

## **FUNDAMENTOS DE SONIDO E IMAGEN**

**Grado en Ingeniería de Tecnologías de Telecomunicación. Universidade de Vigo**

---

### **Práctica 1. Análisis de sonido**

**Objetivos:** En esta práctica se trabajará para que el alumno adquiriera destrezas en el uso de herramientas y funciones MatLab que permiten leer y manejar ficheros de audio, así como realizar el análisis temporal y frecuencial de diversos tipos de señales de audio. Concretamente con la realización de esta práctica se pretende que el alumno aprenda a:

- a) Analizar los efectos de la cuantificación
- b) Realizar el análisis en el tiempo de la señal de sonido: existencia o no de periodicidades (cálculo del periodo fundamental y frecuencia fundamental), aleatoriedad, variaciones de energía...
- c) Analizar el efecto del enventanado temporal de las señales de audio. Tipos de ventanas.
- d) Analizar el espectro haciendo uso de la DFT (Discrete Fourier Transform).
- e) Analizar el comportamiento espectral de las señales de audio a lo largo del tiempo utilizando el espectrograma: espectrograma de banda ancha y espectrograma de banda estrecha.
- f) Analizar el comportamiento espectral y temporal de las señales de voz: pitch, patrones armónicos, formantes y tracto vocal.

**Duración:**

Tres sesiones de 2 horas cada una.

**Evaluación:**

Recuerda que los contenidos del laboratorio son materia a evaluar.

---

### **Introducción**

El paquete matemático MatLab será el empleado en las clases prácticas de la asignatura por lo que deberéis recordar los conocimientos básicos de MatLab obtenidos en otras asignaturas. El material necesario para realizar la práctica se encuentra disponible en faiTIC. Al inicio de la sesión se deberán descargar, al ordenador en el que se va a realizar la práctica, los ficheros necesarios para esta práctica, de este modo se garantiza que en todo caso la práctica se realiza utilizando la versión original de dichos ficheros. Al finalizar la sesión se deberán salvar los cambios importantes en el soporte adecuado (pendrive, cloud, teléfono móvil, etc.) y luego borrar los directorios y ficheros creados en el ordenador.

#### **1. Cuantificación de señales**

El paso de analógico a digital implica varios procesos donde uno de los principales es la cuantificación. Entendemos por cuantificar convertir en bits (en número binario) una cantidad analógica (un número real) de la que conocemos su rango (normalmente será un voltaje comprendido entre dos valores  $V_{min}$  y  $V_{max}$ ). El procedimiento más simple de cuantificación consiste en asignar un número de bits  $B$  y dividir el rango posible de la

señal continua en  $2^B$ . En el caso sencillo los intervalos son todos iguales y de longitud  $\Delta = \frac{V_{\max} - V_{\min}}{2^B}$ , denominado *escalón de cuantificación* (ver figura 1 izquierda). Los sistemas digitales almacenarán, procesarán y enviarán cada muestra como una palabra de B bits que realmente representa en qué intervalo está la muestra. Si es necesario convertir la señal de nuevo a analógica (eso siempre es necesario en audio para su reproducción), se tomará normalmente el valor central de cada intervalo, cometándose un error  $\epsilon$ , llamado *error de cuantificación* o *ruido de cuantificación*. El error máximo para algún valor será, para este tipo de cuantificación, la mitad del escalón de cuantificación.

