The cosine-I transform

The cosine-I transform is an alternative to Fourier series; it is an expansion in an orthobasis for functions on [0, 1] (or any interval on the real line) where the basis functions look like sinusoids. There are two main differences that make it more attractive than Fourier series for certain applications:

- 1. the basis functions and the expansion coefficients are real-valued;
- 2. the basis functions have different symmetries.

Definition. The cosine-I basis functions for $t \in [0, 1]$ are

$$\psi_k(t) = \begin{cases} 1 & k = 0\\ \sqrt{2}\cos(\pi kt) & k > 0 \end{cases}$$
(1)

It is easy to check that the basis functions are orthonormal:

$$\langle \psi_j, \psi_k \rangle = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$$

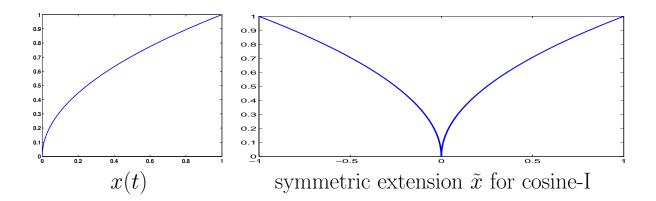
The question is whether or not they are complete; that is, can we build up any function on [0, 1] as a linear combination of the $\{\psi_k\}$. The completeness of cosine-I can be argued as follows. Let x(t)be an arbitrary real-valued function on [0, 1]. Let $\tilde{x}(t)$ be its symmetric extension on [-1, 1]

$$\tilde{x}(t) = \begin{cases} x(-t) & -1 \le t \le 0\\ x(t) & 0 \le t \le 1 \end{cases}$$

We can write the Fourier series (in \cos/\sin) form of $\tilde{x}(t)$ as

$$\tilde{x}(t) = \alpha_0 + \sum_{k=1}^{\infty} \alpha_k \cos(\pi kt) + \sum_{k=1}^{\infty} \beta_k \sin(\pi kt)$$

Notes by J. Romberg – January 8, 2012 – 16:47



(the interval has a length of 2, so the harmonics are $0, \pi, 2\pi, 3\pi, \ldots$), where

$$\alpha_k = \frac{1}{2} \int_{-1}^1 \tilde{x}(t) \cos(\pi kt) \, dt, \qquad \beta_k = \frac{1}{2} \int_{-1}^1 \tilde{x}(t) \sin(\pi kt) \, dt.$$

Since $\tilde{x}(t)$ is even, and $\sin(\pi kt)$ is odd, $\beta_k = 0$ for all $k \ge 1$. Thus

$$\tilde{x}(t) = \alpha_0 + \sum_{k=1}^{\infty} \alpha_k \cos(\pi k t).$$

Since x(t) is just the part of $\tilde{x}(t)$ on [0, 1], we can also write

$$x(t) = \alpha_0 + \sum_{k=1}^{\infty} \alpha_k \cos(\pi k t),$$

and so any function on [0, 1] can be written as a linear combinations of the ψ_k in (1) (the factor of $\sqrt{2}$ is there just to make the basis functions normalized).

Of course, the construction above is easily extended to an arbitrary interval $[T_1, T_2]$ with length $L = T_2 - T_1$. In this case, we simply take

$$\psi_k(t) = \begin{cases} \frac{1}{\sqrt{L}} & k = 0\\ \sqrt{\frac{2}{L}} \cos\left(\pi k \left(\frac{t - T_1}{L}\right)\right) & k > 0 \end{cases}.$$

Notes by J. Romberg – January 8, 2012 – 16:47

The discrete cosine transform (DCT)

The discrete version of the cosine-I transform is call the DCT: **Definition:** The DCT basis functions for \mathbb{R}^N are

$$\psi_k[n] = \begin{cases} \sqrt{\frac{1}{N}} & k = 0\\ \sqrt{\frac{2}{N}} \cos\left(\frac{\pi k}{N} \left(n + \frac{1}{2}\right)\right) & k = 1, \dots, N-1 \end{cases}, \quad n = 0, 1, \dots, N-1.$$
(2)

