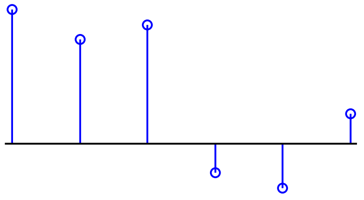
3: Discrete Cosine Transform

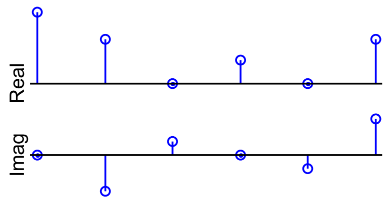
For processing 1-D or 2-D signals (especially coding), a common method is to divide the signal into "frames" and then apply an invertible transform to each frame that compresses the information into few coefficients.

The DFT has some problems when used for this purpose:

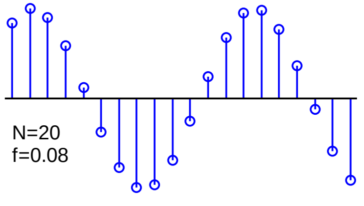
- N real $x[n] \leftrightarrow N$ complex $X[k]$: 2 real, $\frac{N}{2}-1$ conjugate pairs



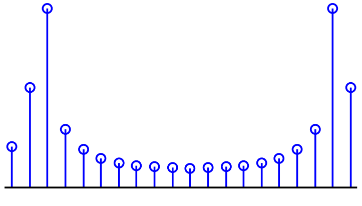
$\rightarrow$



- DFT $\propto$ the DTFT of a periodic signal formed by replicating $x[n]$.
  $\Rightarrow$ Spurious frequency components from boundary discontinuity.

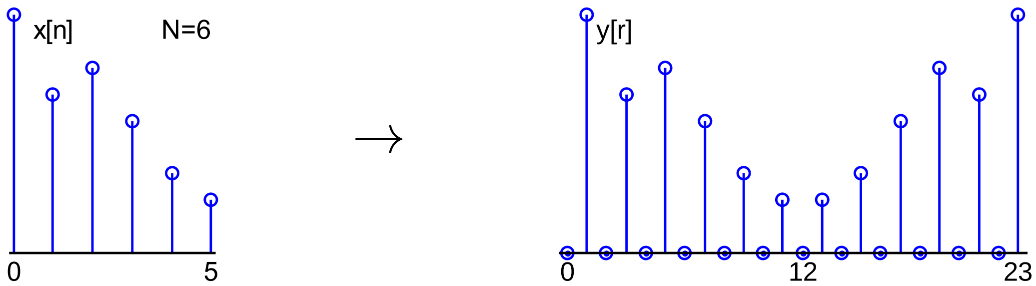


$\rightarrow$

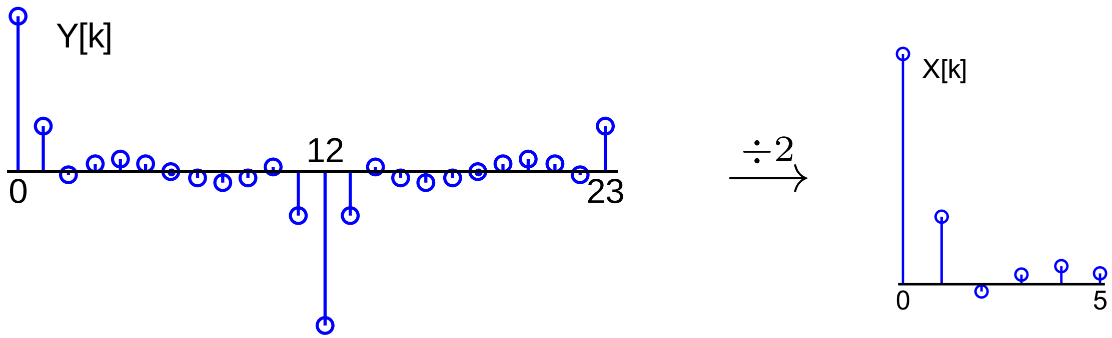


The Discrete Cosine Transform (DCT) overcomes these problems.

To form the Discrete Cosine Transform (DCT), replicate $x[0:N-1]$ but in reverse order and insert a zero between each pair of samples:



Take the DFT of length $4N$ real, symmetric, odd-sample-only sequence.
Result is real, symmetric and anti-periodic: only need first N values



Forward DCT: $X_C[k] = \sum_{n=0}^{N-1} x[n] \cos \frac{2\pi(2n+1)k}{4N}$ for $k = 0 : N-1$

Inverse DCT: $x[n] = \frac{1}{N} X[0] + \frac{2}{N} \sum_{k=1}^{N-1} X[k] \cos \frac{2\pi(2n+1)k}{4N}$

DCT formula derivation

This proof is not examinable.

