



86.09

PROCESOS ESTOCÁSTICOS

Sintetización de fonemas mediante procesos AR

Trabajo Práctico N°3

Grupo N°4

2C 2024

Autores:

Gonzalo Antahuara
Marco Brischetto
Ignacio Cavicchioli
Tiago Sandoval

Padrón:

109965
110008
109428
104169

Correo:

gantahuara@gmail.com
mbrischetto@fi.uba.ar
icavicchioli@fi.uba.ar
tsandoval@fi.uba.ar

Índice

1. Introducción	2
2. Desarrollo	2
2.1. Ejercicio 1	2
2.1.1. Inciso a	2
2.1.2. Inciso b	3
2.1.3. Inciso c	4
2.1.4. Inciso d	5
2.2. Ejercicio 2	6
2.2.1. Inciso a	6
2.2.2. Inciso b	7
2.3. Ejercicio 3	9
2.3.1. Inciso a	9
2.3.2. Inciso b	10
2.4. Ejercicio 4	12
2.4.1. Inciso a	12
2.4.2. Inciso b	13
3. Conclusión	14
4. Bibliografía	14
5. Apéndice - Detalles de la implementación	15
5.1. Ejercicio 2	15
5.2. Ejercicio 3	15
5.3. Ejercicio 4	15
6. Apéndice - Imágenes	16

1. Introducción

Este trabajo está enfocado en la síntesis de voz artificial por medio de sistemas AR configurados con parámetros estimados de voz humana real.

2. Desarrollo

2.1. Ejercicio 1

2.1.1. Inciso a

Para demostrar la ecuación 2 partimos de la expresión 1, que es la ecuación en diferencias del sistema.

$$x(n) = \sum_{i=1}^P a_i \cdot x(n-i) + G \cdot u(n) \quad (1)$$

Se multiplica a ambos términos por $x(n-k)$ y se distribuye.

$$x(n) \cdot x(n-k) = \left(\sum_{i=1}^P a_i \cdot x(n-i) + G \cdot u(n) \right) \cdot x(n-k)$$

$$x(n) \cdot x(n-k) = \sum_{i=1}^P a_i \cdot x(n-i) \cdot x(n-k) + G \cdot u(n) \cdot x(n-k)$$

Se aplica esperanza y se distribuye.

$$E[x(n) \cdot x(n-k)] = E \left[\sum_{i=1}^P a_i \cdot x(n-i) \cdot x(n-k) + G \cdot u(n) \cdot x(n-k) \right]$$

$$E[x(n) \cdot x(n-k)] = E \left[\sum_{i=1}^P a_i \cdot x(n-i) \cdot x(n-k) \right] + E[G \cdot u(n) \cdot x(n-k)]$$

$$R_X(n, n-k) = \sum_{i=1}^P a_i \cdot E[x(n-i) \cdot x(n-k)] + E[G \cdot u(n) \cdot x(n-k)]$$

$$R_X(n, n-k) = \sum_{i=1}^P a_i \cdot R_X(n-i, n-k) + E[G \cdot u(n) \cdot x(n-k)]$$

Las autocorrelaciones se pueden reescribir como función de la diferencia entre los instantes.

$$R_X(k) = \sum_{i=1}^P a_i \cdot R_X(k-i) + E[G \cdot u(n) \cdot x(n-k)]$$

Ahora se expande el término cruzado.

$$\begin{aligned}
 E[G \cdot u(n) \cdot x(n-k)] &= E \left[G \cdot u(n) \cdot \sum_{i=1}^P a_i \cdot x(n-i-k) + G \cdot u(n) \cdot G \cdot u(n-k) \right] \\
 &= G^2 R_U(n, n-k) + E \left[G \cdot u(n) \cdot \sum_{i=1}^P a_i \cdot x(n-i-k) \right] = \\
 &G^2 \delta(k) + G \cdot \sum_{i=1}^P a_i \cdot R_{U,X}(n, n-i-k)
 \end{aligned}$$

Como la entrada $u(n)$ no está correlacionada con muestras de $X(n)$ de tiempos pasados a n , el término a la derecha se anula y la expresión toma la forma:

$$E[G \cdot u(n) \cdot x(n-k)] = G^2 \cdot \delta(k)$$

Y uniendo todo queda:

$$R_X(k) = \sum_{i=1}^P a_i \cdot R_X(k-i) + G^2 \cdot \delta(k) \quad (2)$$

2.1.2. Inciso b

Para encontrar el valor de la variable G partimos de la ecuación ya demostrada:

$$R_X(k) = \sum_{i=1}^P a_i \cdot R_X(k-i) + G^2 \cdot \delta(k)$$

Se acomoda todo en función de los valores conocidos, es decir, la autocorrelación y los coeficientes del filtro.

$$\sqrt{R_X(k) - \sum_{i=1}^P a_i \cdot R_X(k-i)} = G \cdot \delta(k)$$

Evalando en $k = 0$ queda:

$$G = \sqrt{R_X(0) - \sum_{i=1}^P a_i \cdot R_X(-i)}$$

Como $R_X(-i) = R_X(i)$, la expresión final es:

$$G = \sqrt{R_X(0) - \sum_{i=1}^P a_i \cdot R_X(i)} \quad (3)$$

Que es igual a la indicada en las consignas.

2.1.3. Inciso c

Se parte de la ecuación 2 pero se escriben las ecuaciones para $k > 0$ de forma genérica.

$$\begin{bmatrix} R_X(1) = \sum_{i=1}^P a_i \cdot R_X(1-i) \\ R_X(2) = \sum_{i=1}^P a_i \cdot R_X(2-i) \\ \dots \\ R_X(P) = \sum_{i=1}^P a_i \cdot R_X(P-i) \end{bmatrix} \quad (4)$$

Como las a_i se pueden extraer a derecha como el vector “a”, las sumatorias pueden ser escritas como una matriz (en las consignas se llama R). Así, la expresión 4 se puede reescribir como:

$$\begin{bmatrix} R_X(1) \\ R_X(2) \\ \dots \\ R_X(P) \end{bmatrix} = \begin{bmatrix} R_X(1-1) & R_X(1-2) & R_X(1-3) & \dots & R_X(1-P) \\ R_X(2-1) & R_X(2-2) & R_X(2-3) & \dots & R_X(2-P) \\ R_X(3-1) & R_X(3-2) & R_X(3-3) & \dots & R_X(3-P) \\ \dots & \dots & \dots & \dots & \dots \\ R_X(P-1) & R_X(P-2) & R_X(P-3) & \dots & R_X(P-P) \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_P \end{bmatrix} \quad (5)$$

Antes de seguir conviene explicitar las siguientes estructuras ya definidas en las consignas:

$$a = [a_1 \quad a_2 \quad \dots \quad a_P]^T$$

$$r = [R_X(1) \quad R_X(2) \quad \dots \quad R_X(P)]^T \quad (6)$$

$$R = \begin{bmatrix} R_X(0) & R_X(-1) & R_X(-2) & \dots & R_X(1-P) \\ R_X(1) & R_X(0) & R_X(-1) & \dots & R_X(2-P) \\ R_X(2) & R_X(1) & R_X(0) & \dots & R_X(3-P) \\ \dots & \dots & \dots & \dots & \dots \\ R_X(P-1) & R_X(P-2) & R_X(P-3) & \dots & R_X(0) \end{bmatrix}$$

Estas expresiones se pueden usar para resumir la expresión 5 de la siguiente manera:

$$r = \begin{bmatrix} R_X(0) & R_X(-1) & R_X(-2) & \dots & R_X(1-P) \\ R_X(1) & R_X(0) & R_X(-1) & \dots & R_X(2-P) \\ R_X(2) & R_X(1) & R_X(0) & \dots & R_X(3-P) \\ \dots & \dots & \dots & \dots & \dots \\ R_X(P-1) & R_X(P-2) & R_X(P-3) & \dots & R_X(0) \end{bmatrix} \cdot a$$

En principio la matriz “R” es invertible por lo que se multiplica a izquierda a ambos lados de la ecuación por dicha inversa, resultando en:

$$a = \begin{bmatrix} R_X(0) & R_X(-1) & R_X(-2) & \dots & R_X(1-P) \\ R_X(1) & R_X(0) & R_X(-1) & \dots & R_X(2-P) \\ R_X(2) & R_X(1) & R_X(0) & \dots & R_X(3-P) \\ \dots & \dots & \dots & \dots & \dots \\ R_X(P-1) & R_X(P-2) & R_X(P-3) & \dots & R_X(0) \end{bmatrix}^{-1} \cdot r$$

