Section 25: Approximations in Reproducing Kernel Hilbert Spaces

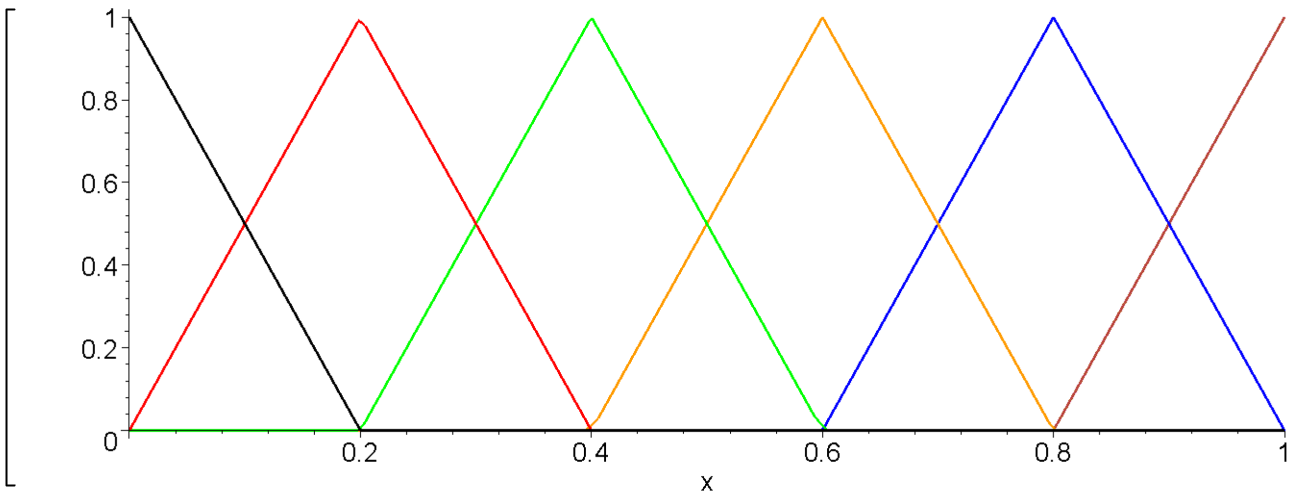
In this section, we address two concepts. One is the wish that if {E, < , >} is an innerproduct space of real valued functions on the interval [0,1], then there should be a function **K** from [0,1]x[0,1] to the real numbers, such that if e is a number in [0,1], then g(x) = **K**(x,e) is in E, and if f is also in E, then

< f(x), **K**(x,e) > = f(e).

The second is the wish that if f is a reasonably nice function in E and ϕ_n, n = 0 .. 5, is a sequence of functions which have graphs that are triangles with peaks at t_p = p/5, p=0..5, (See the next graphs.) then there should be a reasonably good approximation for f(x) as

$$f(x) \sim \sum_{p=0}^{5} f(tp)\, \phi_p(x).$$

In some circumstances these wishes come true.



Both these wishes are related to notions concerning use of the Dirac Delta function. The first idea is related to the often used integral

$$\int_0^1 f(x)\, \delta(x, e)\, dx = f(e)$$

where $\delta(x, e)$ is the Dirac Delta. The second is the common notion that reasonably nice functions can be written as sums of multiples of the Dirac Delta. We will show that the commonly used Dirac Delta cannot exist in the Hilbert Space of square integrable functions on [0,1]. The formulation of such a function is most often properly put in the context of distributions.

<u>Definition:</u> A Hilbert Space of functions on [0,1] has a *reproducing kernel* if there is a function **K** such that, for each e in [0,1], the function g(x) = **K**(x,e) is in E and, for each f in E,

f(e) = < f(x), **K**(x,e) >.

Theorem. If { E, < , >} is a Hilbert Space of functions on [0,1] which has a reproducing kernel, then normed convergence in E implies pointwise convergence on [0,1]. Moreover, if there is a number B such that $|\mathbf{K}(x,x)| < B$, then normed convergence in E implies uniform convergence on [0, 1].

