

# Proyecto 2.

Reconocimiento de voz

Series de Fourier

---



**Integrantes:**

Maximiliano Uribe  
Dante Valenzuela  
Nicolas Jara  
Fernando Guerra  
Jose Luis Calderon  
Boris Berrios

**Asignatura:** Matemáticas IV

**Profesor:** Eric Jeltsch

# Índice.

- 1.- Introducción
- 2.- Objetivos.
- 3.- Estructura del proyecto
- 4.- Información sobre este proyecto
- ...
- 8.- Bibliografía

## **Introducción.**

La voz humana es una de las formas más naturales y complejas de comunicación. Cada persona posee un timbre, tono e intensidad característicos que dependen de las frecuencias presentes en su emisión vocal. El análisis de estas propiedades permite estudiar y clasificar distintos tipos de voces, información que puede aplicarse en áreas como la música, la ingeniería acústica o el reconocimiento de patrones.

A través de un micrófono vamos a capturar voces y en base a los datos que se puedan sacar con las fórmulas matemáticas será posible clasificar el tipo de voz (por ejemplo, grave o aguda) y más características que nos ayuden a guardar un archivo de voz en su mejor modo. La idea de esto es poder desbloquear un archivo privado identificando el tipo de voz ya guardada para ese archivo.

## **Objetivos del proyecto:**

El objetivo principal de nuestro proyecto es analizar la voz de una persona utilizando **Python** y las **transformadas de fourier**, con el fin de crear un reconocimiento de voz que ayude a resguardar algún tipo de objeto privado y/o archivo. buscando determinar la frecuencia fundamental, los armónicos, y la energía de la señal, lo cual nos dará el tono, timbre e intensidad de la voz.

Con el objetivo de que resguarde la voz reconocida y hacer que esta guarde algún archivo y/o documento importante.

Con esta información será posible distinguir diferentes tipos de voz, compararla con el reconocimiento original y así que solo tome el reconocimiento del usuario original para que la contraseña sea correcta.

## ESTRUCTURA DEL PROYECTO:

### PROCESOS:

- Conexión del micrófono.
- Obtención y Recepción de los datos hacia Python.
- Aplicación de la Transformada de Fourier a los datos de voz.
- Análisis de las frecuencias para identificar tono, timbre y energía.
- Comparación de audios y reconocimiento de voz

### Métodos:

- **Investigación:** recopilar información sobre el funcionamiento del micrófono y la Transformada de Fourier.
- **Procesamiento digital:** utilizar Python con bibliotecas como `numpy` y `matplotlib` para realizar la FFT y visualizar resultados.
- **Simulación y prueba:** probar distintas voces para verificar la respuesta del sistema.
- **Análisis comparativo:** comparar los espectros de los dos audios para hacer un análisis y su respectiva verificación

## Este proyecto:

Para poder hacer realidad el objetivo de este proyecto primero debemos conocer algunos conceptos como:

### **¿Que es una señal de voz?**

Es una representación acústica de los sonidos producidos por la voz humana, que presentan variaciones de presión que se propagan a través del aire como ondas sonoras. Puede capturarse y convertirse en una señal eléctrica a través de un micrófono para procesamiento, transmisión y reproducción.

### **¿Qué características son importantes de ella?**

Se caracteriza por propiedades como:

**El timbre:** es la cualidad que nos permitirá diferenciar una voz de otra, incluso si tienen el mismo tono y volumen. Está relacionada con la complejidad de la onda sonora y la resonancia.

**El ritmo:** Es la cadencia y el patrón temporal del habla, incluyendo pausas y la duración de las sílabas

**El volumen:** Se relaciona con la energía de la señal y puede variar según lo que intente transmitir.

**La duración:** longitud del sonido al igual que la velocidad del habla

**La entonación:** Es la variación del tono de la voz a lo largo del tiempo y es crucial para la expresividad y la emoción.

Características temporales:

**Energía de la señal:** indica la potencia de la señal en un momento dado.

**Tasa de cruce por cero:** el número de veces que la forma de onda de la señal cruza el eje horizontal en un intervalo de tiempo.