The DCT maps a real-valued vector in \mathbb{R}^N to another real-valued vector in \mathbb{R}^N , and can be computed efficiently using the fast Fourier transform (FFT). Note that

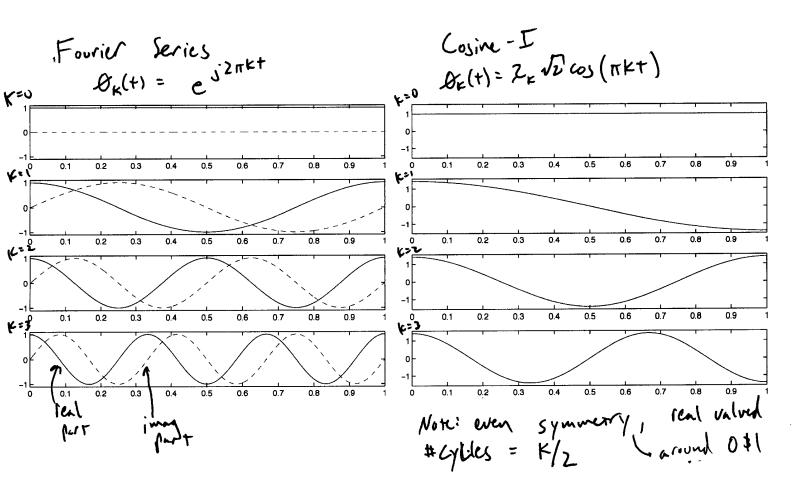
$$\cos\left(\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right) = \operatorname{Re}\left\{e^{-jk\pi/2N}e^{-jk\pi n/N}\right\},\,$$

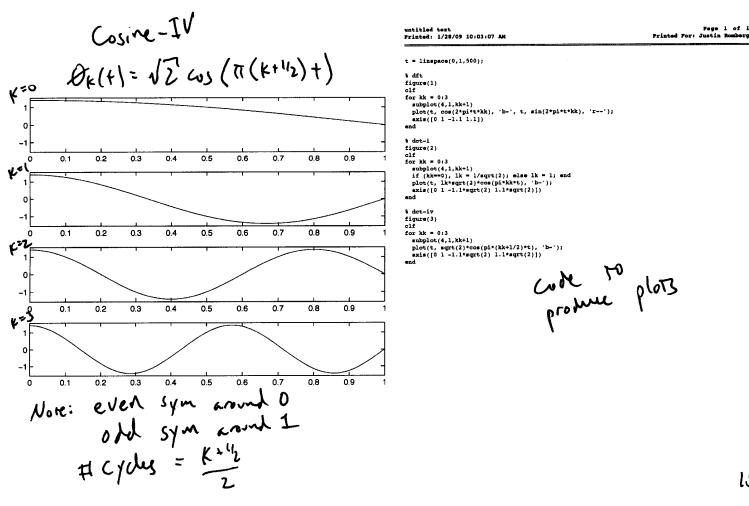
and so if x[n] is real-valued

$$\sum_{n} x[n] \cos\left(\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right) = \operatorname{Re}\left\{e^{-jk\pi/2N}\sum_{n} x[n]e^{-jk\pi n/N}\right\}.$$

Thus in MATLAB, to take the DCT of \mathbf{x} , we could use

This is just meant as an illustration, as there are more efficient ways of doing this than the above (and MATLAB has a nice built-in dct function).





Notes by J. Romberg – January 8, 2012 – 16:47

The cosine-I and DCT for 2D images

Just as for Fourier series and the discrete Fourier transform, we can leverage the 1D cosine-I basis and the DCT into separable bases for 2D images.

Definition. Let $\{\psi_k(t)\}_{k\geq 0}$ be the cosine-I basis in (1). Set

$$\psi_{k_1,k_2}^{\text{2D}}(s,t) = \psi_{k_1}(s)\psi_{k_2}(t).$$

Then $\{\psi_{k_1,k_2}^{2\mathrm{D}}(s,t)\}_{k_1,k_2\in\mathbb{N}}$ is an orthonormal basis for $L_2([0,1]^2)$

This is just a particular instance of a general fact. It is straightforward to argue (you can do so at home) that if $\{\psi_{\gamma}(t)\}_{\gamma\in\Gamma}$ is an orthonormal basis for $L_2([0, 1])$, then $\{\psi_{\gamma_1}(s)\psi_{\gamma_2}(t)\}_{\gamma_1,\gamma_2\in\Gamma}$ is an orthonormal basis for $L_2([0, 1]^2)$.