Proof: A proof follows from this inequality: Suppose $f_n(x)$, n=1, 2, 3,... is a sequence in E. Then

$$| f_n(e) - f_m(e) | = | < f_n - f_m , \mathbf{K}(\ ,e) > | \ \leq \ | f_n - f_m | \ |\mathbf{K}(\ , e)| = | f_n - f_m | \sqrt{\mathrm{K}(e, e)} \ .$$

Comments:

1. It follows that in the Hilbert Space of square integrable functions on [0, 1], if e is in [0,1], then g(x) = δ(x,e) is **not** in E. This follows because we know that there are sequences of functions in that space which converge in the norm of E, but do not converge pointwise on [0,1].
2. We will show later that the property of having normed convergence to imply pointwise convergence in a Hilbert Space of functions on [0,1] is also a sufficient condition to assure that the Hilbert Space has a reproducing kernel.

Example: Let E consist of all functions f on [0,1] that are continuous and for which the derivative f ' is continuous except possibly for a finite number of jumps. Also, take f(0) to be zero for all functions in E. Define the inner product by

$$< f, g > = \int_0^1 f\,'(x)\, g\,'(x) \ dx.$$

Take {E, < , >} to be the Hilbert Space formed from the completion of the innerproduct space defined above. Define the function **K** by

$$\mathbf{K}(x,y) = \min(x,y).$$

Suppose f has a continuous derivative. Then

$$< f(x), \mathbf{K}(x,y) > = \int_0^y \left(\frac{\partial}{\partial x} f\right) 1 \ dx = f(y) - f(0) = f(y).$$

Consequently, this Hilbert Space has a reproducing kernel. It is the **K** as defined above.

Observation 1: With **K** as defined for this example, if a < b ≤ s < t, then

$$< \mathbf{K}(x,b) - \mathbf{K}(x,a), \mathbf{K}(x,t) - \mathbf{K}(x,s) > = 0.$$

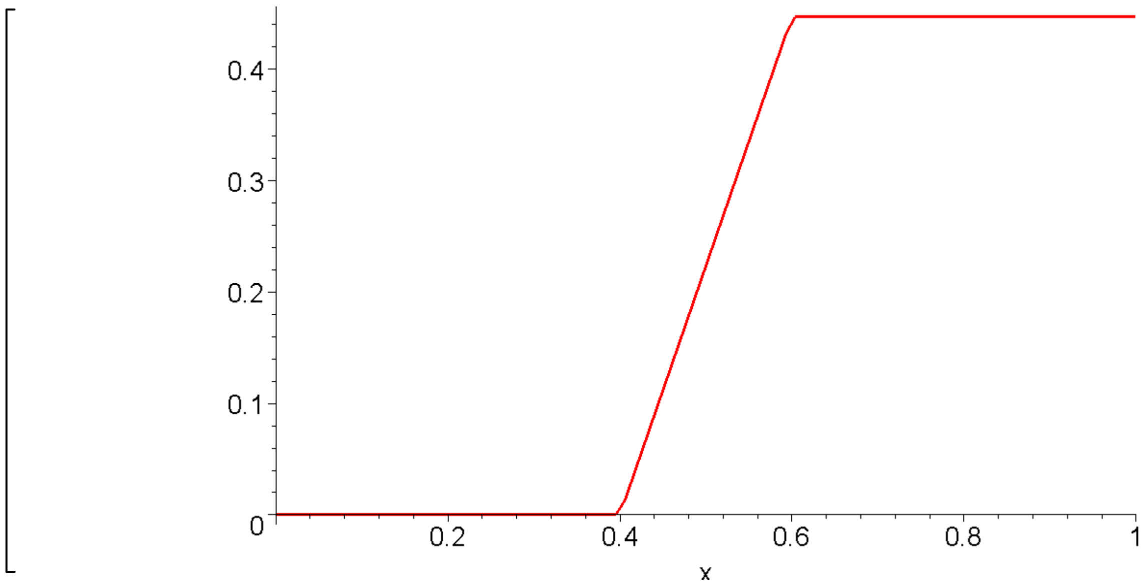
To see this, note that if c < d, then the derivative of **K**(x,d) - **K**(x,c) is the characteristic function on [c, d]. That is, it is the function that is 1 on [c, d] and zero elsewhere. Thus, in the space of the example, the dot product of

$$< \mathbf{K}(x,b) - \mathbf{K}(x,a), \mathbf{K}(x,t) - \mathbf{K}(x,s) >$$

is zero for non overlapping intervals.

