see slides 03.2-(23-28)

Orthogonal transforms beyond the DFT: the Discrete Cosine Transform (DCT) the Walsh-Hadamard Transform the Haar Transform

Gianni Ramponi

DIA, University of Trieste, Italy

ramponi@units.it

The DFT is not the only possible way to represent a linear system or a finite-duration or periodic signal in a "frequency domain".
Other *basis functions* generate different transforms.
For the DFT, the direct and inverse basis functions respectively are:

$$f(n,k) = e^{-j2\pi nk/N} \qquad g(n,k) = e^{j2\pi nk/N}/N$$

but in general a transform can be expressed as

$$X(k) = \sum_{n=0}^{N-1} x(n) f(n,k)/N \qquad \mathbf{X} = \mathbf{F}\mathbf{x}/N \qquad \text{(analysis eq.)}$$

The matrix **F** must be *unitary*:

$$\mathbf{F}^{-\mathbf{1}} = \mathbf{F}^{*\mathbf{T}}$$

and the inverse transform is

$$x(n) = \sum_{k=0}^{N-1} X(k) f^*(n,k)) \qquad \mathbf{x} = \mathbf{F}^{*\mathbf{T}}\mathbf{X} \qquad \text{(synthesis eq.)}$$

$\mathbf{x}$ keeps its length, and the energy is conserved (Parseval's theorem):

$$||\mathbf{X}||^2 = \mathbf{X}^{*\mathbf{T}}\mathbf{X} = \mathbf{x}^{*\mathbf{T}}\mathbf{F}^{*\mathbf{T}}\mathbf{F}\mathbf{x} = \mathbf{x}^{*\mathbf{T}}\mathbf{x} = ||\mathbf{x}||^2$$

$\mathbf{X}$ is a rotated version of $\mathbf{x}$ in an N-dimensional space, i.e. the transform is a rotation of the coordinates, and the terms in $\mathbf{X}$ are the projections of $\mathbf{x}$ in the new space.

A suitable choice of the basis functions permits to exploit the *energy compaction property* of an orthogonal transform: i.e., most of the content of a signal $\mathbf{x}$ is represented in a *subset* of the coefficients $\mathbf{X}$.

If $\mathbf{x}$ is a random signal having autocorrelation $\mathbf{R_x}$, the most efficient transform is the Karhunen-Loeve Transform (KLT) (or Hotelling transform), that takes as bases the eigenvectors of the matrix $\mathbf{R_x}$.

The DFT has generally good compaction properties; for signals with strong autocorrelation, better performances are often provided by the **Discrete Cosine Transform** (DCT):

$$
\begin{aligned}
X(k) &= \alpha(k) \sum_{n=0}^{N-1} x(n) cos(\frac{\pi(2n+1)k}{2N}); \quad \alpha(0) = \sqrt{1/N},\ \alpha(k) = \sqrt{2/N} \\
x(n) &= \sum_{k=0}^{N-1} \alpha(k) X(k) cos(\frac{\pi(2n+1)k}{2N})
\end{aligned}
$$

In matrix form:

$$\mathbf{X} = \mathbf{C}\mathbf{x} \qquad \mathbf{x} = \mathbf{C}^{*\mathbf{T}}\mathbf{X} = \mathbf{C}^{\mathbf{T}}\mathbf{X}$$

The DCT of a real sequence is *real*. It is *not* symmetrical around π.

The DCT can be related to the DFT.
Let us define the *half-sample, symmetric, right periodic extension* of x:

$$y(n) = \{\dots x(0), x(1), \dots, x(N-1), x(N-1), x(N-2), \dots, x(1), x(0) \dots\}$$

$$\text{i.e.} \qquad \begin{aligned} y(n) &= x(n) & 0 \le n \le N-1 \\ &= x(2N-1-n) & N \le n \le 2N-1 \end{aligned}$$

(Note: 8 different versions of the DCT are obtained according to the type of periodic extension that is selected).