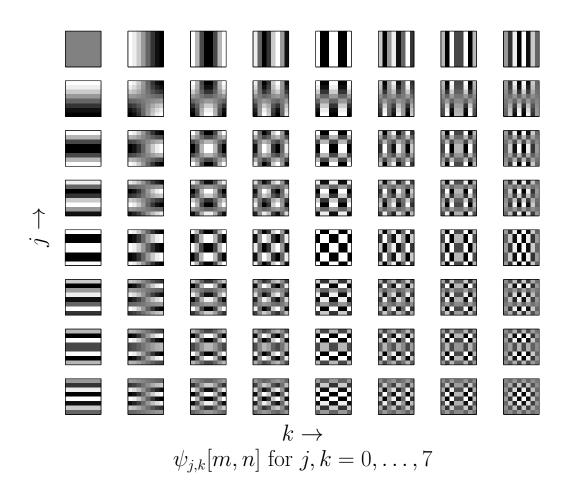
The DCT extends to 2D in the same way.

Definition. Let $\{\psi_k[n]\}_{0 \le k \le N-1}$ be the DCT basis in (2). Set

$$\psi_{j,k}^{\mathrm{2D}}[m,n] = \psi_j[m]\psi_k[n].$$

Then $\{\psi_{j,k}^{2\mathrm{D}}[m,n]\}_{0 \le j,k \le N-1}$ is an orthonormal basis for $\mathbb{R}^N \times \mathbb{R}^N$.

The 64 DCT basis functions for N = 8 are shown below:



2D DCT coefficients are indexed by two integers, and so are naturally arranged on a grid as well:

$lpha_{0,0}$	$lpha_{0,1}$	•••	$lpha_{0,N-1}$
$lpha_{1,0}$	$lpha_{1,1}$	•••	$\alpha_{1,N-1}$
:	÷	:	÷
$\alpha_{N-1,0}$	$\alpha_{N-1,1}$	•••	$\alpha_{N-1,N-1}$

Notes by J. Romberg – January 8, 2012 – 16:47

The DCT in image and video compression

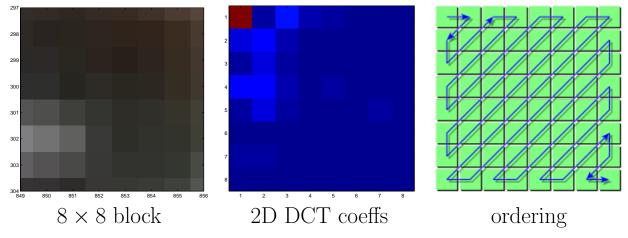
The DCT is basis of the popular JPEG image compression standard. The central idea is that while energy in a picture is distributed more or less evenly throughout, in the DCT transform domain it tends to be *concentrated* at low frequencies.

JPEG compression work roughly as follows:

- 1. Divide the image into 8×8 blocks of pixels
- 2. Take a DCT within each block
- 3. Quantize the coefficients the rough effect of this is to keep the larger coefficients and remove the samller ones
- 4. Bitstream (losslessly) encode the result.

There are some details we are leaving out here, probably the most important of which is how the three different color bands are dealt with, but the above outlines the essential ideas.

The basic idea is that while the energy within an 8×8 block of pixels tends to be more or less evenly distributed, the DCT concentrates this energy onto a relatively small number of transform coefficients. Moreover, the significant coefficients tend to be at the same place in the transform domain (low spatial frequencies).