<u>Observation 2:</u> If n is a positive integer, then the family { $\sqrt{n}$ (**K**(x, p/n) - **K**(x, (p-1)/n) }, p=1, 2, ... n, is an orthonormal sequence in the Example above.

To see that this observation is true, the above Observation 1 establishes that the terms of the sequence are orthogonal. We have only to see that they have norm 1. But, this is verified by integrating the constant function 1 from (p-1)/n to p/n and multiply by $\sqrt{n}$. The result is 1. For illustration, we draw the graph of one of these functions in case n = 5.



<u>Observation 3:</u> Take ϕ_p, p=1, 2, ...n, to be the orthonormal sequence defined above and suppose that f is in the Hilbert Space of the Example. We know that the best approximation for f with the ϕ_p's is given by

$$f \sim \sum_{p=1}^{n} c_p \, \phi_p$$

and that the coefficients c_p are the Fourier coefficients:

$$c_p = < f , \phi_p >.$$

Further, we know how good the approximation is

$$| f - \sum_{p=1}^{n} c_p \, \phi_p |^2 \le |f|^2 - \sum_{p=1}^{n} |c_p|^2 .$$

<u>Observation 4</u>: We calculate both terms of the right hand side of the above inequality in case f ' is continuous except for a finite number f jumps:

$$|f|^2 = \int_0^1 f\,'(x)^2 \, dx,$$

and

$$\sum_{p=1}^{n} |c_p|^2 = 5\sum_{p=1}^{n}\left|\int_{\frac{p-1}{n}}^{\frac{p}{n}} f'(x)\,dx\right|^2 .$$

<u>Observation 5:</u> Recall that if f ' is continuous on the interval [(p-1)/n , p/n] then

c_p = f(p/n) - f((p-1)/n)

so that, in this case, we have an easy way to calculate the coefficients.

<u>Two Approximation Examples:</u>

<u>Example 1</u>. Suppose that f(x) = x^2 on [0,1]. We compute (a) $|f|^2$, (b) $\sum_{p=1}^{n} |c_p|^2$, (c) the difference in these two, and (d) draw a graph to illustrate the approximation.

Here is the computation for (a)

$$f := x \to x^2$$

$$\int_0^1 \left(\frac{d}{dx}(x^2)\right)^2 dx = \frac{4}{3}$$

Here is the computation for (b)

$$\sqrt{n}\int_{\frac{p-1}{n}}^{\frac{p}{n}} \frac{d}{dx}(x^2)\,dx = \frac{2\,p-1}{n^{(3/2)}}$$

$$\sum_{p=1}^{n} \frac{(2\,p-1)^2}{n^3} = \frac{-1+4\,n^2}{3\,n^2}$$

We can make the computation of (c):

$|f|^2$ = 4/3,

$$\sum_{p=1}^{n} |c_p|^2 = \frac{4}{3} - \frac{1}{3\,n^2},$$

and $$| f - \sum_{p=1}^{n} c_p\,\phi_p |^2 \le \frac{1}{3\,n^2} .$$

For (d) we draw the graphs with n = 5 to show the closeness of the approximation. The graph of f is black and the approximation is red.

