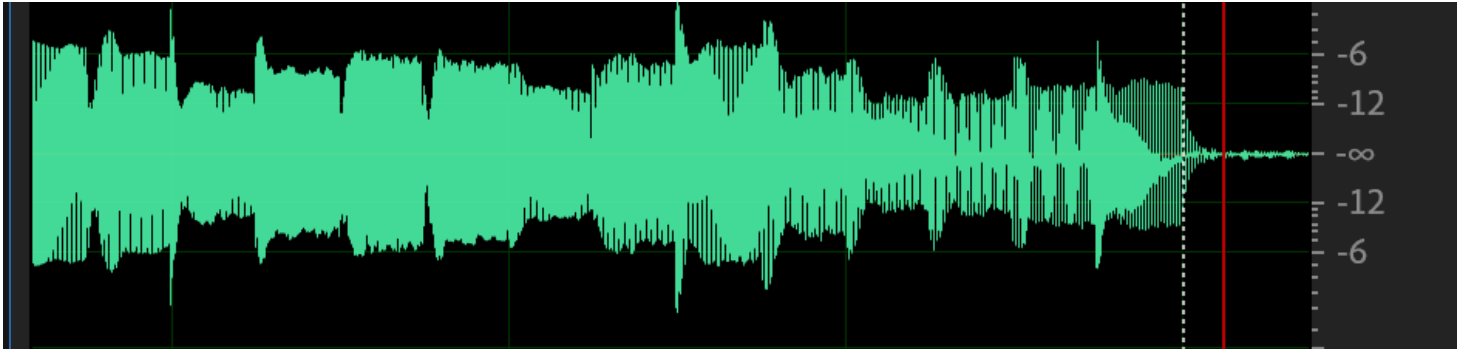# Segment 1 Fourier Transform (continuous)

This is a ten second audio clip by piano.

The x-axis is for time, and y-axis is for intensity.

Let's zoom in a little bit, we see different frequencies happen through time.

Let's just focus on the left hand side part. They appear to be similarly repeated.

From what we learn from the lectures, we know they are formed by sum of perfect sine or cosine waves. And the wave components can be solved by the **Fourier Transform**.

$$F\left(\omega_k\right) \equiv \int_{-\infty}^{\infty} f(t)e^{-2\pi ikt}\ \mathrm{d}t, \quad k \in (-\infty, \infty)$$

Intuitively, this method winds the wave around in circular ways,

```
Animation:
indicate
```

