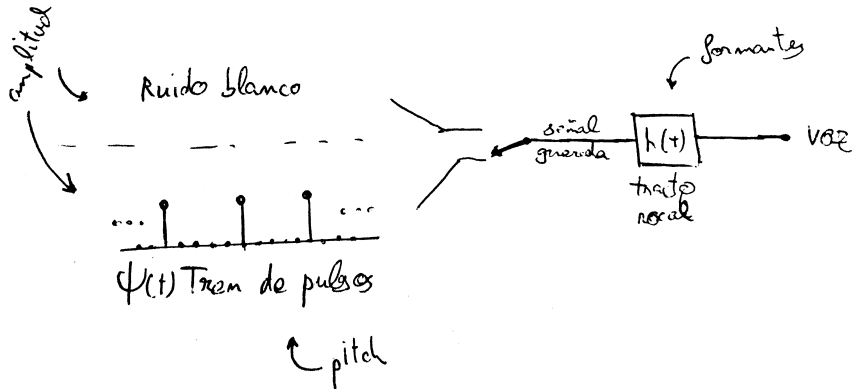


Problema 1.



Problema 2. Los formantes son dos frecuencias a las que el tracto vocal ofrece una ganancia especialmente alta. Se sitúan entre los 200 Hz y los 2 kHz y su combinación determina el tipo de fonema que se percibe.

Si representamos la combinación de los dos formantes en un plano, y anotamos las posiciones de las vocales, veremos que tienen una disposición similar a un triángulo. Esto se llama *triángulo vocálico* y sucede en el español y en la mayoría de idiomas, si bien la forma puede ser algo distinta.

Problema 3. Los bins de la DFT correspondientes a las formantes son $n_1 \simeq 62$ y $n_2 \simeq 135$. Según $f = n/N f_s$ con $f_s = 8 \text{ kHz}$ y $N = 1000$, tenemos que los dos formantes son:

$$f_1 \simeq 496 \text{ Hz}$$

$$f_2 \simeq 1080 \text{ Hz}$$

Según la figura 2, esto corresponde a una vocal «o».

Problema 4.

Apartado A. Respecto al pitch, la distancia entre los «deltas» (armónicos) que se observan en la DFT es de unos 41 bins y por tanto 200 Hz sería la frecuencia fundamental.

Apartado B. Siguiendo el mismo procedimiento que en el problema 2, la envolvente alcanza valores altos en aproximadamente $n_1 \simeq 110$; $n_2 \simeq 360$ que corresponden a las frecuencias $f_1 \simeq 537$ Hz; $f_2 \simeq 1758$ Hz.

Apartado C. En base a los dos formantes, la vocal parece ser una «e». La distancia es más grande de lo habitual.

Problema 5.

Apartado A. Sí, porque podemos apreciar múltiples sincs (armónicos) y separarlas con suficiente claridad.

Apartado B. Si la vocal es una «u», los formantes deben ser aproximadamente 380 Hz y 900 Hz, y dado que $N = 500$ y corresponden a los bins 18 y 63, la frecuencia de muestreo que más cuadra es 8 kHz.

Apartado C. La frecuencia fundamental es de 9 bins y por tanto 144 Hz. Es muy arriesgado decir si corresponde a un hombre o a una mujer, aunque es ligeramente más probable lo primero.

Problema 6.

Apartado A. La frecuencia fundamental 100 Hz se situa en el bin 206, y $N = 2048$, con lo cual aislamos f_s de $f = n/N f_s$ y nos queda $f_s = fN/n \simeq 994 \text{ Hz} \approx 1 \text{ kHz}$.

Apartado B. Por el enunciado deducimos que el segmento de senoide ha sido enventanado con una ventana rectangular. La longitud de la ventana se relaciona con la anchura del lóbulo principal (ambos en muestras) según la expresión $L = \frac{2N}{k} \simeq \frac{2 \cdot 2048}{8} \simeq 512$ muestras.

La duración equivalente en segundos según la frecuencia de muestreo anterior es $\frac{512}{1000 \text{ Hz}} \approx 0.5 \text{ s}$.

Apartado C. La amplitud de las dos sincs de una senoide de amplitud A para una ventana rectangular es $A \frac{L}{2}$. Según esa expresión y lo que observamos en la DFT, la amplitud de la senoide es $A \simeq \frac{1000 \cdot 2}{512} \approx 3.9$.