Inner Product Spaces

In $\mathbb{R}^n$ we have an inner product $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T \boldsymbol{v} = u_1 v_1 + ... + u_n v_n$. Another notation sometimes used is

$$< u, v > \; = \boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T \boldsymbol{v} = u_1 v_1 + ... + u_n v_n$$

The inner product in $< u, v >$ in $\mathbb{R}^n$ has several essential properties (*see Theorem 1, p. 331*) that we have used repearedly:

a) $< \boldsymbol{u}, \boldsymbol{v} > \; = \; < \boldsymbol{v}, \boldsymbol{u} >$
b) $< \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} > \; = \; < \boldsymbol{u}, \boldsymbol{w} > + < \boldsymbol{v}, \boldsymbol{w} >$
c) $< c\boldsymbol{u}, \boldsymbol{v} > \; = c < \boldsymbol{u}, \boldsymbol{v} >$
d) $< \boldsymbol{u}, \boldsymbol{u} > \; \geq 0$ and $< \boldsymbol{u}, \boldsymbol{u} > = 0$ if and only if $\boldsymbol{u} = \boldsymbol{0}$.

We defined the "length" of a vector $\boldsymbol{u}$ by $\|\boldsymbol{u}\| = \sqrt{< \boldsymbol{u}, \boldsymbol{v} >}$, and the distance between two vectors $\boldsymbol{u}, \boldsymbol{v}$ as $|| \boldsymbol{u} - \boldsymbol{v} ||$, that is, as $\sqrt{< \boldsymbol{u} - \boldsymbol{v}, \boldsymbol{u} - \boldsymbol{v} >}$.

Earlier in the course, we used the "essential properties of vectors in $\mathbb{R}^n$" as the starting point to define more general vector spaces V (*p. 190*). In the same spirit, we now use the properties a)-d) to describe, in any vector space, "how an inner product should behave." For a vector space V with real scalars: any rule $< *, * >$ that creates a scalar from each pair of vectors in V and satisfies a)-d) will be called an inner product in V. (*Properties a)-d) are modified slightly when complex scalars are allowed.*) A vector space V with an inner product defined is called an inner product space. Because such an inner product "acts just like" the inner product from $\mathbb{R}^n$, many of the same theorems remain true. You can look at a little of this material in the textbook: Section 6.7.

Here is a little more detail involving one specific example

Let $C[-\pi, \pi]$ be vector space of all continuous real valued functions defined on the interval $[-\pi, \pi]$: call it C, for short.

For vectors (functions) f, g in C, define an inner product by

$$< f, g > \; = \int_{-\pi}^{\pi} f(x) g(x) \, dx$$

This definition satisfies all the essential properties for an inner product listed above:

a) $< f, g > \; = \; < g, f >$
because $\int_{-\pi}^{\pi} f(x) g(x) dx = \int_{-\pi}^{\pi} g(x) f(x) \, dx$

b) $< f + g, h > \; = \; < f, h > + < g, h >$
because $\int_{-\pi}^{\pi} (f(x) + g(x)) h(x) \, dx = \int_{-\pi}^{\pi} f(x) h(x) \, dx + \int_{-\pi}^{\pi} g(x) h(x) \, dx$

c) $< cf, g > \; = c < f, g >$
d) $< f, f > \; \geq 0$ and $< f, f > \; = 0$ if and only if $f = 0$ (the constant function 0)
You should check c) and d). The last part of d) is the only one that uses that the functions $f \in C$ are continuous.

Continuing to parallel to our definitions in $\mathbb{R}^n$, we define the norm (or "length") of f by

$$||f|| = \sqrt{< f, f >} = \sqrt{\int_{-\pi}^{\pi} f^2(x)\, dx}$$

and the distance between f and g as

$$||\, f - g\,|| = \sqrt{\int_{-\pi}^{\pi} (f(x) - g(x))^2\, dx}$$

and we say that f and g are <u>orthogonal</u> $(f \perp g)$ if $< f, g > = \int_{-\pi}^{\pi} f(x)g(x)\, dx = 0$

For example: on $[-\pi, \pi]$, the functions sin and cos are orthogonal because

$$\int_{-\pi}^{\pi} (\sin x)(\cos x)\, dx = \frac{1}{2}\int_{-\pi}^{\pi} \sin(2x)\, dx = -\frac{1}{2}\cos(2x)|_{-\pi}^{\pi} = 0$$

Many of the techniques we developed using inner products still work. For example:

If $\{f_1, f_2, ..., f_n\}$ is a basis for a subspace W of C, we can convert the basis into an orthogonal basis using the same Gram Schmidt process.

