναι". Τα όρια του δ αντιστοιχούν σε ελάχιστα που και διαχωρίζονται στην εν λόγω αναπαράσταση, όχι όμως και στην αναπαράσταση του Σχ. 14

Ένας παράγοντας που περιορίζει την αποτελεσματικότητα του entropy-based segmentation, είναι η διασπορά των πλατών των κορυφών του φάσματος στις έμφωνες περιοχές του σήματος.

Για τον λόγο αυτόν, μία εκτενέστερη ανάλυση των δυνατοτήτων της μεθόδου θα πρέπει να περιλαμβάνει τρόπους μείωσης της διασποράς. Ένας τρόπος είναι η προεργασία του σήματος με την απόρριψη των συντελεστών Malvar που δεν αντιστοιχούν σε ακρότατα και την εξίσωση των πλατών όσων συντελεστών απομένουν (peak-picking and clipping). Εναλλακτικά, θα μπορούσε κανείς να εφαρμόσει την μέθοδο ξεχωριστά σε κάθε κλίμακα ενός wavelet packet decomposition του σήματος.

Σε κάθε περίπτωση πάντως, η δυναμική της εν λόγω προσέγγισης είναι σημαντική, κυρίως εάν αναλογιστούμε την απλότητα των σχετικών υπολογισμών.

////

3.5 ΠΑΡΑΡΤΗΜΑ

ΔΙΑΚΡΙΤΟΙ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΙ ΣΥΝΗΜΙΤΟΝΟΥ

Ο διακριτός μετασχηματισμός συνημίτονου έχει τέσσερις μορφές DCT-I, II, III, IV. Οι μορφές I και IV αντιστοιχούν στις βάσεις 4 και 3 αντιστοίχως της (3.4).

Για λόγους πληρότητας, περιγράφουμε συνοπτικά όλες τις μορφές DCT ξεκινώντας από τις III και IV. Για τις μορφές I, II και IV, δείχνουμε πώς μπορούν να υπολογιστούν μέσω του γρήγορου μετασχηματισμού Fourier εάν το υπό ανάλυση σήμα είναι πραγματικό.

Έστω $f \in \mathbb{C}^N$, $n = 0, \dots, N-1$

DCT-III

Επεκτείνουμε το f συμμετρικά γύρω από το $N-1$, στο $\{0, \dots, 2N-3\}$ θέτοντας

$$f(2N-2-n) = f(n), \quad n = \{1, \dots, N-2\}.$$

Λόγω της παραπάνω συμμετρίας, ο DFT $2N-2$ σημείων του εκτεταμένου διανύσματος γράφεται

$$f_{\text{II}}(k) = \frac{1}{\sqrt{2(N-1)}} \left[f(0) + (-1)^k f(N-1) \right] + \sqrt{\frac{2}{N-1}} \sum_{n=1}^{N-2} f(n) \cos\left(\frac{\pi}{N-1} kn\right) \quad (3.40)$$

όπου $k = 0, \dots, N-1$. Τα διανύσματα

$$\text{II}_k = \left(\frac{1}{\sqrt{2(N-1)}}, \sqrt{\frac{2}{N-1}} \cos\left(\frac{\pi}{N-1} k\right), \dots, \frac{1}{\sqrt{2(N-1)}} (-1)^k \right), \quad k = 0, \dots, N-1$$

είναι σχετικά εύκολο ν' αποδειχθεί ότι είναι ορθοκανονικά, επομένως συνιστούν μία ορθοκανονική βάση του \mathbb{C}^N :

Θέτουμε στην (3.40)

$$f(n) = \text{II}_{k'}(n) = \left(\frac{1}{\sqrt{2(N-1)}}, \frac{1}{\sqrt{2(N-1)}} \left(e^{\frac{\pi}{N-1} k'} + e^{-\frac{\pi}{N-1} k'} \right), \dots, \frac{1}{\sqrt{2(N-1)}} (-1)^k \right)$$

και την ξαναγράφουμε για $n = 0, \dots, 2N-3$, οπότε η ορθοκανονικότητα των II_k έπεται από την ορθοκανονικότητα των μιγαδικών εκθετικών στο $\{0, \dots, 2N-3\}$.