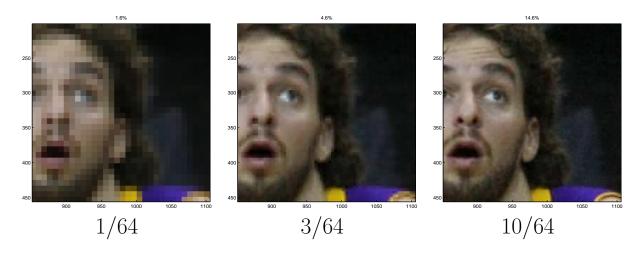


Notes by J. Romberg – January 8, 2012 - 16:47

To get a rough feel for how closely this model matches reality, let's look at a simple example. Here we have an original image 2048×2048 , and a zoom into a 256×256 piece of the image:



Here is the same piece after using 1 of the 64 coefficients per block $(1/64 \approx 1.6\%)$, $3/64 \approx 4.6\%$ of the coefficients, and $10/64 \approx 15/62\%$:



So the "low frequency" heuristic appears to be a good one.

JPEG does not just "keep or kill" coefficients in this manner, it quantizes them using a fixed quantization mask. Here is a common example:

$$Q = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

The quantization simply maps $\alpha_{j,k} \to \tilde{\alpha}_{j,k}$ using

$$\tilde{\alpha}_{j,k} = Q_{j,k} \cdot \text{round}\left(\frac{\alpha_{j,k}}{Q_{j,k}}\right)$$

You can see that the coefficients at low frequencies (upper left) are being treated much more gently than those at higher frequencies (lower right).

Video compression

The DCT also plays a fundamental role in video compression (e.g. MPEG, H.264, etc.), but in a slightly different way. Video codecs are complicated, but here is essentially what they do:

- 1. Estimate, describe, and quantize the motion in between frames.
- 2. Use the motion estimate to "predict" the next frame.
- 3. Use the (block-based) DCT to code the residual.

Cosine-IV transform

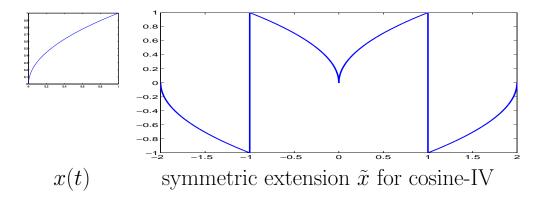
The cosine-IV transform is similar to cosine-I in that it is a basis of cosines at equally spaced frequencies (half harmonics). However, the frequencies used are offset to give the basis functions different symmetries — even at the left end point and odd at the right.

Definition. The cosine-IV basis functions for $t \in [0, 1]$ are

$$\psi_k(t) = \sqrt{2}\cos\left(\left(k + \frac{1}{2}\right)\pi t\right), \qquad k = 0, 1, 2, \dots$$
(3)

It is again an easy exercise to check that the basis functions are orthonormal. The completeness of the basis set can be argued in a similar manner to the cosine-I, but with a different symmetric extension of the signal. Let x(t) be an arbitrary real-valued function on [0, 1]. Let $\tilde{x}(t)$ be an extension defined on [-2, 2] as follows:

$$\tilde{x}(t) = \begin{cases} x(t) & t \in [0,1] \\ x(-t) & t \in [-1,0] \\ -x(2-t) & t \in (1,2] \\ -x(2+t) & t \in [-2,-1] \end{cases}$$



Notes by J. Romberg – January 8, 2012 – 16:47

The Fourier series (in cos/sin form) of $\tilde{x}(t)$ is

$$\tilde{x}(t) = \alpha_0 + \sum_{k=1}^{\infty} \alpha_k \cos\left(\frac{\pi kt}{2}\right) + \sum_{k=1}^{\infty} \beta_k \sin\left(\frac{\pi kt}{2}\right).$$

The extended signal $\tilde{x}(t)$ has support size 4, so the frequencies above are space $2\pi/4 = \pi/2$ apart. Since $\tilde{x}(t)$ is symmetric around zero, all of the β_k are again equal to zero. However, because $\tilde{x}(t)$ has odd symmetry around ± 1 , $\alpha_k = 0$ for k even. Thus

$$\tilde{x}(t) = \sum_{\substack{k \ge 1 \\ k \text{ odd}}} \alpha_k \cos\left(\frac{\pi kt}{2}\right)$$
$$= \sum_{k \ge 0} \alpha_k \cos\left(\left(k + \frac{1}{2}\right)\pi t\right).$$

