Ανάλυση Fourier και Μουσική

Βασιλική Κούνη

Περιεχόμενα

1 Πρόλογος

Οι σειρές Fourier είναι το κλειδί στην ιδέα της αποσύνθεσης ενός σήματος[1] σε ημιτονοειδείς συνιστώσες και η χρησιμότητα της περιγραφικής τους δύναμης είναι εντυπωσιακή. Αυτές οι σειρές έχουν πολλές εφαρμογές στο φυσικό κόσμο, συμπεριλαμβανόμενης αυτής της μοντελοποίησης του ήχου. Οι καθαροί τόνοι[2] έχουν συχνότητα και πλάτος, τα οποία καθορίζουν αυτό που ονομάζουμε pitch (η ποιότητα ενός ήχου που διέπεται από το ποσοστό των δονήσεων που τον παράγουν), καθώς και τη δύναμη του ήχου, αντίστοιχα. Όλα αυτά είναι κύματα, συνεπώς μπορούν να αναπαρασταθούν από ημιτονοειδείς εξισώσεις. Οι ήχοι είναι φτιαγμένοι από καθαρούς τόνους, συνδυασμένους σε γραμμικούς συνδυασμούς προκειμένου να δημιουργήσουν πιο περίπλοκους ήχους, όπως οι συγχορδίες.

Οι παλλόμενες χορδές και τα μεταξύ τους κενά στα όργανα, υπακούουν στην κυματική εξίσωση. Η κυματική εξίσωση, όπως ανακαλύφθηκε από τον D' Alembert $(u(x,t) = \frac{1}{2}(f(x+t) + f(x-t)) + \frac{1}{2}\int_{x-t}^{x+t} g(y)dy)$ είναι μια διαφορική εξίσωση που εξετάζει τη συμπεριφορά ενός σημείου της χορδής βάσει αρχικών συνθηκών. Όπως υπέθεσε ο Bernoulli, οι σειρές Fourier αναπαριστούν μια λύση της κυματικής εξίσωσης. Αυτό σημαίνει ότι οι σειρές αυτές μπορούν να χρησιμοποιηθούν για να μοντελοποιηθούν τα ηχητικά κύματα που παράγονται από τις παλλόμενες χορδές και τα μεταξύ τους κενά.

1.1 Θεμελιώδεις και αρμονικές συχνότητες

Όπως προαναφέραμε, το βολικό σχετικά με τις σειρές Fourier είναι ότι μπορούν να «σπάσουν σε κομμάτια» τον ήχο μέσω τριγωνομετρικών συναρτήσεων με συχνότητες και πλάτη. Αυτές οι συναρτήσεις είναι οι θεμελιώδεις και οι αρμονικές τους. Οι θεμελιώδεις συχνότητες μπορούν να αναπαρασταθούν σαν τον 1ο όρο μετά τη σταθερά a_0 στη λύση μέσω σειρών Fourier για την κυματική εξίσωση, ο οποίος συνιστά έναν όρο με ημίτονο και συνημίτονο. Μάλιστα, ο ένας από τους 2 όρους μπορεί να είναι μηδέν. Ο 1ος μη-σταθερός όρος που αναπαριστά τη θεμελιώδη είναι ο $a_1 \cos x + b_1 \sin x$. Παρατηρήστε ότι αν ο ήχος είναι καθαρός τόνος, οι σειρές Fourier είναι ένα κύμα με μία συχνότητα, οπότε η θεμελιώδης θα είναι η μόνη μη-σταθερά. Η 1η αρμονική μπορεί να αναπαρασταθεί από το 2ο όρο στη σειρά Fourier: $a_2 \cos 2x + b_2 \sin 2x$ κοκ, μέχρι το n-οστό όρο που είναι ο $a_n \cos nx + b_n \sin nx$. Οι συντελεστές των αρμονικών δίνουν το πλάτος κάθε αρμονικής και προσδιορίζουν την τονική ποιότητα, ή αλλιώς το τέμπο. Εν κατακλείδι, η συχνότητα του θεμελιώδους τόνου προσδιορίζει το pitch που ακούμε, και η αρμονική συνεισφορά, όπως φαίνεται από τις σειρές Fourier, προσδιορίζει το τέμπο.