DCT-IV

Το f επεκτείνεται συμμετρικά γύρω από το $\frac{1}{2}$ και αντισυμμετρικά γύρω από το $N-\frac{1}{2}$ όπως στο 3 της (3.4). Η ορθοκανονικότητα των στοιχείων της βάσης DCT-IV έχει αποδειχθεί στην §3.1.

Γράφουμε $f_e(n) = f(2n)$ για τους $N/2$ άρτιους όρους του f και $f_o(n) = f(N-1-2n)$ για τους $N/2$ περιττούς όρους, διατεταγμένους αντίστροφα.

Έχουμε:

$$\begin{aligned}
 f_{\text{IV}}(k) &= \sum_{n=0}^{N/2-1} f_e(n) \cos\left[\frac{\pi}{N}\left(2n + \frac{1}{2}\right)\left(k + \frac{1}{2}\right)\right] + \\
 &\quad \sum_{n=0}^{N/2-1} f_o(n) \cos\left[\frac{\pi}{N}\left(N - 1 - 2n + \frac{1}{2}\right)\left(k + \frac{1}{2}\right)\right] \\
 &= \sum_{n=0}^{N/2-1} f_e(n) \cos\left[\frac{2\pi}{N}\left(n + \frac{1}{4}\right)\left(k + \frac{1}{2}\right)\right] + \\
 &\quad (-1)^k \sum_{n=0}^{N/2-1} f_o(n) \sin\left[\frac{2\pi}{N}\left(n + \frac{1}{4}\right)\left(k + \frac{1}{2}\right)\right]
 \end{aligned} \tag{3.41}$$

Εάν το f είναι πραγματικό, για άρτια k η (3.41) γράφεται σαν το πραγματικό μέρος α-θροίσματος,

$$f_{\text{IV}}(2k) = \text{real}\left\{e^{-i\pi k/N} \sum_{n=0}^{N/2-1} \left[(f_e(n) + if_o(n))e^{-i(n+1/4)\pi/N}\right] e^{-i2\pi kn/(N/2)}\right\} \tag{3.42}$$

και για περιττά k

$$f_{\text{IV}}(N-1-2k) = -im\left\{e^{-i\pi k/N} \sum_{n=0}^{N/2-1} \left[(f_e(n) + if_o(n))e^{-i(n+1/4)\pi/N}\right] e^{-i2\pi kn/(N/2)}\right\} \tag{3.43}$$

Για $0 \leq n < N/2$ θέτουμε

$$g(n) = (f_e(n) + if_o(n))e^{-i(n+1/4)\pi/N}$$

Ο DFT \hat{g} του g υπολογίζεται από έναν FFT $N/2$ σημείων. Οι εξισώσεις (3.42) και (3.43) δείχνουν ότι

$$\begin{aligned}
 f_{\text{IV}}(2k) &= \text{real}\left\{e^{-i\pi k/N} \hat{g}(k)\right\} \text{ και} \\
 f_{\text{IV}}(N-1-2k) &= -im\left\{e^{-i\pi k/N} \hat{g}(k)\right\}.
 \end{aligned}$$

DCT-I

Για k άρτιο η 4 της (3.4) γράφεται

$$f_{\text{I}}(2k) = \lambda_k \sum_{n=0}^{N/2-1} [f(n) + f(N-1-n)] \cos\left(\frac{\pi}{N}k\left(n + \frac{1}{2}\right)\right)$$

όπου $\lambda_0 = \sqrt{\frac{1}{N}}$ και $\lambda_k = \sqrt{\frac{2}{N}}$ αν $k \neq 0$.

Για περιττά k , η 4 της (3.4) γράφεται

$$f_{\text{I}}(2k+1) = \sum_{n=0}^{N/2-1} [f(n) + f(N-1-n)] \cos\left(\frac{\pi}{N}\left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right)$$

που είναι ο DCT-IV $N/2$ σημείων.

Επομένως ο DCT-I του f μπορεί να υπολογιστεί εφαρμόζοντας αναδρομικά την παραπάνω διάσπαση για άρτια k και περιττά k ξεχωριστά. Κατασκευάζουμε πρώτα τα σήματα $f_j(n) = f_{j-1}(n) + f_{j-1}(N/2^{j-1} - 1 - n)$ για $j = 1, \dots, \log N - 1$. Στο j -στο αναδρομικό βήμα οι άρτιοι όροι του DCT-I $N/2^j$ σημείων υπολογίζονται από τον DCT-I του f_{j+1} και οι περιττοί όροι από τον DCT-IV του f_{j+1} .