$$\frac{1}{N} \sum_{k=1}^{N} g(t_k) e^{-2\pi i f t_k}$$
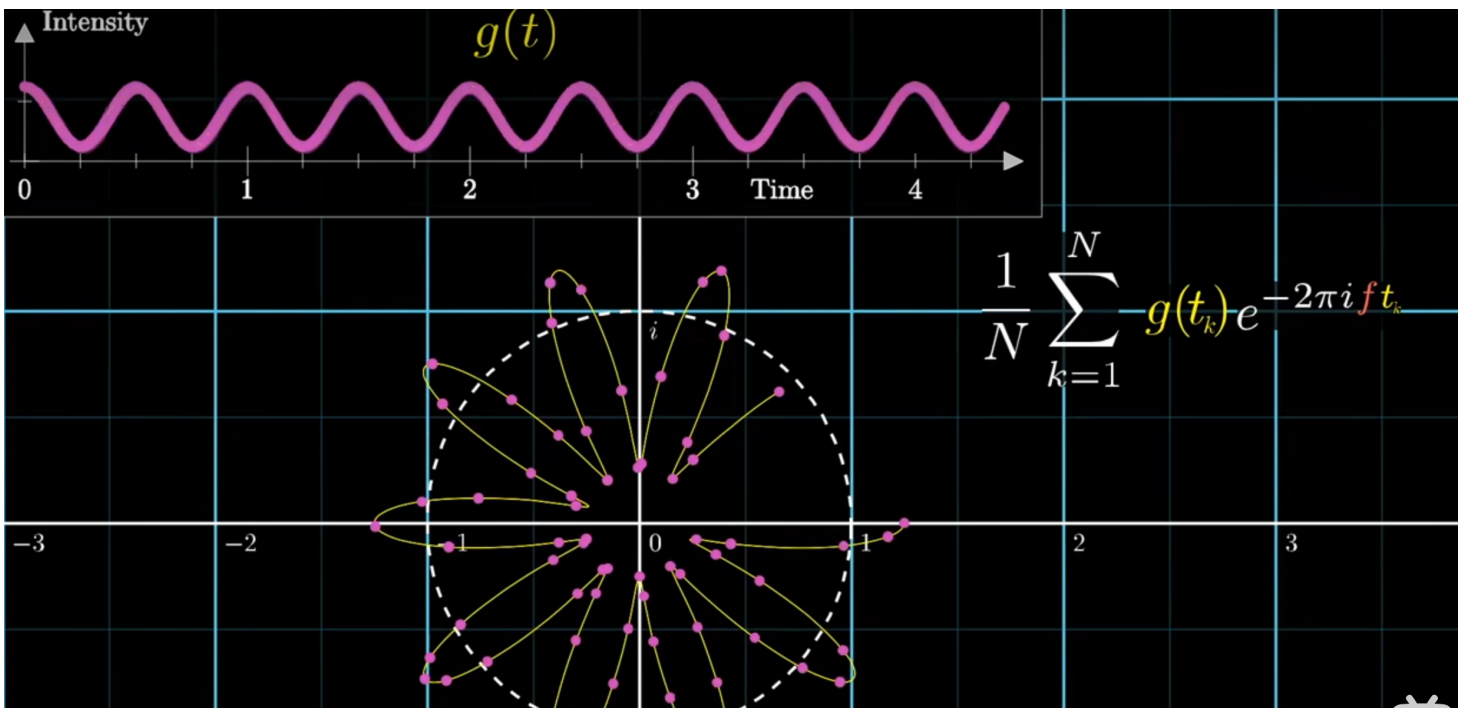
then traces the CM of the resultant graph.

If we record all results of winding frequency about the positijnm on of CM, we get the graph for **frequency domain**.

Segment 2 DFT

However, if we really want to use it in practice, problems arise.

We never get continuous signals. Instead, we get discrete samples from the signal source.

If we continue to zoom in on our sound, we see these dots.
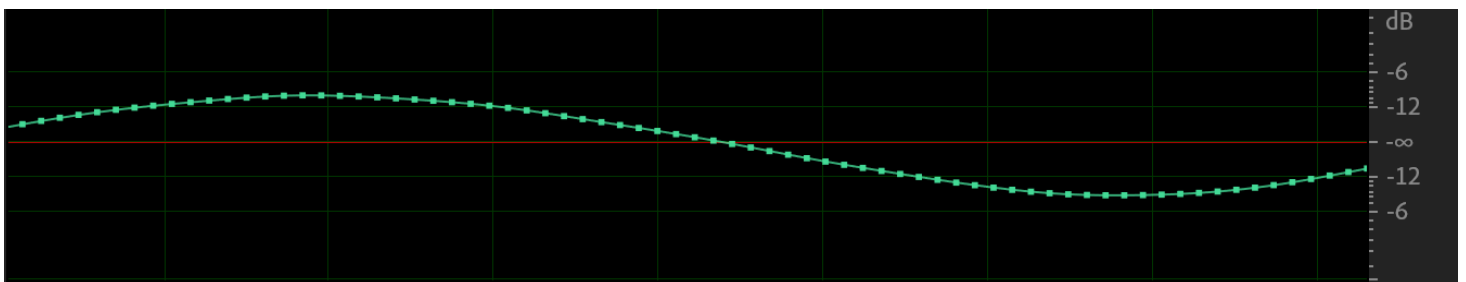
Animation

This is how we store the audio in computer. And we require the **discrete version of Fourier Transform**,

Here is the formula for dicrete Fourier Transform.

$$\hat{f}_k = \sum_{j=0}^{n-1} f_j e^{-i2\pi kj/n}$$

$$k = 0, 1, 2, ..., n - 1$$

Let's see what the indexes mean and we will find it similar to our continuous version.

Let's show this with an example. We encounter a 30 Hz cosine sound wave. We assume 100Hz of sampling rate, meaning we would get 100 samples if we record one second.

And we actually sample for 0.1s, i.e we have 10 samples.

We treat them as a vector.