And since x(t) is just the part of $\tilde{x}(t)$ on [0, 1], we have the same expansion

$$x(t) = \sum_{k \ge 0} \alpha_k \cos\left(\left(k + \frac{1}{2}\right)\pi t\right).$$

Again, the construction is easily extended to any interval $[T_1, T_2]$. With $L = T_2 - T_1$,

$$\psi_k(t) = \sqrt{\frac{2}{L}} \cos\left(\left(k + \frac{1}{2}\right)\pi\left(\frac{t - T_1}{L}\right)\right), \qquad k = 0, 1, 2, \dots$$

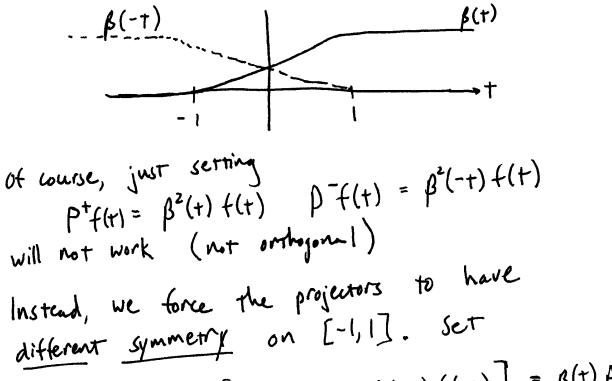
is an orthonormal basis for $L_2([T_1, T_2])$.

Notes by J. Romberg – January 8, 2012 – 16:47

Lapped Orthogonal Transforms

Lapped Projectors
How can we have overlapping windows which
Keep things orthogonal ?
To start, we will find projectors
$$\rho^+$$
, ρ^- for $fel_2(IR)$
such that
 $\bigcirc \rho^+f(t)$ is supported on $[-1, \infty)$
 $\bigcirc \rho^-f(t)$ is supported on $(-\infty, 1]$
 $\bigcirc \langle \rho^+f, \rho^-f \rangle = 0$ (orthogonal)
 $\bigcirc \rho^+f(t) + \rho^-f(t) = f(t)$
 $(\rho^+ + \rho^- = I)$
 $f(t) = \rho^-f(t)$ projection, we need an
auxiliary function $\beta(t)$ that satisfies
 $\beta(t) = \begin{cases} 0 + < -1 \\ 1 + > 1 \\ monotrial increasing for -1 s t = 1 \\ \beta^-(t) + \beta^2(-t) = I \end{cases}$

Notes by J. Romberg – January 8, 2012 - 16:47



$$\begin{array}{l} \rho^{+}f(t) = \beta(t) \left[\beta(t)f(t) + \beta(-t)f(-t) \right] = \beta(t)\beta(t) \\ \rho^{-}f(t) = \beta(-t) \left[\beta(-t)f(t) - \beta(t)f(-t) \right] = \beta(t)q(t) \end{array}$$

Note:

$$p(t) = \beta(t)f(t) + \beta(-t)f(-t) \quad is \quad \underline{even}$$

$$q(t) = \beta(-t)f(t) - \beta(t)f(-t) \quad is \quad \underline{odd}$$

$$q(t) = \beta(-t)f(t) - \beta(t)f(-t)$$

$$(also \quad \beta(t)\beta(-t) \quad is \quad trivially \quad even$$

$$Thus$$

$$\langle p^{+}f, p^{-}f \rangle = \int_{-1}^{1} \beta(t) \beta(-t) p(t) q^{*}(t) dt$$