We want to show that $X_C[k] = \sum_{n=0}^{N-1} x[n] \cos \frac{2\pi(2n+1)k}{4N}$ is equivalent to replicating $x[n]$ in reverse order, inserting alternate zeros, taking DFT, dividing by 2 and keeping first N values:

Replicating + zero insertion gives $y[r] = \begin{cases} 0 & r \text{ even} \\ x\left[\frac{r-1}{2}\right] & r \,\text{odd}, 1 \le r \le 2N-1 \\ x\left[\frac{4N-1-r}{2}\right] & r \,\text{odd}, 2N+1 \le r \le 4N-1 \end{cases}$

$Y_F[k] = \sum_{r=0}^{4N-1} y[r] W_{4N}^{kr} \overset{\text{(i)}}{=} \sum_{n=0}^{2N-1} y[2n+1] W_{4N}^{(2n+1)k}$ where $W_a^b = e^{-j\frac{2\pi b}{a}}$

$\overset{\text{(ii)}}{=} \sum_{n=0}^{N-1} y[2n+1] W_{4N}^{(2n+1)k} + \sum_{m=0}^{N-1} y[4N-2m-1] W_{4N}^{(4N-2m-1)k}$

$\overset{\text{(iii)}}{=} \sum_{n=0}^{N-1} x[n] W_{4N}^{(2n+1)k} + \sum_{m=0}^{N-1} x[m] W_{4N}^{-(2m+1)k}$

$= 2\sum_{n=0}^{N-1} x[n] \cos \frac{2\pi(2n+1)k}{4N} = 2X_C[k]$

(i) odd r only: $r = 2n+1$

(ii) reverse order for $n \ge N$: $m = 2N-1-n$

(iii) substitute y definition & $W_{4N}^{4Nk} = e^{-j2\pi\frac{4Nk}{4N}} \equiv 1$

IDCT formula derivation

This proof is not examinable.

We want to show that $x[n] = \frac{1}{N}X[0] + \frac{2}{N}\sum_{k=1}^{N-1} X[k]\cos\frac{2\pi(2n+1)k}{4N}$

Since $Y[k] = 2X[k]$ we can write $y[r] = \frac{1}{4N}\sum_{k=0}^{4N-1} Y[k]W_{4N}^{-rk} = \frac{1}{2N}\sum_{k=0}^{4N-1} X[k]W_{4N}^{-rk}$

So we can write,

$x[n] = y[2n+1] = \frac{1}{2N}\sum_{k=0}^{4N-1} X[k]W_{4N}^{-(2n+1)k}$ where $W_a^b = e^{-j\frac{2\pi b}{a}}$

$\stackrel{\text{(i)}}{=} \frac{1}{2N}\sum_{k=0}^{2N-1} X[k]W_{4N}^{-(2n+1)k} - \frac{1}{2N}\sum_{l=0}^{2N-1} X[l]W_{4N}^{-(2n+1)(l+2N)}$

$\stackrel{\text{(ii)}}{=} \frac{1}{N}\sum_{k=0}^{2N-1} X[k]W_{4N}^{-(2n+1)k}$

$\stackrel{\text{(iii)}}{=} \frac{1}{N}X[0] + \frac{1}{N}\sum_{k=1}^{N-1} X[k]W_{4N}^{-(2n+1)k}$
$\quad + \frac{1}{N}X[N]W_{4N}^{-(2n+1)N} + \frac{1}{N}\sum_{r=1}^{N-1} X[2N-r]W_{4N}^{-(2n+1)(2N-r)}$

