



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ  
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

---

## Ψηφιακή Επεξεργασία Φωνής

Άσκηση 3η

Στυλιανού Ιωάννης

Τμήμα Επιστήμης Υπολογιστών

---

# ΗΥ578: 3<sup>η</sup> Σειρά Ασκήσεων

1. Κατά την παραγωγή ηχηρών (voiced) ήχων, η ταχύτητα των μορίων του αέρα στη γλωττίδα (glottal airflow velocity) σε μια περίοδο μπορεί να περιγραφεί από την εξίσωση:

$$g[n] = (a^{-n}u[-n]) \star (b^{-n}u[-n]) \quad 0 < a < 1, \quad 0 < b < 1$$

όπου  $\star$  δηλώνει την πράξη της συνέλιξης και  $a, b$  είναι πραγματικοί αριθμοί.

- (α') Υπολογίστε και σχεδιάστε το αποτέλεσμα της συνέλιξης. Μοιάζει το αποτέλεσμα με το ιδανικό σήμα της ταχύτητας των μορίων του αέρα στη γλωττίδα;
- (β') Υπολογίστε τον μετ.Ζ,  $G(z)$  και σχεδιάστε στο  $z$ -επίπεδο τους πόλους και μηδενικά της  $G(z)$ . Ποιο είναι το πεδίο σύγκλισης;
- (γ') Θεωρώντας ότι για  $-\infty < n < \infty$ , η ταχύτητα του αέρα στη γλωττίδα περιγράφεται από την εξίσωση:

$$u[n] = g[n] \star \delta_P[n]$$

όπου

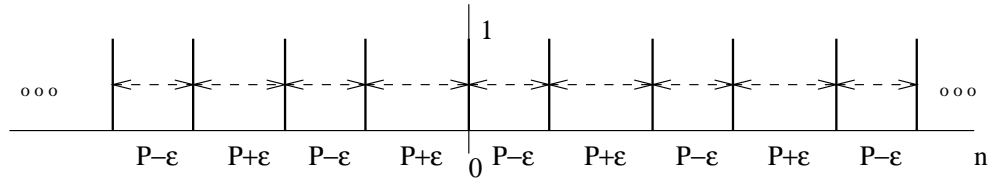
$$\delta_P[n] = \sum_{k=-\infty}^{\infty} \delta[n - kP]$$

και ότι το σήμα  $u[n]$  το παρατηρούμε μ'ένα τετραγωνικό παράθυρο με κέντρο στη χρονική στιγμή  $n = \tau$ :

$$u[n, \tau] = w[n, \tau] (g[n] \star \delta_P[n])$$

να υπολογίσετε τον μετ. Fourier  $U(\omega, \tau)$  του σήματος  $u[n, \tau]$ .

- (δ') Σχεδιάστε χονδρικά το φάσμα πλάτους  $U(\omega, \tau)$  για  $0 \leq \omega \leq \pi$  για μήκους παραθύρου  $P$  και  $2P$ . Ποιο από τα δύο παράθυρα θα χρησιμοποιούσατε για στενής ζώνης φασματογραφική ανάλυση (narrowband spectrogram) και ποιο για ευρείας ζώνης φασματογραφική ανάλυση (wideband spectrogram);
2. Μελέτη του φαινομένου jitter: Θεωρούμε έναν ομιλητή ο οποίος παράγει έναν ηχηρό (έμφωνο) ήχο αλλά δεν μπορεί να διατηρήσει σταθερή τη θεμελιώδη περίοδο. Συγκεκριμένα θεωρούμε ότι η θεμελιώδης περίοδο,  $P$ , εναλλάσσεται μεταξύ των τιμών  $P - \epsilon$  και  $P + \epsilon$  όπως φαίνεται στο Σχήμα 1



Σχήμα 1: Μελέτη φαινομένου jitter.

(α') Δείξτε ότι η σειρά των μοναδιαίων παλμών που φαίνεται στο Σχήμα 1 μπορεί να περιγραφεί από την εξίσωση:

$$p[n] = \sum_{k=-\infty}^{\infty} \delta[n - 2kP] + \sum_{k=-\infty}^{\infty} \delta[n - P + \epsilon - 2kP]$$

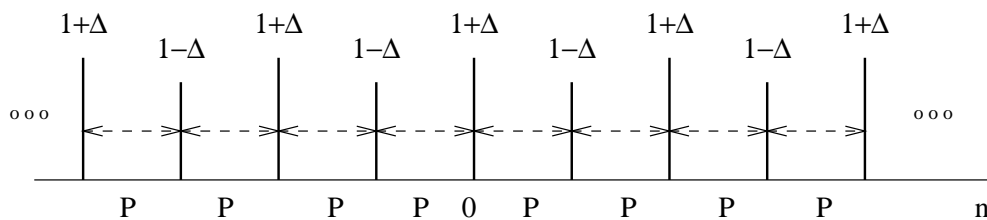
(β') Δείξτε επίσης ότι:

$$|P(\omega)|^2 = \left(\frac{2\pi}{P}\right)^2 \sum_{k=-\infty}^{\infty} \cos^2 \left[ \frac{\pi}{P} \left( \frac{P - \epsilon}{2} \right) k \right] \delta\left(\omega - \frac{\pi}{P}k\right)$$

(γ') Σχεδιάστε (χονδρικά) το φάσμα ισχύος  $|P(\omega)|^2$  για  $\epsilon = 0$ ,  $0 < \epsilon \ll P$ , και  $\epsilon = 1$  θεωρώντας ότι παρατηρούμε τους παλμούς μέσα από ένα παράθυρο μήκους  $4P$ .