$$c_1 := \frac{\sqrt{5}}{25}$$

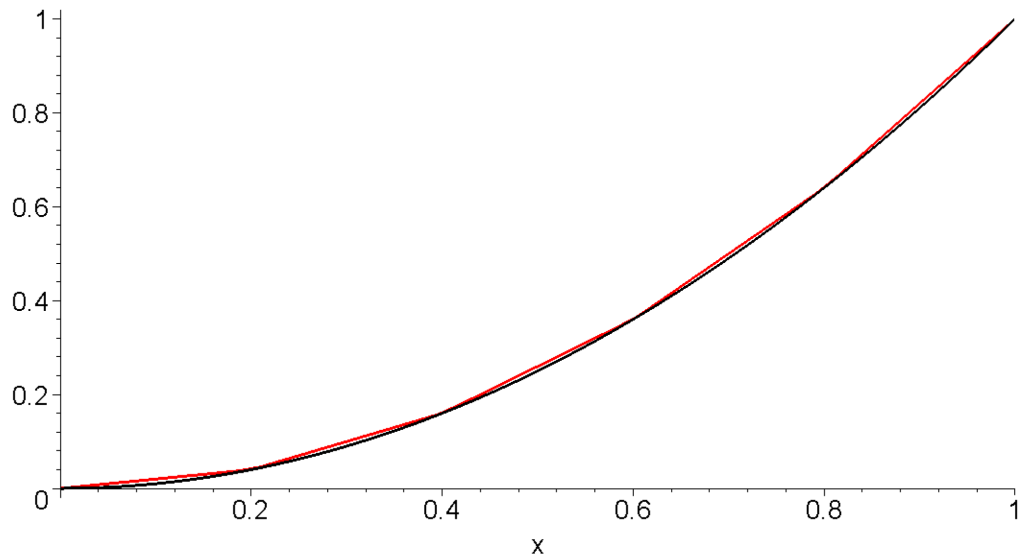
$$c_2 := \frac{3\sqrt{5}}{25}$$

$$c_3 := \frac{\sqrt{5}}{5}$$

$$c_4 := \frac{7\sqrt{5}}{25}$$

$$c_5 := \frac{9\sqrt{5}}{25}$$

$$approx := x \to \sum_{p=1}^{n} c_p \sqrt{5} \left(\mathrm{K}\left(x, \frac{p}{n}\right) - \mathrm{K}\left(x, \frac{p-1}{n}\right) \right)$$



Example 2. Suppose that f(x) = $x\,(1-x)$ on [0,1]. We make the same computations.
Here is the computation for (a)

$$\int_0^1 \left(\frac{d}{dx}(x\,(1-x)) \right)^2 dx = \frac{1}{3}$$

Here is the computation for (b)

$$\sqrt{n}\int_{\frac{p-1}{n}}^{\frac{p}{n}} \frac{d}{dx}(x(1-x))\,dx = \frac{n-2p+1}{n^{(3/2)}}$$

$$\sum_{p=1}^{n} \frac{(n-2p+1)^2}{n^3} = \frac{n^2-1}{3n^2}$$

We can make the computation of (c):

$$|f|^2 = 1/3,$$

$$\sum_{p=1}^{n} |c_p|^2 = \frac{1}{3} - \frac{1}{3n^2},$$

and $$|f - \sum_{p=1}^{n} c_p \phi_p|^2 \le \frac{1}{3n^2}.$$

For (d) we draw the graphs with n = 5 to show the closeness of the approximation. The graph of f is black and the approximation is red.