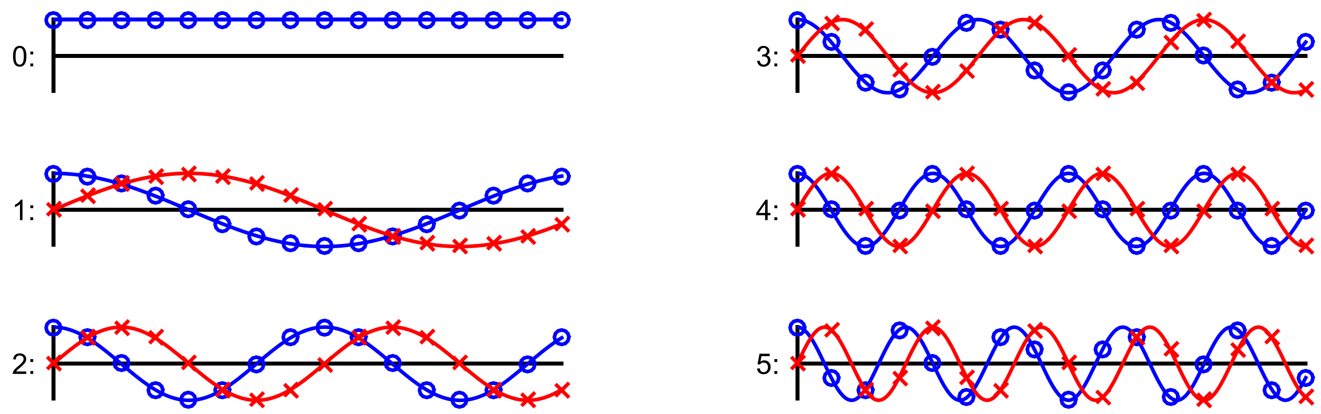
$\stackrel{\text{(iv)}}{=} \frac{1}{N}X[0] + \frac{1}{N}\sum_{k=1}^{N-1} X[k]W_{4N}^{-(2n+1)k} + \frac{1}{N}\sum_{r=1}^{N-1} -X[r]W_{4N}^{(2n+1)r+2N}$

$= \frac{1}{N}X[0] + \frac{2}{N}\sum_{k=1}^{N-1} X[k]\cos\frac{2\pi(2n+1)k}{4N}$

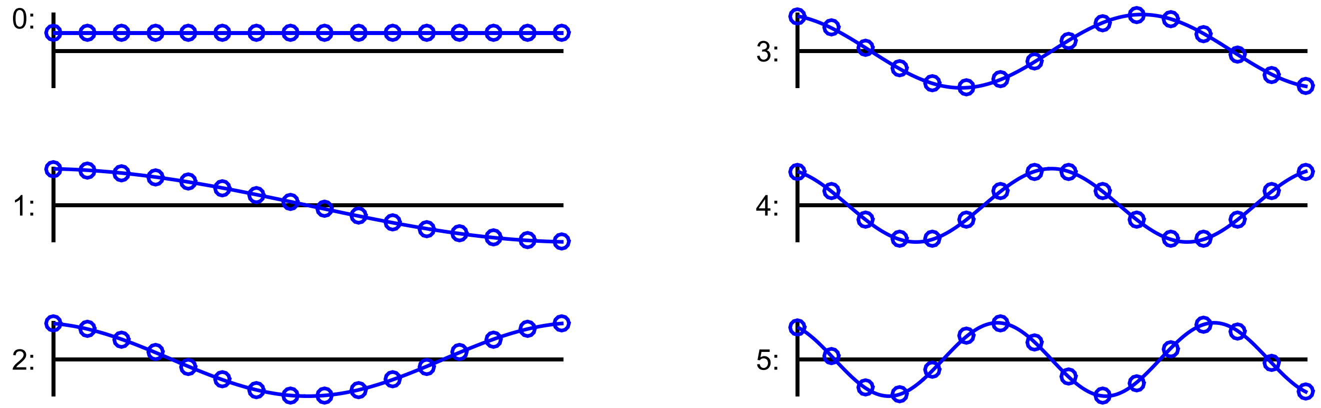
Notes: (i) $k = l + 2N$ for $k \ge 2N$ and $X[k+2N] = -X[k]$
(ii) $\frac{(2n+1)(l+2N)}{4N} = \frac{(2n+1)l}{4N} + n + \frac{1}{2}$ and $e^{j2\pi\left(n+\frac{1}{2}\right)} = -1$
(iii) $k = 2N - r$ for $k > N$
(iv) $X[N] = 0$ and $X[2N-r] = -X[r]$

Basis Functions

DFT basis functions: $x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k]e^{j2\pi\frac{kn}{N}}$



DCT basis functions: $x[n] = \frac{1}{N}X[0] + \frac{2}{N}\sum_{k=1}^{N-1} X[k]\cos\frac{2\pi(2n+1)k}{4N}$



