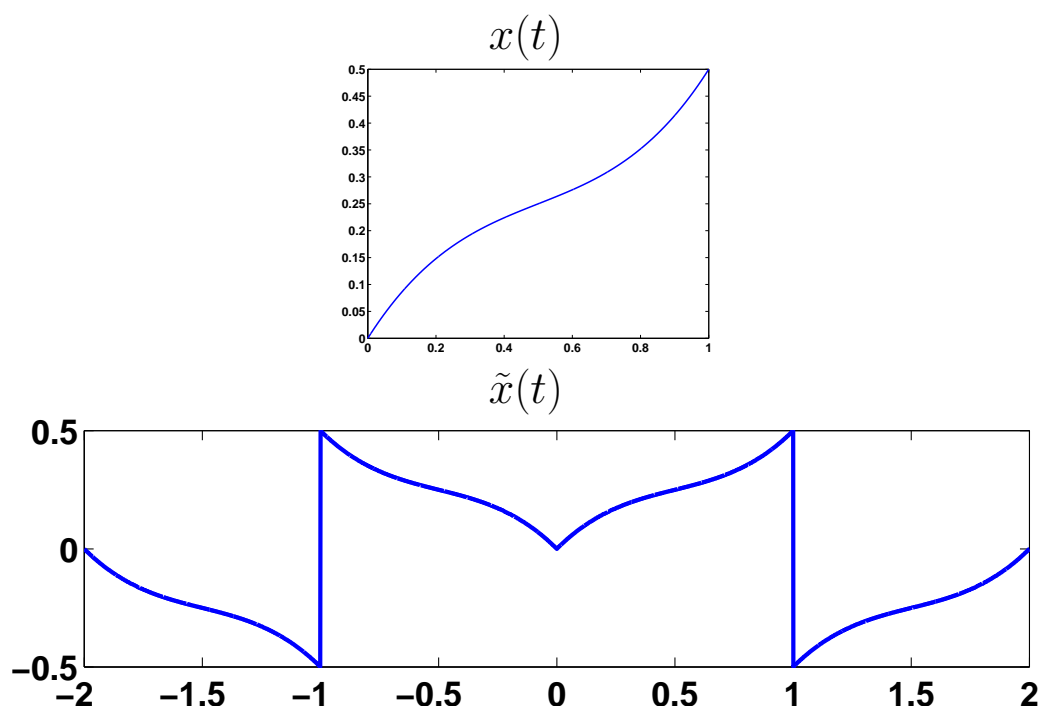# The cosine-IV basis

We can derive a transform which is related (but different) from the cosine-I by starting from a different type of extension for $x(t)$. Let $x(t)$ be a continuous-time signal on $[0, 1]$, and let $\tilde{x}(t)$ be an extension on $[-2, 2]$ defined by

$$\tilde{x}(t) = \begin{cases} -x(t+2) & -2 \leq t < -1 \\ x(-t) & -1 \leq t \leq 0 \\ x(t) & 0 \leq t \leq 1 \\ -x(2-t) & 1 < t \leq 2 \end{cases}$$

Example:

$x(t)$

$\tilde{x}(t)$

102

Notice that $\tilde{x}(t)$ has even symmetry around $t = 0$ and odd symmetry around $t = -1$ and $t = 1$. Since $\tilde{x}(t)$ is a continuous-time signal on $[-2, 2]$, we can represent it using the Fourier series:

$$\tilde{x}(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos\left(\frac{\pi k t}{2}\right) + \sum_{k=1}^{\infty} b_k \sin\left(\frac{\pi k t}{2}\right).$$

We will have $b_k = 0$ for all k above, since $\tilde{x}(t)$ has even symmetry around $t = 0$. In addition, we will also have $a_k = 0$ for k even, thanks to the odd symmetry around $t = \pm 1$. Thus the $\{\cos(\pi k t/2) : k \text{ odd}\}$ are a basis for functions on $[-2, 2]$ with these types of symmetries, and also for arbitrary functions on $[0, 1]$.
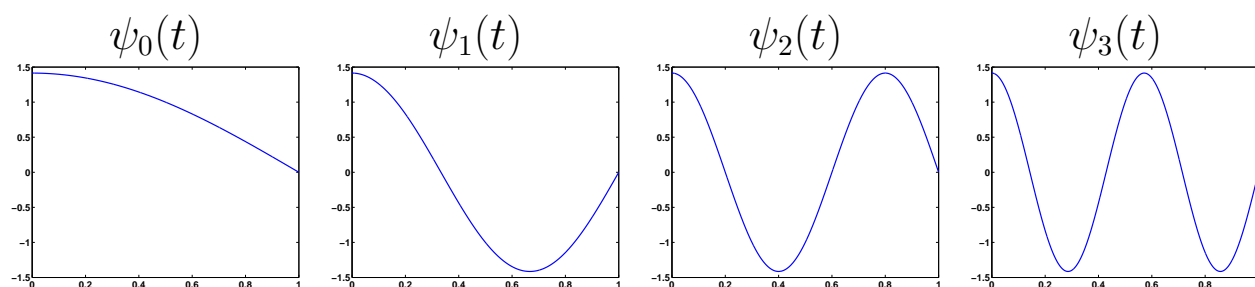
Cosine-IV Orthobasis:

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

form an orthobasis for $L_2([0, 1])$. We argued above that any signal on $[0, 1]$ be be synthesis using a linear combination of these $\boldsymbol{\psi}_k$; it is a straightforward exercise to check that $\langle \boldsymbol{\psi}_k, \boldsymbol{\psi}_\ell \rangle = 0$ for $k \neq \ell$ and $\|\psi_k(t)\|_2 = 1$.

One way to thing about the cosine-IV expansion is that we are taking a 4x oversampled Fourier Series (with frequencies spaced $\pi/2$ apart instead of 2π), then only using the real part, and then throwing away every other coefficient.

Here are the first four cosine-IV basis functions:

There is also a discrete version of the cosine-IV for $\mathbb{R}^N$. The vectors

$$\psi_k[n] = \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{N}\left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right), \quad k = 0, 1, \ldots, N-1$$

with sample indices $n = 0, 1, \ldots, N-1$ form and orthobasis for $\mathbb{R}^N$. This is sometimes referred to as the DCT-IV. Just as with the regular DCT, there is a fast algorithm for computing the DCT-IV transform that uses the FFT.

Block transforms

In general, we interested in signals that live on long time intervals. To represent these in an orthobasis, we can break the real line into intervals, and build up each interval separately. This was essentially what was happening when we talked about JPEG — it broke the image into 8×8 blocks, a represented each using a 2D DCT.

To make this more precise, suppose that $\{\psi_k(t), \ k \geq 0\}$ is an orthobasis for $L_2([0,1])$. Then it should be clear that for a fixed integer n, $\{\psi_k(t-n), \ k \geq 0\}$ is an orthobasis for $L_2([n, n+1])$. If we define $\psi_{n,k}(t) = \psi_k(t-n)$, then

$$\{\psi_{n,k}(t), \ n \in \mathbb{Z}, \ k \geq 0\} \quad \text{is an orthobasis for } L_2(\mathbb{R}),$$

and so we can write any $x(t) \in L_2(\mathbb{R})$ as

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{k=0}^{\infty} \langle \boldsymbol{x}, \boldsymbol{\psi}_{n,k} \rangle \, \psi_{n,k}(t).$$

For this type of **block transform**, the basis signals $\boldsymbol{\psi}_{n,k}$ and the expansion coefficients $\alpha_{n,k} = \langle \boldsymbol{x}, \boldsymbol{\psi}_{n,k} \rangle$ are naturally arranged on a grid with horizontal index n (which we might call the **time frame** index) and vertical index k (which in this case corresponds to a frequency).

If we use the DCT in each block, we can grid the basis functions as below:
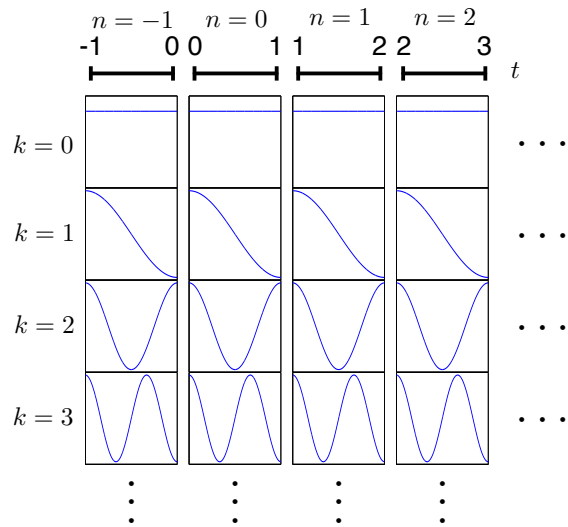
Here is a signal along with its block DCT expansion coefficients:

$$x(t)$$

$$\{\alpha_{n,k}\}$$

$$k \rightarrow$$

106
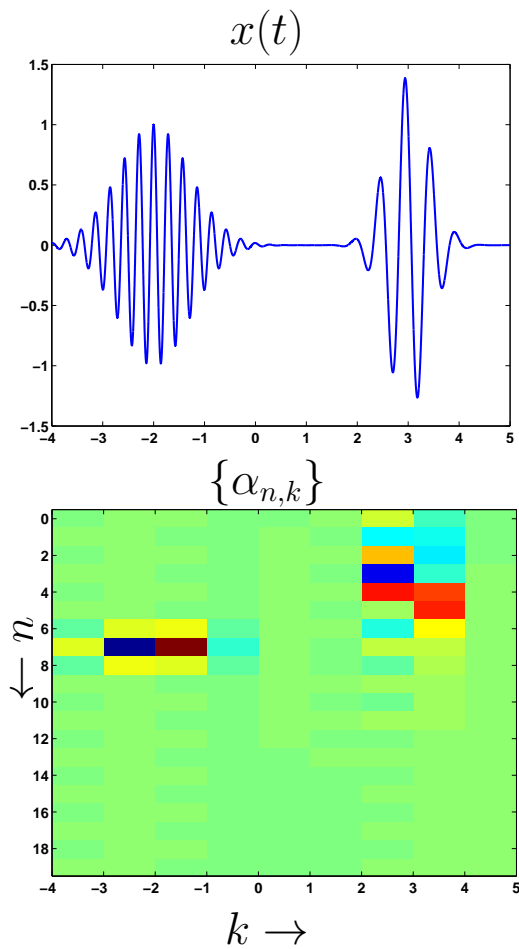
Of course, we can use intervals of any length, not just 1. In fact, each window could be a different length. In the case of the block DCT (or Fourier series), the width of the interval determines how densely the frequencies are sampled — wider windows will have more closely spaced frequencies.

One problem with block transforms is that the continuity of the signal is disrupted at the edges of each block. For example, suppose for the two pulses above, I perturb the $k = 0$ coefficients by adding ± 0.1 — the reconstructed signal looks like
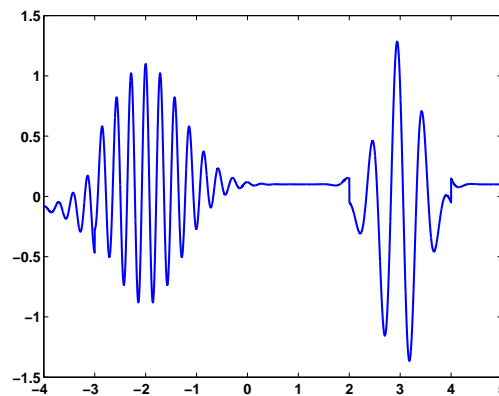

Notice the sharp discontinuities at $t = 2$ and $t = 4$.

In the next section, we will see a more graceful way to "blend" the intervals together.

107

The Lapped Orthogonal Transform

The lapped orthogonal transform (or LOT) is a modified version of the cosine-IV transform that breaks the signal into **overlapping** frames. It essentially works by multiplying the signal by a smooth window (centered on the current frame being analyzed), then taking a cosine transform of the result. The tricky part of doing this is maintaining orthogonality. This is done by imposing certain symmetry properties on the window and exploiting the symmetries in the cosine-IV transform.

The LOT (and variations thereof) are used extensively in audio codecs[1], including mp3. It is also called the *modified discrete cosine transform*.

The LOT starts with a smooth windowing function $g(t)$ which we use to "pull out" $x(t)$ around the interval $[0, 1]$.

$g(t)$
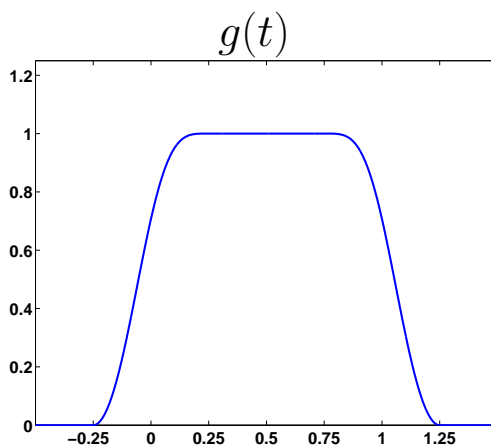
The specific properties we need from $g(t)$ are

1. $g(t) = 0$ for $t < -1/4$ and $t > 5/4$,

---

[1]The LOT was developed by Henrique Malvar, a researcher at Microsoft, specifically for encoding music.

2. $g(t) = 1$ for $1/4 \leq t \leq 3/4$,

3. $g(t)$ is symmetric around $t = 1/2$,

4. $g(t)$ is monotonically increasing for $-1/4 \leq t \leq 1/4$ (and so with symmetry monotonically decreasing for $3/4 \leq t \leq 5/4$),

5. $|g(t)|^2 + |g(t-1)|^2 = 1$ for $3/4 \leq t \leq 5/4$.