Y en su forma resumida la expresión queda tal como en las consignas.

$$a = R^{-1} \cdot r \quad (7)$$

2.1.4. Inciso d

Este punto pide demostrar la relación 8. Convenientemente, la ecuación 9 indica la relación entre la PSD de entrada y salida de un sistema LTI y la transferencia (ecuación 10) está indicada en las consignas.

$$S_X(\omega) = \frac{G^2}{|1 - \sum_{k=1}^P a_k \cdot e^{-j\omega k}|^2} \quad (8)$$

$$S_X(\omega) = |H(e^{j\omega})|^2 S_U(\omega) \quad (9)$$

$$H(z) = \frac{G}{1 - \sum_{k=1}^P a_k \cdot z^{-k}} \quad (10)$$

La transferencia $H(z)$ se puede escribir en función de ω porque $z = e^{j\omega}$.

$$H(e^{j\omega}) = \frac{G}{1 - \sum_{k=1}^P a_k \cdot e^{-j\omega k}} \quad (11)$$

Juntando las expresiones 9 y 11 se llega a:

$$S_X(\omega) = \left| \frac{G}{1 - \sum_{k=1}^P a_k \cdot e^{-j\omega k}} \right|^2 S_U(\omega)$$

No hay que olvidar que $u(n)$ es ruido blanco de media nula y varianza unitaria, por lo que su autocorrelación es una delta unitaria en el origen. La expresión sigue:

$$S_X(\omega) = \left| \frac{G}{1 - \sum_{k=1}^P a_k \cdot e^{-j\omega k}} \right|^2$$

Así quedando demostrado lo pedido.

2.2. Ejercicio 2

2.2.1. Inciso a

A partir de las muestras provistas y del desarrollo matemático previo, se estimarán para cada uno de los fonemas los coeficientes de un sistema AR que al ser excitado por ruido blanco produzca el fonema en cuestión. Para ello, se generó una función que recibe la muestra de audio y retorna los coeficientes del sistema a_i y su ganancia G . La implementación se detalla a continuación:

- Las muestras de audio x se usan para estimar la autocorrelación del proceso $R_X(k)$ mediante su estimador sesgado.
- Se busca el índice del vector correspondiente a la autocorrelación evaluada en cero $R_X(0)$.
- Se conservan únicamente los valores de R_X entre cero y P , donde P indica el orden del modelo a estimar. Recordamos que al considerar las muestras como provenientes de un proceso ESA, R_X es una función par.
- Se genera el vector r definido en la ecuación (6) a partir de R_X .
- Se genera la matriz R aprovechando que el valor en cada una de sus diagonales es constante. Para ello, se genera un vector del largo de cada diagonal para cada valor de R_X y luego se lo inserta en la diagonal que corresponda.
- Teniendo R y r , calculamos el vector a según la expresión (7).
- Por otra parte, la ganancia G se obtiene según la ecuación (3).

La función pedida en las consignas se encuentra en el archivo de *MATLAB* asociado a a la entrega.

2.2.2. Inciso b

A continuación se muestran los resultados obtenidos de la estimación para los fonemas [a] y [s] con un modelo de orden 20. La autocorrelación fue limitada a 200 muestras para mejorar el entendimiento del gráfico. Los gráficos realizados para los demás fonemas se encuentran en el apéndice de este documento por conveniencia.

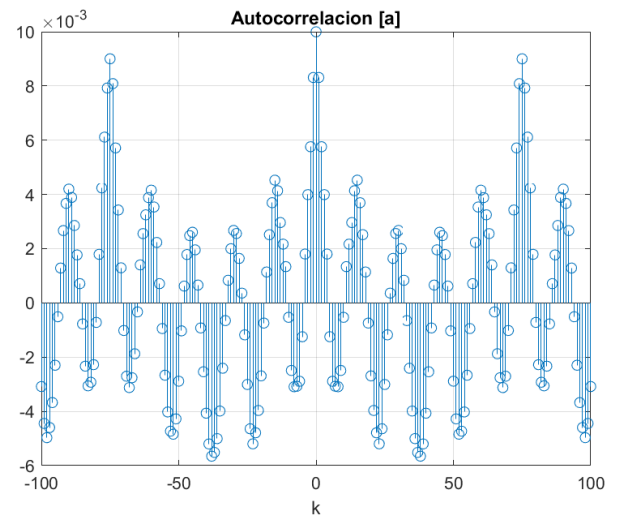
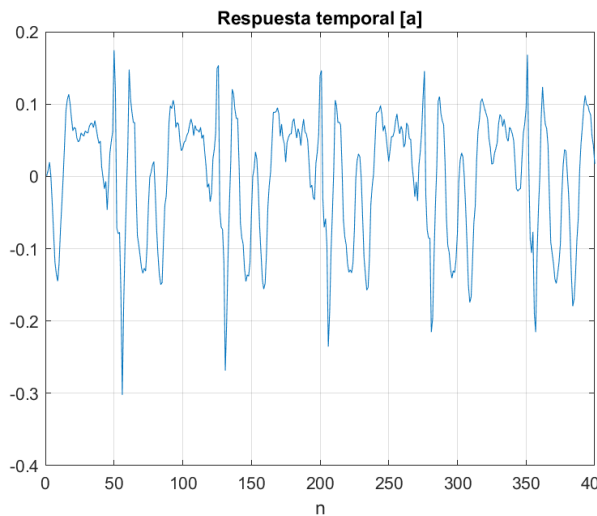


Figura 1: Respuesta temporal del fonema [a]

Figura 2: Autocorrelación del fonema [a]

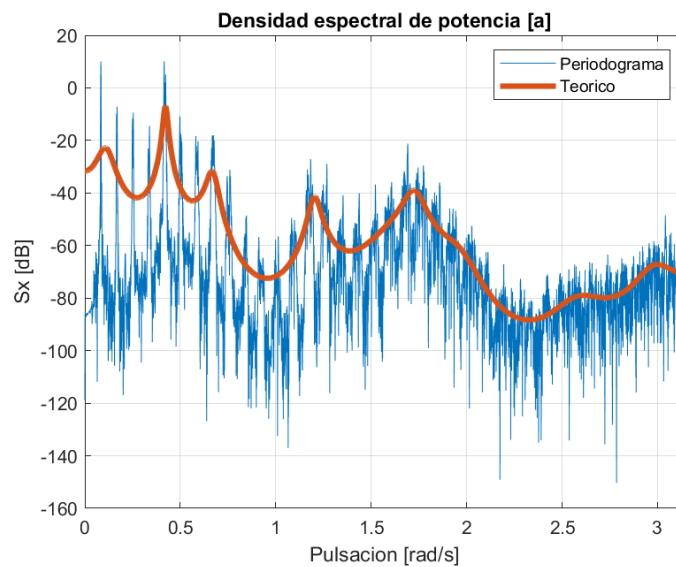


Figura 3: Periodograma y PSD teórica del fonema [a]

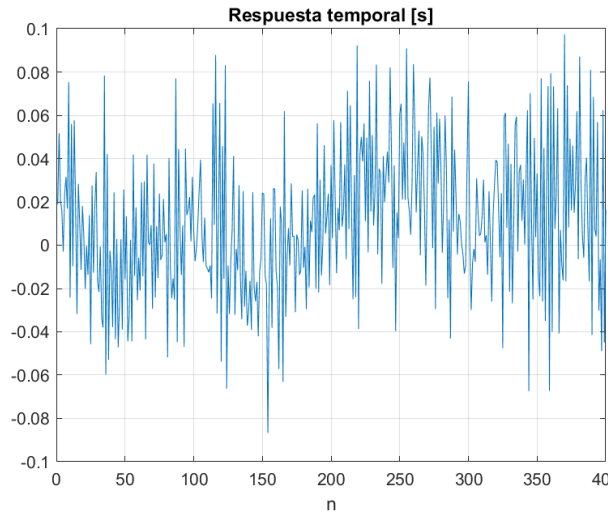


Figura 4: Respuesta temporal del fonema [s]