$$c_1 := \frac{4\sqrt{5}}{25}$$

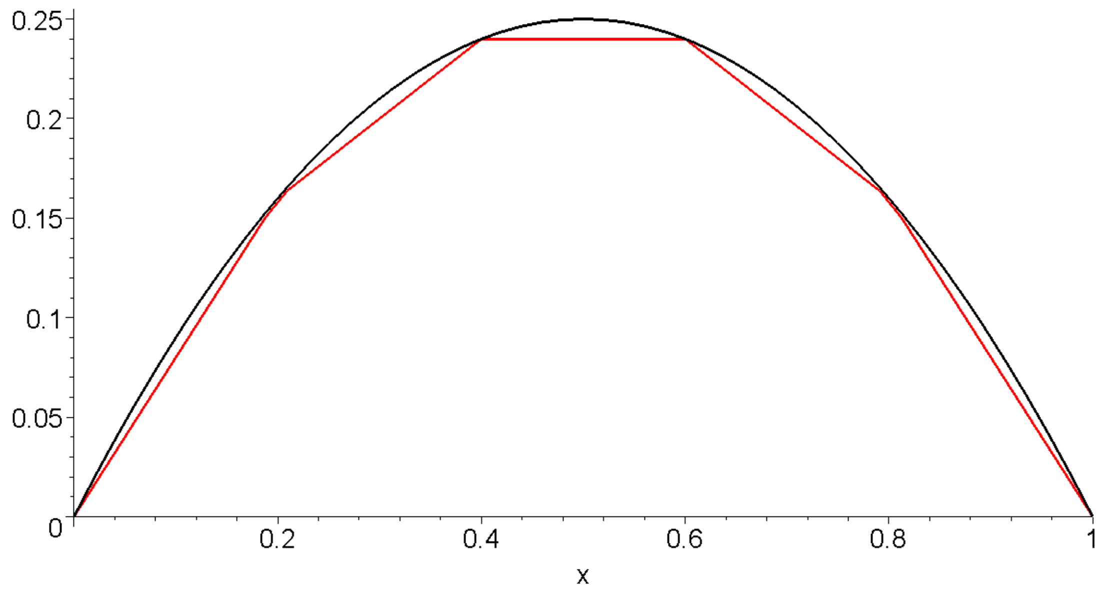
$$c_2 := \frac{2\sqrt{5}}{25}$$

$$c_3 := 0$$

$$c_4 := -\frac{2\sqrt{5}}{25}$$

$$c_5 := -\frac{4\sqrt{5}}{25}$$

$$approx := x \to \sum_{p=1}^{n} c_p \sqrt{5}\left(\mathrm{K}\left(x, \frac{p}{n}\right) - \mathrm{K}\left(x, \frac{p-1}{n}\right)\right)$$



Alternate Computations.

We present here computations for the approximation that are simpler and make an interesting geometric construction. Here is the idea. Suppose we have a function h which has a continuous derivative. Then the Fourier Coefficients are easily computed:

$$\int_a^b h\,'(x)\,dx = \text{h(b) - h(a).}$$

In this case, we can write the approximation as

$$\sum_{p=1}^{n}\left(\mathrm{h}\left(\frac{p}{n}\right)-\mathrm{h}\left(\frac{p-1}{n}\right)\right)n\left(\mathrm{K}\left(x,\frac{p}{n}\right)-\mathrm{K}\left(x,\frac{p-1}{n}\right)\right).$$

Observe that h(p/n) is multiplied by n (2 **K**(x,p/n) - **K**(x,(p-1)/n) - **K**(x,(p+1)/n)). Hence, an alternate representation for the approximation that involves no computation of coefficients and the associate integration can be created as follows:

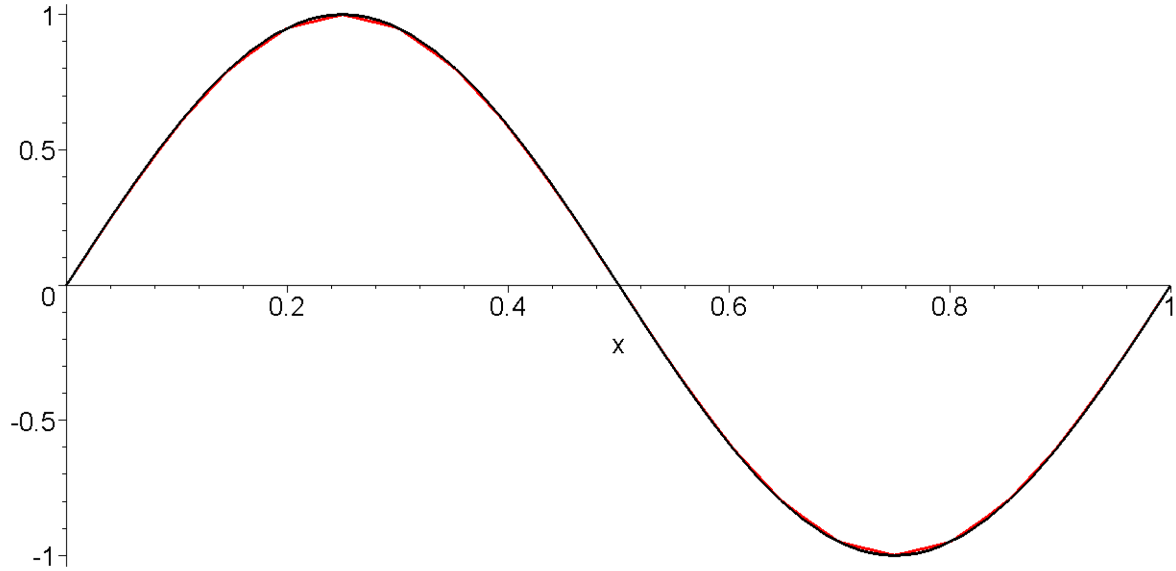
define $\quad T_p(x)$ = n (2 **K**(x,p/n) - **K**(x,(p-1)/n) - **K**(x,(p+1)/n)), p=1, 2, ...n.