(δ') Σε ποια περίπτωση από τις παραπάνω θα έχουμε το πρόβλημα να εμφανιστούν 'ψεύτικα formants' και γιατί;

3. Μελέτη του φαινομένου shimmer: Θεωρούμε έναν ομιλητή ο οποίος παράγει έναν ηχηρό (έμφωνο) ήχο και ότι οι παλμοί στη γλωττίδα είναι ιδανικοί, απέχουν σταθερή περίοδο  $P$  αλλά έχουν πλάτος που εναλλάσσεται μεταξύ των τιμών  $1 + \Delta$  και  $1 - \Delta$ , όπως φαίνεται στο Σχήμα 2



Σχήμα 2: Μελέτη φαινομένου shimmer.

(α') Δείξτε ότι η σειρά των μοναδιαίων παλμών που φαίνεται στο Σχήμα 2 μπορεί να περιγραφεί από την εξίσωση:

$$p[n] = (1 + \Delta) \sum_{k=-\infty}^{\infty} \delta[n - 2kP] + (1 - \Delta) \sum_{k=-\infty}^{\infty} \delta[n - P - 2kP]$$

(β') Δείξτε επίσης ότι:

$$|P(\omega)|^2 = \left(\frac{2\pi}{P}\right)^2 \left[ \Delta^2 \sum_{k=-\infty}^{\infty} \delta(\omega - (2k+1)\frac{\pi}{P}) \sum_{k=-\infty}^{\infty} \delta(\omega - 2k\frac{\pi}{P}) \right]$$

(γ') Σχεδιάστε (χονδρικά) το φάσμα ισχύος  $|P(\omega)|^2$  για  $\Delta = 0$ ,  $0 < \Delta \ll 1$  και  $\Delta = 1$  θεωρώντας ότι παρατηρούμε το  $p[n]$  μέσα από ένα παράθυρο μήκους  $4P$ .

(δ') Σε ποια περίπτωση από τις παραπάνω θα έχουμε το πρόβλημα να εμφανιστούν 'ψεύτικα formants' και γιατί;

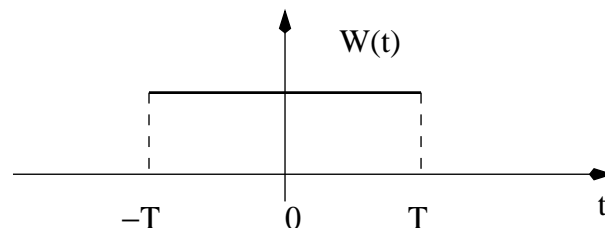
4. Θεωρούμε ότι στην ελληνική γλώσσα έχουμε 60 φωνήματα (όπου θεωρούμε και κάποιους ιδιοματισμούς της γλώσσας μας) και ότι επίσης προφέρουμε κατά μέσο όρο 4, 5 συλλαβές ανά δευτερόλεπτο. Κάθε συλλαβή θεωρούμε ότι περιέχει κατά μέσο όρο 2 φωνήματα.

(α') Ποιος είναι ο ρυθμός πληροφορίας (bit rate που μετριέται σε bits/sec) που μεταδίδουμε καθώς μιλάμε αν θεωρήσουμε απλή ακολουθία φωνημάτων;

(β') Αν δειγματοληπτήσουμε το σήμα της φωνής με συχνότητα δειγματοληψίας  $f_s = 8000 \text{ Hz}$  και κάθε δείγμα αναπαρασταθεί με ακρίβεια 16 bits ποιο θα είναι το bit rate αν θα θέλαμε να μεταδώσουμε κάθε δείγμα;