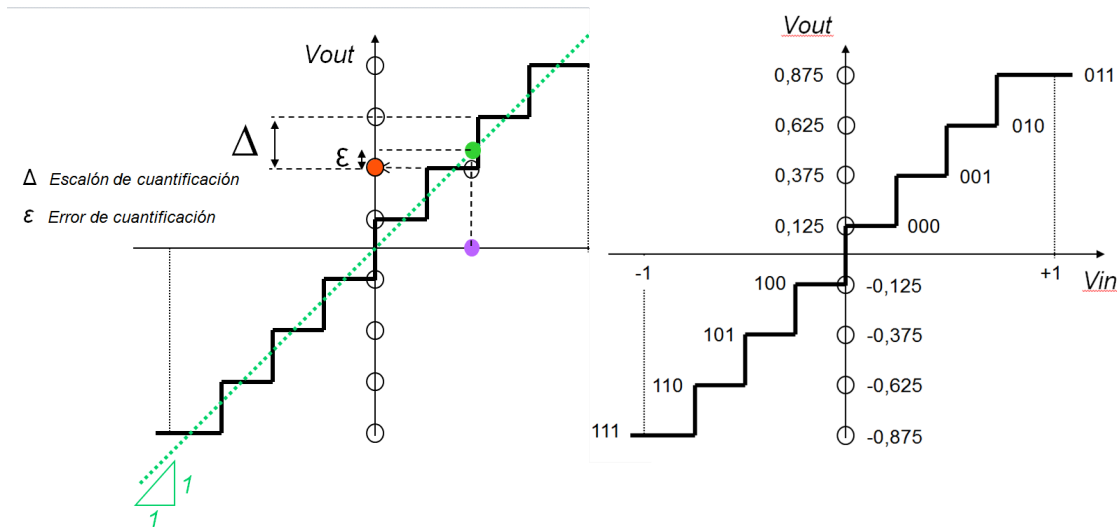


Figura 1: Explicación teórica y ejemplo práctico de la cuantificación

**Ejemplo numérico:** si una señal tiene un voltaje entre -1 V y +1 V y la representamos con  $B=3$  bits, el intervalo de cuantificación será  $\Delta = \frac{2}{8} = 0.250$  V. Habitualmente uno de los 3 bits de cada muestra representa el signo (+/-) y en la figura 1 derecha se corresponde con el más significativo.

Si la señal tiene un valor de, por ejemplo, 0.40 V, caerá en el segundo intervalo de valores positivos, como se indica en la Figura 1, por lo que será representada con el nivel de valor 0.375, con representación binaria 001, que es el valor del nivel de cuantificación disponible más próximo. El error de cuantificación que se comete:  $\epsilon = 0.4 - 0.375 = 0.025$  V.

Pasamos a ahora a analizar utilizando MatLab las señales almacenadas en los ficheros `senal1.mat` y `senal2.mat`. Para ello las leemos desde el MatLab con el comando `load` y las incorporamos al espacio de trabajo como `senal1` y `senal2`.

Sabiendo que estas señales se han muestreado a una frecuencia de muestreo de 1000 Hz y sus muestras se han almacenado utilizando 16 bits para `senal1` y 3 bits para `senal2`, responde a las siguientes cuestiones para cada una de las anteriores señales:

- ¿Cuántas muestras tiene la señal?
- ¿Cuál es la duración en segundos de la señal?
- ¿Cuál es el tamaño del escalón de cuantificación?
- ¿Cuál es el mayor valor del error de cuantificación?

Utilizando el comando plot de MatLab haz una representación gráfica de las dos señales.

```
>> plot(1:length(senal1),senal1, 1:length(senal2),senal2)
fx >> |
```

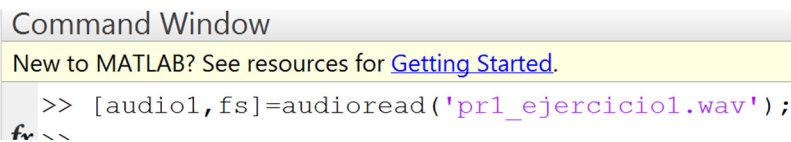
Analiza y justifica las diferencias entre ambas.

## 2. Lectura de ficheros de audio

MatLab dispone de funciones específicas para leer, escribir y escuchar ficheros de audio. Estas son [audioread](#), [audiowrite](#), [sound](#) y [soundsc](#). Utiliza el comando [help](#) para tener una descripción detallada de las capacidades que ofrece cada una de estas funciones.

Vamos a analizar primero señales que representan tonos. En primer lugar trabajaremos con la señal de audio del archivo pr1\_ejercicio1.wav.

- a) Lee la señal de audio del archivo pr1\_ejercicio1.wav utilizando el comando [audioread](#):



```
Command Window
New to MATLAB? See resources for Getting Started.
>> [audio1,fs]=audioread('pr1_ejercicio1.wav');
fx \>
```

La variable fs contiene la frecuencia con la que se ha muestreado. ¿Cuál es dicha frecuencia de muestreo?

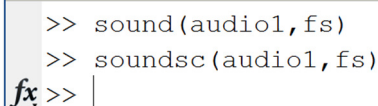
La función [audioinfo](#) nos proporciona información sobre el número de bits utilizado en la representación de cada muestra lo que nos permite responder a las siguientes cuestiones sobre la señal de audio:

- b) ¿Cuál es la frecuencia de muestreo?
- c) ¿Cuántas muestras tiene la señal?
- d) ¿Cuál es el número de bits/muestra?
- e) ¿Cuál es la duración en segundos de la señal?
- f) ¿Es una señal estéreo o mono?

