Discrete Cosine Transform

\[ X_C[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \cos \left( \frac{\pi k}{N} \left( n + \frac{1}{2} \right) \right) \]

\[ x[n] = X_C[0] + 2 \sum_{k=1}^{N-1} X_C[k] \cos \left( \frac{\pi k}{N} \left( n + \frac{1}{2} \right) \right) \]
The Discrete Fourier Transform (DFT) implicitly represents the frequencies that are contained in a *periodically extended* version of the input signal, and periodic extension can generate frequencies that are *not present* in the original signal.

Consider the following $8 \times 8$ example.

![Diagram of an $8 \times 8$ grid with different brightnesses on the left and right edges.](image)

The brightnesses of left and right edges are different and will generate a sequence of large transitions when periodically extended.

![Diagram of a periodically extended $8 \times 8$ grid with sequences of large transitions.](image)
Consider a single row from the previous image. The DFT implicitly extends the signal (here a ramp) periodically.

\[ x[n] = x[n + 8] \]

Although the function is smooth from \( n = 0 \) to 7, the periodic extension contains a series of steps.
We can eliminate the step discontinuities by first replicating one period in reverse order and then extending the result periodically.

\[ y[n] = y[n + 16] \]

The resulting signal is continuous across the edges (however the slope is still discontinuous).

Finally, insert zeros between successive samples.

\[ z[n] = z[n + 32] \]

The resulting signal is real-valued, symmetric about \( n = 0 \), periodic in \( 4N \), and contains only odd numbered samples.

The DFT of this signal is real, symmetric about \( k = 0 \) and anti-periodic. It is completely characterized by \( N \) values: \( Z[0] \) to \( Z[7] \).

This process is captured in the **Discrete Cosine Transform**.
Use this construction to relate the DCT and DFT.

\[ x[m] = x[m + N] \]

\[ z[n] = z[n + 4N] \]

\[ Z[k] = \frac{1}{4N} \sum_{n=-2N}^{2N-1} z[n] e^{-j \frac{2\pi k}{4N} n} = \frac{1}{4N} \sum_{m=0}^{N-1} x[m] \left( e^{j \frac{2\pi k}{4N} (2m+1)} + e^{-j \frac{2\pi k}{4N} (2m+1)} \right) \]

\[ = \frac{1}{2N} \sum_{m=0}^{N-1} x[m] \cos \left( \frac{2\pi k}{4N} (2m + 1) \right) = \frac{1}{2N} \sum_{m=0}^{N-1} x[m] \cos \left( \frac{\pi k}{N} (m + \frac{1}{2}) \right) \]

\[ = \frac{1}{2} X_C[k] \]
Compare the DCT and DFT of the following signal. $N = 32$.

Which transform is “more compact”? Why?

If $N = 32$, how many unique numbers are in a DCT? in a DFT?
DCT Expansions

Each arrow in the log plot indicates a value that is smaller than 0.01 (i.e., below the x axis). The red arrows indicate values of $-\infty$.

The DCT is clearly more compact than the DFT.

The DFT represents the step as the sum of sinusoidal components that spread across all frequencies.

The DCT is based on an expansion that does not have a step, and therefore does not require as much energy at high frequencies. Some energy is needed however because there is a slope discontinuity.

Number of unique numbers?

There are $N$ samples in time.

The DCT has $N$ real-valued coefficients $X_C[k]$.

The DFT has $N$ complex-valued coefficients $X[k]$. However, half of those numbers are redundant if $x[n]$ is real, since $X[-k] = X^*[k]$. 
DCT Expansions

Compare the DCT and DFT of the following signal. $N = 32$.

The DCT is not much more compact than the DFT. Why?

$X_C[k] = 0$ for $k = 2, 4, 6, \ldots$. Why? How about the DFT?
DCT Expansions

There is a step in both $x[n]$ and in the mirrored signal upon which the DCT is based. Therefore both transforms require high frequency coefficients.

$X_C[k] = 0$ for $k = 2, 4, 6, \ldots$ because of the symmetry of the DCT basis functions (see next slide). The sum of samples in the first (or second) half-period of the even-numbered basis functions is always zero. Since $x[n]$ is constant over each half-period, the resulting inner product is zero.