The DFT of y is

$$\begin{aligned} Y(k) &= \sum_{n=0}^{2N-1} y(n) e^{-j\frac{2\pi}{2N}kn} \\ &= \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{2N}kn} + \sum_{n=N}^{2N-1} x(2N-1-n) e^{-j\frac{2\pi}{2N}kn} \end{aligned}$$

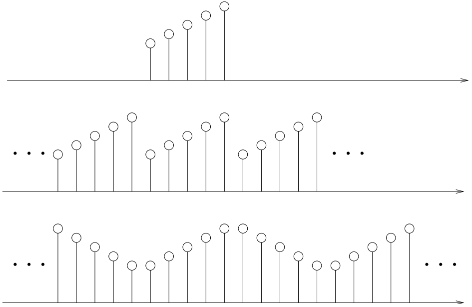
Let $m = 2N - 1 - n$. We get:

$$\begin{aligned} n &= N \rightarrow m = N-1 \\ n &= 2N-1 \rightarrow m = 0 \\ n &= 2N-1-m \end{aligned}$$

$$\begin{aligned} Y(k) &= \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{2N}kn} + \sum_{m=0}^{N-1} x(m) e^{-j\frac{2\pi}{2N}k(2N-1-m)} \\ &= \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{2N}kn} + \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{2N}k(-1-n)} \\ &= \sum_{n=0}^{N-1} x(n) (e^{-j\frac{2\pi}{2N}kn} + e^{-j\frac{2\pi}{2N}k(-1-n)}) \qquad \cdot\, e^{j\frac{2\pi}{2N}\frac{k}{2}} \quad e^{-j\frac{2\pi}{2N}\frac{k}{2}} \\ &= e^{j\frac{2\pi}{2N}\frac{k}{2}} \sum_{n=0}^{N-1} x(n) (e^{-j\frac{2\pi}{2N}k(n+1/2)} + e^{+j\frac{2\pi}{2N}k(n+1/2)}) \\ &= e^{j\frac{2\pi}{2N}\frac{k}{2}}\, 2 \sum_{n=0}^{N-1} x(n) (cos(\frac{2\pi}{2N}k(n+1/2)) \end{aligned}$$

Which is the DCT of **x** apart from some scaling factors. Then, it is possible to calculate the DCT using an FFT.

The periodicity artifacts are smaller in the DCT than in the DFT:



A signal and its DFT and DCT periodic extensions

- An **MDCT** (lapped) is e.g. used in MP3 coding.
- A **2-D DCT** is used in JPEG.

Another extremely simple and *real* orthogonal transform is:

$$\left\{ \begin{array}{lcl} X(0) & = & (x(0)+x(1))/\sqrt{(2)} \\ X(1) & = & (x(0)-x(1))/\sqrt{(2)} \end{array} \right. \qquad \left\{ \begin{array}{lcl} x(0) & = & (X(0)+X(1))/\sqrt{(2)} \\ x(1) & = & (X(0)-X(1))/\sqrt{(2)} \end{array} \right.$$

(for convenience, a single scaling by /2 can be done in either the direct or inverse transform)

It has the matrix formulation

$$F = \left\| \begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array} \right\|, \quad F_{inv} = \left\| \begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array} \right\| = F^{-1} = F^{T}$$

and is the elementary (N = 2) version of the **Walsh-Hadamard** and **Haar** transforms.