Una vez leído el fichero de audio lo podemos escuchar utilizando la función [sound](#) o la función [soundsc](#).

- g) ¿Qué diferencias existen entre estas dos funciones?

Escucha el fichero utilizando ambos comandos:



```
>> sound(audio1,fs)
>> soundsc(audio1,fs)
fx >> |
```

- h) ¿Qué tipo de sonidos has escuchado?

### 3. Análisis de tonos individuales

En esta sección vamos a seguir trabajando con la señal de audio del archivo `pr1_ejercicio1.wav`.

#### 3.1 Forma de onda

Para representar la forma de onda de la señal podemos hacer uso de la función MatLab `plot`.

Representa en una figura la señal de audio del ejercicio anterior con el eje de tiempos en segundos. Para ello primero tienes que definir el vector `t` que contiene los instantes de tiempo en segundos de cada muestra:

```
>> %construimos el eje de tiempos:
t=0:1/fs:(length(audio1)-1)/fs;
```

y luego utilizando el comando `plot` hacer su representación temporal.

```
>> %hacemos una representación temporal de la señal
figure(1)
plot(t,audio1);
```

Vemos que la señal tiene una gran duración y que la cantidad de muestras es muy grande para observarlas en una figura utilizando un solo `plot`, pero podemos ver claramente que hay espacios más saturados (a la derecha) que otros (los de la izquierda). Con ayuda del comando o función `axis` (utiliza el comando `help axis` para comprender su funcionamiento) se puede elegir de forma sencilla una zona de observación de la señal (hacer un zoom) lo que permite hacer un análisis más detallado de la misma.

Utilizando el comando `axis` (modificando los límites de forma conveniente) visualiza en una misma figura, utilizando el comando `subplot`, dos segmentos de 0.02 segundos de duración de la señal, uno al comienzo de la señal y otro al final.

```
>> %un trozo de 0.02 segundos de la parte inicial.
figure(2)
subplot(211); plot(t,audio1);
axis([0 0.02 -1 1])
>> %segmento de 0.02 del final
subplot(212); plot(t,audio1);
axis([2.98 3 -1 1])
%...
```

Puedes utilizar el comando `axis` modificando los límites de forma conveniente para los análisis y visualizaciones que necesites realizar sobre la forma de onda de la señal.

- En las dos gráficas anteriores tenemos la misma cantidad de muestras, sin embargo, lo que vemos es bastante diferente. ¿Cuál es el periodo fundamental y la frecuencia fundamental en cada parte de la señal?

Todas estas acciones también se pueden realizar con la aplicación “signal analyzer” de MatLab. Para ello ir al menú APPS y seleccionar la aplicación “signal analyzer” pasar la señal audio del Workspace de MatLab al panel de “filter signals”. Para ver detalles de la señal seleccionar panner del menú superior y hacerlo muy estrecho para observar los detalles de la señal.

### 3.2 Análisis del espectro

A continuación vamos a realizar el análisis del espectro de la señal de audio utilizando la función `ver_espectro` que os proporcionamos. Esta función cuya llamada desde MatLab se realiza como:

`ver_espectro(senal,ventana,Nfft,fs)`

Esta función calcula y representa el espectro de una señal almacenada en la variable `senal`, utilizando el tipo de ventana indicado en `ventana` y una FFT de `Nfft` puntos (`senal`, `ventana` y `Nfft` son los parámetros de entrada a la función).

Abre con el editor el fichero `ver_espectro.m` que contiene la función `ver_espectro` y analiza con cuidado el código para entender como realiza el cálculo y representación del espectro.