=
$$\int_{-1}^{1} even \cdot even \cdot odd dt$$

=
$$\int_{-1}^{1} odd dt$$

=
$$\int_{-1}^{1} odd dt$$

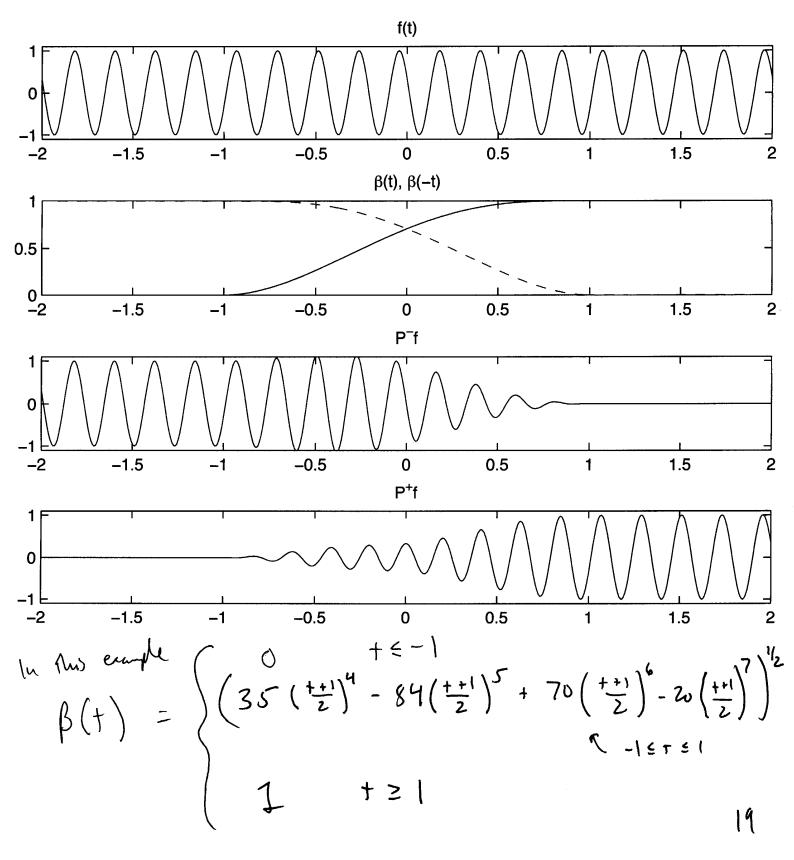
Notes by J. Romberg – January 8, 2012 - 16:47

Establish that

$$p^{+}f(t) + p^{-}f(t) = f(t)$$

proof:

.



Notes by J. Romberg – January 8, 2012 – 16:47

Pt and P⁻ are orthogonal projectors onto
the spaces
$$W^+, W^-$$

 $W^+ = space of f(t)$ in $L_2(IR)$ such that
 $\textcircledightingleftingl$

Using the same
$$\beta(t)$$
, we can decompose signals
into \bot - components with supports $[a-n, \infty)$ and $(-\infty, a+n]$
 $p_{a,n}^{+} f(t) = \beta(\frac{t-\alpha}{n}) \left[\beta(\frac{t-\alpha}{2}) f(t) + \beta(\frac{a-t}{2}) f(2a-t) \right]$
 $p_{a,n}^{-} f(t) = \beta(\frac{a-t}{2}) \left[\beta(\frac{a-t}{2}) f(t) - \beta(\frac{t-\alpha}{2}) f(2a-t) \right]$
(this is just shifting t scaling $\beta(t)$ and reflecting
 $f(t)$ across α) - spaces $W_{a,n}^{+}$, $W_{a,n}^{-}$ are
defined accordingly
 $projecting onto intervals$
Divide the real-line using an
 $\lim_{n\to\infty} a_n = -\infty$, $\lim_{n\to\infty} a_n = \infty$
 $\sum_{n\to\infty} \sum_{n\to\infty} \sum_{n\to\infty}$

Notes by J. Romberg – January 8, $2012-16{:}47$

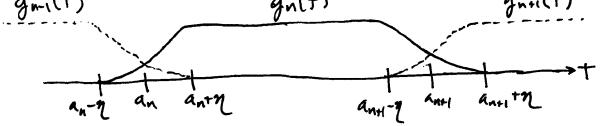
We can define the space

$$W^{n} = W^{+}_{an,n} \cap W^{-}_{an+1,n}$$
with orthogonal projecton

$$P_{n} = P^{+}_{an,n} P^{-}_{an+1,n}$$
The space W^{n} is characterized using a
window function $g_{n}(t)$

$$g_{n}(t) = \begin{cases} 0 \quad t < a_{n} - \eta \\ \beta(\frac{t-a_{n}}{\eta}) \quad a_{n} - \eta \le t \le a_{n+1} - \eta \\ 1 \quad a_{n} + \eta \le t \le a_{n+1} - \eta \\ \beta(\frac{a_{n+1} - t}{\eta}) \quad a_{n+1} - \eta \le t \le a_{n+1} + \eta \\ 0 \quad t > a_{n+1} + \eta \end{cases}$$

