Capstone D: Tying it with Music - Fourier Analysis

Table of Contents

Sound as a Vector: The Digital Signal

How can we analyze sound with linear algebra? First, we represent it as a vector.

- **Analog to Digital**: A microphone captures a sound wave, which is a continuous function of pressure over time, `p(t)`.
- **Sampling**: To make it digital, we sample this function at regular intervals. If we take `N` samples, we get a sequence of numbers: `(s_1, s_2, ..., s_N)`.
- **The Sound Vector**: This sequence can be treated as a vector `s` in an `N`-dimensional vector space, `R^N`.

This vector `s` lives in the **time domain**. Its components represent the amplitude of the sound at each moment in time.

The Big Idea: Decomposing Sound into Harmonics

Any complex sound, like the note from a violin or a human voice, is actually a combination of many simple, pure sine waves of different frequencies and amplitudes. These pure tones are called **harmonics** or **overtones**.

- **Fundamental Frequency**: The lowest frequency, which determines the pitch we hear (e.g., A4 at 440 Hz).
- **Harmonics**: Integer multiples of the fundamental frequency (e.g., 880 Hz, 1320 Hz, etc.).
- **Timbre**: The unique "color" or quality of a sound is determined by the relative strengths (amplitudes) of its harmonics.

Fourier Analysis is the mathematical tool that allows us to take a complex sound vector `s` and decompose it into its constituent sine and cosine waves. It's like taking a smoothie and figuring out the exact recipe of fruits that went into it.

Fourier Analysis as a Change of Basis

This is where linear algebra provides a beautiful insight. Fourier analysis is simply a **change of basis**.

- **The Time Basis (Standard Basis)**: Our original sound vector `s` is represented in the standard basis, where each basis vector `e_i` represents a single point in time.

$$s = s_1e_1 + s_2e_2 + \cdots + s_Ne_N$$

- **The Frequency Basis (Fourier Basis)**: The Fourier basis is a set of **orthogonal** sine and cosine wave vectors of different frequencies. Let's call them `f_k`.
  - These basis vectors are themselves samples of pure sine and cosine waves.
- **The Transformation**: The **Discrete Fourier Transform (DFT)** is a linear transformation (represented by a matrix, `F`) that converts a vector from the time basis to the frequency basis.

$$c = Fs$$

  - `s` is the vector in the time domain.
  - `c` is the vector of coefficients in the **frequency domain**. The components of `c` tell us the amplitude and phase of each frequency component in the original sound.

The core idea: We are rewriting our sound vector `s` as a linear combination of pure frequency vectors.

$$s = c_1f_1 + c_2f_2 + \cdots + c_Nf_N$$

Unitarity of the DFT: If `F` is the DFT matrix with entries `F_{k\ell} = e^{-2\pi i k\ell / N}`, then the scaled matrix `\tilde F = \frac{1}{\sqrt{N}}F` is **unitary**:

$$\tilde{F}^* \tilde{F} = I.$$

This guarantees that energy (the inner product) is preserved when moving between time and frequency domains.

Connection to the Spectral Theorem

The Fourier basis is not just any basis; it's an **orthonormal eigenbasis** for a specific class of operators related to time-invariance and frequency.

- **The Operator**: Consider the **circular shift operator** `C`, which shifts the components of a vector cyclically. This operator represents a time delay.

$$C \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_N \end{pmatrix} = \begin{pmatrix} s_N \\ s_1 \\ \vdots \\ s_{N-1} \end{pmatrix}$$

- **The Eigenvectors**: The eigenvectors of this shift operator are precisely the **complex exponentials** (which are combinations of sines and cosines) that form the Fourier basis.
- **The Spectral Theorem**: The shift operator `C` is a **normal matrix** (`C^†C = CC^†`), and the Spectral Theorem guarantees that it can be diagonalized by a unitary matrix. That unitary matrix is the Fourier matrix `F`!

In essence, the Spectral Theorem tells us that we can find a special basis (the Fourier basis) where the complicated action of a time shift becomes a simple multiplication. This is why Fourier analysis is so powerful for analyzing signals.

- **Superposition = Chords**: A musical chord is a literal superposition of notes. Fourier analysis decomposes the resulting complex waveform back into its constituent notes.
- **Spectral Theorem = Decomposing Timbre**: The timbre of an instrument is its harmonic spectrum. The Spectral Theorem provides the mathematical guarantee that we can decompose any sound into this unique spectrum.

Applications: Compression, Synthesis, and Effects

Once a sound is in the frequency domain, we can easily manipulate it.

- **Audio Compression (MP3, AAC)**: Human hearing is less sensitive to certain high frequencies. In the frequency domain, we can simply set the coefficients for these frequencies to zero (or use fewer bits to store them). This is called lossy compression.
- **Audio Synthesis**: To create the sound of a specific instrument, synthesizers add together sine waves with the correct harmonic amplitudes, effectively building the sound from its frequency-domain representation.
- **Equalizers (EQ) and Effects**: An equalizer is a tool that boosts or cuts the amplitudes of specific frequency ranges. This is a simple multiplication in the frequency domain. Effects like reverb or chorus are also implemented by manipulating the signal in this domain.

Practice Problems

1. If you sample a sound for 1 second at 8 Hz (8 samples per second), what is the dimension of the vector space containing your sound vector?
2. A C major chord is made of the notes C, E, and G. How would you represent this chord from a Fourier analysis perspective?
3. Why is it more efficient to apply an EQ (e.g., boosting the bass) in the frequency domain rather than the time domain?
4. MP3 compression is "lossy." In the context of Fourier analysis, what information is being "lost" and why is it generally acceptable?
5. **Conceptual**: How is the idea of decomposing a sound into harmonics similar to the idea of finding the principal components of a dataset in PCA?