DCT–II

Η βάση DCT–I είναι τα διανύσματα $\lambda_k \cos \frac{\pi}{N} k \left(n + \frac{1}{2} \right)$, $n = 0, \dots, N-1$. Ο DCT–II προκύπτει αντιμεταθέτοντας στον DCT–I τον ρόλο των n και k .

ΑΝΑΦΟΡΕΣ

- [ATAL83] Atal B. S. "Efficient Coding of LPC Parameters by Temporal Decomposition" International Conference on Audio, Speech and Signal Processing 1983 p.81-84.
- [BD] Brockwell P.J. and Davis R. A. "Time Series: Theory and Methods" 2nd Edition, Springer 2006.
- [BENFRA] Benedeto J. and Frazier M. Eds. "Wavelets – Mathematics and Applications" CRC Press, 1994.
- [CHRIS] Christensen O. "Frames and Bases – an Introductory Course" Birkhauser 2008.
- [COH] Cohen L. "Time-Frequency Analysis", Prentice-Hall Signal Processing Series, 1995.
- [COIF92] Coifman R. and Wickerhauser M. "Entropy-based Algorithms for Best Basis Selection" IEEE Transactions on Information Theory Vol. 38, No. 2, pp. 713-718, 1992.
- [COIF92] Coifman R. R., Wickerhauser M. V. "Entropy-Based Algorithms for Best Basis Selection" IEEE Trans. on Information Theory, Vol. 38, No. 2, pp. 1992.
- [DAU92] Daubechies I. "Ten Lectures on Wavelets" SIAM 1992.
- [DAVIS] Davis S., Mermelstein P. "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences" IEEE Transactions on ASSP Vol. 28, No. 4, 1980.
- [DELLER] Deller J. et al "Discrete-time Processing of Speech Signals" Macmillan Publications NY. 1993.
- [GOL3] Golub G. H. and Van Loan C. F. "Matrix Computations" 3rd Ed., Johns Hopkins University Press 1996.
- [GROH] Grohenig K. "Foundations of Time-Frequency Analysis" Birkhauser 2001.
- [HW96] Hernandez E. and Weiss G. "A First Course on Wavelets" CRC Press 1996.
- [KATZ3] Katznelson Y. "An Introduction to Harmonic Analysis" 3rd Ed., Cambridge University Press 2004.
- [LI99] Li, B.N.L. Liu, J.N.K. "A Comparative Study of Speech Segmentation and Feature Extraction on the Recognition of Different Dialects", Proceedings of the IEEE Conference on Systems, Man and Cybernatics, 1999, Vol. 1, pp. 538-542.
- [MALLAT2] Mallat S. "A Wavelet Tour of Signal Processing" 2nd Ed. Academic Press 1999.
- [MEYER93] Meyer Y. and Ryan R. "Wavelets – Algorithms and Applications" SIAM Publications 1993.

- [PARK] Park S. and Kim N. "Automatic Speech Segmentation Based on Boundary-Type Candidate Selection", IEEE Signal Processing Letters, Vol. 13, pp. 640-643, 2006.
- [RS78] Rabiner L.R. and Schafer R.W. "Digital Processing of Speech Signals" Prentice-Hall 1978.
- [RUDINRCA] Rudin W. "Real and Complex Analysis" 3rd Ed. McGraw-Hill 1987.
- [RUDPMA] Rudin W. "Principles of Mathematical Analysis" 3rd Ed. McGraw-Hill 1976.
- [SHA49] Shannon C. E. "Communications in the Presence of Noise", Proceedings of the IRE, vol. 37, pp. 10-21, 1949.
- [SVEN] Svedsen T., Soong F. K. "On the Automatic Segmentation of Speech Signals" Trans. IC-CASP 1987, pp. 693-696.
- [VM90] Vidal E., Marzal A. "A Review and New Approaches for Automatic Segmentation of Speech Signals" in Torres et al (eds.) "Signal Processing V: Theories and Applications" Elsevier 1990.
- [ZUE] Zue V. et al. "Multilevel Acoustic Segmentation of Continuous Speech" Trans. ICASSP 1988, pp. 429-432.

////