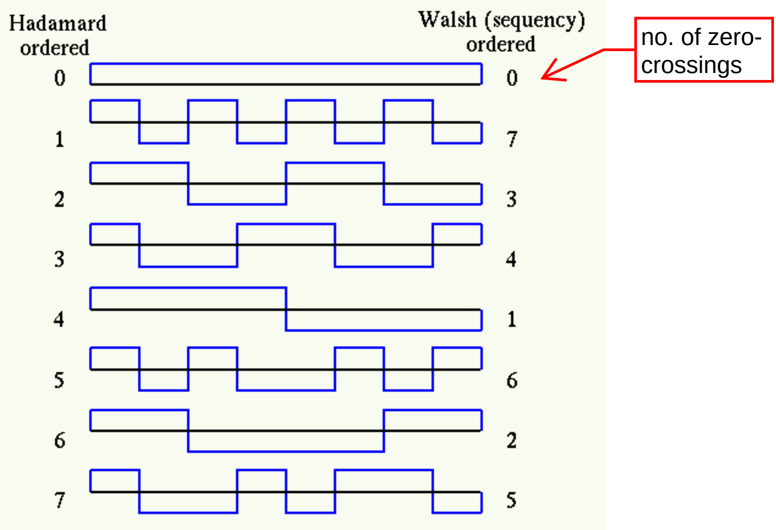
By the way, even the DFT in the elementary case N = 2 has the same matrix: its only element ≠ 1 is $F(2,2) = \exp(-j\ 2\ \pi\ (N-1)\ /\ N) = \exp(-j\ 2\ \pi\ 1/2) = -1$

The matrices for the higher-order transforms can be derived using a recursion. E.g., for N = 4 and N = 8 the **Walsh-Hadamard transform** matrices are

$$\mathbf{H}_2 = \mathbf{H}_1 \otimes \mathbf{H}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_1 \\ \mathbf{H}_1 & -\mathbf{H}_1 \end{bmatrix} = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

Kronecker product

$$\mathbf{H}_3 = \mathbf{H}_1 \otimes \mathbf{H}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_2 & -\mathbf{H}_2 \end{bmatrix} = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$



R. Wang, Introduction to orthogonal transforms, Cambridge University Press 2011

The matrix **H** is real, symmetric, and orthogonal:

$$\mathbf{H} = \mathbf{H}^* = \mathbf{H}^\mathbf{T} = \mathbf{H}^{-1}$$

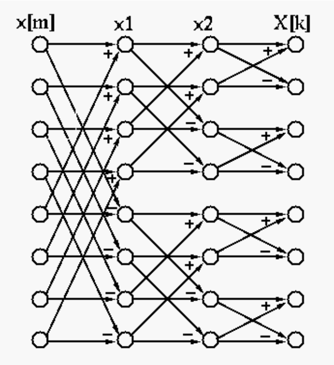
The forward and inverse transforms are

$$\mathbf{X} = \mathbf{H}\mathbf{x}$$

$$\mathbf{x} = \mathbf{H}\mathbf{X}$$

and they are identical.

Fast transform methods have been devised

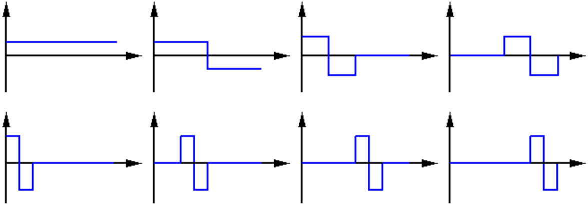


A different recursion leads to the matrices used in the **Haar transform**.
E.g., for $N = 8$ the matrix is

$$\mathbf{H}_8 = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix} \begin{array}{l} \varphi_0(t) \\ \psi_0(t) \\ \psi_{1,0}(t) \\ \psi_{1,1}(t) \\ \psi_{2,0}(t) \\ \psi_{2,1}(t) \\ \psi_{2,2}(t) \\ \psi_{2,3}(t) \end{array}$$

shift

scale



R. Wang, Introduction to orthogonal transforms, Cambridge University Press 2011

The Haar functions contain a single prototype shape: a square wave. The parameters specify the width (or scale) of the shape and its position (or shift).

In this way they represent not only the details in the signal but also their *locations* in time.

The Haar transform is the simplest *wavelet* transform.

Fast Haar Transform techniques have been devised.