Then $\quad \mathrm{h}(x) \sim \sum_{p=1}^{n} \mathrm{h}\left(\frac{p}{n}\right) T_p(x)$

We present two illustrations.

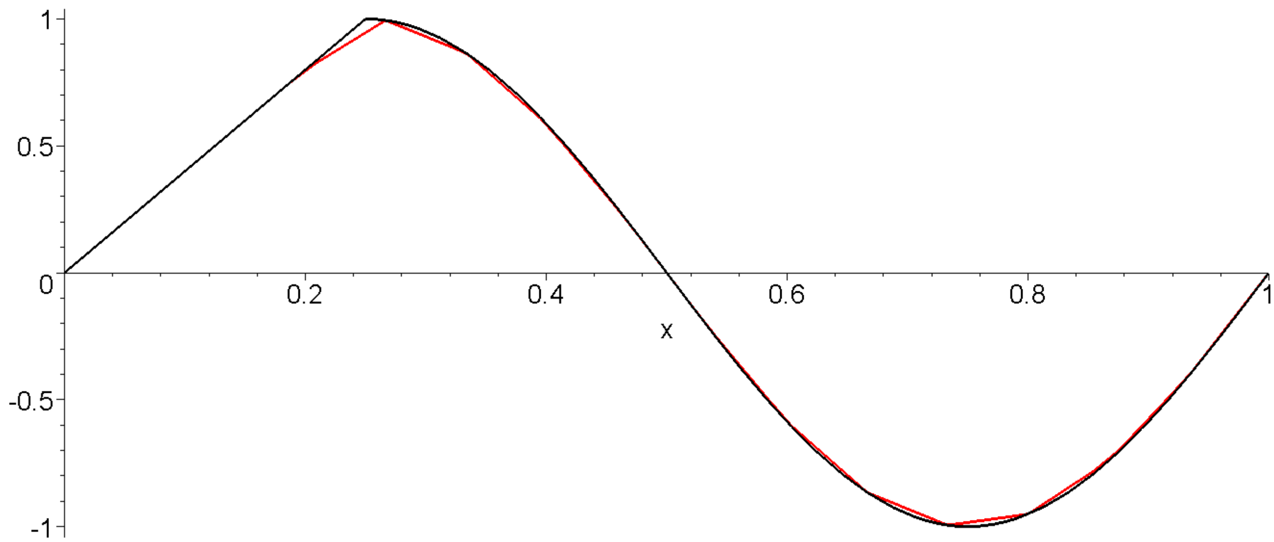
Example 3: Let h(x) = sin(2 π x). We draw the graphs using the approximation for h given above and with n = 20.