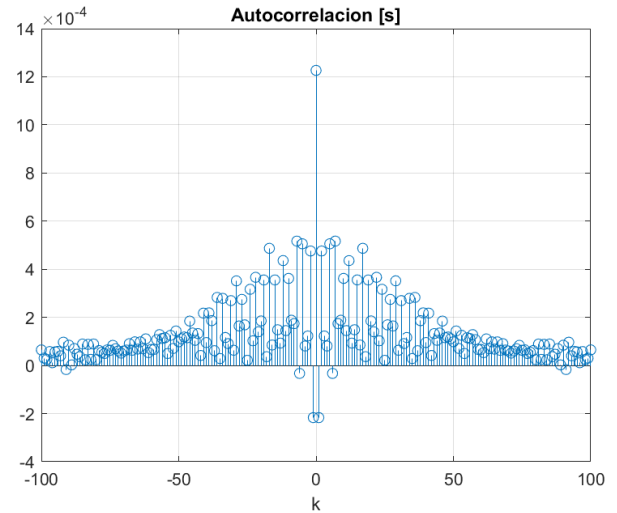


Figura 5: Autocorrelación del fonema [s]

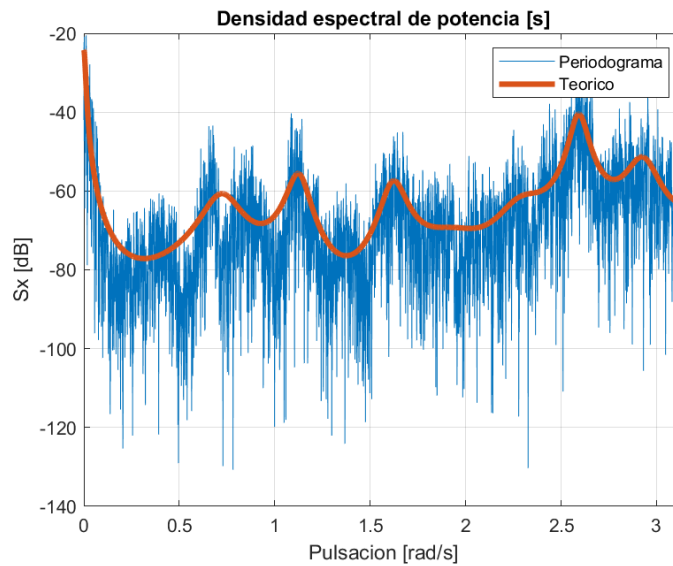


Figura 6: Periodograma y PSD teorica del fonema [s]

Todas las vocales estudiadas (acá y en el apéndice) presentan una estructura periódica obvia en el dominio temporal, mientras que las consonantes y el ruido “sh” parecen ruido. Al observar los periodogramas estimados (que están entre 0 y π rad/seg), se nota que las vocales tienen picos claramente distinguibles mientras que las consonantes fricativas no. Estos picos se denominan formantes, y son concentraciones de energía en bandas de frecuencias particulares. Las fricativas no presentan estos obvios picos en baja frecuencia, y su potencia parecería estar más distribuida en el espectro, coincidiendo con los artículos consultados (ver bibliografía).

2.3. Ejercicio 3

2.3.1. Inciso a

En este inciso se estimará el periodograma para los fonemas [e] y [sh] a partir del método de Welch y se los comparará con la PSD teórica del modelo hallado anteriormente. El método de Welch será realizado para distintos tamaños de ventana y con un solapamiento del 50 %. A continuación se explica la implementación del método:

- Se fracciona la señal en realizaciones más pequeñas de longitud M , haciendo un solapamiento del 50 % entre realizaciones consecutivas.
- Para disminuir la correlación entre los extremos de las realizaciones consecutivas se multiplica cada realización por una ventana de hamming, la cual reduce el peso de las muestras en los extremos.
- Se calcula la potencia de la ventana, usada para normalizar el resultado.
- Se computa el periodograma de cada realización y se toma la media ajustando por la potencia de la ventana. En la FFT usada en la estimación del periodograma se aplicó un zero-padding para obtener más puntos en el gráfico final, esto no agrega información que ya no estuviera en la señal.

La matemática de la implementación fue tomada directamente de la página 50, sección 2.7.2 del libro “Spectral analysis of signals” citado en la bibliografía.

2.3.2. Inciso b

A continuación se muestran los resultados obtenidos para $M = \{10, 100, 1000\}$ de ambos fonemas utilizando el método de Welch.

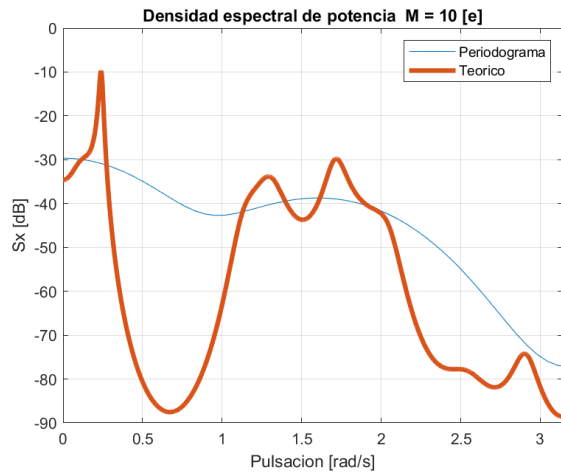


Figura 7: Periodograma $M = 10$ y PSD teórica del fonema [e]

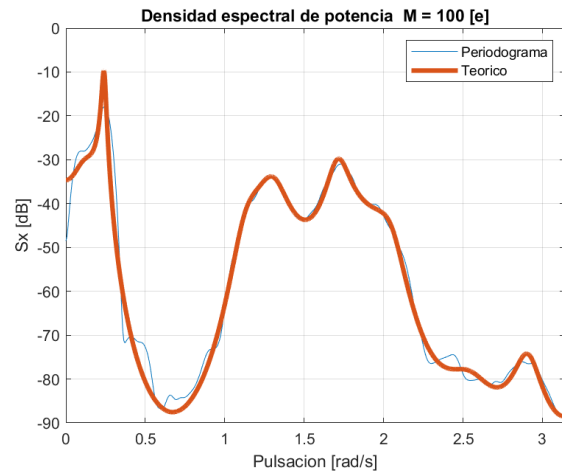


Figura 8: Periodograma $M = 100$ y PSD teórica del fonema [e]

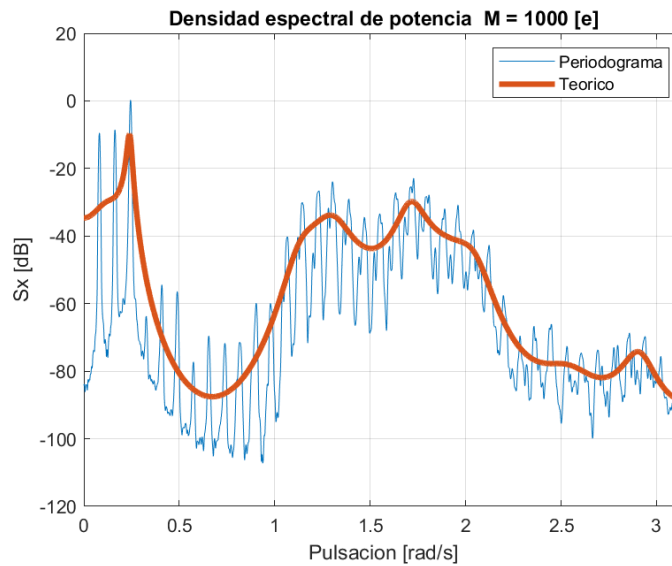


Figura 9: Periodograma $M = 1000$ y PSD teórica del fonema [e]

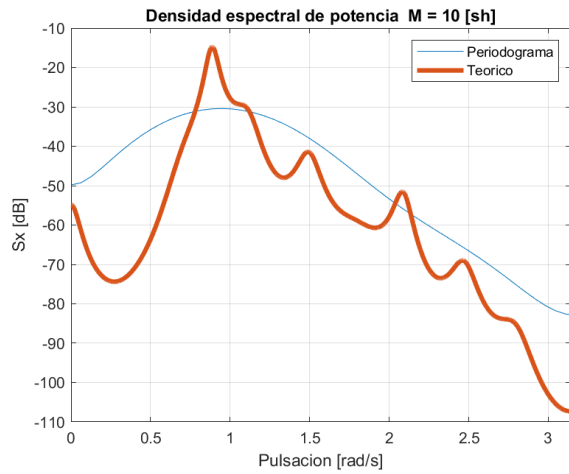


Figura 10: Periodograma $M = 10$ y PSD teórica del fonema [e]