- a) Explica brevemente como se calcula y representa el espectro.
- b) Realiza un análisis del espectro de la señal de audio `pr1_ejercicio1.wav` del apartado anterior utilizando una ventana hamming de 1024 muestras. ¿Qué observas? ¿Es correcto este espectro para el tramo inicial? ¿Y para el tramo final?
- c) Realiza ahora un análisis del espectro de los últimos 0.5 segundos de la señal de audio utilizando una ventana hamming de 1024 muestras. ¿Cuál es la diferencia respecto al resultado anterior? A la vista del espectro, ¿qué señal estamos analizando? ¿Cuál es su frecuencia fundamental? ¿Coincide con la obtenida en el apartado 2.1?
- d) Prueba distintos tamaños de la ventana de análisis temporal (64, 256, 512, 1024, 2048). ¿De qué duración temporal en segundos es la ventana de 64 muestras? ¿Y la de 1024? ¿Esta última es correcta para el trozo de señal que has seleccionado en la forma de onda temporal?
- e) Realiza ahora un análisis del espectro del tramo inicial de la señal utilizando los tamaños de ventana del apartado anterior. Utiliza en la representación del espectro un eje logarítmico de frecuencias (modificación en la función `ver_espectro`). ¿Qué sucede cuando utilizamos una ventana de 64 muestras? ¿Por qué?
- f) Vamos a analizar ahora los efectos de distintos tipos de ventana de análisis temporal: rectangular, hamming, hanning. Utiliza en primer lugar una ventana rectangular de 1024 muestras, fíjate en el valor en dB del pico, ¿a qué valor de potencia equivale en unidades naturales? ¿Y el valor de los lóbulos secundarios? Repite para una ventana hanning y comenta las diferencias.
- g) Supón que el enventanado de la señal se realiza con un solapamiento del 50%. Haz un cálculo del número aproximado de ventanas de 1024 muestras que se emplean para analizar 1.5 s de la señal.

### 3.3 Espectrograma

Como habéis visto en los apuntes del Tema 1 el espectrograma es una gráfica que permite visualizar cómo varía el comportamiento espectral de una señal con el tiempo. En el eje horizontal, se representa el paso del tiempo mientras que en el eje vertical, se representa la frecuencia. La mayor o menor concentración de energía se visualiza con la calidez del color; cuanto más cálido, mayor es la energía en esa zona. Si el espectrograma es en escala de grises, cuanto más oscura es la tonalidad, mayor es la energía.

En MatLab se dispone de la función `spectrogram` que realiza el cálculo y representación del espectrograma de una señal que se le pasa como parámetro. Analizar y estudiar que hace esta función, cuáles son sus parámetros de entrada y de salida utilizando el comando `help` de MatLab o abriendo la página de referencia de la función `spectrogram` desde el Help.

```
[s,w,t] = spectrogram(x>window,noverlap,f,fs)
% Cálculo del espectrograma utilizando la transformada de Fourier "short-time"
% Parámetros de entrada
% x: señal de entrada
% window: ventana que se utiliza para dividir la señal y hacer en enventanado
% noverlap: número de muestras de la señal que se solapan entre ventanas consecutivas.
% f: vector de frecuencias donde se calculará el espectro
% fs: frecuencia de muestreo utilizada
%
% Parámetros de salida
% s: señal de entrada
% w: vector con las frecuencias normalizadas en las que se ha calculado el espectrograma
% t: vector con los instantes de tiempo en los que se ha calculado el espectrograma
```

Si ejecutáis esta función sin parámetros de salida, se dibuja el espectrograma en la ventana de figure que está activa.

Hay dos tipos de espectrogramas, dependiendo de la longitud temporal de la ventana. Cuando la ventana temporal es larga, el espectrograma se denomina de banda estrecha mientras que si la ventana temporal es corta, el espectrograma se denomina de banda ancha. Vamos a analizar cada uno de ellos.

### 3.3.1 Espectrograma de banda estrecha

Crea un script MatLab con las siguientes líneas de código:

```
1 window=hamming(2048);
2 noverlap=0;
3 Nfft = 2048;
4 spectrogram(audio1>window,noverlap,Nfft,fs,'yaxis');
5
```

- a) ¿Qué se representa al ejecutar el script? A la vista de la gráfica que se obtiene, ¿qué tipo de sonidos contiene la señal almacenada en el vector audio?

- b) Al igual que hicimos en el apartado 2.1 con el comando axis podemos controlar lo que se ve: `axis([0 2.5 0 2])`
- c) Haz un zoom de la parte final, del espectrograma. ¿Cuál es la frecuencia del tono?
- d) Haz un zoom de la parte inicial, ¿puedes obtener con precisión la frecuencia del tono?
- e) Observa qué sucede en torno a 1.5s y relaciona este efecto con el tamaño de la ventana.

### **3.3.2 Espectrograma de banda ancha**

Elige ahora ventanas de 128 y de 256 puntos para el cálculo y representación del espectrograma.