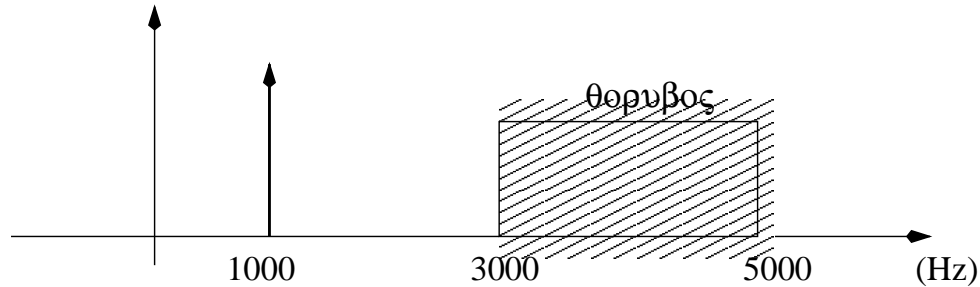
(γ') Χρησιμοποιώντας μαθηματικά μοντέλα για την παραγωγή της φωνής, για την ίδια συχνότητα δειγματοληψίας μπορούμε να αναπαραστήσουμε κάθε 160 δείγματα της φωνής με 46 bits με σχετικά καλή ποιότητα αναπαραγωγής της φωνής. Ποιο είναι το bit rate των μοντέλων αυτών και πόσο συμπίεση προσφέρουν σχετικά με την παραπάνω MH μαθηματική αναπαράσταση του σήματος της φωνής;

5. Θεωρούμε ένα τετραγωνικό παράθυρο στον συνεχή χρόνο μήκους  $2T$  όπως φαίνεται στο Σχήμα 3



Σχήμα 3: Τετραγωνικό παράθυρο μήκους  $2T$ .

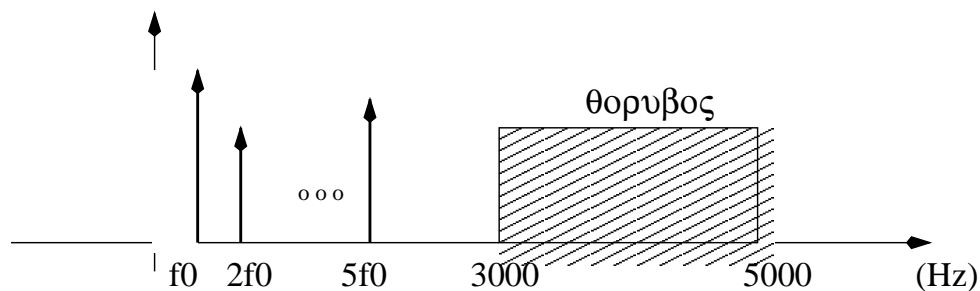
- (α') Ορίζοντας τη διάρκεια του παράθυρου ως  $D = 2T$  και το εύρος ζώνης συχνοτήτων,  $B$ , να είναι η απόσταση μεταξύ των δύο πρώτων μηδενισμών του φάσματος πλάτους του μετ. Fourier του παράθυρου εκατέρωθεν της μηδενικής συχνότητας, δείξτε ότι το γινόμενο  $DB$  είναι σταθερό.
- (β') Θεωρούμε ένα πολύ απλό σήμα φωνής το οποίο έχει φάσμα πλάτους όπως φαίνεται στο Σχήμα. 4 (μόνο οι θετικές συχνότητες εμφανίζονται) Ποια θα πρέπει να είναι η ελάχιστη



Σχήμα 4: Φάσμα πλάτους ενός απλού και ιδανικού σήματος φωνής.

διάρκεια,  $D$ , του παράθυρου για να μην υπάρχουν αλληλοεπιδράσεις του ημιτονοειδούς φάσματος και του φάσματος του θορύβου αν θεωρήσουμε μόνο το βασικό λοβό του φάσματος πλάτους του παράθυρου (ο οποίος έχει εύρος συχνοτήτων  $B$ );

- (γ') Θεωρούμε ένα πιο πολύπλοκο σήμα φωνής το οποίο έχει φάσμα πλάτους όπως φαίνεται στο Σχήμα. 5 (μόνο οι θετικές συχνότητες εμφανίζονται) Ποια θα πρέπει να είναι η ελά-



Σχήμα 5: Φάσμα πλάτους ενός πιο πολύπλοκου (σχετικά με το σήμα στο προηγούμενο Σχήμα) και ιδανικού σήματος φωνής.

χιστη διάρκεια,  $D$ , του παράθυρου ώστε και να μπορούν να διαχωριστούν οι συχνότητες των αρμονικών και να μην υπάρχουν αλληλοεξαρτήσεις μεταξύ των αρμονικών συνιστωσών και του θορυβώδους φάσματος; Όπως και πριν θεωρούμε κατά την αναλύσή μας ότι

μόνο ο βασικός λοβός του φάσματος πλάτους του μετ. Fourier επηρεάζει το συνολικό φάσμα πλάτους.

# Σημειώματα

## Σημείωμα αναφοράς

Copyright Πανεπιστήμιο Κρήτης, **Στυλιανού Ιωάννης**. «Ψηφιακή Επεξεργασία Φωνής. Άσκηση 3η». Έκδοση: 1.0. **Ηράκλειο/Ρέθυμνο** 2015. Διαθέσιμο από τη δικτυακή διεύθυνση: <http://www.csd.uoc.gr/~hy578>.

## Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση, Όχι Παράγωγο Έργο 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

## Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.

## Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Κρήτης» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