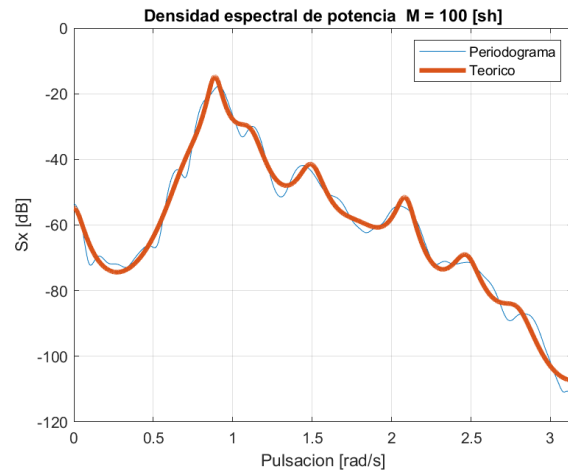


Figura 11: Periodograma $M = 100$ y PSD teórica del fonema [e]

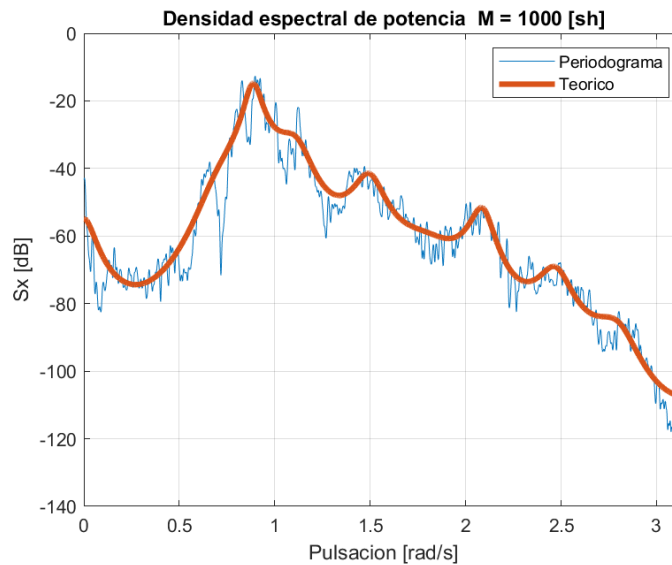


Figura 12: Periodograma $M = 1000$ y PSD teórica del fonema [e]

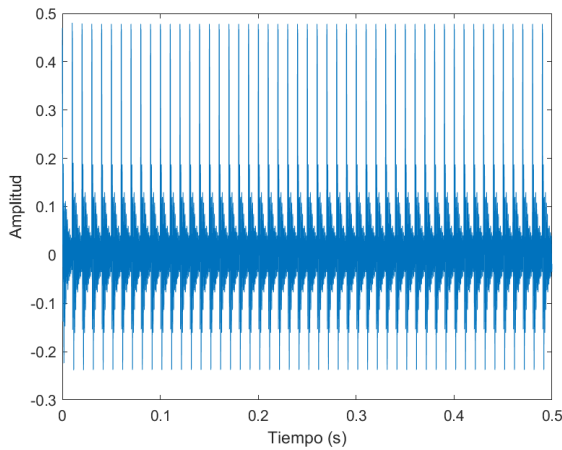
En las figuras se observa que el ruido del periodograma disminuye a medida que la ventana se achica. A su vez, cuando se toma una ventana grande, el periodograma se torna ruidoso. Este efecto aparentaría estar ligado principalmente a la cantidad de secciones en las que se divide la señal original. Cuando el tamaño de ventana es elevado hay menos secciones, lo que significa que hay menos muestras para promediar, que llevaría a un resultado más ruidoso. Recíprocamente, una ventana de menor tamaño implicaría más muestras para promediar y, por ende, un periodograma más suave.

Además, una ventana pequeña tiene menos muestras con información espectral extraíble, que causaría una menor resolución espectral comparado con periodogramas calculados con ventanas más grandes. Esto se ve claramente al contrastar la figura 10 contra la figura 11 o 12. El zero-padding permite interpolar más puntos en el dominio frecuencial pero no agrega información. Consideramos que, para este caso, el $M=100$ da el mejor trade-off.

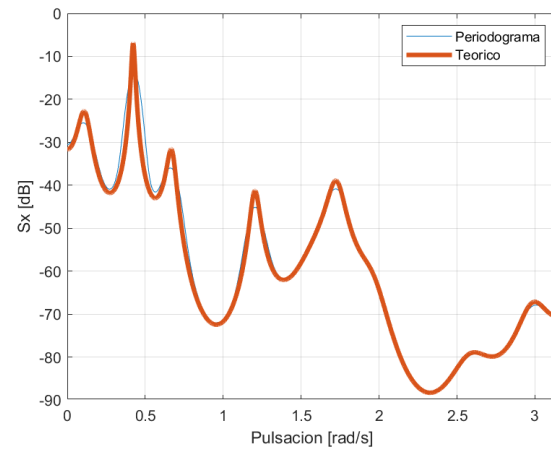
2.4. Ejercicio 4

2.4.1. Inciso a

Utilizando los parámetros LPC obtenidos en el inciso 2 se sintetizaron las señales de cada fonema, cada uno con una duración de 500ms. Para la generación del habla sonora se empleó como entrada del sistema un tren de impulsos periódicos con una frecuencia de pitch de 100 Hz, mientras que para el habla sorda se usó un proceso de ruido blanco gaussiano $\sim \mathcal{N}(0, 1)$. A modo de ejemplo, abajo se muestran las señales sintetizadas de los fonemas [a] y [s] junto con las imágenes de los periodogramas superpuestos a la PSD teórica.

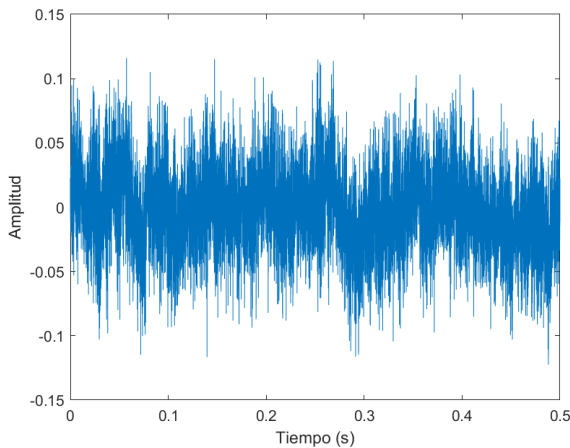


(a) Señal sintetizada.

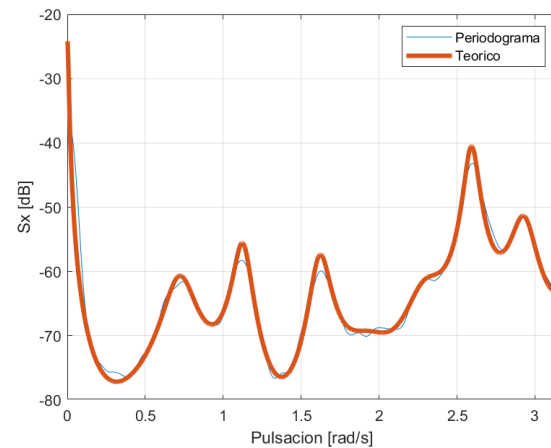


(b) Periodograma y PSD teórica.

Gráfico 13: Habla sonora [a].



(a) Señal sintetizada.



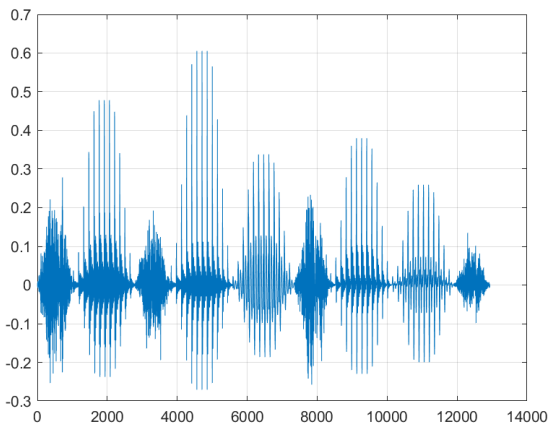
(b) Periodograma y PSD teórica.

Gráfico 14: Habla sorda [s].