The last property, along with the symmetries, tells us that the sum of the squares of all of the integer shifts of the window function is identically equal to one:

$$\sum_{n=-\infty}^{\infty} |g(t-n)|^2 = 1 \quad \text{for all } t.$$

There are many windows which obey these properties. One popular one (and in fact the one which is shown) is

$$g(t) = \begin{cases} \sin\left(\frac{\pi}{2}\sin^2(\pi(t + \frac{1}{4}))\right) & -1/4 \leq t \leq 1/4 \\ 1 & 1/4 \leq t \leq 3/4 \\ \sin\left(\frac{\pi}{2}\sin^2(\pi(\frac{9}{4} - t))\right) & 3/4 \leq t \leq 5/4 \end{cases}.$$

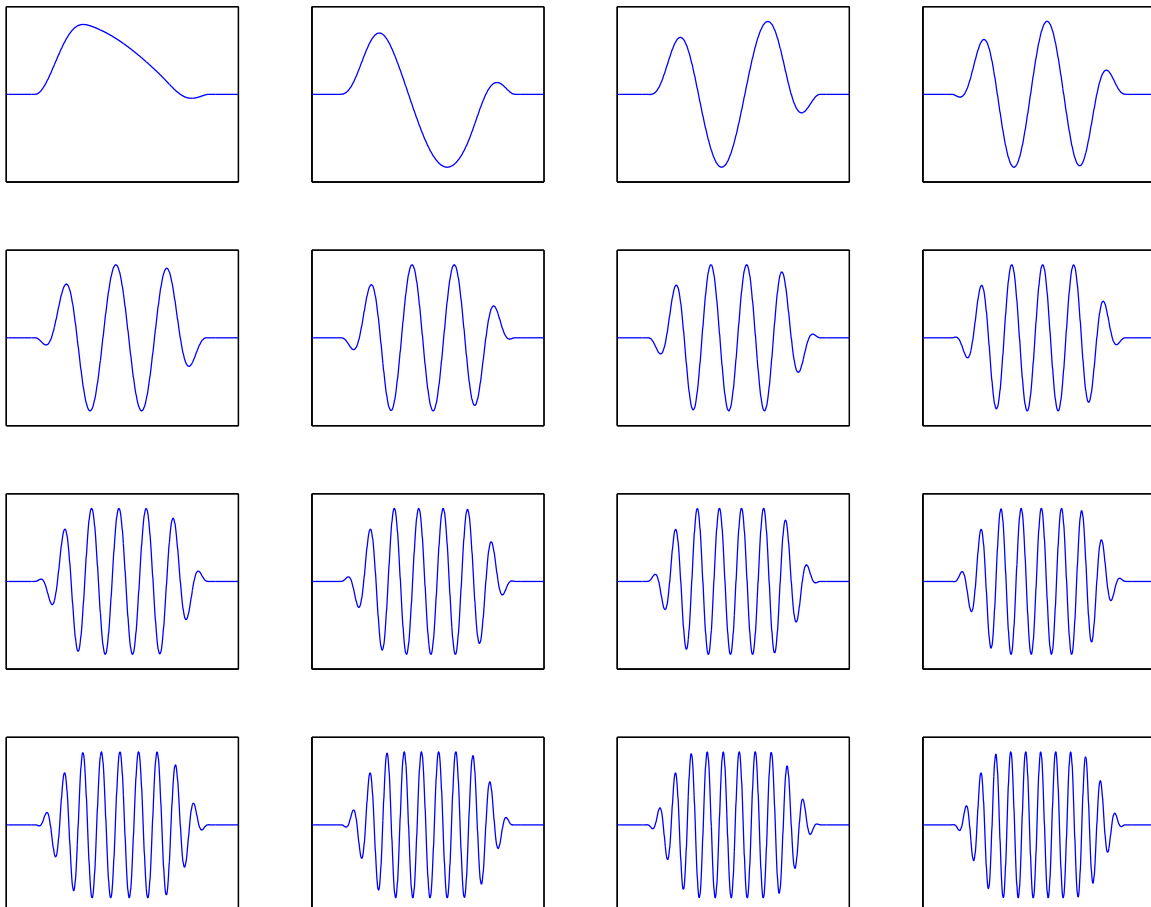We shift this window around to analyze different frames of the signal:

$$g(t-n), \ n = -2, \ldots, 2$$

109
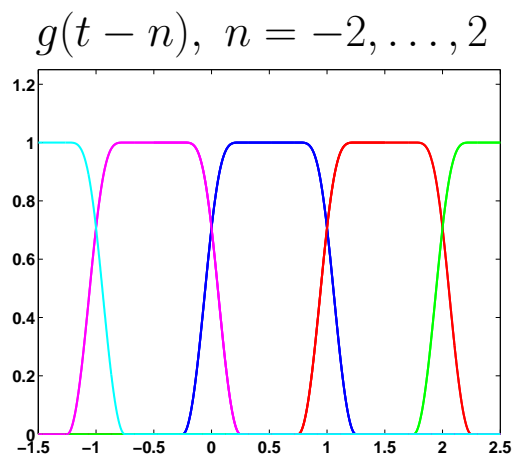
We create an orthobasis by modulating this window by the cosine-IV basis functions. The LOT orthobasis is