Then, for k=2 for example, we apply the formula

—   —

$$
\begin{bmatrix} \vdots \\ \vdots \\ \hat{f}_2 \\ \vdots \\ \hat{f}_9 \end{bmatrix} = \left(\begin{array}{ccccc} \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ e^{-i2\pi(0)/N} & e^{-i2\pi(2)/N} & e^{-i2\pi(4)/N} & \cdots & e^{-i2\pi(2(N-1))/N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{array} \right) \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix}
$$

```
Animation:
for j=0:9
    Write e^{-2*pi*k*j/n} next to f_j,
    Draw this point on the polar co-ordinate system.
end;
Find center of mass on the graph;
Obtain and write f_bar_k on the left;
e^xxxxxxxx vertical, then move into the matrix here
```

What is the meaning of $\hat{f}_2$? Well, we see from animation that it is similar as continous version where we find the center of mass. But it is slightly different. We will go back to this later.

Now we continue the process for $k = 0, 1, ..., 9$.

```
Animation:
for all k :
    write row k in the matrix,
    plot 9 dots on the graph simultaniously,
    find CM,
    write f_hat_k
-> obtain a full matrix
```

$$
\begin{bmatrix} \hat{f}_0 \\ \hat{f}_1 \\ \hat{f}_2 \\ \vdots \\ \hat{f}_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w_n & \omega_n^2 & \cdots & w_n^{n-1} \\ 1 & w_n^2 & w_n^4 & \cdots & w_n^{2(n-1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & w_n^{n-1} & w_n^{2(n-1)} & \cdots & w_n^{(n-1)^2} \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ \vdots \\ f_{n-1} \end{bmatrix}
$$

$$
w_n = e^{-2\pi i/n}
$$

Now we come back to the meaning of $\hat{f}_k$. $\hat{f}_2/10$ is the amplitute of (2)*(100/10)=20 Hz cosine wave.

Furthermore, we can plot the amplitute of $\hat{f}_k$'s. We obtain the graph of frequency domain.
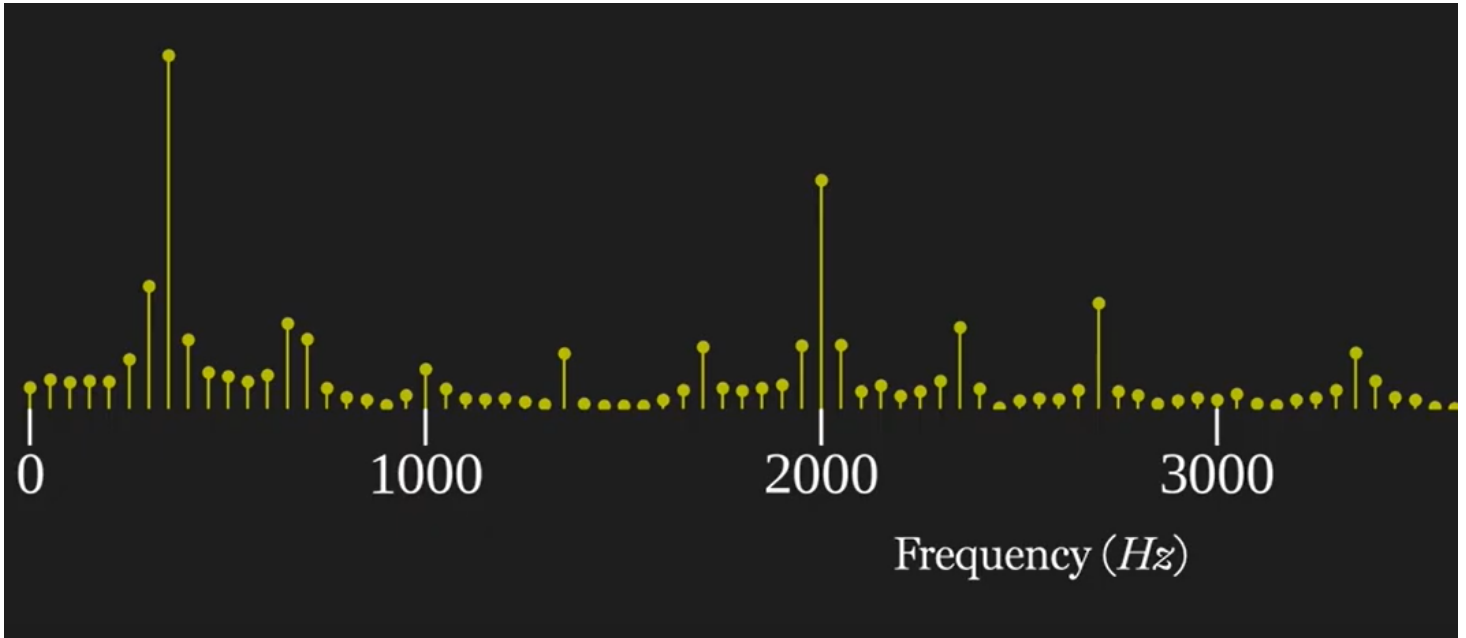
We see that when k=3, the $\hat{f}_3$ is large, meaning that 3*(100/10)=30 Hz cosine wave is a large components of the signal.