For a subspace W, we can define $W^{\perp} = \{f : < f, g > = 0$ for all g in $W\}$

If W is a subspace of C with an <u>orthogonal</u> basis* $\{g_1, ..., g_n\}$ and $f \in C$

then we can uniquely write $f = \widehat{f} + g$ where $\widehat{f} \in W$ and $g \in W^{\perp}$.

$\widehat{f}$ is called the projection of f on W and $\widehat{f}$ is given by the formula

$$\widehat{f} = \frac{<f, g_1>}{<g_1, g_1>} g_1 + ... + \frac{<f, g_n>}{<g_n, g_n>} g_n$$

Then $\widehat{f}$ is the function in W <u>closest</u> to f, that is

$$\|\, f - \widehat{f}\,\| < \|\, f - g\| \text{ for all } g \text{ in } W \text{ different from } \widehat{f}$$

<u>Note</u>: Unlike $\mathbb{R}^n$, C is <u>not</u> finite dimensional. But for the results just listed, that doesn't matter. What <u>does</u> matter is that the subspace W <u>is</u> finite dimensional.

<u>A calculation in C</u> : find the polynomial of degree ≤ 5 in C that is closest to the function sin.

We use the subspace $W = \text{Span}\{1, x, x^2, x^3, x^4, x^5\}$. The polynomial we want is $q = \widehat{\sin} = \text{proj}_W \sin$. This polynomial q is the closest in W to sin $-$ closest in the sense of the distance we defined:

$$||\, q - \sin\,|| = \left(\int_{-\pi}^{\pi} |q(x) - \sin x|^2\, dx\right)^{1/2} \text{ is as small as possible.}$$

<u>In that sense</u>, q is the “best available approximation in W for sin.”

It's easy to compute $\text{Proj}_U \sin$ if we choose an <u>orthonormal</u> basis for W. So we convert the standard basis

$\{v_1, v_2, ..., v_6\} = \{1, x, x^2, ..., x^5\}$ into an orthonormal basis

$\{e_1, e_2, e_3, ..., e_6\}$

We use the Gram Schmidt process (and normalize, to get a vector of length 1 at each step).

> *Note: the integrations below were done using Matlab. Although <u>every</u> integration needed to find the e_i's is very easy, the constants that arise are messy and pile up fast; they can easily lead to errors when the computation is done by hand. Try to compute at least e_1, e_2, e_3 for yourself (with or without computer assistance) to be sure you understand what's going on. The steps are the same as for the usual Gram Schmidt process in $\mathbb{R}^n$.*

We start with the first basis vector $v_1 = 1$. But v_1 is <u>not</u> a unit vector in C because $||v_1||^2 = ||1||^2 = \int_{-\pi}^{\pi} 1 \cdot 1\, dx = 2\pi$. So we normalize and use

$$e_1 = \frac{v_1}{||\, v_1\, ||} = \frac{1}{||1||} = \frac{1}{\sqrt{2\pi}}$$

For $j = 2, ..., 6$ in turn we use the Gram Schmidt formula, normalizing at each step to get a unit vector:

$$e_j = \frac{v_j - <v_j,e_1>e_1 - <v_j,e_2>e_2 - ... - <v_j,e_{j-1}>e_{j-1}}{||\, v_j - <v_j,e_1>e_1 - <v_j,e_2>e_2 - ... - <v_j,e_{j-1}>e_{j-1}\, ||}$$

So $$e_2 = \frac{v_2 - <v_2,e_1>e_1}{||v_2 - <v_2,e_1>e_1||} = \frac{x - \left(\int_{-\pi}^{\pi} x \cdot \frac{1}{\sqrt{2\pi}}\, dx\right) \cdot \frac{1}{\sqrt{2\pi}}}{||x - \left(\int_{-\pi}^{\pi} x \cdot \frac{1}{\sqrt{2\pi}}\, dx\right) \cdot \frac{1}{\sqrt{2\pi}}||}$$

$$= \frac{x}{||x||} \qquad (\text{since } \int_{-\pi}^{\pi} x\, dx = 0)$$

$$= \frac{x}{\left(\int_{-\pi}^{\pi} x \cdot x\, dx\right)^{1/2}} = \frac{x}{\frac{1}{3}\,\pi(6\pi)^{1/2}} = \frac{\sqrt{6}\, x}{2\pi\sqrt{\pi}}$$

