The discrete Fourier transform sends the data into discrete frequency space.

- If \( x_{0}, \ldots, x_{N-1} \) is the data vector \( x \), then \( \hat{x} = DFT(x) = F_{x} \) where \( F_{ij} = \frac{1}{\sqrt{N}} e^{-2\pi i j k/N} \).

- With \( c_{0} = e^{\frac{2\pi i}{N}} \), we have

\[
\hat{x}_{j} = \sum_{k=0}^{N-1} x_{k} e^{-2\pi i j k/N}
\]

In formulas,

\[
x_{j} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_{k} e^{2\pi i j k/N}
\]

- The \( x_{j} \) are usually complex numbers, so we plot \( |x_{j}|^{2} \) vs \( j \).

Example
Proof:

(a) \[ \hat{x}_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k e^{-2\pi i k j/N} \]

Using the formula for the DFT.

(b) \[ x_0 = \frac{1}{N} \sum_{k=0}^{N-1} x_k \]

and so is real.

Fact 1: If \( x \) is a real data vector, \text{\( \hat{x} \)} is even.

(a) \[ \hat{x}_j = \hat{x}_{N-j} \]

(b) \[ \hat{x}_j = \hat{x}_{N-j} \]

Note: Be symmetry about \( \frac{N}{2} \) (assumed new \( k \) is even)
We now want to connect the DFT to the amplitudes of various frequencies present in the data.

The key is the formula for the inverse DFT.
\[ x_j = \sqrt{\frac{\mu}{\pi}} \]

\[ f_j(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} & \text{if } j = 1, 2, \ldots, N \end{cases} \]

\[ W_{x_k} = \sum_{k=0}^{\infty} r_k e^{2\pi i k + j} \]

where \( \pi = 3.14159 \) is a sample point.

So let \( y_{jk} = e^{2\pi i k t_j} \) where \( j = 1, 2, \ldots, N \) and \( k = 0, 1, 2, \ldots \). Then

\[ y_{jk} = \begin{cases} e^{2\pi i k t_j} & \text{if } j = 1, 2, \ldots, N \end{cases} \]

for \( j = 0, 1, \ldots, p-1 \)

where \( t_j = \frac{1}{\mu} \) is the time.

The formula above sees for \( \mu > 1 \) as

\[ f(x_j) = y_j \]

So \( y_j \) interpolates the data exactly.
Theorem 2: If \( x = x_0, x_1, \ldots \) is a data stream (time series) sampled at the points \( t_0, t_1, \ldots \) with \( t_j = j/n \) then

\[
Y(t) = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} x_k e^{2\pi i k t}
\]

interpolates the data.

- For this reason since power is amplitude squared the collection \( |\hat{x}_k|^2 \) (or \( \frac{1}{n} |Y_k|^2 \)) is called the power spectrum.

- Recall \( x_{n-j} = \hat{x}_j \) so \( \|X_{n-j}\|^2 = \|\hat{x}_j\|^2 \).
**Theorem 3**

If $\hat{X} = \text{DFT}(x)$ then

$$\sum_{k=0}^{n-1} |\hat{X}_k|^2 = \sum_{k=0}^{n-1} |X_k|^2$$

Proof: Recall $F$ is unitary and unitary matrices preserve inner products and thus norms so

$$\sum_{k=0}^{n-1} |\hat{X}_k|^2 = \|x\|_2^2 = \|Fx\|_2^2 = \|x\|_2^2 = \sum_{k=0}^{n-1} |X_k|^2$$
We get that

\[
\sum_{k=1}^{n} \frac{3}{x + 7k} \leq \frac{3}{x - 1} + \sum_{k=1}^{n} \frac{1}{x + 7k}
\]

So, in theorem 2, we substitute

\[
\frac{3}{x - 1} + \sum_{k=1}^{n} \frac{1}{x + 7k} = 1 \quad (7)
\]

For \( 0 < x \leq \frac{3}{2} \).

If valued at our same points \( \tau = y \).

There is a relation in terms of the associated frequencies.

In addition to the relation \( \frac{1}{x-k} = 1 \times \frac{1}{x} \).
Since \( y_0(t) = 1 \)

\[
y(t) = \frac{1}{N} \left[ y_0 + \sum_{k=1}^{\frac{N}{2}-1} \left( \hat{x}_k y_k(t) + \hat{x}_k y_{-k}(t) \right) + \hat{x}_{\frac{N}{2}} y_{\frac{N}{2}}(t) \right]
\]

So associated with frequency \( k \), \( y_k(t) = \cos 2\pi kt + \sin 2\pi kt \)

we have amplitudes \( \left| \hat{x}_k \right|^2 + \left| \hat{x}_{-k} \right|^2 = 2 \left| \hat{x}_k \right|^2 \)

This leads to the periodogram
\[ P_0 = \frac{1}{\nu} \left| x_0 \right|^2 \]
\[ P_k = \frac{2}{\nu} \left| \hat{x}_k \right|^2 \quad k = 1, \ldots, \frac{\nu}{2} - 1 \]
\[ P_{\nu/2} = \frac{1}{\nu} \left| x_{\nu/2} \right|^2 \]

[DEMO]

This is the first step in Power Spectrum estimation. And there are many other considerations:
- Sampling rate, Nyquist frequency
- Averaging spectrum on overlapping windows
- Etc.
Now $N/4$ still has complex functions in it.

To simplify further, notice that for any complex number $z$, $z + \bar{z} = 2 \text{Re}(z)$ ($\text{Re} = \text{real part}$)

\[
\sum_{N} x_k \psi_k(z) + \sum_{N} \psi_k(\bar{z}) = 2 \text{Re}(\sum_{N} x_k \psi_k(z))
\]

\[
\Rightarrow x_k \psi_k(z) + \bar{x}_k \psi_k(\bar{z}) = 2 \text{Re}(\sum_{N} x_k \psi_k(z))
\]

\[
= 2 \text{Re}((a_k + ib_k)(\cos 2\pi k t + i\sin 2\pi k t))
\]

\[
= 2[a_k \cos 2\pi k t - b_k \sin 2\pi k t]
\]

where $x_k = a_k + ib_k$

We also have a term $\sum_{N} x_{\frac{N}{2}} \psi_{\frac{N}{2}}(z)$

\[
= \sum_{N} x_{\frac{N}{2}} \left(\cos \frac{N \pi}{2} + i \sin \frac{N \pi}{2}\right)
\]

Since the data is real and $\sum_{N} x_{\frac{N}{2}}$ is real (FACT 1) we just need $\sum_{N} x_{\frac{N}{2}} \cos (N \pi t)$ for this term. So in the end

\[
\sum_{N} x_{\frac{N}{2}} \cos (N \pi t)
\]
Theorem: Given real data \( \{x_0, x_1, \ldots, x_{N-1}\} \)

sampled at \( \{t_0, t_1, \ldots, t_{N-1}\} \) with \( t_j = j/N \) and

Then if \( \hat{x}_k = a_k + ib_k \)

\[
\hat{\nabla} = \frac{1}{\sqrt{N}} \left[ \hat{a}_0 + 2 \sum_{k=1}^{N/2-1} \hat{a}_k \cos 2\pi kt - b_k \sin 2\pi kt \right] + \frac{a_{N/2}}{2} \cos \pi t
\]

Interpolates the data, \( \hat{\nabla}(t_j) = x_j \) for \( j = 0, \ldots, N-1 \)

- Notice that \( \hat{\nabla}(t) \) is a real function now.
- To get the best order \( M \) approximation in least squares, you truncate the sum and eliminate the \( N/2 \)-term.