2 Η κυματική εξίσωση

2.1 Εισαγωγικά

Θέλουμε λοιπόν να ερευνήσουμε πώς μετατρέπονται τα μουσικά σήματα σε συχνότητες και πλάτη. Αυτό επιτυγχάνεται με αυτό που αποκαλούμε μετασχηματισμό Fourier. Πιο συγκεκριμένα, η ανάπτυξή του οφείλεται στην ανάγκη επίλυσης του προβλήματος της κίνησης μιας απομονωμένης χορδής της οποίας τα άκρα είναι στερεωμένα σε 2 σημεία, πρόβλημα που σχετίζεται άμεσα με τις παλλόμενες

[1] Βλ. παράρτημα

[2] Καθαρά και απλά ηχητικά κύματα, που μπορούν να μοντελοποιηθούν από μια τριγωνομετρική συνάρτηση

χορδές στα όργανα. Η παλλόμενη χορδή θα χρησιμοποιηθεί ως ένα περιοδικό σύστημα μέσω του οποίου θα συνδεθούν οι 2 αναπαραστάσεις των συναρτήσεων: η αναπαράσταση στον τομέα του χρόνου και στον τομέα της συχνότητας.

2.2 Περιγραφή της κυματικής εξίσωσης

Η κίνηση μιας δονούμενης χορδής ικανοποιεί την εξίσωση

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} \tag{1}$$

την οποία καλούμε «η μονοδιάστατη κυματική εξίσωση». Μια φυσική γενίκευσή της σε d μεταβλητές είναι η

$$\frac{\partial^2 u}{\partial {x_1}^2} + \cdots + \frac{\partial^2 u}{\partial {x_d}^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} \tag{2}$$

Για την ακρίβεια, είναι γνωστό ότι στην περίπτωση $d = 3$ αυτή η εξίσωση προσδιορίζει τη συμπεριφορά ηλεκτρομαγνητικών κυμάτων στο κενό (με το c να αντιπροσωπεύει την ταχύτητα του φωτός). Επιπλέον, αυτή η εξίσωση περιγράφει τη διάδοση των ηχητικών κυμάτων, γι' αυτό την αποκαλούμε «η d-διάστατη κυματική εξίσωση». Η κυματική εξίσωση επιδέχεται δύο μεθόδους επίλυσης: είτε μέσω οδευόντων κυμάτων είτε μέσω της υπέρθεσης στάσιμων κυμάτων. Η πρώτη, αν και απλούστερη και κομψότερη, δε δίνει άμεσα πλήρη εικόνα του προβλήματος, γι' αυτό θα κάνουμε μια αναφορά στη δεύτερη μέθοδο.

Θεωρούμε το πρόβλημα αρχικών και συνοριακών συνθηκών για την κυματική εξίσωση με ομογενείς συνοριακές συνθήκες Dirichlet σε ένα φραγμένο διάστημα:

$$u_{tt} = c^2 u_{xx}, \qquad 0 < x < \ell, \quad t > 0$$

$$\begin{aligned} &u(0,t) = u(\ell,t) = 0, \quad t \geqslant 0 \\ &u(x,0) = \phi(x), \quad 0 \leqslant x \leqslant \ell \\ &u_t(x,0) = \psi(x), \quad 0 \leqslant x \leqslant \ell. \end{aligned}$$

Χωρίζουμε τη ζητούμενη $u(x,t)$ σε 2 συναρτήσεις μιας μεταβλητής, δηλαδή θέτουμε $u(x,t) = f(x)g(t)$[3]. Παραγωγίζουμε ως προς τη x και μετά ως προς την t μεταβλητή, αντικαθιστούμε στην (1) και διαιρούμε με το γινόμενο $c^2 f(x)g(t)$. Θέτουμε το κάθε μέλος ίσο με μια σταθερά $-\lambda$ και λύνουμε το x και το t πρόβλημα, αντίστοιχα. Το x πρόβλημα (γνωστό και ως πρόβλημα ιδιοτιμών Sturm-Liouville) έχει λύσεις της μορφής $F_n(x) = \sin\frac{n\pi x}{\ell}$ και ιδιοτιμές τις $\lambda_n = (\frac{n\pi}{\ell})^2$. Το t πρόβλημα έχει λύσεις της μορφής $G_n(t) = a_n \cos\frac{n\pi ct}{\ell} + b_n \sin\frac{n\pi ct}{\ell}$. Από τα παραπάνω έχουμε ένα άπειρο πλήθος λύσεων της μορφής

$$u_n(x,t) = \left(a_n \cos\frac{n\pi ct}{\ell} + b_n \sin\frac{n\pi ct}{\ell}\right)\sin\frac{n\pi x}{\ell}, \qquad n \in \mathbb{N},$$

με αυθαίρετες σταθερές a_n και b_n. Χρησιμοποιώντας την αρχή της υπέρθεσης (λόγω γραμμικότητας της εξίσωσης) και τις αρχικές συνθήκες καταλήγουμε στην

$$u(x,t) = \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi ct}{\ell} - \frac{1}{c} b_n \sin\frac{n\pi ct}{\ell}\right)\sin\frac{n\pi x}{\ell}. \tag{3}$$