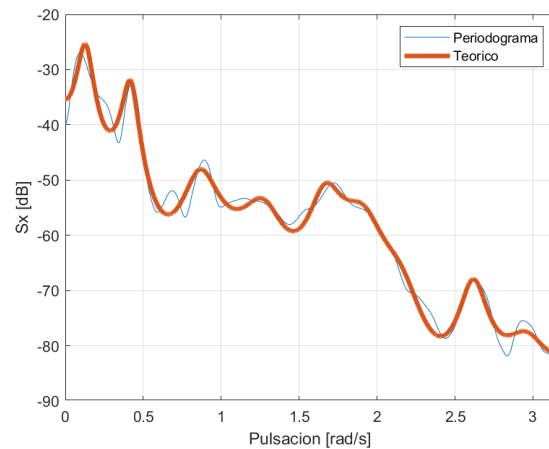
En los dos casos mostrados arriba el gráfico del espectrograma de las señal sintetizada se condice con la PSD teórica. Debido a las características propias de cada fonema ya mencionadas en el inciso 2, el habla sonora [a] presenta mayor potencia espectral para bajas frecuencia con picos distinguibles. En cambio, en el habla sorda [s], la PSD indica que la potencia se encuentra más distribuida aunque es mayor en frecuencias cercanas a 2,5 rad/seg.

2.4.2. Inciso b

Luego de sintetizar los fonemas en el inciso anterior, se procede a concatenarlos para generar una señal que se asemeje a una palabra. Para ello, se decidió concatenarlas en un orden tal que quede la pseudopalabra *shafeijous*. La duración de cada fonema de este arreglo se varía de tal forma que la señal obtenida se asemeje al idioma español (i.e. los fonemas sordos duran menos que los fonemas sonoros). Abajo se muestran las imágenes pertinentes:

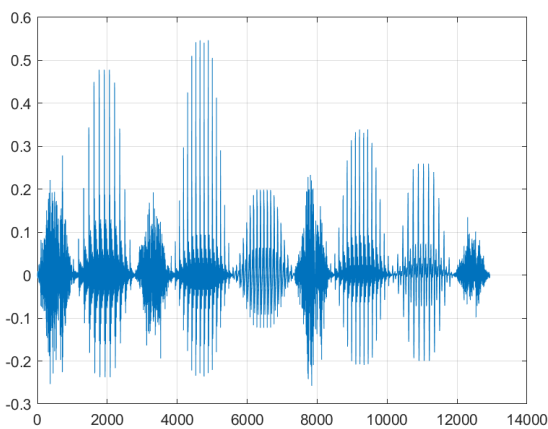


(a) Señal temporal.

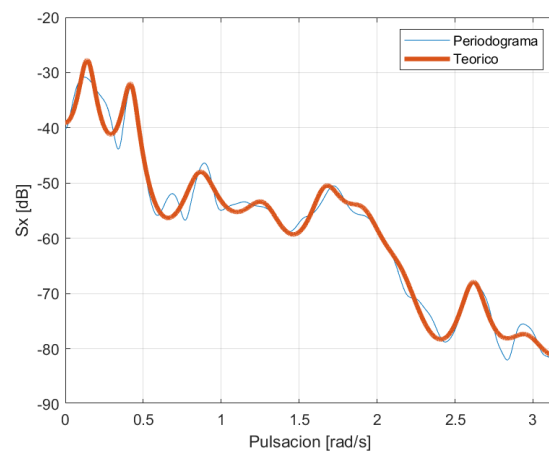


(b) Periodograma y PSD teórica.

Gráfico 15: Palabra con el mismo pitch.



(a) Señal temporal.



(b) Periodograma y PSD teórica.

Gráfico 16: Palabra con distinto pitch.

En los gráficos de arriba los periodogramas de las señales sintetizadas se condicen con las PSD teóricas correspondientes. En la figura 15a se observa la señal resultante en función del tiempo. Luego, se repite el procedimiento anterior variando el pitch de las vocales para emular el acento que puede tener una persona al hablar. Esta señal se ve en la figura 16a. Los cambios entre periodogramas son prácticamente imperceptibles a simple vista pero audibles en la muestra.

EL sonido obtenido suena robótico, que parece razonable dada la falta de post-procesamiento y el approach directo usado en la generación, con señales de excitación comparativamente simples. Sin embargo, destacamos que el audio se entiende, un gran logro.

3. Conclusión

El método propuesto en este documento es una manera de generar sonido que emula el habla humana a partir de un filtrado con coeficientes previamente estimados. Algunas características interesantes de resaltar son:

- El método solo requiere de muestras de audio en la fase de estimación de los parámetros. Es decir, una vez estimados los coeficientes, las muestras de audio originales se vuelven innecesarias. Esto podría ser un beneficio ya que, considerando que los vectores de coeficientes ocupan menos espacio que los audios, el hardware en el que se implemente tendrá un requerimiento de memoria menor.
- Las palabras generadas tienen una tonada robótica, antinatural, pero se mantienen inteligibles, comprensibles.
- El método de estimación por medio de las ecuaciones de Y-W planteadas matricialmente no tiene una complejidad de implementación notable y se puede usar para obtener tantos coeficientes como se desee. Mientras más coeficientes del filtro se tengan, mejor se va a poder recrear la voz humana.

Se estuvo investigando y algunas posibles mejoras podrían ser las siguientes:

- En vez de usar un tren de deltas determinístico como fuente de excitación para los fonemas de habla sonora se podría hacer estocástico. Es decir, que el espaciado entre deltas y potencia varíen ligeramente, emulando mejor lo que sería la excitación real. Otra alternativa a lo anterior es usar señales de excitación estimadas a partir del habla humana. En nuestras pruebas por fuera de las consignas estimamos el filtro que mapea el habla humana a la excitación sonora determinística, usándolo para generar una excitación aproximada. Esta se terminó usando para sintetizar algunas vocales, que suenan más humanas que las sintetizadas directamente con los deltas.
- Los audios generados podrían ser refinados para emular mejor el habla humana. Se podrían filtrar las altas frecuencias o aplicar ponderaciones particulares para resalta ciertos tonos.

Por otra parte, no hubo dificultades particularmente relevantes que ameriten ser mencionadas.

Con respecto a los resultados, se considera que los audios obtenidos transmiten el sonido deseado pero podrían ser refinados más como ya fue indicado. Desde un punto de vista educativo, el trabajo permitió ahondar en los temas de estimación de filtros AR-N por el método de Y-W, en el método de Welch de estimación del periodograma y en el uso de herramientas como *MATLAB*.

4. Bibliografía

- *Spectral analysis of signals*, Petre Stoica and Randolph Moses, editorial Prentice Hall, 2005, versión digital.
- [Formante](#), Wikipedia.
- [Consonante fricativa](#), Wikipedia.

5. Apéndice - Detalles de la implementación

Por completitud se decidió agregar este apéndice describiendo un poco más los varios algoritmos usados en el trabajo.

5.1. Ejercicio 2

El pseudocódigo de la implementación de *param_ar* ya fue explicado en el desarrollo del inciso. El archivo de *MATLAB* entregado para el ejercicio 2 tiene otra función llamada *generar_graficos* que encapsula la estimación y la generación de gráficos. En esta también se hace el gráfico de la PSD teórica, siendo la idea muestrear en frecuencia la potencia de la transferencia del filtro.

5.2. Ejercicio 3

Este inciso usa la función *param_ar* (ya explicada), la función *Welch_alt* (explicada en el desarrollo - realiza la estimación del periodograma por método de Welch) y otra función que genera los gráficos. Es crítico no olvidar que en el método de Welch hay que normalizar por la potencia de la ventana usada, sino aparecen discrepancias de potencia entre la PSD teórica y la estimada.

Recordando, el método de Welch (libro de Spectral analysis of signals) plantea:

$$\hat{\phi}_j(\omega) = \frac{1}{MP} \left| \sum_{t=1}^M v(t) y_j(t) e^{-i\omega t} \right|^2$$

$$P = \frac{1}{M} \sum_{t=1}^M |v(t)|^2$$

P es la potencia de la ventana temporal $v(t)$ y ϕ es el periodograma ventaneado de $y_j(t)$. El estimador de Welch de la PSD es el promedio de los periodogramas ventaneados:

$$\hat{\phi}_W(\omega) = \frac{1}{S} \sum_{j=1}^S \hat{\phi}_j(\omega)$$

El código desarrollado sigue estas ecuaciones.

5.3. Ejercicio 4

En este inciso se usan todas las funciones mencionadas antes (excepto las de graficación) y se usan otras para sintetizar los sonidos pedidos. El pitch y duración fueron determinados por las consignas mientras que la frecuencia de sampling se tomó de los archivos de audio provistos. La función provista se usó para suavizar los bordes de las señales.

Los sonidos se crean a partir del filtrado de las señales de excitación con los coeficientes correspondientes. El ruido blanco se genera con las funciones de *MATLAB* y los deltas se generan a partir de un vector de ceros al que se le agregan los deltas normalizados.

