The data into discrete frequency space.

- If $x[0, N-1]$ is the data vector $x$, then $X[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N}$.

With $w = e^{2\pi j/N}$,

- The $X[k]$ are usually complex numbers, so we plot $|X[k]|^2$ vs $k$.

Example
Notice the symmetry about $\frac{N}{2}$ (assume now $N$ is even)

**Fact 4:** If $x$ is a real data vector, $N$ is even

(a) $X_{\nu-j} = \hat{x}_j$, $j = 0, \ldots, \frac{N}{2} - 1$

(b) $\hat{x}_0 = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k$ and so is real

(c) $\hat{x}_{N/2} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k e^{-2\pi i k / N}$ and so is real

**Proof:** Using the formula for the DFT

(a) $X_{\nu-j} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k e^{-2\pi i k (\nu-j) / N}$

= $\frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k e^{-2\pi i k \nu / N}$

= $\frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{x}_k e^{2\pi i k j / N}$

= $\frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{x}_k e^{2\pi i k \nu / N}$

Since $\hat{x}_k$ is real ($\bar{x}_k = x_k$) and $e^{2\pi i k / N} = 1$
\( (b) \quad \hat{x}_0 = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k e^{-2\pi i k \cdot 0/N} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k \)

\( (c) \quad \hat{x}_k = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k e^{-2\pi i k \cdot k/N} \)

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

---

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.
\[ x_j = \frac{1}{\sqrt{\nu}} \sum_{k=0}^{\nu-1} \hat{x}_k \, e^{2\pi i k j / \nu} \]
\[ = \frac{1}{\sqrt{\nu}} \sum_{k=0}^{\nu-1} \hat{x}_k \, e^{2\pi i k t_j} \]
\[ = \frac{1}{\sqrt{\nu}} \sum_{k=0}^{\nu-1} \hat{x}_k \, e^{2\pi i k t} \]

where \( t_j = j/\nu \) is a sample point.

So let \( \psi_k(t) = e^{2\pi i k t} = \cos 2\pi k t + i \sin 2\pi k t \)

and \( \hat{\psi}_k(t_j) = \frac{1}{\sqrt{\nu}} \sum_{k=0}^{\nu-1} \hat{x}_k \, \psi_k(t_j) \)

The formula above says

\( \hat{\psi}(t_j) = x_j \) for \( j = 0, \ldots, \nu-1 \)

so \( \hat{\psi} \) 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

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

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 \( \hat{X}_{N-j} = \hat{X}_j \) so \( \| \hat{X}_{N-j} \|^2 = \| \hat{X}_j \|^2 \).
"Conservation of Energy"

**Theorem 3**

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

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

**Proof**

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

$$
\sum_{k=0}^{n-1} |x_k|^2 = \|x\|^2 = \|Fx\|^2 = \|\hat{x}\|^2 = \sum_{k=0}^{n-1} |\hat{x}_k|^2
$$
In addition to the relation \[ \| \hat{x}_{\mu} \|^2 = \| x_{\mu} \|^2 \]

there is a relation in terms of the associated frequencies. Evaluated at our sample points \( z_j = j / \nu \)

\[
\psi_{\nu - \kappa}(z_j) = e^{2 \pi i (\nu - \kappa) j / \nu} = e^{2 \pi i j} e^{-2 \pi i \kappa j / \nu}
\]

\[
\psi_{\nu - \kappa}(z_j) = 1 \psi_{\kappa}(z_j)
\]

So, in Theorem 2, we substitute

\[
\frac{1}{\sqrt{2}} \sum_{\kappa=1}^{\nu-1} \psi_{\kappa}(z_j)
\]

for

\[
\frac{1}{\sqrt{2}} \sum_{\kappa=\frac{\nu}{2} + 1}^{\nu-1} \psi_{\kappa}(z_j)
\]

we get that
Since \( \psi_0(t) = 1 \)

\[
\psi(t) = \frac{1}{N} \left[ \psi_0 + \sum_{k=1}^{n-1} \left( \hat{\psi}_k \psi_k(t) + \hat{x}_k \hat{\psi}_k(t) \right) + \frac{\hat{x}_k}{2} \hat{\psi}_k(t) \right]
\]

So associated with frequency \( k \), \( \hat{\psi}_k(t) = \cos 2\pi kt + \sin 2\pi kt \)

we have amplitudes

\[
|\hat{\psi}_k|^2 + |\hat{x}_k|^2 = 2|\hat{x}_k|^2
\]

This leads to the periodogram
\[ p_0 = \frac{1}{n} |X_0|^2 \]
\[ p_k = \frac{2}{n} |\hat{X}_k|^2 \quad k = 1, \ldots, \frac{n}{2} - 1 \]
\[ p_{\frac{n}{2}} = \frac{1}{n} |X_{\frac{n}{2}}|^2 \]

[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 $\Psi/\pi$ still has complex functions in it. To simplify further, notice that for any complex number $z$, $z + \bar{z} = 2\text{Re}(z)$ ($\text{Re} = \text{real part}$).

So $x_k \psi_k/\pi + \overline{x_k \psi_k/\pi} = 2\text{Re}(x_k \psi_k/\pi)$

$= 2 \text{Re} (\bar{a_k} + ib_k)(\cos 2\pi x_k + \text{i}\sin 2\pi x_k)$

$= 2[\bar{a_k} \cos 2\pi x_k - b_k \sin 2\pi x_k]$ where $x_k = a_k + ib_k$

We also have a term $\frac{x_k}{\pi} \psi_k/\pi$

$= \frac{x_k}{\pi} (\cos \frac{\pi}{2} x_k + \text{i}\sin \frac{\pi}{2} x_k)$. Since the data is real and $x_k/\pi$ is real (FACT 1) we just need $\frac{x_k}{\pi} \cos (\pi x_k)$ for this term. So in the end
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 evenly

Then \(f_k = a_k + ib_k\)

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

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

Notice that \(\hat{f}(t)\) is a real function now.

To get the best order \(N\) approximation in least squares, you truncate the sum and eliminate the \(N/2\)-term.