Características espectrales:

**Frecuencia fundamental:** Determina la altura del tono y es el resultado de la frecuencia de vibración de las cuerdas vocales.

**Espectro:** la distribución de la energía de la señal a través de las diferentes frecuencias que la componen.

## **¿Qué fórmulas ocuparemos y en qué?**

En este proyecto deseamos implementar en específico 2 fórmulas con respecto a Fourier las cuales son la Transformada rápida de Fourier y la Transformada Inversa de Fourier las cuales en pocas palabras nos ayudarán a encontrar las frecuencias y eliminar ruido externo al que queremos.

### **¿Qué es la transformada de fourier?**

Es una herramienta matemática que descompone una señal o función en una suma de funciones sinusoidales (oscilación suave y periódica) de diferentes frecuencias, lo que permite ver la señal desde el dominio de la frecuencia en lugar del dominio del tiempo.

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt$$

### **¿Qué muestran sus resultados?**

Muestra los componentes de frecuencia de una señal, esto representa gráficamente las frecuencias en el eje X y las amplitudes de cada frecuencia en el eje Y. En otra manera de explicar, transforma una señal del dominio del tiempo al dominio de la frecuencia.

**Componentes de una frecuencia:** como el sonido o una vibración.

**Amplitud de cada frecuencia:** muestra que tan fuerte es cada componente de frecuencia individual en la señal original.

**Información oculta en el tiempo:** en lugar de ver cómo cambia una señal en el tiempo, se ve la composición frecuencial, lo que permite un análisis más detallado.

**Análisis de ruido y vibraciones:** permite la identificación de ruido y vibraciones no deseados en una señal, facilitando su eliminación.

### **¿Qué es la transformada inversa de Fourier?**

La Transformada inversa de Fourier sirve para lo contrario a la TFF ya que esta permite reconstruir una señal en el dominio del tiempo a partir de su representación en el dominio de la frecuencia. Ayudando en procesos de filtrado, como la eliminación de ruido no deseado en grabaciones, hace posible modificar el espectro de la señal y luego obtener nuevamente la señal

original limpia, De esta manera, se puede analizar con mayor precisión la energía y las frecuencias reales de la voz.

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega$$

Y no puede faltar el conocimiento del código:

**¿Cómo funciona la recepción del audio y generación de datos?**

...

**¿Qué hará Python con los datos que se obtengan?**

los datos que se obtengan gracias al micrófono serán serán analizados y guardados en un archivo .wav para poder probar que se guarde o se escuche, luego limpiara el ruido externo de el archivo siendo más más limpia la señal. con ello podremos pedir autenticación autenticación de un archivo que crearemos luego para que python reconozca la señal de voz y pueda abrir el archivo.

**¿Qué bibliotecas usaremos?**

usaremos numpy, sound device, soundfile, scipy.signal, pickle, os

finalizando con las preguntas:

**¿Qué resultados esperamos?**

**¿Para qué nos sirve este proyecto en la vida cotidiana?**

**¿Para qué nos sirve en el ámbito matemático?**

## BIBLIOGRAFÍA.

Ideas del proyecto:

[https://scholar.google.com/scholar?hl=es&as\\_sdt=0%2C5&q=A+software+tool+for+the+analysis+of+vocal+performances+m+howard&btnG=#d=gs\\_qabs&t=1761673852905&u=%23p%3DKvzcC4sV3fAJ](https://scholar.google.com/scholar?hl=es&as_sdt=0%2C5&q=A+software+tool+for+the+analysis+of+vocal+performances+m+howard&btnG=#d=gs_qabs&t=1761673852905&u=%23p%3DKvzcC4sV3fAJ)

<https://www.espaicoriveu.com/diferentes-tipos-de-voz-y-como-distinguirlos>

<https://rua.ua.es/server/api/core/bitstreams/a631d1ee-5df5-4896-baee-da9ca6fd3e41/content>