DCT of sine wave

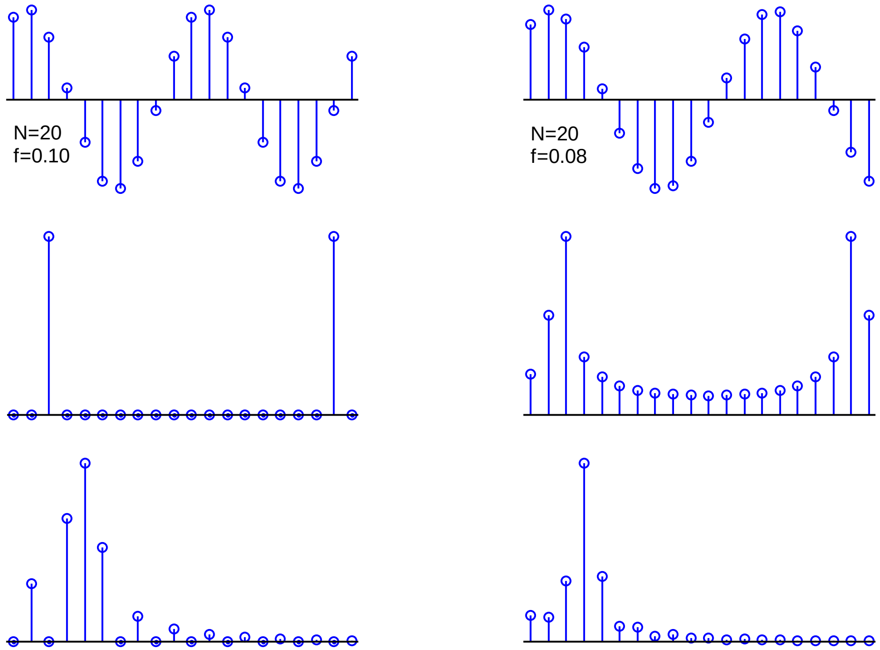
DCT: $X_C[k] = \sum_{n=0}^{N-1} x[n] \cos \frac{2\pi(2n+1)k}{4N}$

$f = \frac{m}{N}$ $\qquad$ $f \neq \frac{m}{N}$

$x[n]$

$|X_F[k]|$

$|X_C[k]|$



DFT: Real→Complex; Freq range $[0, 1]$; Poorly localized unless $f = \frac{m}{N}$; $|X_F[k]| \propto k^{-1}$ for $Nf < k \ll \frac{N}{2}$

DCT: Real→Real; Freq range $[0, 0.5]$; Well localized $\forall f$; $|X_C[k]| \propto k^{-2}$ for $2Nf < k < N$

Definition: $X[k] = \sum_{n=0}^{N-1} x[n] \cos \frac{2\pi(2n+1)k}{4N}$

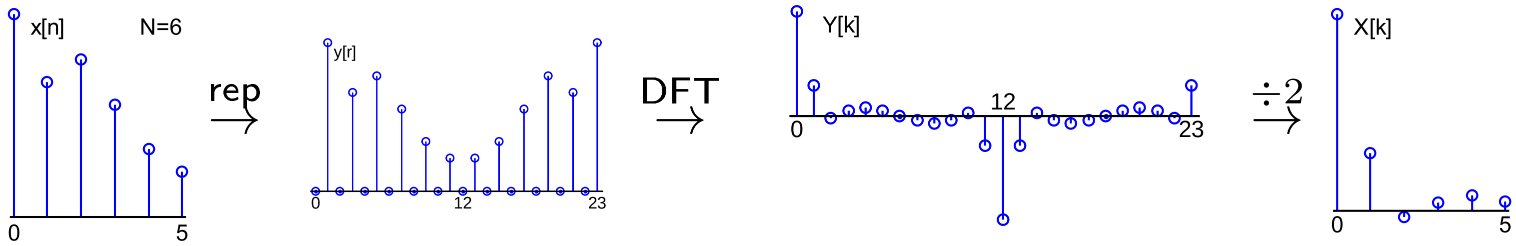
- Linear: $\alpha x[n] + \beta y[n] \to \alpha X[k] + \beta Y[k]$
- "Convolution⟷Multiplication" property of DFT does not hold ☹
- Symmetric: $X[-k] = X[k]$ since $\cos -\alpha k = \cos +\alpha k$
- Anti-periodic: $X[k+2N] = -X[k]$ because:
  - $2\pi(2n+1)(k+2N) = 2\pi(2n+1)k + 8\pi Nn + 4N\pi$
  - $\cos(\theta + \pi) = -\cos\theta$

  $\Rightarrow X[N] = 0$ since $X[N] = X[-N] = -X[-N+2N]$
- Periodic: $X[k+4N] = -X[k+2N] = X[k]$