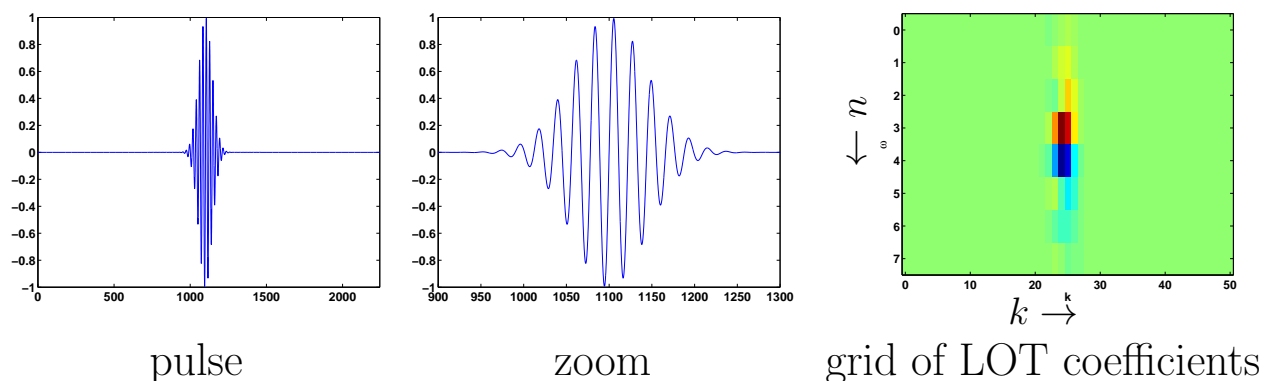$$\phi_{n,k}(t) = g(t-n)\psi_k(t-n), \quad n \in \mathbb{Z}, \ k \geq 0$$

where

$$\psi_k(t-n) = \sqrt{2}\cos\left(\pi\left(k+\frac{1}{2}\right)(t-n)\right)$$

Here are what the LOT basis functions look like for a fixed n and $k = 0, \ldots, 15$:

Here is the LOT of a modulated pulse:

| pulse | zoom | grid of LOT coefficients |

LOT: Technical Details

Let's look a little more closely at why the $\{\phi_{n,k}(t),\ n \in \mathbb{Z},\ k \in \mathbb{Z}^+\}$ form an orthobasis for $L_2(\mathbb{R})$. We will show that any $x(t) \in L_2(\mathbb{R})$ can be written using the reproducing formula

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{k=0}^{\infty} \langle \boldsymbol{x}, \boldsymbol{\phi}_{n,k} \rangle \phi_{n,k}(t), \quad \text{for all} \ -\infty < t < \infty.$$

We start with the inner product

$$\langle \boldsymbol{x}, \boldsymbol{\phi}_{n,k} \rangle = \int_{-\infty}^{\infty} x(t)\phi_{n,k}(t) \ dt = \int_{n-1/4}^{n+5/4} x(t)g(t-n)\psi_k(t-n) \ dt,$$

(1)

where $\psi_k(t)$ the kth cosine-IV basis vector extended to $[-1/4, 5/4]$:

$$\psi_k(t) = \sqrt{2} \, \cos\left(\left(k + \frac{1}{2}\right)\pi t\right), \quad t \in [-1/4, 5/4].$$

Recall that $\psi_k(t)$ has even symmetry around $t = 0$ and odd symmetry around $t = 1$:

$$\psi_k(-t) = \psi(t), \qquad \psi_k(t+1) = -\psi_k(-t+1).$$

111

The odd symmetry condition around $t = 1$ can also be written as $\psi_k(2 - t) = -\psi_k(t)$.

Looking at an inner product in (1) for $n = 0$, we break it down into three parts:

$$\int_{-1/4}^{5/4} x(t)g(t)\psi_k(t) \ \mathrm{d}t = \int_{-1/4}^{1/4} x(t)g(t)\psi_k(t) \ \mathrm{d}t \qquad (2)$$
$$+ \int_{1/4}^{3/4} x(t)g(t)\psi_k(t) \ \mathrm{d}t$$
$$+ \int_{3/4}^{5/4} x(t)g(t)\psi_k(t) \ \mathrm{d}t.$$

Now we manipulate the first and last terms using the symmetry properties of $g(t)$ and $\psi_k(t)$. For the first term:

$$\int_{-1/4}^{1/4} x(t)g(t)\psi_k(t) \ \mathrm{d}t = \int_{-1/4}^{0} x(t)g(t)\psi_k(t) \ \mathrm{d}t + \int_{0}^{1/4} x(t)g(t)\psi_k(t) \ \mathrm{d}t$$
$$= \int_{0}^{1/4} \left(x(t)g(t) + x(-t)g(-t) \right) \psi_k(t) \ \mathrm{d}t,$$
$$(3)$$

where the last step follows from the fact that $\psi_k(-t) = \psi_k(t)$.