Then e_3

$$= \frac{v_3 - <v_3,e_1>e_1 - <v_3,e_2>e_2}{||\, v_3 - <v_3,e_1>e_1 - <v_3,e_2>e_2 \,||} = \frac{x^2 - (\int_{-\pi}^{\pi} x^2 \cdot e_1\, dx)e_1 - (\int_{-\pi}^{\pi} x^2 \cdot e_2\, dx)e_2}{||\, x^2 - (\int_{-\pi}^{\pi} x^2 \cdot e_1\, dx)e_1 - (\int_{-\pi}^{\pi} x^2 \cdot e_2\, dx)e_2 \,||}$$

$$= \frac{x^2 - (\int_{-\pi}^{\pi} x^2 \cdot \frac{1}{\sqrt{2\pi}}\, dx) \cdot \frac{1}{\sqrt{2\pi}} - (\int_{-\pi}^{\pi} x^2 \cdot \frac{\sqrt{6}\,x}{2\pi\sqrt{\pi}}\, dx) \cdot \frac{\sqrt{6}\,x}{2\pi\sqrt{\pi}}}{||\, x^2 - (\int_{-\pi}^{\pi} x^2 \cdot \frac{1}{\sqrt{2\pi}}\, dx) \cdot \frac{1}{\sqrt{2\pi}} - (\int_{-\pi}^{\pi} x^2 \cdot \frac{\sqrt{6}\,x}{2\pi\sqrt{\pi}}\, dx) \cdot \frac{\sqrt{6}\,x}{2\pi\sqrt{\pi}} \,||}$$

Notice that each integration requires nothing harder than $\int x^n\, dx$.
But already the constants are becoming a headache.

$$= \quad \ldots \quad = \frac{1}{8}\,\frac{\sqrt{8}\,\sqrt{45}\,(x^2 - \frac{1}{3}\pi^2)}{\sqrt{\pi^5}}$$

Continuing in this way gives:

$$e_4 = \quad \ldots \quad = \frac{1}{8}\,\frac{\sqrt{175}\,\sqrt{8}\,(x^3 - \frac{3}{5}\pi^2\, x)}{\sqrt{\pi^7}}$$

$$e_5 = \quad \ldots \quad = \frac{1}{128}\,\frac{\sqrt{11025}\,\sqrt{128}\left(x^4 - \frac{1}{5}\pi^4 - \frac{6}{7}\pi^2(x^2 - \frac{1}{3}\pi^2)\right)}{\sqrt{\pi^9}}, \text{ and finally}$$

$$e_6 = \quad \ldots \quad = \frac{1}{128}\,\frac{\sqrt{128}\,\sqrt{43659}\left(x^5 - \frac{3}{7}\pi^4 x - \frac{10}{9}\pi^2\,(x^3 - \frac{3}{5}\pi^2\, x)\right)}{\sqrt{\pi^{11}}}$$

Then $e_1, \ldots, e_6$ are orthogonal polynomials of degree ≤ 5; we use them as our orthogonal basis for W. With them, we can compute

$$q(x) = \text{Proj}_W \sin x = \; < \sin x, e_1 > e_1 + \; < \sin x, e_2 > e_2 + \ldots + \; < \sin x, e_6 > e_6$$

$$= (\int_{-\pi}^{\pi} (\sin x) e_1\, dx) e_1 + (\int_{-\pi}^{\pi} (\sin x) e_2\, dx) e_2 + \ldots + (\int_{-\pi}^{\pi} (\sin x) e_6\, dx) e_6$$

$$= (\int_{-\pi}^{\pi} (\sin x) \frac{1}{\sqrt{2\pi}}\, dx) \frac{1}{\sqrt{2\pi}} + (\int_{-\pi}^{\pi} (\sin x) \frac{\sqrt{6}\,x}{2\pi\sqrt{\pi}}\, dx) \cdot \frac{\sqrt{6}\,x}{2\pi\sqrt{\pi}}$$

$$+ (\int_{-\pi}^{\pi} (\sin x) \frac{1}{8}\,\frac{\sqrt{8}\,\sqrt{45}\,(x^2 - \frac{1}{3}\pi^2)}{\sqrt{\pi^5}}\, dx) \cdot \frac{1}{8}\,\frac{\sqrt{8}\,\sqrt{45}\,(x^2 - \frac{1}{3}\pi^2)}{\sqrt{\pi^5}}$$