6. Apéndice - Imágenes

En este apéndice se adjuntan el resto de los gráficos del ejercicio 2.

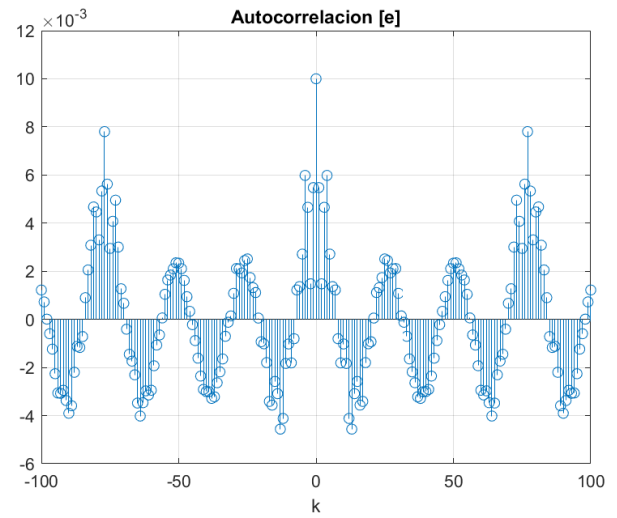
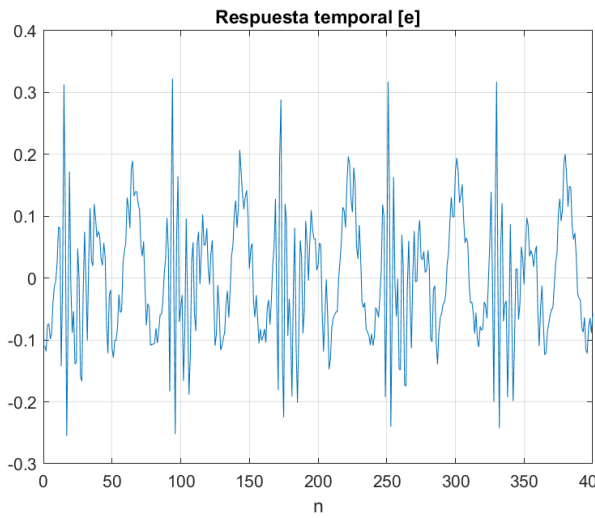


Figura 17: Respuesta temporal del fonema [e]

Figura 18: Autocorrelación del fonema [e]

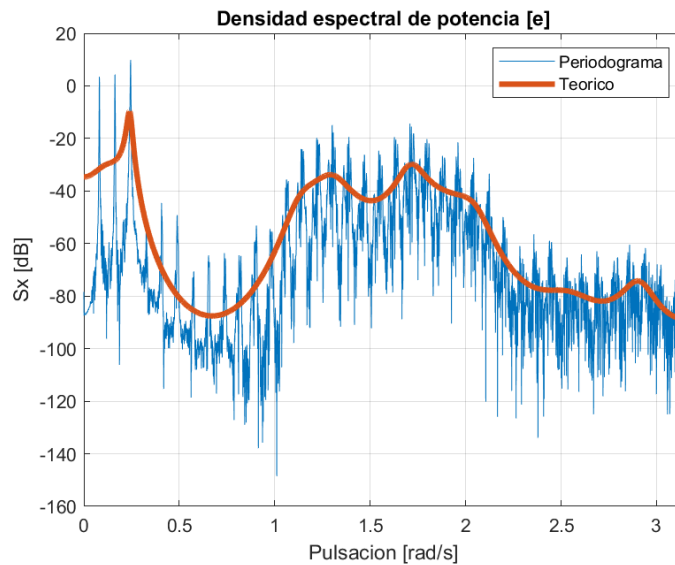


Figura 19: Periodograma y PSD teórica del fonema [e]

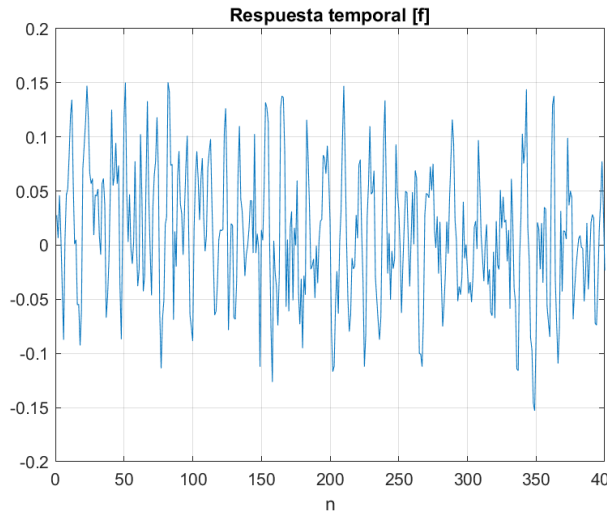


Figura 20: Respuesta temporal del fonema [f]

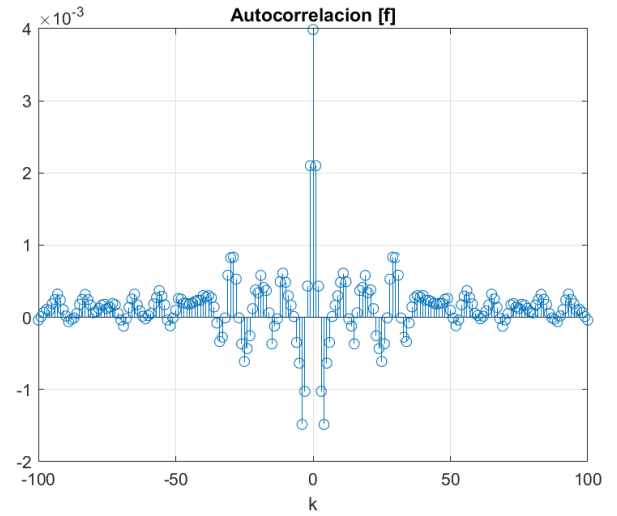


Figura 21: Autocorrelación del fonema [f]

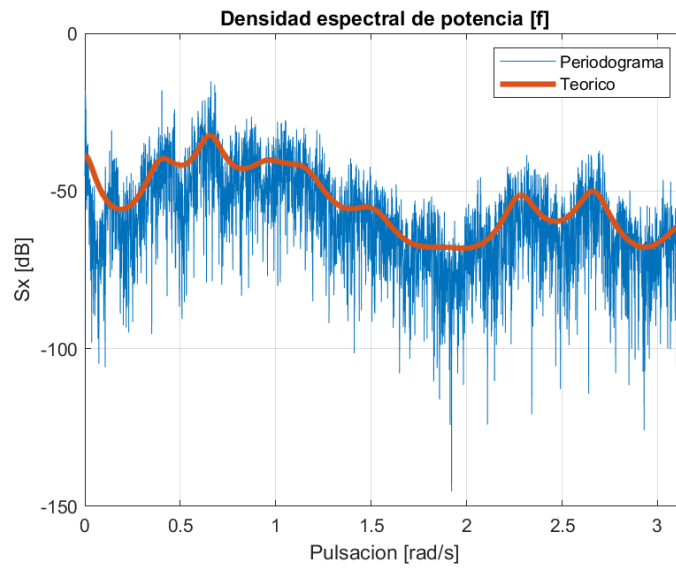


Figura 22: Periodograma y PSD teórica del fonema [f]

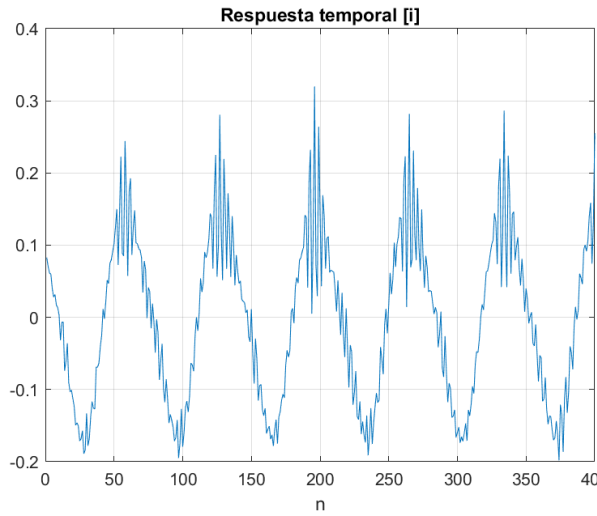


Figura 23: Respuesta temporal del fonema [i]

