12.8 Orthogonality

The notion of orthogonality originates in $\mathcal{R}^3$, where nonzero vectors $\mathbf{v}_1$, $\mathbf{v}_2$ are said to be **orthogonal**, written $\mathbf{v}_1 \perp \mathbf{v}_2$, provided $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. The dot product in $\mathcal{R}^3$ is defined by

$$\mathbf{x} \cdot \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = x_1y_1 + x_2y_2 + x_3y_3.$$

Similarly, $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$ defines the dot product in $\mathcal{R}^n$. Literature uses the notation $(\mathbf{x}, \mathbf{y})$ as well as $\mathbf{x} \cdot \mathbf{y}$. Modern terminology uses **inner product** instead of **dot product**, to emphasize the use of functions and abstract properties. The inner product satisfies the following properties.

$(\mathbf{x}, \mathbf{x}) \geq 0$	Non-negativity.
$(\mathbf{x}, \mathbf{x}) = 0$ implies $\mathbf{x} = \mathbf{0}$	Uniqueness.
$(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$	Symmetry.
$k(\mathbf{x}, \mathbf{y}) = (k\mathbf{x}, \mathbf{y})$	Homogeneity.
$(\mathbf{x} + \mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z})$	Additivity.

The storage system of choice for answers to differential equations is a real vector space V of functions f. A **real inner product space** is a vector space V with real-valued inner product function $(\mathbf{x}, \mathbf{y})$ defined for each $\mathbf{x}$, $\mathbf{y}$ in V, satisfying the preceding rules.

Dot Product for Functions. The extension of the notion of dot product to functions replaces $\mathbf{x} \cdot \mathbf{y}$ by average value. Insight can be gained from the approximation

$$\frac{1}{b-a} \int_a^b F(x)dx \approx \frac{F(x_1) + F(x_2) + \cdots + F(x_n)}{n}$$

where $b - a = nh$ and $x_k = a + kh$. The left side of this approximation is called the **average value of** F **on** $[a, b]$. The right side is the classical average of F at n equally spaced values in $[a, b]$. If we replace F by a product fg, then the average value formula reveals that $\int_a^b fgdx$ acts like a dot product:

$$\frac{1}{b-a} \int_a^b f(x)g(x)dx \approx \frac{\mathbf{x} \cdot \mathbf{y}}{n}, \quad \mathbf{x} = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} g(x_1) \\ g(x_2) \\ \vdots \\ g(x_n) \end{pmatrix}.$$

The formula says that $\int_a^b f(x)g(x)dx$ is approximately a constant multiple of the dot product of samples of f, g at n points of $[a,b]$.

Given functions f and g integrable on $[a,b]$, the formula

$$(f,g) = \int_a^b f(x)g(x)dx$$

defines a dot product satisfying the abstract properties cited above. When dealing with solutions to differential equations, this dot product, along with the abstract properties of a dot product, provide the notions of *distance* and *orthogonality* analogous to those in $\mathcal{R}^3$.

Orthogonality, Norm and Distance. Define nonzero functions f and g to be **orthogonal** on $[a,b]$ provided $(f,g)=0$. Define the **norm** or the **distance** from f to 0 to be the number $\|f\| = \sqrt{(f,f)}$ and the distance from f to g to be $\|f-g\|$. The basic properties of the norm $\|\cdot\|$ are as follows.

$\|f\| \ge 0$	Non-negativity.
$\|f\| = 0$ implies $f = 0$	Uniqueness.
$\|cf\| = \|c\|\|f\|$	Homogeneity.
$\|f\| = \sqrt{(f,f)}$	Norm and the inner product.
$\|f+g\| \le \|f\| + \|g\|$	The triangle inequality.
$\|(f,g)\| \le \|f\|\,\|g\|$	Cauchy-Schwartz inequality.