- a) ¿Qué sucede con la precisión de cálculo de las frecuencias de los tonos?
- b) ¿Qué sucede en la zona de transición de ambos tonos?

El espectrograma de banda estrecha se consigue tomando ventanas de análisis de gran longitud por lo que la precisión espectral es muy grande mientras que el de banda ancha utiliza ventanas temporales de longitud pequeña con lo que su precisión espectral es baja pero hace un buen seguimiento de los cambios temporales de la señal.

## **4. Análisis de una señal de audio con frecuencia fundamental y amplitud variable en el tiempo**

Realiza el análisis temporal y del espectro de la señal de audio contenida en el fichero `pr1_ejercicio2.wav` y responde a las siguientes cuestiones. Edita para ello un script en el que vayas escribiendo los comandos que vayas empleando para resolver los distintos apartados.

- a) Escucha la señal y comenta los cambios de volumen respecto a la máxima amplitud de cada zona. Este efecto se explicará con detalle en el Tema 3 de Percepción Sonora.
- b) Obtén a partir de la ventana de la forma de onda, la frecuencia fundamental del tramo inicial, la del tramo entre 0.8 y 1 s y la del tramo final. ¿Qué sucede en este último caso?
- c) Selecciona el tramo final y mediante un análisis del espectro obtén la frecuencia fundamental, ajustando convenientemente los parámetros del análisis.
- d) Repite el apartado anterior para el tramo inicial.
- e) Utilizando el espectrograma ajusta de forma conveniente los parámetros para su cálculo de forma que puedas obtener las frecuencias fundamentales de cada intervalo.

## **5. Procedimiento para el análisis de señales reales**

Un análisis correcto de señales de audio reales requiere analizar:

- ✓ La variación de la energía de la señal en el tiempo. Si es más o menos cambiante, si es constante o no. Además es importante identificar en el tiempo periodos de repetición, si existen.

- ✓ La variación de la energía en frecuencia. Si hay componentes tonales o no, si existen zonas ruidosas, si hay patrones armónicos o no, etc. También es importante identificar la posición precisa de los tonos en caso de haberlos.
- ✓ Variación de la energía en frecuencia a lo largo del tiempo (espectrogramas). Analizar los intervalos temporales con patrones armónicos, los de apariencia ruidosa, con bajas, medias o altas frecuencias. Variación de la frecuencia fundamental a lo largo del tiempo (como la entonación en la voz o la melodía en un instrumento) es otro elemento a tener en consideración.

En los siguientes apartados de la práctica vamos a analizar una señal de voz femenina y otra de voz masculina. Utilizaremos la metodología estudiada en los anteriores apartados para analizar su contenido frecuencial y temporal.

## **6. Análisis de una señal de voz femenina**

Vamos a trabajar en primer lugar con la señal de audio del archivo pr1\_ejercicio3.wav. En primer lugar lee y carga en el Workspace de MatLab la señal contenida en este fichero.

- a) ¿Cuál es la frecuencia de muestreo?
- b) ¿Cuántas muestras tiene la señal? ¿Cuál es la duración en segundos de la señal?
- c) ¿Cuál es el número de bits/muestra?
- d) ¿Es una señal estéreo o mono?

Escucha la señal.

Utilizando el comando `axis` haz un zoom de la vocal “a” y apunta su duración aproximada en ms y en muestras.

- e) Observa su cuasi-periodicidad y obtén la frecuencia fundamental.
- f) Elige una ventana de 256 muestras del tramo central de la “a” y visualiza su espectro (no el espectrograma) con ventana Hanning. ¿Hay componentes tonales? ¿Y patrones armónicos?
- g) Apunta los valores de la frecuencia fundamental (pitch) y de los 5 tonos siguientes, así como sus amplitudes (energía). ¿Son armónicos puros de la frecuencia fundamental?

Utilizando el comando `axis` haz un zoom de la última “s” pronunciada.

- h) Compara su amplitud media respecto a la de la vocal “a”. Observa su forma de onda. Teniendo en cuenta los apuntes del tema de “Análisis de sonido”, ¿a qué se asemeja esta forma de onda?
- i) Elige el intervalo de 1.15 a 1.30 s. ¿Cuál es su duración en muestras? Visualiza su espectro con ventana Hanning de 1024 muestras. Compara el valor de potencia máximo con el de la vocal “a”. Empleando el eje de frecuencias lineal, ¿observas componentes tonales?