$+ \ldots +$ <u>three more terms</u> corresponding to $e_4, e_5,$ and e_6

(*Notice that the integrations involved are now more challenging because they involve terms like $\int x^n \sin x\; dx$. But they are manageable, with enough patience, using integration by parts.*)

When all the smoke clears, Matlab gives

$$q(x) = \frac{21}{8\pi^{10}} \Big((33\pi^4 - 3465\pi^2 + 31185)x^5 + (3750\pi^4 - 30\pi^6 - 34650\pi^2)x^3 + (5\pi^8 - 765\pi^6 + 7425\pi^4)x) \Big) \quad (*)$$

If we convert the <u>exact</u> coefficients in (*) to <u>approximate</u> decimal coefficients we have

$$q(x) \approx 0.98786213557467x - 0.15527141063343\, x^3 + 0.00564311797635\, x^5$$

In the sense of the distance $|| \ ||$ that we defined in C, q is the closest polynomial in W to sin.

<u>For comparison</u>: there is better known 5th degree polynomial approximation for $\sin x$: use the 5th degree "Taylor polynomial:"

$$T_5(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + ... + \frac{f^{(v)}(x)}{5!}x^5.$$

For $f(x) = \sin x$, $\sin x \approx T_5(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}$

With approximate decimal coefficients,

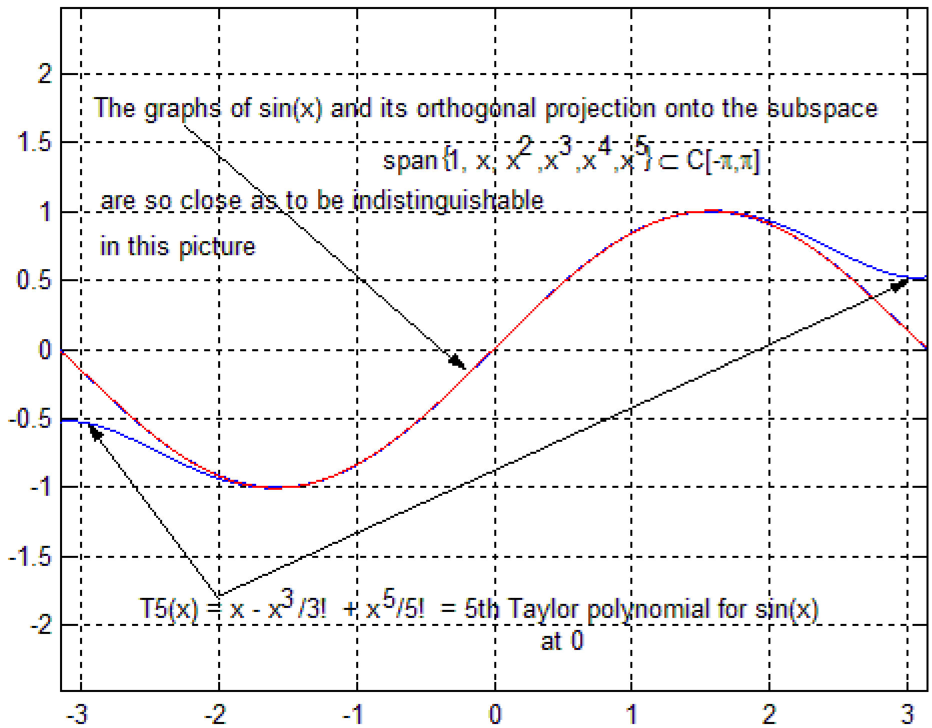
$$T_5(x) = x - 0.16666666666667x^3 + 0.00833333333333x^5$$

Because $T_5(x)$ is constructed using derivatives of f <u>at 0</u>, it turns to be the better approximation for $\sin x$ <u>near 0</u>, but it produces a larger and larger error as x gets further and further from 0.

We can see this in the figure on the next page.

The figure below shows the graphs of $\sin x$, $T_5(x)$ and $q(x)$ on the interval $[-\pi, \pi]$.

<u>Across the whole interval</u> $[-\pi, \pi]$: the graphs of $\sin x$ and $q(x)$ are so close that you can't see any difference between them at this scale. You can see that $T_5(x)$ and $\sin x$ also are indistinguishable near 0 but get fairly far apart as x gets nearer to the endpoints $\pm\pi$.