$$g_{n}(t) \qquad g_{n}(t) \qquad g_{n}(t)$$



Notes by J. Romberg – January 8, 2012 - 16:47

Applying what we have from before,
Wⁿ is the space of signals in L2(IR)
such that
$$f(t)$$
 can be written
 $f(t) = g_n(t) h(t)$
for some $h(t)$ that is
 $() \text{ symmetric with an} \\ h(t) = h(2a_n - t) \quad a_n - 1 \le t \le a_n + 1$
 $() \text{ anti-symmetric with anti-1} \quad a_{n-1} - 1 \le t \le a_{n+1} + 1$
 $h(t) = -h(2a_{n+1} - t) \quad a_{n-1} - 1 \le t \le a_{n+1} + 1$
The projection $P_n f(t)$ can be written
 $P_n f(t) = \begin{cases} P_{a_n,n} f(t) & a_n - 1 \le t \le a_{n+1} + 1 \\ h(t) = -h(2a_{n+1} - t) & a_{n+1} - 1 \le t \le a_{n+1} + 1 \\ h(t) = -h(2a_{n+1} - t) & a_{n+1} - 1 \le t \le a_{n+1} + 1 \\ h(t) = f(t) & a_{n+1} - 1 \le t \le a_{n+1} + 1 \\ h_n(t) = \begin{cases} q_n(t) f(t) + q_n(2a_n - t) f(2a_n - t) \\ f(t) \\ q_n(t) f(t) - q_n(2a_{n+1} - t) f(2a_{n+1} - t) \\ f(t) \\ q_n(t) f(t) - q_n(2a_{n+1} - t) f(2a_{n+1} - t) \end{cases}$

.

It is easy to show that

$$W^{n} \perp W^{m} \quad n \neq m$$

and $\sum_{n=-\infty}^{\infty} P_{n}f(t) = f(t) \quad \forall f(t) \in L_{2}(IR)$
 $(i.e. \sum_{n=-\infty}^{\infty} P_{n} = identity)$

Now that we have an "orthogonal partition"
of
$$L_2(IR)$$
, we construct an orthobasis for
each of the W?
We do this simply by taking an orthobasis
for $L_2[an, anti]$ and extending it in the
same way we did for cosine-IV:
Let $\{\mathcal{O}_{k}(t)\}_{k\in\mathbb{T}}$ be an orthobasis for $L_2[0,1]$
Set
 $\mathcal{O}_{k}(t) + \mathcal{O}_{k}(t) + \mathcal{O}_{k}(t)$

$$\widetilde{\mathscr{O}}_{k}(+) = \begin{cases} \mathscr{O}_{k}(+) & | e \in [-1, 0] \\ \mathscr{O}_{k}(-+) & +e(-1, 0) \\ -\mathscr{O}_{k}(2-+) & +e(-1, 2] \\ -\mathscr{O}_{k}(2++) & +e(-1, 2) \end{cases}$$

Notes by J. Romberg – January 8, $2012-16{:}47$

Then

$$\begin{cases} g_n(t) \frac{1}{\sqrt{2n}} \quad \tilde{\mathscr{O}}_{K} \left(\frac{t-a_n}{2n} \right) \\ \text{KeT} \end{cases}$$
is an orthonormal busis for W^n .
To see this, we will establish the following useful fact:
Suppose $f_i(t), f_2(t) \in W^n$; we can write $f_i(t) = g_n(t) \cdot h_i(t)$
 $f_2(t) = g_n(t) \cdot h_2(t)$
where h_i, h_2 obey the symmetry properties discussed above. Then
 $\langle f_{i,1}, f_2 \rangle = \int_{a_n}^{a_{n+1}} h_i(t) h_2^*(t) dt$

profi
by definition

$$\langle f_{11}, f_{2}, 7 = \int_{a_{n-1}}^{a_{n+1}+\eta} h_{1}(t) h_{2}^{*}(t) g_{1n}^{2}(t) dt$$

break the integral into three parts
 $\int_{a_{n-1}}^{a_{n+1}+\eta} \int_{a_{n+1}}^{a_{n+1}+\eta} \int_{a_{n+1}}^{a_{n+1}+\eta} f_{1n}^{a_{n+1}+\eta} + \int_{a_{n+1}}^{a_{n+1}+\eta} f_{1n}^{a_{n+1}+\eta}$
exist 25