Calcula y representa un espectrograma utilizando ventanas hanning de 1024 muestras.

- j) ¿Puedes distinguir las zonas donde vibran las cuerdas vocales de la locutora?



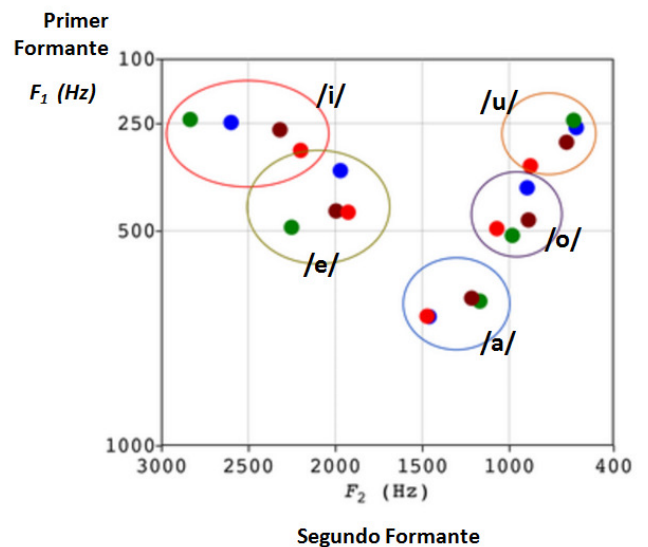
- k) Amplía la zona temporal de la /a/, ¿distingues los armónicos que aparecían en el espectro? ¿Coinciden sus frecuencias? ¿La intensidad de los colores cálidos se corresponde con los picos de mayor/menor energía vistos anteriormente en el espectro? Puedes usar la opción “insert color bar”.
- l) Haz un seguimiento temporal del pitch y comenta la evolución de la entonación.

Calcula y representa un espectrograma con ventanas de 128 muestras

- m) Analiza la evolución de la energía en la banda comprendida entre 200 a 700Hz.
- n) Ídem para la banda entre 700 a 3000 Hz.

La evolución de estas 2 bandas (**formantes**) en el espectrograma de banda ancha nos va indicando el movimiento del **tracto vocal** cuando pronunciamos vocales.

*El tracto vocal comprende la cavidad oral (desde la laringe hasta los labios) y la cavidad nasal acoplada a la oral a través del velo y puede considerarse como un filtro con cinco zonas de resonancias (5 formantes) por el que pasa el aire que viene de los pulmones. Si las cuerdas vocales están vibrando, aparecen las rayas horizontales en el espectrograma de banda estrecha. Para las vocales, los formantes principales están en las bandas analizadas anteriormente. En la siguiente figura, se observan unos promedios de las frecuencias de los dos primeros formantes de las cinco vocales del español.*



## 7. Análisis de una señal de voz masculina

Vamos a trabajar ahora con la señal de audio del archivo pr1\_ejercicio4.wav.

En primer lugar lee y carga en el Workspace de MatLab la señal contenida en este fichero.

- a) ¿Cuál es la frecuencia de muestreo?
- b) ¿Cuántas muestras tiene la señal? ¿Cuál es la duración en segundos de la señal?
- c) ¿Cuál es el número de bits/muestra?
- d) ¿Es una señal estéreo o mono?

Escucha la señal.

Utilizando el comando `axis` haz un zoom del intervalo que comprende “aba”. Todos estos fonemas son sonoros, es decir, las cuerdas vocales vibran al pronunciarlos.

- a) Observa su cuasi-periodicidad. Obtén el periodo fundamental y la frecuencia fundamental.

Calcula y representa un espectrograma utilizando ventanas hamming de 2048 muestras.

- b) ¿La duración de estas ventanas en segundos es mayor que las ventanas aplicadas para el espectrograma de banda estrecha de la voz femenina?
- c) Haz un zoom en el espectrograma del tramo “*aba*”. Obtén la frecuencia del pitch. ¿Hay más o menos energía en esta frecuencia fundamental o en sus (cuasi) armónicos?

Calcula y representa un espectrograma de 256 muestras y haz un zoom de la frase hasta las sílabas “*nismo*”.

- d) Escucha los tramos donde se observa energía por encima de 5kHz. ¿Con qué consonantes se corresponden?
- e) En el espectrograma identificas algún sonido de tipo “transitorio”.