We can see mathematically why the DCT low-pass filter works so well.
For simplicity consider 1-d, so the interpolating trigonometric polynomial is

$$p_n(x) = \sqrt{\tfrac{2}{n}} \sum_{k=0}^{n-1} a_k \sigma_k \cos(kx), \quad \text{where} \quad \sigma_k = \begin{cases} 1/\sqrt{2} & \text{if } k=0, \\ 1 & \text{if } k>0. \end{cases}$$

Where $\vec{a} = C_n^{-1}\vec{f}$.

Suppose we want to approximate the same data using fewer terms.

Thm. 3.4 — (Least-squares property) Let $\vec{a} = C_n^{-1}\vec{f}$ be the (1-d) interpolation coefficients using the DCT at n points. If $m<n$ then the truncated trigonometric polynomial

$$p_m(x) = \sqrt{\tfrac{2}{n}} \sum_{k=0}^{m-1} c_k \sigma_k \cos(kx)$$

sum of squared error

minimises $\sum_{j=0}^{n-1} (p_m(x_j) - f_j)^2$ at the n nodes x_j if $c_0 = a_0, c_1 = a_1, \ldots, c_{m-1} = a_{m-1}$.

Recall (2H Numerical Analysis):-

Suppose we want to find $\vec{x}$ to minimise $\|A\vec{x}-\vec{b}\|_2^2$, where A has more rows than $\vec{x}$ has entries.

e.g. $\begin{pmatrix} 3 & 1 \\ 1 & 1 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ $\longrightarrow$ overdetermined so can't solve exactly.

This is equivalent to minimising $(A\vec{x}-\vec{b})^T(A\vec{x}-\vec{b}) = \vec{x}^TA^TA\vec{x} - 2\vec{x}^TA^T\vec{b} + \vec{b}^T\vec{b}$.

This is a non-negative quadratic function of $\vec{x}$, so a minimum exists and is found by setting the partial derivatives $\partial/\partial x_i$ to zero, giving

$$A^TA\vec{x} = A^T\vec{b} \quad \text{— normal equations.}$$

If A is orthogonal, then $A^T = A^{-1}$ so these reduce to $\vec{x} = A^T\vec{b}$.

Proof of Thm 3.4:-

We are trying to solve $C_n\vec{c} = \vec{f}$ with C_n orthogonal and $\vec{c} \in \mathbb{R}^m$, so the least-squares solution is simply

$$c_j = (C_n^{-1}\vec{f})_j = a_j \quad \text{for } j = 0, \ldots, m-1. \qquad \blacksquare$$

Example:- Interpolate the data $(\frac{\pi}{8}, 1), (\frac{3\pi}{8}, 0), (\frac{5\pi}{8}, -1), (\frac{7\pi}{8}, 0)$ with the DCT.

We have

$$\vec{a} = C_n^{-1}\vec{f} = \sqrt{\tfrac{1}{2}} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \cos(\frac{\pi}{8}) & \cos(\frac{3\pi}{8}) & \cos(\frac{5\pi}{8}) & \cos(\frac{7\pi}{8}) \\ \cos(\frac{2\pi}{8}) & \cos(\frac{6\pi}{8}) & \cos(\frac{10\pi}{8}) & \cos(\frac{14\pi}{8}) \\ \cos(\frac{3\pi}{8}) & \cos(\frac{9\pi}{8}) & \cos(\frac{15\pi}{8}) & \cos(\frac{21\pi}{8}) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \simeq \begin{bmatrix} 0 \\ 0.9239 \\ 1 \\ -0.3827 \end{bmatrix}$$

So the interpolant is

$$p_4(x) = \tfrac{1}{2}a_0 + \tfrac{1}{\sqrt{2}}a_1\cos(x) + \tfrac{1}{\sqrt{2}}a_2\cos(2x) + \tfrac{1}{\sqrt{2}}a_3\cos(3x)$$
$$= \tfrac{1}{\sqrt{2}}\Big(0.9239\cos(x) + \cos(2x) - 0.3827\cos(3x)\Big).$$

Thm 3.4 implies that the least-squares approximation using the basis $1, \cos(x), \cos(2x)$ is

$$p_3(x) = \tfrac{1}{\sqrt{2}}\Big(0.9239\cos(x) + \cos(2x)\Big).$$

Also the least-squares approximation using only $1, \cos(x)$ is

$$p_2(x) = \tfrac{1}{\sqrt{2}}(0.9239)\cos(x)$$

and using only a constant is

$$p_1(x) = 0.$$