Notes by J. Romberg – January 8, $2012-16{:}47$

First integral

$$\int_{a_{n+1}}^{a_{n+1}} h_{1}(t) h_{2}^{*}(t) g_{n}^{2}(t) dt$$

$$= \int_{a_{n}}^{a_{n+1}} h_{1}(t) h_{2}^{*}(t) \left[g_{n}^{2}(t) + g_{n}^{2}(2a_{n}-t) \right] dt$$
Since h_{1}, h_{2} are symmetric w.r.t. an
We explicitly designed $g_{n}(t)$ so that
 $g_{n}^{2}(t) + g_{n}^{2}(2a_{n}-t) = 1$ $t + e[a_{n}, a_{n}+1]$

$$\int_{a_{n}}^{a_{n}+1} h_{1}(t) h_{2}^{*}(t) g_{n}^{2}(t) dt = \int_{a_{n}}^{a_{n}+1} h_{1}(t) h_{2}^{*}(t) dt$$
Second integral
 $\int_{a_{n}+1}^{a_{n}+1} h_{1}(t) h_{2}^{*}(t) g_{n}^{2}(t) dt$

$$= \int_{a_{n}+1}^{a_{n}+1} h_{1}(t) h_{2}^{*}(t) dt$$
since $g_{n}(t) = 1$ on this interval

•

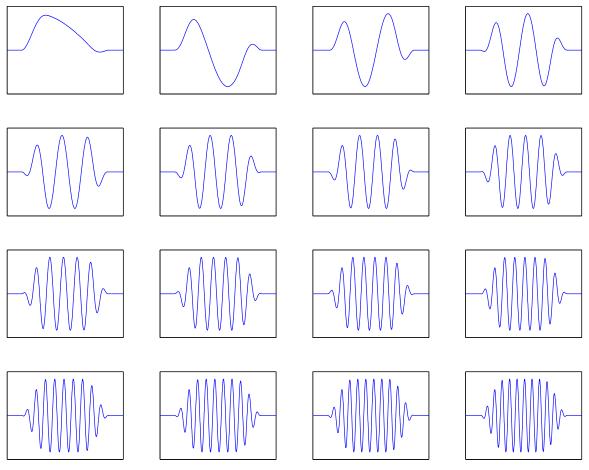
Third integral

$$\int_{a_{n+1}+n_{2}}^{a_{n+1}+n_{2}} h_{1}(t)h_{2}^{*}(t) q_{n}^{2}(t)dt$$

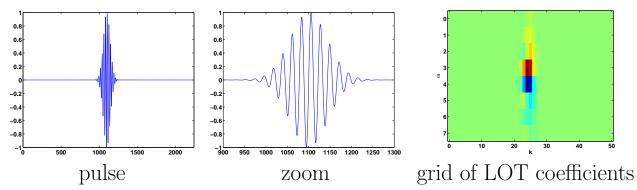
$$= \int_{a_{n+1}-n_{2}}^{a_{n+1}} h_{1}(t)h_{2}^{*}(t)dt$$
where we argue just as we did for the
first integral. Combining the three integrals,
we have
 $\langle f_{1}, f_{2} \rangle = \int_{a_{n}}^{a_{n+1}} h_{1}(t)h_{2}^{*}(t)dt$

Finally, the
$$\{q_n(t), \frac{1}{\sqrt{L_n}}, \tilde{\mathcal{O}}_{\mathcal{K}}(\frac{t-a_n}{2n})\}$$
 $n\in\mathbb{Z}, k\in\mathbb{N}$
form an orthobasis for $L_2(\mathbb{IR})$.

Plots of the LOT basis functions, single window, first 16 frequencies:



LOT of a modulated pulse:



Notes by J. Romberg – January 8, 2012 - 16:47