Series of Orthogonal Functions. Let (f,g) denote a dot product defined for functions f, g. Especially, we include $(f,g) = \int_a^b fg dx$ and a weighted dot product $(f,g) = \int_a^b fg\rho dx$. Let $\{f_n\}$ be a sequence of nonzero functions orthogonal with respect to the dot product (f,g), that is, a system $\{f_n\}_{n=1}^{\infty}$ satisfying the **orthogonality relations**

$$(f_i, f_j) = 0, \quad i \ne j, \quad (f_i, f_i) > 0, \quad i = 1, 2, \ldots.$$

A **generalized Fourier series** is a convergent series of such orthogonal functions

$$F(x) = \sum_{n=1}^{\infty} c_n f_n(x).$$

The coefficients $\{c_n\}$ are called the **Generalized Fourier Coefficients** of F. Convergence is taken in the sense of the norm $\|g\| = \sqrt{(g,g)}$, defined as follows:

$$F = \sum_{n=1}^{\infty} c_n f_n \quad \text{means} \quad \lim_{N\to\infty} \left\| \sum_{n=1}^{N} c_n f_n - F \right\| = 0.$$

For example, when $\|g\| = \sqrt{(g,g)}$ and $(f,g) = \int_a^b fg dx$, then series convergence is called **mean-square convergence**, defined by

$$\lim_{N\to\infty} \sqrt{\int_a^b \left| \sum_{n=1}^{N} c_n f_n(x) - F(x) \right|^2 dx} = 0.$$

Orthogonal Series Method. The coefficients $\{c_n\}$ in an orthogonal series are determined by a technique called the **orthogonal series method**, described in words as follows.

> The coefficient c_n in an orthogonal series is found by taking the dot product of the equation with the orthogonal function that multiplies c_n.

The details of the method:

$(F, f_n) = \left(\sum_{k=1}^{\infty} c_k f_k, f_n \right)$	Dot product the equation with f_n.
$(F, f_n) = \sum_{k=1}^{\infty} c_k (f_k, f_n)$	Apply dot product properties.
$(F, f_n) = c_n (f_n, f_n)$	By orthogonality, just one term remains from the series on the right.

Division after the last step leads to the **Fourier Coefficient Formula**

$$c_n = \frac{(F, f_n)}{(f_n, f_n)}.$$

Orthogonal Projection. The **shadow projection** of vector $\vec{X}$ onto the direction of vector $\vec{Y}$ is the number d defined by

$$d = \frac{\vec{X} \cdot \vec{Y}}{|\vec{Y}|}.$$

The triangle determined by $\vec{X}$ and $d\dfrac{\vec{Y}}{|\vec{Y}|}$ is a right triangle.

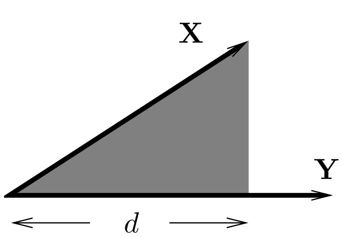

Figure 1. Shadow projection d of vector X onto the direction of vector Y.

The **vector shadow projection** of $\vec{X}$ onto the line L through the origin in the direction of $\vec{Y}$ is defined by

$$\mathbf{proj}_{\vec{Y}}(\vec{X}) = d\frac{\vec{Y}}{|\vec{Y}|} = \frac{\vec{X} \cdot \vec{Y}}{\vec{Y} \cdot \vec{Y}}\vec{Y}.$$

Shadow Projection and Fourier Coefficients. The term $c_n f_n$ in a generalized Fourier series can be expanded as

$$c_n f_n = \frac{(F, f_n)}{(f_n, f_n)} f_n = \text{ Shadow projection of } F \text{ onto } f_n.$$

This formula appears in the Gram-Schmidt formulas and the Least Squares formulas, because those formulas also involve orthogonal projections. The complexity of such formulas is removed by thinking of the results as sums of shadow projections or as subractions of shadow projections.