[3] Η μέθοδος είναι γνωστή ως «χωρισμός μεταβλητών»

2.3 Κίνηση της χορδής

Μπορούμε να περιγράψουμε την κίνηση μιας χορδής με το να εστιάσουμε στις θέσεις που βρίσκονται κάθε φορά τα σημεία της σχετικά με το χρόνο. Μοντελοποιείται δε, με τη γνωστή διαφορική εξίσωση, την κυματική εξίσωση, δεδομένων κάποιων αρχικών συνθηκών. Η λύση της είναι μια συνάρτηση του χρόνου $f(t)$, γνωστή ως η χρονική αναπαράσταση της κίνησης της χορδής. Έστω τώρα ότι τοποθετούμε τη χορδή με τρόπο ώστε να δονείται μόνο στη θεμελιώδη αρμονική. Ως γνωστόν, αυτό το μοτίβο αναπαρίσταται στον τομέα του χρόνου από μία μόνο ημιτονοειδή συνάρτηση με συχνότητα -ας πούμε- v_0. Συνεπώς, η αναπαράσταση του τομέα της συχνότητας $F(v)$ αυτής της χορδής έχει ένα μόνο άλμα στο $v = v_0$ με το ύψος του άλματος ίσο με το πλάτος του κύματος. Αυτό το παράδειγμα βοηθάει στην οπτικοποίηση του τομέα της συχνότητας, αλλά στην πραγματικότητα υπάρχουν περισσότερες από μία συχνότητες. Κατασκευάζουμε λοιπόν τη χρονική αναπαράσταση μέσω άπειρων σειρών αυτών των αρμονικών φτιαγμένων έτσι ώστε να αναπαριστούν την κίνηση της χορδής. Αυτές οι σειρές είναι οι σειρές Fourier και σε αυτές στηρίζεται ο μετασχηματισμός Fourier. Μέσω του μετασχηματισμού Fourier μπορούμε να βρούμε την αναπαράσταση της συχνότητας μιας χορδής από μια συνάρτηση χρόνου και επειδή είναι αντιστρέψιμος, ο αντίστροφος επιστρέφει μια συνάρτηση χρόνου, δεδομένης της συνάρτησης συχνότητας.

Τα μουσικά σήματα ηχογραφούνται και ακούγονται στον χρονικό τομέα, αφού περιέχουν την ένταση του ήχου σαν συνάρτηση του χρόνου. Εντούτοις, τα pitches και κατ¨ επέκταση οι συγχορδίες αναπαρίστανται από συχνότητες. Ο μετασχηματισμός Fourier χρησιμοποιείται για να μετασχηματίσει το εισερχόμενο - συναρτήσει χρόνου - σήμα σε μια αναπαράσταση συχνότητας που μπορεί να αναλυθεί για τις εντάσεις συγκεκριμένων pitches. Η σχέση μεταξύ των pitches σε δεδομένο χρόνο αναλύεται μετά, ώστε να προσδιοριστεί η συγχορδία που παίζεται. Πρέπει να επισημανθεί ότι ο μετασχηματισμός Fourier πραγματικών συναρτήσεων μπορεί να αναπαριστά τη συχνότητα σε μιγαδική μορφή.

3 Ο συνεχής μετασχηματισμός Fourier

Αρχικά ορίζουμε ως ω_k τη γωνιακή συχνότητα μέσω της σχέσης της με τη συνηθισμένη συχνότητα v ως εξής: $\omega_k = 2\pi k v$. Τότε η σχέση μεταξύ της συνάρτησης χρόνου f και της αντίστοιχης συνάρτησης συχνότητας F ορίζεται ως ο μετασχηματισμός Fourier:

$$F(\omega_k) = \int_{-\infty}^{\infty} f(t) e^{-2i\pi k t} dt, \qquad k \in (-\infty, \infty). \tag{4}$$

Προσέξτε την παρουσία των ημιτονοειδών συντελεστών μέσα στη μιγαδική εκθετική $e^{i\omega t} = \cos \omega t + i \sin \omega t$. Από την (4) προκύπτει ο αντίστροφος μετασχηματισμός Fourier :

$$f(t) = \int_{-\infty}^{\infty} F(\omega_k) e^{2i\pi k t} dt, \qquad k \in (-\infty, \infty). \tag{5}$$

Πρακτικά, ψηφιακές εφαρμογές απαιτούν το διακριτό μετασχηματισμό Fourier,γνωστό ως Discrete Fourier Transform (DFT).