The table on the following page illustrates three things (the third is not clear from the graphs above):

i) As we move away from 0 toward $\pm\pi$, the approximation from $T_5(x)$ is not as good as the $q(x)$ approximation

ii) Linear algebra gives us the $q(x)$ approximation. It has the advantage that it gives us a good approximation for $\sin x$ <u>over the whole interval</u> $[-\pi, \pi]$.

iii) The Taylor polynomial $T_5(x)$ is actually a <u>better</u> approximation than $q(x)$ for $\sin x$ when x is near 0 (*in some sense,* $T_5(x)$ *is acutally the best of all 5*th *degree polynomials to approximate sin* x *<u>near</u>* 0).

All table values are rounded to 4 significant digits
so, for example, 0.0000 is not exactly 0

x	$\sin x$	$T_5(x)$	$q(x)$	$\lvert\text{Error}\rvert = \lvert \sin x - T_5(x) \rvert$		$\lvert\text{Error}\rvert = \lvert \sin x - q(x) \rvert$	
-3.1416	-0.0000	-0.5240	-0.0160	0.5240	← **large** T_5	0.0160	← **smallish** q
-2.9416	-0.1987	-0.5347	-0.1966	0.3361	← **|error|**	0.0021	← **|error|**
-2.7416	-0.3894	-0.5979	-0.3827	0.2084	← **near** $-\pi$	0.0067	← **over**
-2.5416	-0.5646	-0.6891	-0.5600	0.1244	⋮	0.0047	← **the whole**
-2.3416	-0.7174	-0.7884	-0.7169	0.0710		0.0005	← **interval**
-2.1416	-0.8415	-0.8800	-0.8447	0.0385		0.0032	⋮
-1.9416	-0.9320	-0.9516	-0.9372	0.0196		0.0052	
-1.7416	-0.9854	-0.9947	-0.9906	0.0092		0.0052	↕
-1.5416	-0.9996	-1.0035	-1.0032	0.0040		0.0036	
-1.3416	-0.9738	-0.9754	-0.9749	0.0015		0.0011	
-1.1416	-0.9093	-0.9098	-0.9077	0.0005		0.0016	
-0.9416	-0.8085	-0.8086	-0.8047	0.0001		0.0038	
-0.5416	-0.5155	-0.5155	-0.5106	0.0000	← **very**	0.0049	← **but |error|**
-0.3416	-0.3350	-0.3350	-0.3313	0.0000	← **small**	0.0037	← **for** q
-0.1416	-0.1411	-0.1411	-0.1394	0.0000	← **|error|**	0.0017	← **near 0**
0.0584	0.0584	0.0584	0.0577	0.0000	← **near 0**	0.0007	← **is larger**
0.2584	0.2555	0.2555	0.2526	0.0000	← **using**	0.0029	← **than for**
0.4584	0.4425	0.4425	0.4380	0.0000	← T_5	0.0045	← **T_5 approx**
0.6584	0.6119	0.6119	0.6068	0.0000	← **approx**	0.0051	⋮
0.8584	0.7568	0.7569	0.7524	0.0001	⋮	0.0044	
1.0584	0.8716	0.8719	0.8690	0.0003		0.0026	
1.2584	0.9516	0.9526	0.9515	0.0010		0.0001	
1.4584	0.9937	0.9964	0.9963	0.0027		0.0026	
1.6584	0.9962	1.0028	1.0009	0.0066		0.0047	
1.8584	0.9589	0.9734	0.9644	0.0145		0.0054	↕
2.0584	0.8835	0.9128	0.8877	0.0293		0.0043	
2.2584	0.7728	0.8282	0.7740	0.0554		0.0012	
2.4584	0.6313	0.7304	0.6283	0.0991		0.0030	← **smallish** q
2.6584	0.4646	0.6336	0.4583	0.1690	⋮	0.0063	← **|error|**
2.8584	0.2794	0.5561	0.2742	0.2767	← **large** T_5	0.0052	← **over**
3.0584	0.0831	0.5204	0.0894	0.4373	← **|error|**	0.0063	← **the whole**
3.1416	9.0000	0.5240	0.0160	0.5240	← **near** π	0.0160	← **interval**

One more example, without many details

We are still working in the vector space C of continuous functions on the interval $[-\pi, \pi]$.