Energy Conservation

DCT: $X[k] = \sum_{n=0}^{N-1} x[n] \cos \frac{2\pi(2n+1)k}{4N}$

IDCT: $x[n] = \frac{1}{N} X[0] + \frac{2}{N} \sum_{k=1}^{N-1} X[k] \cos \frac{2\pi(2n+1)k}{4N}$



Energy: $E = \sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} |X[0]|^2 + \frac{2}{N} \sum_{n=1}^{N-1} |X[n]|^2$

In diagram above: $E \to 2E \to 8NE \to \approx 0.5NE$

Orthogonal DCT (preserves energy: $\sum |x[n]|^2 = \sum |X[n]|^2$)

$$\text{ODCT: } X[k] = \begin{cases} \sqrt{\frac{1}{N}} \sum_{n=0}^{N-1} x[n] & k = 0 \\ \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} x[n] \cos \frac{2\pi(2n+1)k}{4N} & k \neq 0 \end{cases}$$

$$\text{IODCT: } x[n] = \sqrt{\frac{1}{N}} X[0] + \sqrt{\frac{2}{N}} \sum_{k=1}^{N-1} X[k] \cos \frac{2\pi(2n+1)k}{4N}$$

Note: MATLAB dct() calculates the ODCT

Energy Compaction

If consecutive $x[n]$ are positively correlated, DCT concentrates energy in a few $X[k]$ and decorrelates them.

Example: Markov Process: $x[n] = \rho x[n-1] + \sqrt{1-\rho^2}u[n]$

where $u[n]$ is i.i.d. unit Gaussian.

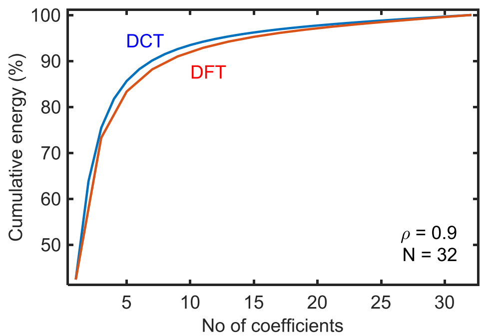
Then $\langle x^2[n] \rangle = 1$ and $\langle x[n]x[n-1] \rangle = \rho$.

Covariance of vector $\mathbf{x}$ is $\mathbf{S}_{i,j} = \left\langle \mathbf{x}\mathbf{x}^H \right\rangle_{i,j} = \rho^{|i-j|}$.

Suppose ODCT of $\mathbf{x}$ is $\mathbf{Cx}$ and DFT is $\mathbf{Fx}$.

Covariance of $\mathbf{Cx}$ is $\left\langle \mathbf{Cxx}^H\mathbf{C}^H \right\rangle = \mathbf{CSC}^H$ (similarly $\mathbf{FSF}^H$)

Diagonal elements give mean coefficient energy.

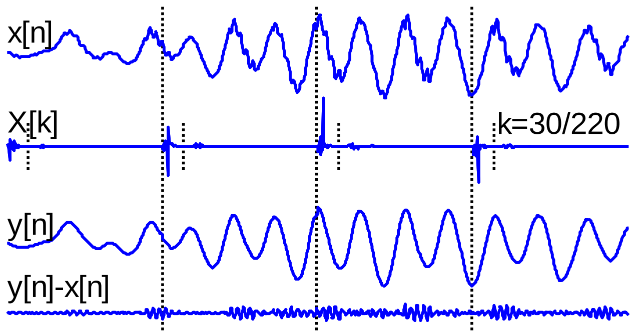


- Used in MPEG and JPEG (superseded by JPEG2000 using wavelets)
- Used in speech recognition to decorrelate spectral coeficients: DCT of log spectrum

Energy compaction good for coding (low-valued coefficients can be set to 0)

Decorrelation good for coding and for probability modelling

- Divide continuous signal into frames
- Apply DCT to each frame
- Encode DCT
  - e.g. keep only 30 $X[k]$
- Apply IDCT $\rightarrow y[n]$



Problem: Coding may create discontinuities at frame boundaries
e.g. JPEG, MPEG use 8×8 pixel blocks

8.3 kB (PNG) 1.6 kB (JPEG) 0.5 kB (JPEG)