For the middle term in (2),

$$\int_{1/4}^{3/4} x(t)g(t)\psi_k(t) \ \mathrm{d}t = \int_{1/4}^{3/4} x(t)\psi_k(t) \ \mathrm{d}t, \qquad (4)$$

since $g(t) = 1$ for $-1/4 \leq t \leq 3/4$.

112

For the last term in (2),

$$\int_{3/4}^{5/4} x(t)g(t)\psi_k(t) \ dt = \int_{3/4}^{1} x(t)g(t)\psi_k(t) \ dt + \int_{1}^{5/4} x(t)g(t)\psi_k(t) \ dt$$

$$= \int_{3/4}^{1} \left(x(t)g(t) - x(2-t)g(2-t) \right) \psi_k(t) \ dt,$$

$$(5)$$

where the last step follows from the fact that $\psi_k(2-t) = -\psi_k(t)$.

By combining (3),(4), and (5), we see that

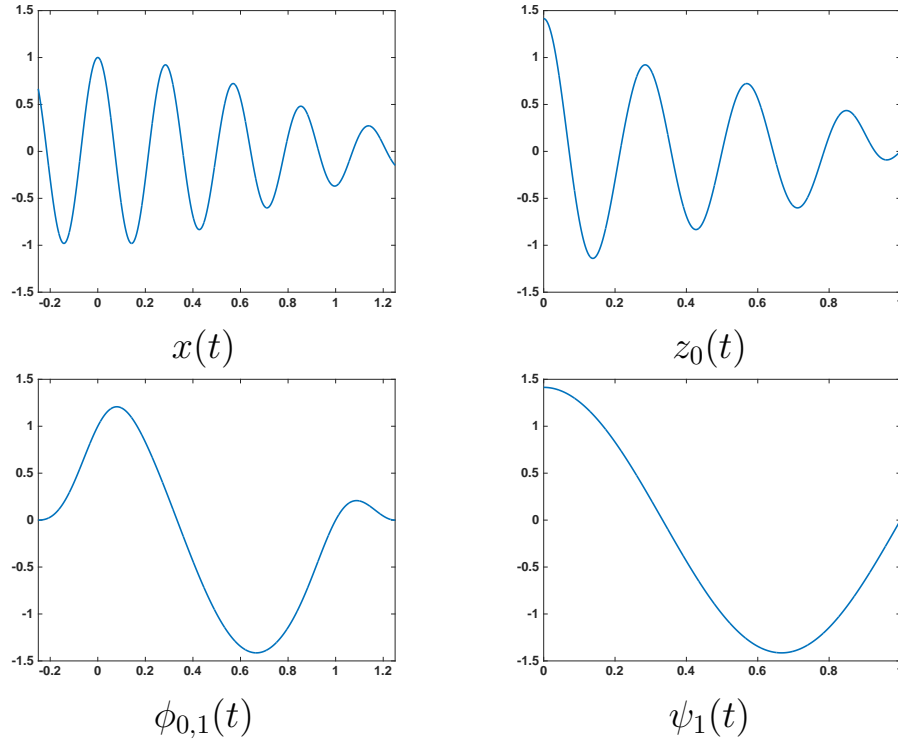$$\langle \boldsymbol{x}, \boldsymbol{\phi}_{0,k} \rangle = \int_{0}^{1} z_0(t)\psi_k(t) \ dt,$$

where $z_0(t)$ is the following signal on $[0, 1]$:

$$z_0(t) = \begin{cases} x(t)g(t) + x(-t)g(-t) & 0 \le t \le 1/4 \\ x(t) & 1/4 \le t \le 3/4 \ . \\ x(t)g(t) - x(2-t)g(2-t) & 3/4 \le t \le 1 \end{cases}$$

That is

$$\langle x(t), \phi_{0,k}(t) \rangle_{L_2(\mathbb{R})} = \langle z_0(t), \psi_k(t) \rangle_{L_2([0,1])}.$$

In pictures, this means that the inner product on $[-0.25, 1.25]$ of the two signals in the left-hand column below is equal to the inner product on $[0, 1]$ of the two signals on the right:

113

$$x(t)$$

$$z_0(t)$$

$$\phi_{0,1}(t)$$

$$\psi_1(t)$$

This tells us that the piece of the representation corresponding to $n = 0$ is equivalent to decomposing $g(t)z_0(t)$ in the cosine-IV orthobasis. More precisely, since the $\{\psi_k(t)\}_k$ are an orthobasis for $[0,1]$, it is clear that when we synthesize for the frame $n = 0$:

$$\sum_{k=0}^{\infty}\langle \boldsymbol{x}, \phi_{0,k}\rangle \phi_{0,k}(t) = \sum_{k=0}^{\infty}\langle \boldsymbol{z}_0, \boldsymbol{\psi}_k\rangle_{L_2([0,1])} g(t)\psi_k(t)$$

$$= g(t)\left(\sum_{k=0}^{\infty}\langle \boldsymbol{z}_0, \boldsymbol{\psi}_k\rangle_{L_2([0,1])}\psi_k(t)\right)$$

$$= g(t)z_0(t) \quad \text{for } t \in [0,1].$$

But since $z_0(t)$ has even symmetry around $t = 0$ and odd symmetry around $t = 1$, the sum above reproduces $z_0(t)$ on the entire interval $t \in [-1/4, 5/4]$ (recall how the cosine-IV basis was constructed

earlier in these notes). Thus we can extend the synthesis interval:

$$\sum_{k=0}^{\infty} \langle \boldsymbol{x}, \phi_{0,k} \rangle \phi_{0,k} = g(t)z_0(t) \quad \text{for } t \in \left[-\frac{1}{4}, \frac{5}{4} \right].$$

Similarly, for the $n = 1$ frame, we have

$$\sum_{k=0}^{\infty} \langle \boldsymbol{x}, \phi_{1,k} \rangle \phi_{1,k}(t) = g(t-1)z_1(t) \quad \text{for } t \in \left[\frac{3}{4}, \frac{9}{4} \right],$$

where
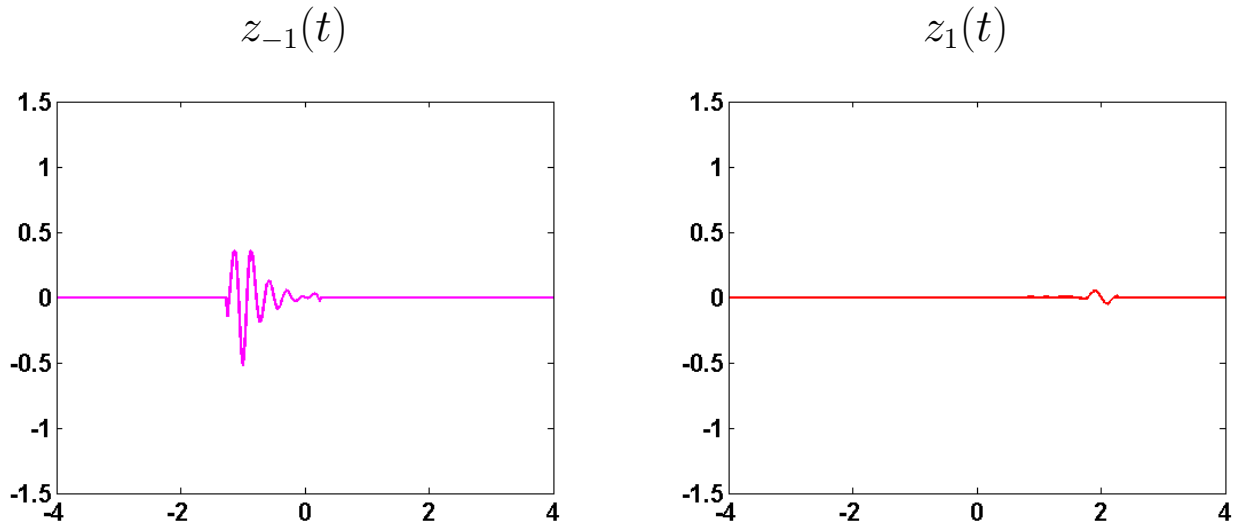
$$z_1(t) = \begin{cases} x(t)g(t-1) + x(2-t)g(1-t) & 3/4 \le t \le 5/4 \\ x(t) & 5/4 \le t \le 7/4 \,, \\ x(t)g(t-1) - x(4-t)g(3-t) & 7/4 \le t \le 9/4 \end{cases}$$

and for the $n = -1$ frame we have

$$\sum_{k=0}^{\infty} \langle \boldsymbol{x}, \phi_{-1,k} \rangle \phi_{-1,k}(t) = g(t+1)z_{-1}(t) \quad \text{for } t \in \left[-\frac{5}{4}, \frac{1}{4} \right],$$

where

$$z_{-1}(t) = \begin{cases} x(t)g(t+1) + x(-2-t)g(-t-1) & -5/4 \le t \le -3/4 \\ x(t) & -3/4 \le t \le -1/4 \,. \\ x(t)g(t+1) - x(-t)g(1-t) & -1/4 \le t \le 1/4 \end{cases}$$