$$approx := x \to \sum_{p=1}^{n} \mathrm{h}\left(\frac{p}{n}\right) T_p(x)$$

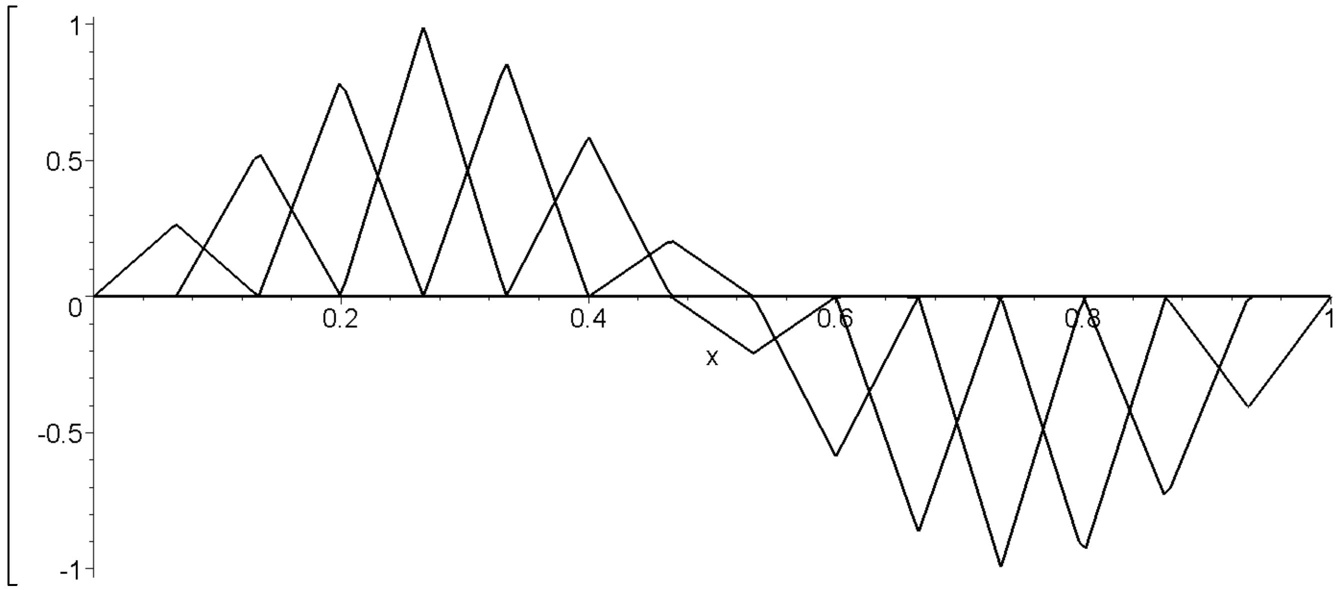


Example 4: In this example, we make a graph of this modified representation for the approximation even when h'(x) is not continuous. In this example, h is continuous, with a break in the derivative at x = 1/4. We use n = 15 so that there is not a *node point* at the break.

$$approx := x \rightarrow \sum_{p=1}^{n} \mathrm{h}\left(\frac{p}{n}\right) T_p(x)$$



In this example, it is interesting to see what are the *supporting functions* for the approximation.



Finally, we present the main theorem of this section.

<u>Theorem.</u> If { E, < , >} is a Hilbert Space of number valued functions on [0,1]. These are equivalent:
(1) There is a reproducing kernel for {E, < , >}.
(2) If x is in [0, 1] and L_x is the function from E to the real number defined by

$$L_x (f) = f(x)$$

for each f in E, then L_x is continuous on E.
(3) Normed convergence in E implies pointwise convergence on [0, 1].

Indication of proof: It is routine to show that (2) and (3) are equivalent. We saw above that (1) implies (3). To see that (2) implies (1), we use the Riesz Representation Theorem as follows. Suppose that x is in [0, 1] and L_x is defined as in (2). To assert that this is a continuous function from E to the numbers implies that there is h in E such that

$$L_x(f) = < f, h >.$$

Clearly, this *Riesz point* changes with x. Thus, we write, for each f in E,

$$f(x) = L_x(f) = < f, h_x > = < f(t), h(t,x) >.$$