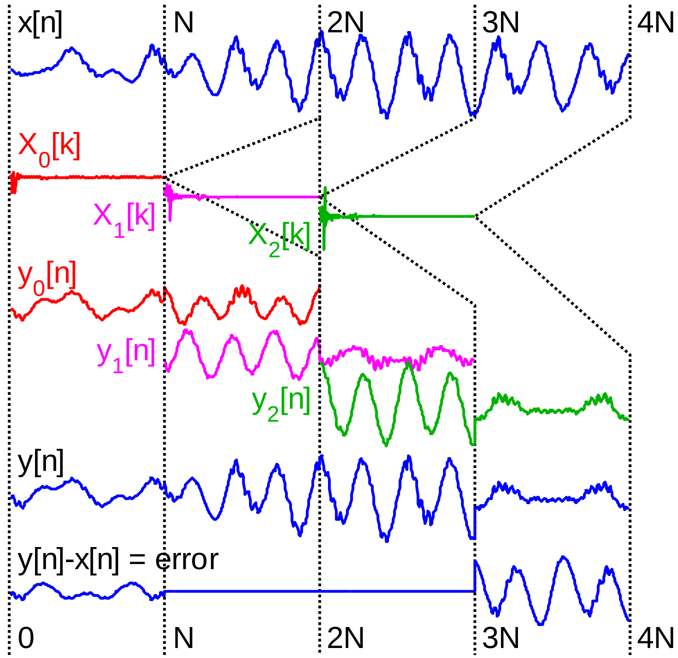
Lapped Transform

Modified Discrete Cosine Transform (MDCT): overlapping frames $2N$ long

$x[0:2N-1] \xrightarrow{\text{MDCT}} X_0[0:N-1] \xrightarrow{\text{IMDCT}} y_0[0:2N-1]$

$x[N:3N-1] \xrightarrow{\text{MDCT}} X_1[N:2N-1] \xrightarrow{\text{IMDCT}} y_1[N:3N-1]$

$x[2N:4N-1] \xrightarrow{\text{MDCT}} X_2[2N:3N-1] \xrightarrow{\text{IMDCT}} y_2[2N:4N-1]$



$y[n] = y_0[n] + y_1[n] + y_2[n]$

MDCT: $2N \to N$ coefficients, IMDCT: $N \to 2N$ samples
Add $y_i[n]$ together to get $y[n]$. Only two non-zero terms far any n.
Errors cancel exactly: Time-domain alias cancellation (TDAC)

MDCT (Modified DCT)

MDCT: $X[k] = \sum_{n=0}^{2N-1} x[n] \cos \frac{2\pi(2n+1+N)(2k+1)}{8N}$ $\qquad 0 \le k < N$

IMDCT: $y[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \cos \frac{2\pi(2n+1+N)(2k+1)}{8N}$ $\qquad 0 \le n < 2N$

If $\mathbf{x}$, $\mathbf{X}$ and $\mathbf{y}$ are column vectors, then $\mathbf{X} = \mathbf{M}\mathbf{x}$ and $\mathbf{y} = \frac{1}{N}\mathbf{M}^T\mathbf{X} = \frac{1}{N}\mathbf{M}^T\mathbf{M}\mathbf{x}$

where $\mathbf{M}$ is an $N \times 2N$ matrix with $m_{k,n} = \cos \frac{2\pi(2n+1+N)(2k+1)}{8N}$.

Quasi-Orthogonality: The $2N \times 2N$matrix, $\frac{1}{N}\mathbf{M}^T\mathbf{M}$, is almost the identity:

$$\frac{1}{N}\mathbf{M}^T\mathbf{M} = \frac{1}{2}\begin{bmatrix} \mathbf{I}-\mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}+\mathbf{J} \end{bmatrix} \text{ with } \mathbf{I} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}, \mathbf{J} = \begin{bmatrix} 0 & \cdots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \cdots & 0 \end{bmatrix}$$

When two consective $\mathbf{y}$ frames are overlapped by N samples, the second half of the first frame has thus been multiplied by $\frac{1}{2}(\mathbf{I}+\mathbf{J})$ and the first half of the second frame by $\frac{1}{2}(\mathbf{I}-\mathbf{J})$. When these $\mathbf{y}$ frames are added together, the corresponding $\mathbf{x}$ samples have been multiplied by $\frac{1}{2}(\mathbf{I}+\mathbf{J}) + \frac{1}{2}(\mathbf{I}-\mathbf{J}) = \mathbf{I}$ giving perfect reconstruction.

Normally the $2N$-long $\mathbf{x}$ and $\mathbf{y}$ frames are windowed before the MDCT and again after the IMDCT to avoid any discontinuities; if the window is symmetric and satisfies $w^2[i] + w^2[i+N] = 2$ the perfect reconstruction property is still true.

[Deriving the value of $\frac{1}{N}\mathbf{M}^T\mathbf{M}$]

This proof is not examinable.