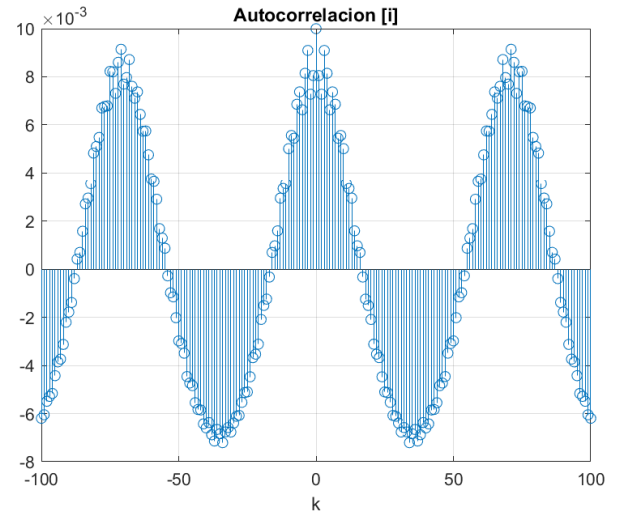


Figura 24: Autocorrelación del fonema [i]

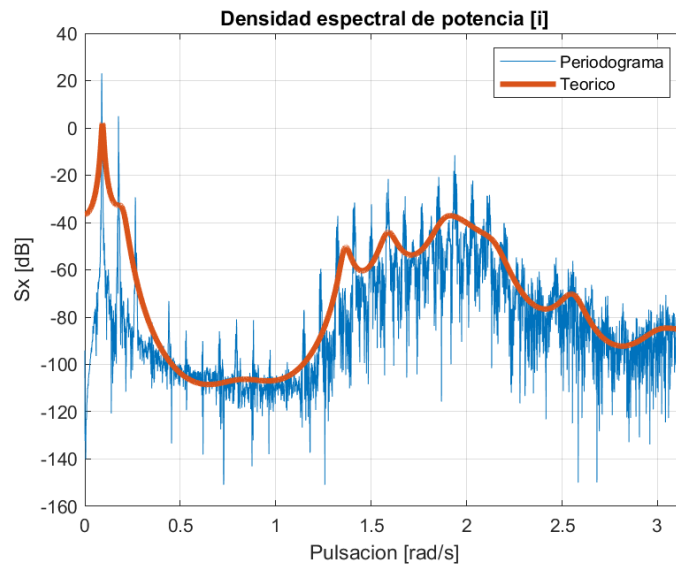


Figura 25: Periodograma y PSD teórica del fonema [i]

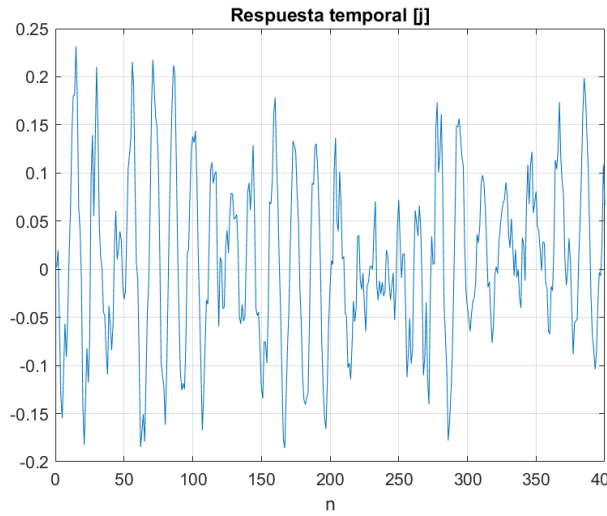


Figura 26: Respuesta temporal del fonema [j]

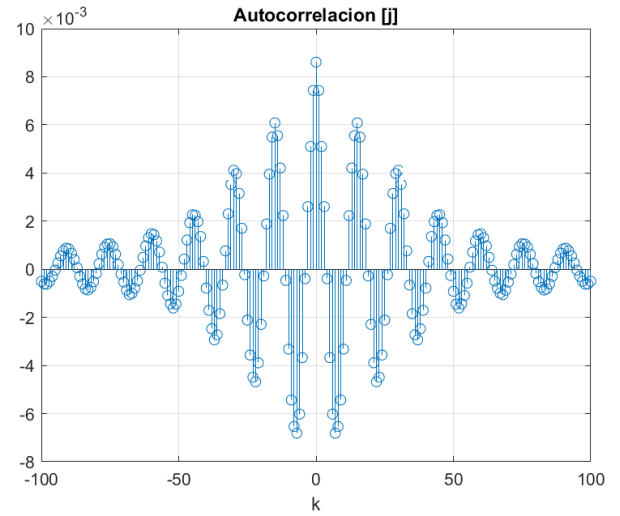


Figura 27: Autocorrelación del fonema [j]

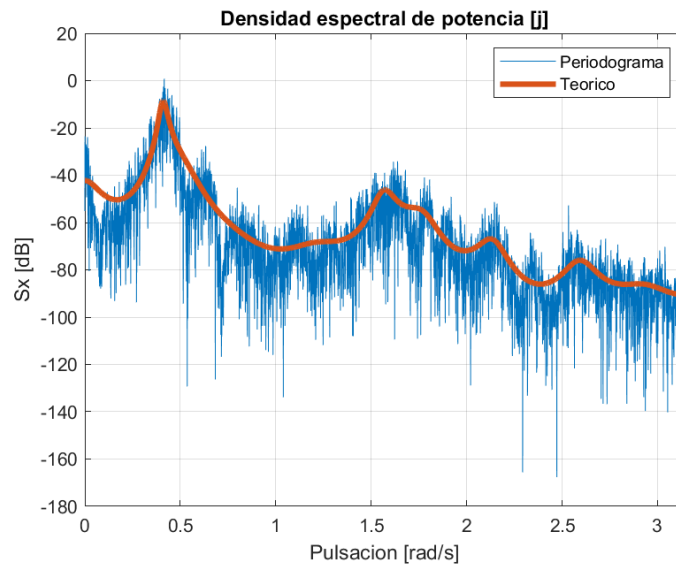


Figura 28: Periodograma y PSD teórica del fonema [j]

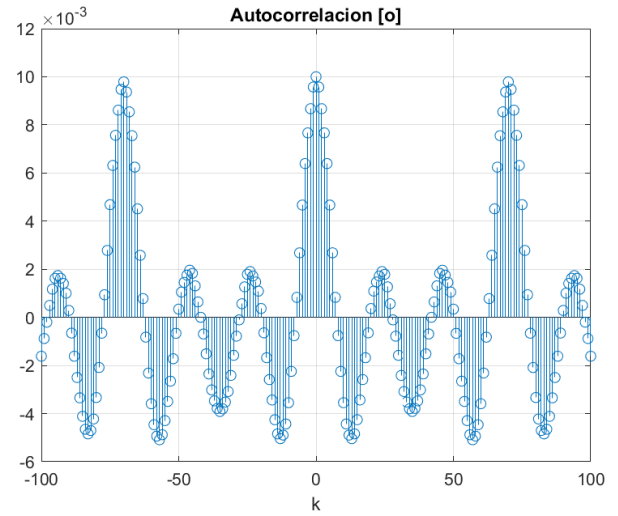
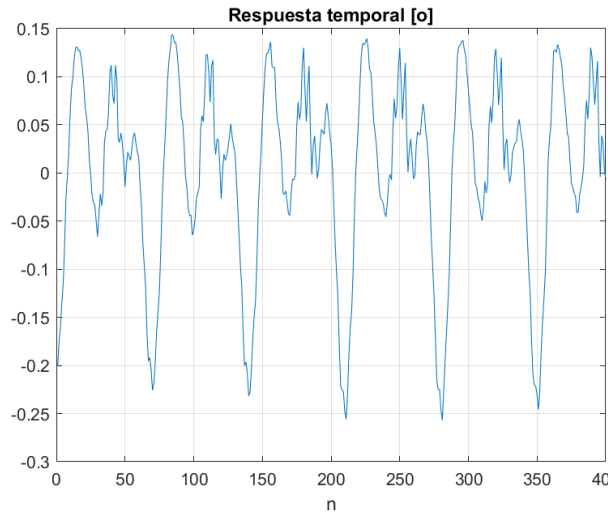


Figura 29: Respuesta temporal del fonema [o]

Figura 30: Autocorrelación del fonema [o]