Conside the subspace $W = \text{Span}\{1, \sin x, \sin 2x, ..., \sin nx, \cos x, ..., \cos nx\}$ for some n.

You can check that these functions form an orthogonal basis for W : .

For example, to show $\sin 2x \perp \cos 3x$:

$$< \sin 2x, \cos 3x > = \int_{-\pi}^{\pi} (\sin 2x)(\cos 3x)\, dx$$

Since $\sin(\alpha + \beta) = \sin\alpha \cos\beta + \cos\alpha \sin\beta$
and $\sin(\alpha - \beta) = \sin\alpha \cos\beta - \cos\alpha \sin\beta$
then $\sin(\alpha + \beta) + \sin(\alpha - \beta) = 2\sin\alpha \cos\beta$

If we let $\alpha = 2x$ and $\beta = 3x$, we

$$\sin(5x) + \sin(-x) = 2\sin(2x)\cos(3x)$$

so $\frac{1}{2}(\sin(5x) - \sin(x)) = \sin(2x)\cos(3x)$

Therefore $\int_{-\pi}^{\pi} (\sin 2x)(\cos 3x)\, dx = \frac{1}{2}\int_{-\pi}^{\pi} \sin(5x) - \sin(x)\, dx$

$$= \frac{1}{2}\left(-\frac{1}{5}\cos 5x + \cos x\, dx\right)\Big|_{-\pi}^{\pi} = 0$$

A function in W is a linear combination of the basis elements.

If f is in C, we can compute $\text{proj}_W f = \widehat{f} =$ the function in W closest to f

$$= \frac{<f(x),1>}{<1,1>}1 + \frac{<f(x), \cos x>}{<\cos x, \cos x>}\cos x + ... + \frac{<f(x), \cos nx>}{<\cos nx, \cos nx>}\cos nx$$

$$+ \frac{<f(x), \sin x>}{<\sin x, \sin x>}\sin x + ... + \frac{<f(x), \sin nx>}{<\sin nx, \sin nx>}\sin nx$$

For the denominators:

$$< 1, 1 > = \int_{-\pi}^{\pi} 1 \cdot 1\, dx = 2\pi$$

$< \cos kx, \cos kx> = \int_{-\pi}^{\pi} \cos^2 kx\, dx = \pi$ (*why?*)

$< \sin kx, \sin kx> = \int_{-\pi}^{\pi} \sin^2 kx\, dx = \pi$ (*why?*)

So $\text{proj}_W\, f = \widehat{f}$

$$= \tfrac{1}{2}\left(\tfrac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cdot 1\,dx\right)1$$

$$+ \tfrac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot \cos x\ dx\right)\cos x + \ldots + \tfrac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot \cos nx\, dx\right)\cos nx$$

$$+ \tfrac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot \sin x\ dx\right)\sin x + \ldots + \tfrac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot \sin nx\, dx\right)\sin nx$$

$$= \frac{a_0}{2} + a_1\cos x + \ldots + a_n\cos nx + b_1\sin x + \ldots + b_n\sin nx$$

So $\text{proj}_W\, f = \widehat{f} = \frac{a_0}{2} + \sum_{k=1}^{n} a_k\cos kx + \sum_{k=1}^{n}\sin kx$ (*)

where $a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cdot 1\,dx$

and $\begin{cases} a_k = \frac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot \cos kx\,dx\right) & \text{for } 1 < \ \leq k \leq n \\ b_k = \frac{1}{\pi}\left(\int_{-\pi}^{\pi} f(x)\cdot \sin nx\,dx\right) & \text{for } 1 < \ \leq k \leq n \end{cases}$

The coefficients used in writing $\text{proj}_W\, f$ are called the Fourier coefficients of f

and the "trigonmetric series" (*) is the n^{th} Fourier approximation for f.

If $n \to \infty$, then $||\, f - \text{proj}_W f\, || \to 0$, that is, the approximation error $\to 0$ (as measured by our distance function $||\, f - \text{proj}_W f\, ||$. (This is called "mean square convergence.")

*"Mean square convergence" is **not** the same as saying:*

for each $x \in [-\pi, \pi]$, $\widehat{f}(x) \to f(x)$ (this is called "pointwise convergence") and it's a much harder problem to determine for which x's this is true. However, pointwise converge is true, for example, at any point x where f is differentiable.