If we define $\mathbf{A} = \frac{\mathbf{1}}{\mathbf{N}}\mathbf{M}^{\mathbf{T}}\mathbf{M}$ with $m_{kn} = \cos\frac{2\pi(2n+1+N)(2k+1)}{8N}$, we want to show that $\mathbf{A} = \frac{1}{2}\begin{bmatrix} \mathbf{I}+\mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}-\mathbf{J} \end{bmatrix}$. To avoid fractions, we write $\alpha = \frac{2\pi}{8N}$ so that $m_{kn} = \cos\left(\alpha(2n+1+N)(2k+1)\right)$. Now we can say

$$
\begin{aligned}
a_{rn} &= \frac{1}{N}\sum_{k=0}^{N-1} m_{kr}m_{kn} \\
&= \frac{1}{N}\sum_{k=0}^{N-1} \cos\left(\alpha(2r+1+N)(2k+1)\right)\cos\left(\alpha(2n+1+N)(2k+1)\right) \\
&= \frac{1}{2N}\sum_{k=0}^{N-1} \cos\left(2\alpha(r-n)(2k+1)\right) + \frac{1}{2N}\sum_{k=0}^{N-1} \cos\left(2\alpha(r+n+1+N)(2k+1)\right)
\end{aligned}
$$

where, in the last line, we used the identity $\cos\theta\cos\phi = \frac{1}{2}\cos\left(\theta-\phi\right) + \frac{1}{2}\cos\left(\theta+\phi\right)$.

We now convert these terms to complex exponentials to sum them as geometric progressions.

$[\frac{1}{2N}\sum_{k=0}^{N-1}\cos\left(2\alpha(r-n)(2k+1)\right)]$

Converting to a the real part ($\Re$) of geometric progression (with $\alpha = \frac{2\pi}{8N}$):

$$\begin{aligned}
\frac{1}{2N}\sum_{k=0}^{N-1}\cos\left(2\alpha(r-n)(2k+1)\right) &= \frac{1}{2N}\Re\left(\sum_{k=0}^{N-1}\exp\left(j2\alpha(r-n)(2k+1)\right)\right) \\
&= \frac{1}{2N}\Re\left(\exp\left(j2\alpha(r-n)\right)\sum_{k=0}^{N-1}\exp\left(j4\alpha(r-n)k\right)\right) \\
&= \frac{1}{2N}\Re\left(\exp\left(j2\alpha(r-n)\right)\frac{1-\exp\left(j4\alpha(r-n)N\right)}{1-\exp\left(j4\alpha(r-n)\right)}\right) \\
&= \frac{1}{2N}\Re\left(\frac{1-\exp\left(j4\alpha(r-n)N\right)}{\exp\left(-j2\alpha(r-n)\right)-\exp\left(j2\alpha(r-n)\right)}\right) \\
&= \frac{1}{2N}\Re\left(\frac{1-\exp\left(j4\alpha(r-n)N\right)}{-2j\sin\left(2\alpha(r-n)\right)}\right) \\
&= \frac{1}{4N}\frac{\sin\left(4\alpha(r-n)N\right)}{\sin\left(2\alpha(r-n)\right)} = \frac{1}{4N}\frac{\sin\left((r-n)\pi\right)}{\sin\left(\frac{r-n}{2N}\pi\right)}
\end{aligned}$$

The numerator is sine of a multiple of π and is therefore 0. Therefore the whole sum is zero unless the denominator is zero or, equivalently, $(r-n)$ is a multiple of $2N$. Since $0 \le r,\ n < 2N$, this only happens when $r = n$ in which case the sum becomes $\frac{1}{2N}\sum_{k=0}^{N-1}\cos 0 = \frac{1}{2}$.

$[\frac{1}{2N}\sum_{k=0}^{N-1}\cos\left(2\alpha(r+n+1+N)(2k+1)\right)]$

$\frac{1}{2N}\sum_{k=0}^{N-1}\cos\left(2\alpha(r+n+1+N)(2k+1)\right)$ is the same as before with $r-n$ replaced by $r+n+1+N$.

We can therefore write

$$\frac{1}{2N}\sum_{k=0}^{N-1}\cos\left(2\alpha(r+n+1+N)(2k+1)\right) \quad = \quad \frac{1}{4N}\frac{\sin\left((r+n+1+N)\pi\right)}{\sin\left(\frac{r+n+1+N}{2N}\pi\right)}$$

The numerator is again the sine of a multiple of π and is therefore 0. Therefore the whole sum is zero unless $(r+n+1+N)$ is a multiple of $2N$. This only happens when $r+n=N-1$ or $3N-1$ since $0\le r,\, n<2N$. The constraint $r+n=N-1$ corresponds to the anti-diagonal of the top left quadrant of the $\mathbf{A}$ matrix, while $r+n=3N-1$ corresponds to the anti-diagonal of the bottom right quadrant.