4 Ο διακριτός μετασχηματισμός Fourier

Ορμώμενοι από τη συνεχή περίπτωση, ορίζουμε το διακριτό μετασχηματισμό Fourier για μια διανυσματική συνάρτηση $f \in \mathbb{C}^N$ ως:

$$F_k \equiv \sum_{n=\frac{-N}{2}+1}^{\frac{N}{2}} f_n e^{\frac{-i2\pi nk}{N}}, \qquad k = 0, 1, \ldots, N-1 \tag{6}$$

όπου ως f_n ορίζουμε τη n-οστή είσοδο της διανυσματικής συνάρτησης f. Όμοια, ο αντίστροφος DFT έχει την παρακάτω μορφή:

$$f_k \equiv \sum_{n=\frac{-N}{2}+1}^{\frac{N}{2}} F_n e^{\frac{-i2\pi nk}{N}}, \qquad k = 0, 1, \ldots, N-1. \tag{7}$$

Αν πάλι υποθέσουμε ότι έχουμε N σημεία σε ένα χρονικό διάστημα μήκους L ξέρουμε[4] ότι το βήμα της απόστασης στον τομέα της συχνότητας πρέπει να είναι $v_k = kDv = \frac{k}{L}$, για $k = -\frac{L}{2}+1, \ldots, \frac{L}{2}$ και εφαρμόζοντάς το στην (6) έχουμε:

$$F_k \equiv \sum_{n=\frac{-N}{2}+1}^{\frac{N}{2}} f_n e^{\frac{-i2\pi nk}{N}} = \sum_{n=\frac{-N}{2}+1}^{\frac{N}{2}} f_n e^{-i2\pi v_k x_n} = F(v). \tag{8}$$

5 Φασματόγραμμα και χρωμόγραμμα

Κοιτώντας στον FFT τα επικαλυπτόμενα χρονικά τμήματα σε ένα ηχητικό αρχείο, μπορούμε να κατασκευάσουμε το Σύντομο Μετασχηματισμό Fourier (Short-Time Fourier Transform-STFT). Κάθε ένας από τους FFTs στον STFT αποκαλύπτει την αναπαράσταση του συχνοτικού τομέα ή το φάσμα ενός μικρού χρονικού διαστήματος του εισερχόμενου σήματος. Συνδυάζοντας όλα αυτά τα χρονικά τμήματα με την κατάλληλη σύνθεση, μπορούμε να κατασκευάσουμε δεδομένα των εντάσεων των συχνοτήτων με την πάροδο του χρόνου. Ο STFT μας επιτρέπει να προσθέσουμε μια χρονική διάσταση στον FFT , το οποίο μας βοηθάει να παρατηρήσουμε πώς ο τομέας της συχνότητας αλλάζει με το χρόνο. Αφού οι συχνότητες είναι αντιπροσωπευτικές των pitches , μπορούμε να χρησιμοποιήσουμε τα φασματογράμματα για να προσδιορίσουμε το αναπαραγόμενο χρώμα του σήματος σε μια χρονική στιγμή. Η δημιουργία αυτών είναι καθοριστική στο να προσδιορίσουμε πώς οι συγχορδίες αλλάζουν τη μουσική. Επιπλέον, η πληροφορία για το σε ποια οκτάβα ανήκει ο ήχος είναι άσχετη. Προσδιορίζουμε την χρωματική ένταση μέσω της συλλογής όλων των εντάσεων μιας νότας ασχέτως της οκτάβας της. Ο αλγόριθμος για τη συλλογή αυτών των χρωμάτων αναφέρεται ως «χρωμόγραμμα».

5.1 Ερμηνεία του χρωμογράμματος μέσω πινάκων

Αφού ο DFT είναι το άθροισμα πινακοποιημένων τιμών, μπορούμε να εκφράσουμε την (8) σε μορφή πίνακα ως γραμμική εξίσωση: $F = Wf$, όπου f είναι το διάνυσμα στον τομέα του χρόνου, F είναι η

[4] Θεωρούμε δεδομένους τους ορισμούς μιας δέλτα συνάρτησης, το μετασχηματισμό Fourier της, καθώς και τη χρήση μιας ακολουθίας αλμάτων στο DFT της

έξοδος στον τομέα της συχνότητας, και W είναι ο αντιστρέψιμος πίνακας

$$W = \begin{pmatrix} e^{\frac{-i2\pi(0)}{N}} & e^{\frac{-i2\pi(0)}{N}} & \dots & e^{\frac{-i2\pi(0)}{N}} \\ e^{\frac{-i2\pi(0)}{N}} & e^{\frac{-i2\pi(1)}{N}} & \dots & e^{\frac{-i2\pi(N-1)}{N}} \\ \vdots & \vdots & \dots & \vdots \\ e^{\frac{-i2\pi(0)}{N}} & e^{\frac{-i2\pi(N-1)}{N}} & \dots & e^{\frac{-i2\pi(N-1)(N-1)}{N}} \\ \end{pmatrix}.$$