$X[k] = 0$ for $k = 2, 4, 6, \ldots$ for a similar reason.
Comparison of DFT and DCT Basis Functions

DFT (real and imaginary parts) versus DCT.

\[
\begin{align*}
\text{Re}(e^{j\frac{2\pi k}{N}n}) & & \text{Im}(e^{j\frac{2\pi k}{N}n}) & & \cos\left(\frac{\pi k}{N} \left(n + \frac{1}{2}\right)\right) \\
 k = 0 & & & & \\
 k = 1 & & & & \\
 k = 2 & & & & \\
 k = 3 & & & & \\
 k = 4 & & & & \\
 k = 5 & & & & \\
 k = 6 & & & & \\
 k = 7 & & & & 
\end{align*}
\]
DCT Expansions

Compare the DCT and DFT of the following signal. $N = 32$.

$$x[n] = \cos\left(\frac{3\pi}{32} n\right)$$

For what value of $k$ is $|X_C[k]|$ largest? Why?
DCT Expansions

The previous signal contains 1.5 cycles of a cosine wave in $0 \leq n < N$. Therefore the frequencies are not harmonically related to the fundamental frequency, and $X[k] \neq 0$ for any $k$ in the DFT.

The signal frequency is at a half-multiple of the fundamental frequency, specifically $\frac{3}{2}$. This $X_C[3]$ has the largest magnitude.

We might expect that $X_C[k]$ would be zero if $k \neq 3$. Why isn’t this true? It’s because $x[n]$ is not quite a basis function. The basis function would be shifted by $n = \frac{1}{2}$, as in the following plot.
DCT Expansions

Compare the DCT and DFT of the following signal. $N = 32$.

$$x[n] = \cos \left( \frac{3\pi}{32} n + \frac{3\pi}{64} \right) = \cos \left( \frac{3\pi}{32} (n + \frac{1}{2}) \right)$$

A small shift in time changes which components $X_C[k]$ are non-zero.
DCT Expansions

This example illustrates a VERY important property of Fourier transforms. Both sines and cosines are basis functions in ALL Fourier transforms. Sinusoids of any phase can be constructed from a linear combination of a sine and cosine, as follows.

\[
\cos(\omega t + \phi) = \cos(\omega t) \cos(\phi) - \sin(\omega t) \sin(\phi)
\]

By contrast, the DCT only contains cosine terms. We can still generate any arbitrary phase of a sinusoid, but it now involves more than just a pair of \(k\)'s.
DCT Expansions

This issue is clear from the time-shift property of Fourier transforms. If

\[ x[n] \overset{\text{DTFT}}{\leftrightarrow} X(\Omega) \]

then

\[ x[n - n_0] \overset{\text{DTFT}}{\leftrightarrow} e^{-j\Omega n_0} X(\Omega) \]

This result gives a recipe for time shifting in terms of a reweighting of the sine and cosine components.
By contrast, shifting time has a more complicated effect on the DCT. Consider a (circular) delay of 2.

\[ x[n - 2] \]

\[ z[n] = z[n + 4N] \]

There is no simple relation between the DCTs of \( x[n] \) and \( x[(n - 2) \mod N] \).
DCT Expansions

Compare the DCT and DFT of the following signal. $N = 32$.

$$x[n] = \sin \left( \frac{3\pi}{32} n \right)$$

Why is the DCT less “compact” for sine than cosine (previous slide)?

Why is the DFT more “compact” for sine than for the cosine?
Mirroring the cosine function (previous example) generates a sinusoidal signal.

However, mirroring the sine function does not. Rather, it generates a signal with large slope discontinuities.

By contrast, the DFT of the sine has less high-frequency energy than the DFT of the (previous) cosine. This is because the periodic extension of the cosine generates a step discontinuity in value, and periodic extension of the sine generates a step discontinuity in slope only.
DCT Expansions

Compare the DCT and DFT of the following signal. \( N = 32 \).

\[ x[n] = \cos \left( \frac{2\pi}{32} n \right) \]

Why is the DCT less compact than the DFT?
For what value of \( k \) is \( |X_C[k]| \) largest? Why?
The DFT is maximally compact here because \( x[n] \) is a basis function of the DFT.

The DCT is less compact because it is not a basis function. The basis function with the same frequency has a bit of a phase shift.

\[ |X(k)| \text{ is greatest when } k = 1. \]

\[ |X_C[k]| \text{ is greatest when } k = 2. \]
Conclusions

The DCT is great for reducing edge artifacts due to the periodic extension that is implicit in the DFT.

The DCT does not have a simple time-shift property. Since the DCT is real-valued, time shifts cannot simply change phase as it does in the DFT. Shifting in time generally changes the magnitudes of multiple $k$'s in the DCT.