Writing $r+n+1+N=x$, we can use L'Hôpital's rule to evaluate $\frac{1}{4N}\frac{\sin(x\pi)}{\sin\left(\frac{x}{2N}\pi\right)}$ at $x=\{2N,\,4N\}$.

Differentiating numerator and denominator gives $\frac{1}{2}\frac{\cos(x\pi)}{\cos\left(\frac{x}{2N}\pi\right)}$ which comes to $\left\{-\frac{1}{2},\,\frac{1}{2}\right\}$ respectively at $x=\{2N,\,4N\}$.

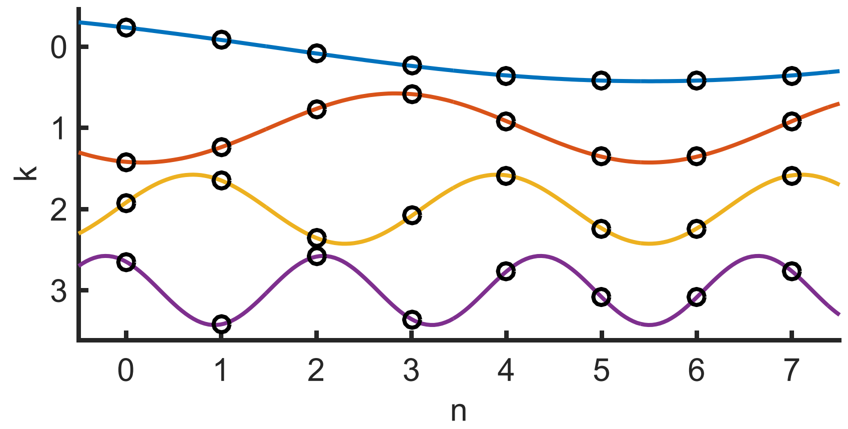
MDCT Basis Elements

MDCT: $X[k] = \sum_{n=0}^{2N-1} x[n] \cos \frac{2\pi(2n+1+N)(2k+1)}{8N}$ $\qquad 0 \le k < N$

IMDCT: $y[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \cos \frac{2\pi(2n+1+N)(2k+1)}{8N}$ $\qquad 0 \le n < 2N$

In vector notation: $\mathbf{X} = \mathbf{M}\mathbf{x}$ and $\mathbf{y} = \frac{1}{N}\mathbf{M}^T\mathbf{X} = \frac{1}{N}\mathbf{M}^T\mathbf{M}\mathbf{x}$

The rows of $\mathbf{M}$ form the MDCT basis elements.



Example $(N = 4)$:

$$\mathbf{M} = \begin{bmatrix} 0.56 & 0.20 & -0.20 & -0.56 & -0.83 & -0.98 & -0.98 & -0.83 \\ -0.98 & -0.56 & 0.56 & 0.98 & 0.20 & -0.83 & -0.83 & 0.20 \\ 0.20 & 0.83 & -0.83 & -0.20 & 0.98 & -0.56 & -0.56 & 0.98 \\ 0.83 & -0.98 & 0.98 & -0.83 & 0.56 & -0.20 & -0.20 & 0.56 \end{bmatrix}$$

The basis frequencies are $\{0.5,\ 1.5,\ 2.5,\ 3.5\}$ times the fundamental.

DCT: Discrete Cosine Transform

- Equivalent to a DFT of time-shifted double-length $\left[\ \mathbf{x}\quad \overleftarrow{\mathbf{x}}\ \right]$
- Often scaled to make an orthogonal transform (ODCT)
- Better than DFT for energy compaction and decorrelation ☺
  - Energy Compaction: Most energy is in only a few coefficients
  - Decorrelation: The coefficients are uncorrelated with each other
- Nice convolution property of DFT is lost ☹

MDCT: Modified Discrete Cosine Transform

- Lapped transform: $2N \to N \to 2N$
- Aliasing errors cancel out when overlapping output frames are added
- Similar to DCT for energy compaction and decorrelation ☺
- Overlapping windowed frames can avoid edge discontinuities ☺
- Used in audio coding: MP3, WMA, AC-3, AAC, Vorbis, ATRAC

For further details see Mitra: 5.

MATLAB routines

dct, idct	ODCT with optional zero-padding