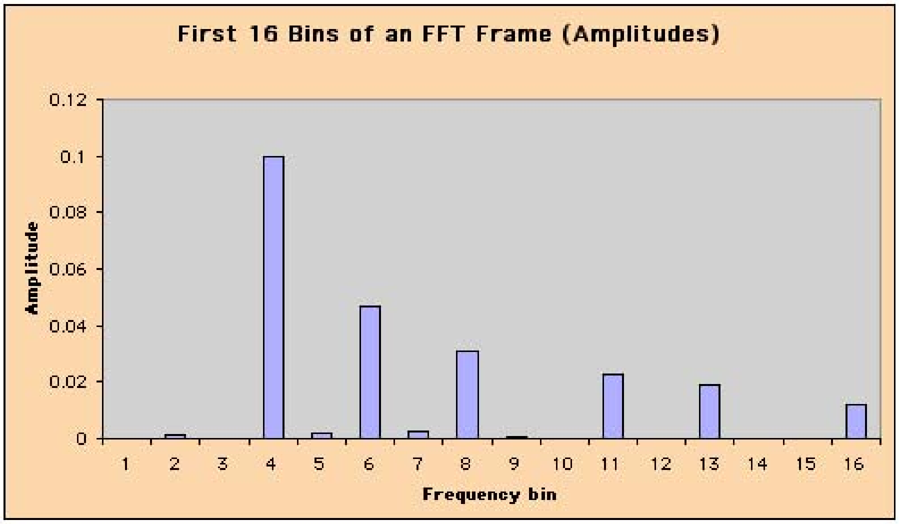
Γνωρίζουμε ότι ο W είναι αντιστρέψιμος αφού οι στήλες είναι ορθογώνιες. Αφού ο W είναι αντιστρέψιμος, μπορούμε να εκφράσουμε τον αντίστροφο μετασχηματισμό Fourier ωσ: $f = W^{-1}F$. Βάσει αυτής της μεθόδου πολλαπλασιασμού πινάκων, μπορούμε να υπολογίσουμε τον DFT σε $O(n^2)$ χρόνο όπου n είναι το μήκος του διανύσματος εισόδου. Εντούτοις, με τη δειγματοποιημένη (sampled) μουσική, τα δεδομένα εισόδου είναι ιδιαίτερα υψηλών διαστάσεων, οπότε και θα θέλαμε να βρούμε μια μέθοδο που να υπολογίζει τον DFT πολύ πιο γρήγορα.

6 Ο ταχύς μετασχηματισμός Fourier

6.1 Εισαγωγικά

Το 1965 οι Cooley & Tukey δημοσίευσαν έναν αλγόριθμο που θεμελιωδώς άλλαξε το τοπίο της διαδικασίας ψηφιοποίησης σημάτων. Αξιοποιώντας τις συμμετρίες του DFT, , κατάφεραν να ελαττώσουν το χρόνο τρεξίματος των DFT από $O(n^2)$ σε $O(n \log n)$. Αυτός ο αλγόριθμος είναι ο πρώτος ταχύς μετασχηματισμός Fourier ή αλλιώς FFT(Fast Fourier Transform) και ονομάστηκε έτσι λόγω της αύξησης της υπολογιστικής ταχύτητας (βλ.παράδειγμα - Ενότητα 6.3), ενώ υπάρχει και ο αντίστροφος μετασχηματισμός Fourier, ο IFFT (Inverse Fast Fourier Transform). Οι DFT,FFT, IFFT γίνανε πολύ σημαντικοί σε πολλούς κλάδους, όταν οι μηχανικοί αντελήφθησαν πώς να παίρνουν δείγματα αρκετά γρήγορα για να παράγουν αρκετά δεδομένα ώστε να αναδημιουργήσουν τον ήχο και άλλα αναλογικά φαινόμενα ψηφιακώς (φωτογραφίες, ραδιοκύματα, σεισμογραφικά δεδομένα κλπ).

Ένας FFT ενός σήματος του χρονικού τομέα παίρνει τα δείγματα και μας δίνει ένα νέο σύνολο αριθμών που αναπαριστούν τις συχνότητες, τα πλάτη και τις φάσεις των ημιτονοειδών κυμάτων που συνθέτουν τον ήχο που έχουμε αναλύσει.