Reason of k=0...,9. The largest possible component is...

Segment 3: STFT

Notice that if we take the Fourier transform of the whole audio, we're going to extract a lot of frequencies. However, we do not know when these frequencies occur.

STFT do DFT on small segments of the original signal. So the results tell only freuencies around particular time. With this we get time information about the frequency.

```
Animation:
show a time window on the frame
then put the graph on the top of the screen
```
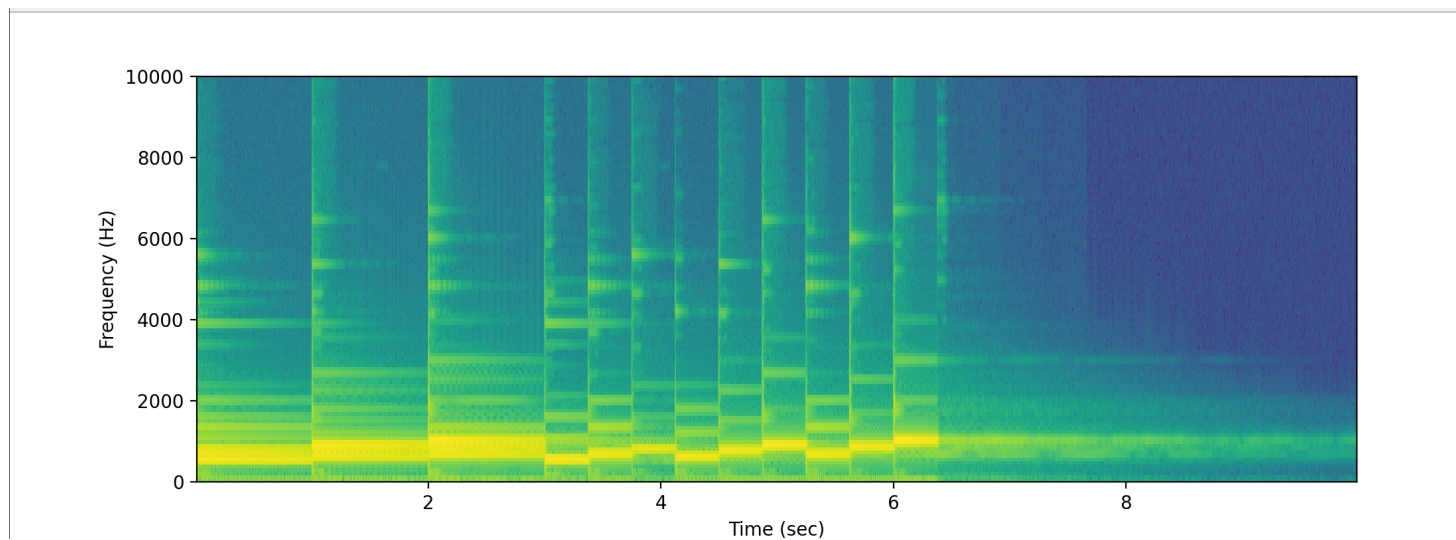
```
Animation:
zoom in on the time window (show the dots of the selected window)
show resultant freuency domain of this window

move window, repeat.
```

It is like scanning on the sound. As a result, we can draw the **spectrogram**.

For t=3s for example,

The height of the frequency domain graph is represented here by the brightness (or color) at a time.