It is clear that when we add these pieces together

$$z_{-1}(t)g(t+1) + z_0(t)g(t) + z_1(t)g(t-1) = x(t),$$

for $t \in [-3/4, -1/4]$ and $t \in [1/4, 3/4]$ and $t \in [5/4, 7/4]$. What about the intervals where the windows overlap? Well, for $t \in [-1/4, 1/4]$, we have

$$
\begin{aligned}
g(t+1)z_{-1}(t) + g(t)z_0(t) &= x(t)g^2(t+1) - x(-t)g(t+1)g(1-t) \\
&\quad + x(t)g^2(t) + x(-t)g(t)g(-t) \\
&= x(t),
\end{aligned}
$$

where the last equality comes from the facts that by design

$$g^2(t) + g^2(t+1) = 1, \quad \text{for } t \in [-1/4, 1/4]$$

and since $g(t)$ has even symmetry around $t = 1/2$,

$$g(t) = g(1-t), \quad \text{and} \quad g(-t) = g(t+1), \quad \text{for } t \in [-1/4, 1/4].$$

116

Similarly for $t \in [3/4, 5/4]$, where the frames for $n = 0$ and $n = 1$ overlap

$$
\begin{aligned}
g(t)z_0(t) + g(t-1)z_1(t) &= x(t)g^2(t) - x(2-t)g(t)g(2-t) \\
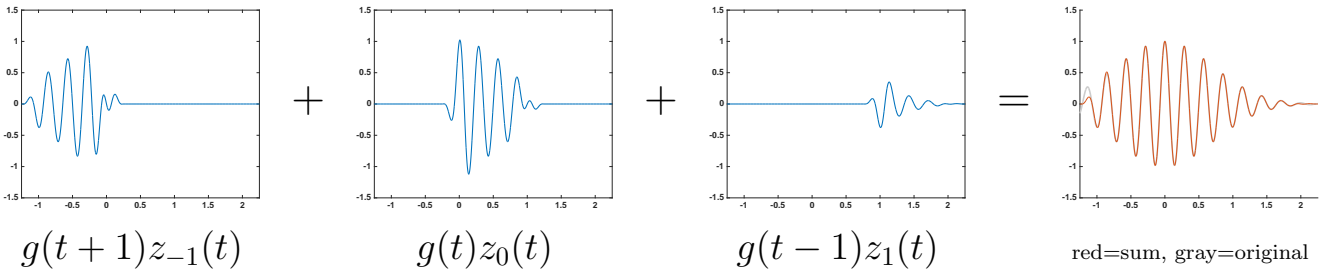&\quad + x(t)g^2(t-1) + x(2-t)g(t-1)g(1-t) \\
&= x(t),
\end{aligned}
$$

since
$$
g^2(t) + g^2(t-1) = 1, \quad \text{for } t \in [3/4, 5/4],
$$

and since $g(t-1)$ is symmetric around $t = 3/2$,

$$
g(t) = g(1-t) \quad \text{and} \quad g(2-t) = g(t-1), \quad \text{for } t \in [3/4, 5/4].
$$

Thus, $g(t+1)z_{-1}(t) + g(t)z_0(t) + g(t-1)z_1(t)$ matches $x(t)$ exactly on $[-0.75, 2.75]$, and tapers off in a smooth, controlled way to 0 on $[-1.25, -0.75]$ and $[2.75, 3.32]$. Here is a picture:

$g(t+1)z_{-1}(t)$ $+$ $g(t)z_0(t)$ $+$ $g(t-1)z_1(t)$ $=$ red=sum, gray=original

Similar arguments hold for the rest of the overlapping regions between frames, where for general n

$$
\sum_{k=0}^{\infty} \langle \boldsymbol{x}, \phi_{n,k} \rangle \phi_{n,k}(t) = g(t-n)z_n(t) \quad \text{for } t \in [n - 1/4, n + 5/4],
$$

117

where

$$z_n(t) = \begin{cases} x(t)g(t-n) + x(2n-t)g(-t+n) & n - \frac{1}{4} \le t \le n + \frac{1}{4} \\ x(t) & n + \frac{1}{4} \le t \le n + \frac{3}{4} \\ x(t)g(t-n) - x(2n+2-t)g(n+2-t) & n + \frac{3}{4} \le t \le n + \frac{5}{4} \end{cases}.$$

Thus we conclude that

$$\sum_{n=-\infty}^{\infty} \sum_{k=0}^{\infty} \langle \boldsymbol{x}, \boldsymbol{\phi}_{n,k} \rangle \phi_{n,k}(t) = \sum_{n=-\infty}^{\infty} g(t-n)z_n(t) = x(t), \ \forall t \in \mathbb{R}.$$

Below, the plot on the left shows five terms $g(t-n)z_n(t)$ in the sum above for $n = 0, \ldots, 4$. The plot on the right shows the sum of those for terms. As expected, the sum matches the original function $x(t)$ (shown in light grey) on the appropriate interval $[-0.25, 3.75]$.

$g(t-n)z_n(t), \ n = 0, \ldots, 4$ $\sum_{n=0}^{4} g(t-n)z_n(t)$