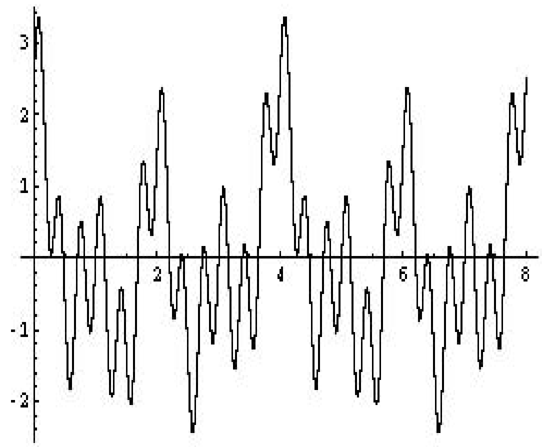
Σχήμα 1: Ζώνες Συχνοτήτων

Το σχήμα 1 μας δείχνει τις πρώτες 16 συχνοτικές ζώνες μιας τυπικής FFT ανάλυσης, αφού έχει γίνει η μετατροπή από πραγματικούς και φανταστικούς αριθμούς σε ζευγάρια πλάτους - φάσης.

6.2 Η εφαρμογή του FFT

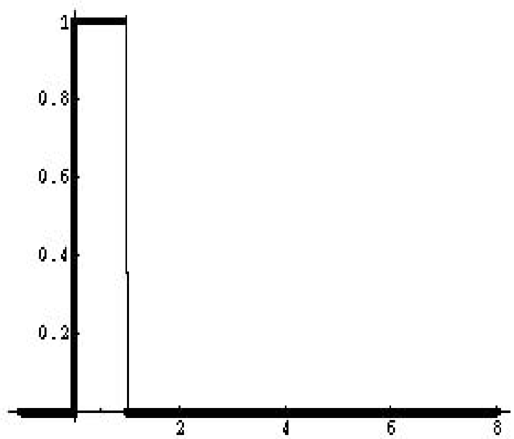
Ο τρόπος που δουλεύει ο FFT είναι αρκετά απλός. Ουσιαστικά παίρνει ένα μεγάλο κομμάτι χρόνου που ονομάζεται «πλαίσιο» ή αλλιώς frame (ένας συγκεκριμένος αριθμός δειγμάτων) και θεωρεί αυτό το κομμάτι σαν μια μοναδική περίοδο μιας επαναλαμβανόμενης κυματομορφής. Ο λόγος για τον οποίο δουλεύει κάτι τέτοιο είναι ότι οι περισσότεροι ήχοι είναι «τοπικά στάσιμοι», εννοώντας ότι πάνω από κάθε σύντομη χρονική περίοδο, ο ήχος μοιάζει πραγματικά σαν μια κανονική επαναλαμβανόμενη συνάρτηση. Παραθέτουμε στη συνέχεια μια σύντομη διαδικασία - 3 βημάτων - για την επεξεργασία των ψηφιακών ήχων.

Θεωρούμε μια συνάρτηση (μπορεί να είναι και περιοδική, όχι όμως υποχρεωτικά) έστω $f(t)$, με το παρακάτω γράφημα:



Σχήμα 2: Γράφημα της $f(t)$

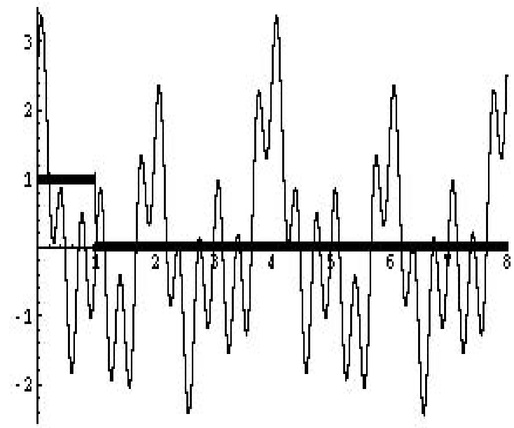
Ας υποθέσουμε ότι ενδιαφερόμαστε μόνο για το τμήμα του γραφήματος μεταξύ $0 \leqslant t \leqslant 1$. Ακολουθεί ένα γράφημα της συνάρτησης εκλέπτυνσης[5], έστω $\omega(t)$, που χρειαζόμαστε να χρησιμοποιήσουμε. Παρατηρήστε ότι η $\omega(t)$ είναι ίση με 1 μόνο στο διάστημα $0 \leqslant t \leqslant 1$ και μηδέν σε κάθε άλλο διάστημα.