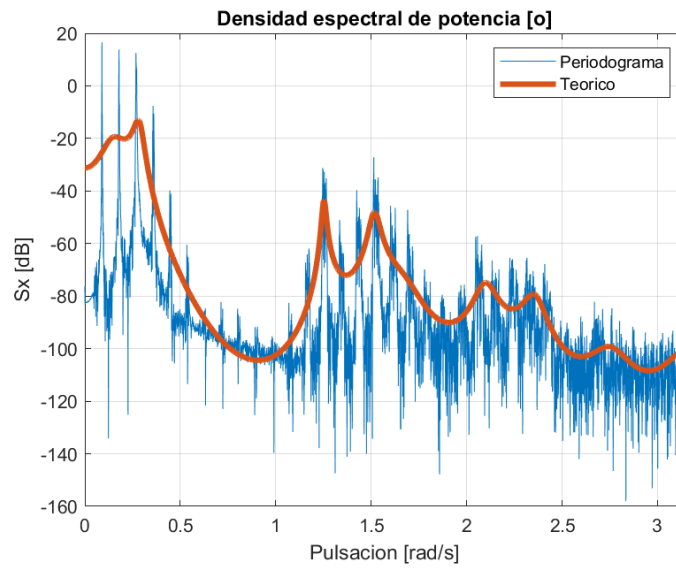


Figura 31: Periodograma y PSD teórica del fonema [o]

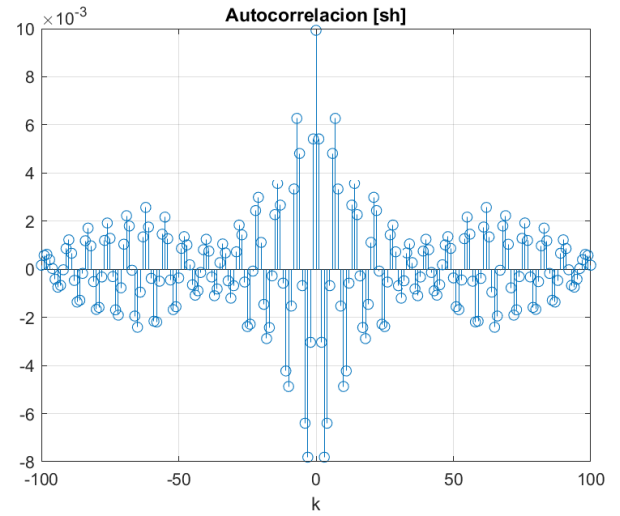
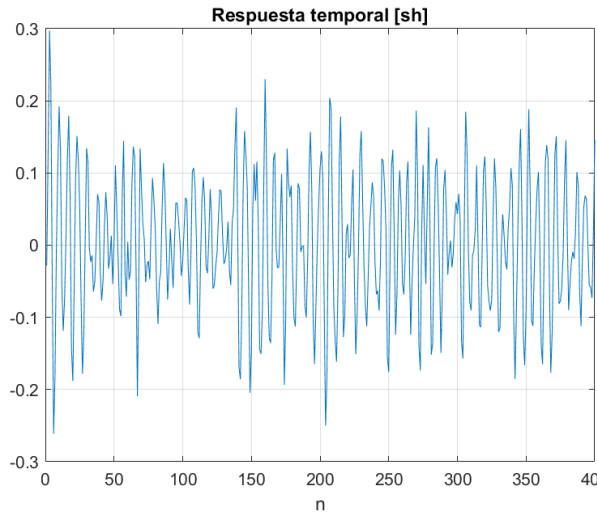


Figura 32: Respuesta temporal del fonema [sh] Figura 33: Autocorrelación del fonema [sh]

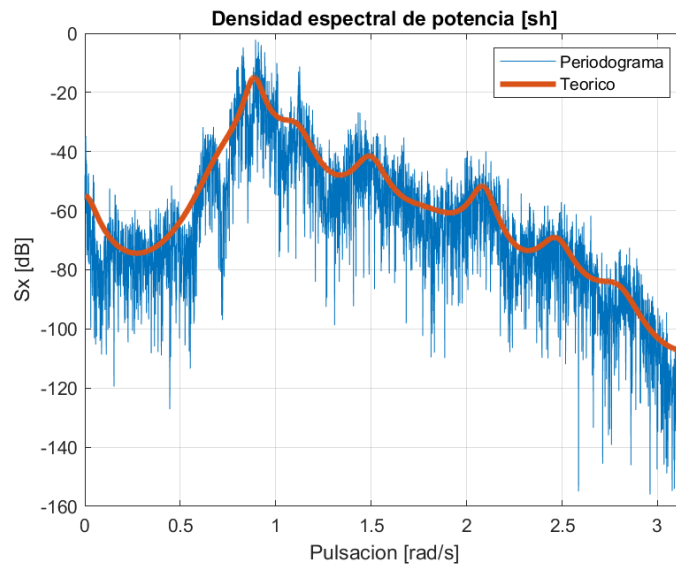


Figura 34: Periodograma y PSD teórica del fonema [sh]

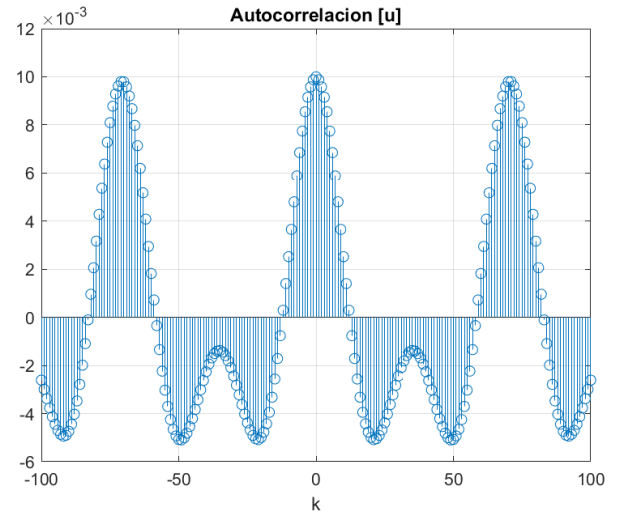
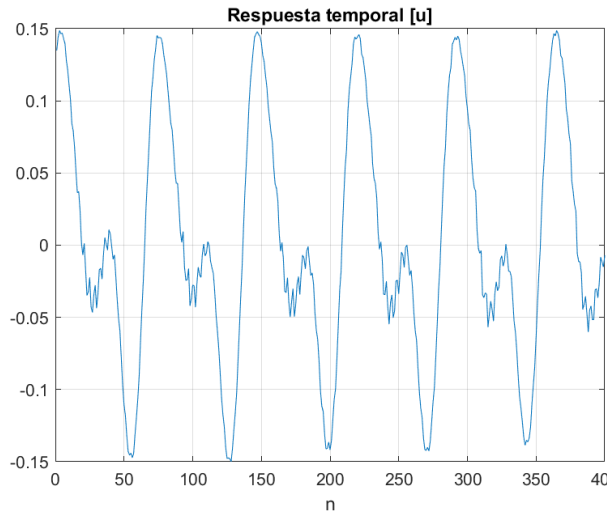


Figura 35: Respuesta temporal del fonema [u]

Figura 36: Autocorrelación del fonema [u]

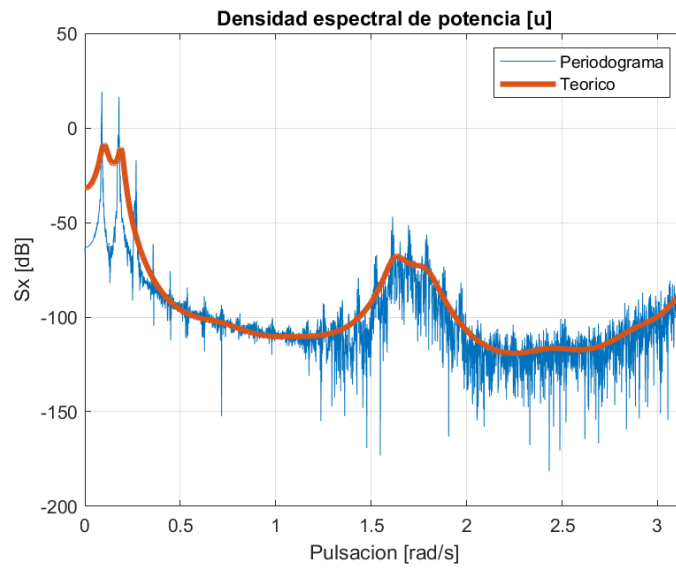


Figura 37: Periodograma y PSD teórica fonema [u]