Σχήμα 3: Γράφημα της $\omega(t)$

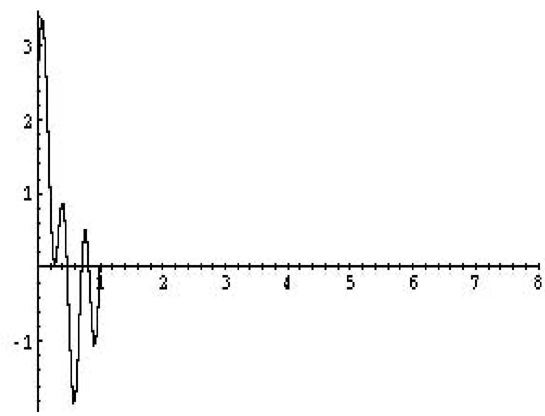
Στο Βήμα 1 χρειάζεται να «εκλεπτύνουμε» τη συνάρτηση. Στο προηγούμενο σχήμα έχουμε σχεδιάσει αμφότερες τις συναρτήσεις εκλέπτυνσης $\omega(t)$ (που είναι μη μηδενική στην περιοχή που μας ενδιαφέρει και σχεδιασμένη με έντονη μαύρη γραμμή) και $f(t)$ στο ίδιο σχήμα. Στο επόμενο σχήμα έχουμε σχεδιάσει τη $f(t)\omega(t)$, που είναι η περιοδική συνάρτηση πολλαπλασιασμένη επί τη συνάρτηση εκλέπτυνσης. Από αυτή την εικόνα, είναι φανερό ποιο τμήμα της $f(t)$ μας ενδιαφέρει,

[5]Μια συνάρτηση εκλέπτυνσης (window function)είναι μια συνάρτηση που έχει μηδενική τιμή εκτός ενός επιλεγμένου διαστήματος



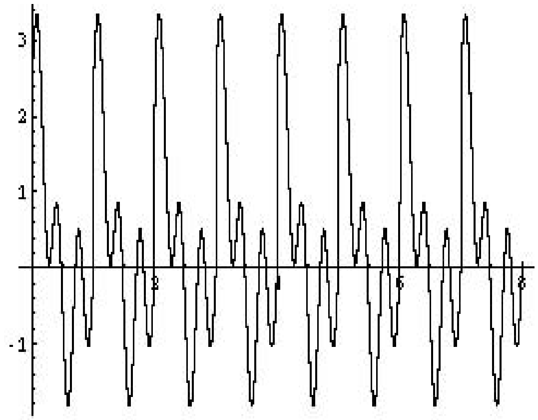
Σχήμα 4: Γράφημα της $f(t) * \omega(t)$

το οποίο φαίνεται στο σχήμα 5.



Σχήμα 5: Περιορισμός της $f(t)$ στο $[0, 1]$

Στο Βήμα 2, πρέπει να επεκτείνουμε περιοδικά την $f(t)\omega(t)$ σε όλο τον t-άξονα.



Σχήμα 6: Επέκταση της $f(t)\omega(t)$ στον t-άξονα

Τώρα λοιπόν, έχουμε μια περιοδική συνάρτηση και, σύμφωνα με το θεώρημα του Fourier , μπορούμε να την αναπαραστήσουμε σαν άθροισμα ημιτόνων και συνημιτόνων. Αυτό είναι το Βήμα 3, όπου χρησιμοποιώντας τον FFT , θα βρούμε τις ημιτονοειδείς συνιστώσες αυτής της κυματομορφής. Στη θεωρία, ο αριθμός ημιτονοειδών συνιστωσών είναι άπειρος, αφού δεν υπάρχει όριο στο πόσες συχνοτικές συνιστώσες μπορεί να έχει ένας ήχος. Πρακτικά ωστόσο, χρειάζεται να περιοριστούμε σε μερικούς προκαθορισμένους αριθμούς. Αυτό το όριο έχει μια σοβαρή επίδραση στην ακρίβεια της ανάλυσής μας.

Αυτό που κάνουμε λοιπόν είναι το εξής: αντί να ψάχνουμε για το περιεχόμενο της συχνότητας του ήχου σε όλες τις πιθανές συχνότητες (που είναι απείρως μεγάλος αριθμός - 100.000000001 Hz, 100.000000002 Hz κλπ), κατανέμουμε το φάσμα συχνοτήτων σε έναν αριθμό ζωνών συχνοτήτων που ονομάζουμε bins. Το μέγεθος αυτών των bins είναι καθορισμένο από τον αριθμό των δειγμάτων στο «πλαίσιο» (frame) της ανάλυσής μας (το χρονικό κομμάτι που αναφέραμε πιο πάνω). Το πλήθος των bins δίνεται από τον παρακάτω τύπο:

(9) $$q_{bins} = \frac{s_{frame}}{2}$$

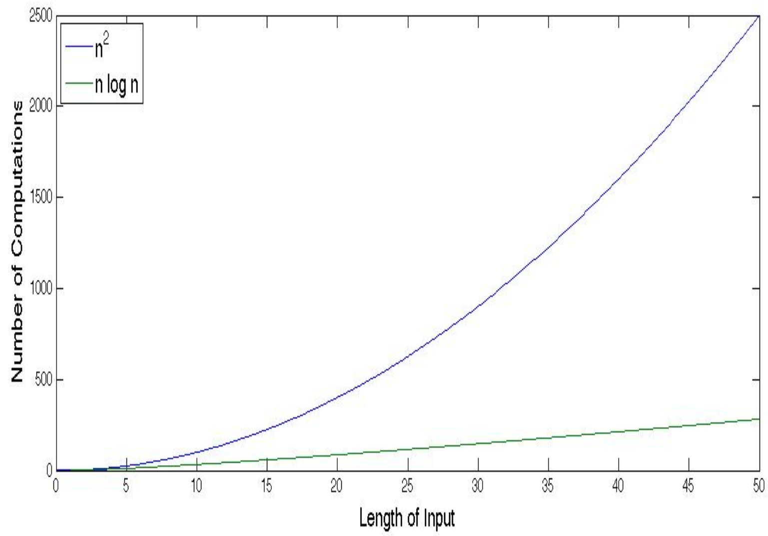
όπου q_{bins} το πλήθος των bins και s_{frame} το μέγεθος του frame .

6.3 Παράδειγμα

Όπως μπορούμε να δούμε στην παρακάτω εικόνα, η ελάττωση στο χρόνο επεξεργασίας είναι ιδιαίτερα σημαντική ακόμα και για ένα διάνυσμα εισόδου μήκους 50. Στο παράδειγμά μας, χρησιμοποιούμε ήχο που έχει δειγματικές συχνότητες της τάξεως των 11025 Hz. Κατά συνέπεια, σε ένα τραγούδι τριών λεπτών, υπάρχουν περίπου 2.000.000 σημεία εισόδου. Σε αυτή την περίπτωση ο $O(n \log n)$ FFT αλγόριθμος, μας παρέχει μια αναπαράσταση της συχνότητας των δεδομένων μασ:

(10) $$\frac{n^2}{n \log_2 n} = \frac{(2 * 10^6)^2}{(2 * 10^6) \log_2 (2 * 10^6)} \approx 100.000$$

φορές γρηγορότερα.



Σχήμα 7: Γρήγορη σύγκριση των $O(n)$ και $O(n \log n)$ ταχυτήτων επεξεργασίας

7 Επίλογος

Σε αυτή την εργασία κάναμε μια πρώτη εισαγωγή στη σχέση μουσικής και μαθηματικών και την περιορίσαμε στον τομέα της Ανάλυσης Fourier . Επιπρόσθετα, αποκτήσαμε μια εκτίμηση για τις εφαρμογές της στην κυματική εξίσωση και την επεξεργασία σήματος. Τελικά, δείξαμε μέσω παραδειγμάτων, πώς μπορεί να χρησιμοποιηθεί ο FFT για τη μετατροπή ενός σήματος από τον τομέα του χρόνου στον τομέα της συχνότητας.

8 Παράρτημα

8.1 Ημιτονοειδής συνάρτηση

Ορίζουμε ως ημιτονοειδές μια συνάρτηση της μορφής: $x(t) = A\sin(2\pi v t + \phi)$. Όταν μιλάμε για ηχητικά σήματα, ορίζουμε ως:

- $A =$ το πλάτος κάθε σήματος
- $v =$ την ακτινική συχνότητα σε rad /sec

- $2\pi v =$ τη συχνότητα σε Hz
- $t =$ το χρόνο σε δευτερόλεπτα (sec)
- $\phi =$ την αρχική φάση σε ακτίνια (radians)
- $2\pi vt + \phi =$ τη στιγμιαία φάση σε ακτίνια (radians)

Οι μετασχηματισμοί Fourier χτίζονται πάνω σε μιγαδικές ιδιότητες των ημιτονοειδών που προκύπτουν από τις ταυτότητες του Euler:

- $e^{i\theta} = \cos\theta + i\sin\theta$
- $e^{\pm i2\pi vx} = \cos 2\pi vx \pm i\sin 2\pi vx$

με την τελευταία να είναι η πιο συχνά χρησιμοποιούμενη στη διαδικασία επεξεργασίας του ηχητικού σήματος.