1

Inner Product Spaces (Pre-Hilbert)

This chapter will focus on inner product spaces, that is, vector spaces with a scalar product, specifically those of *finite dimension*.

1.1. Real and complex inner products

In real Euclidean spaces $\mathbb{R}^2$ and $\mathbb{R}^3$, the inner product of two vectors v, w is defined as the real number:

$$v \bullet w = \langle v, w \rangle = \|v\| \|w\| \cos(\vartheta)$$

where ϑ is the smallest angle between v and w and $\| \ \|$ represents the norm (or the magnitude) of the vectors.

Using the inner product, it is possible to define the orthogonal projection of vector v in the direction defined by vector w. A distinction must be made between:

– the *scalar projection* of v in the direction of w: $\|v\| \cos(\theta) = \frac{\langle v,w \rangle}{\|w\|}$; and

– the *vector projection* of v in the direction of w: $\|v\| \cos(\theta) \frac{w}{\|w\|} = \frac{\langle v,w \rangle}{\|w\|^2} w$;

where $\frac{w}{\|w\|}$ is the unit vector in the direction of w. Evidently, the roles of v and w can be reversed.

The absolute value of the scalar projection measures the "similarity" of the directions of two vectors. To understand this concept, consider two remarkable relative positions between v and w:

– if v and w possess the same direction, then the angle between them ϑ is either null or π, hence $\cos(\vartheta) = \pm 1$, that is, the absolute value of the scalar projection of v in direction w is $\|v\|$;

– however, if v and w are perpendicular, then $\vartheta = \frac{\pi}{2}$ and hence $\cos(\vartheta) = 0$, showing that the scalar projection of v in direction w is null.

When the position of v relative to w falls somewhere in the interval between the two vectors described above, the absolute value of the scalar projection of v in the direction of w falls between 0 and $\|v\|$; this explains its use to measure the similarity of the direction of vectors.

In this book, we shall consider vector spaces which are far more complex than $\mathbb{R}^2$ and $\mathbb{R}^3$, and the measure of vector similarity obtained through projection supplies crucial information concerning the coherence of directions.

Before we can obtain this information, we must begin by moving from Euclidean spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ to abstract vector spaces. The general definition of an inner product and an orthogonal projection in these spaces may be seen as an extension of the previous definitions, permitting their application to spaces in which our representation of vectors is no longer applicable.

Geometric properties, which can only be apprehended and, notably, visualized in two or three dimensions, must be replaced by a set of algebraic properties which can be used in any dimension.

Evidently, these algebraic properties must be necessary and sufficient to characterize the inner product of vectors in a plane or in real space. This approach, in which we generalize concepts which are "intuitive" in two or three dimensions, is a classic approach in mathematics.

In this chapter, the symbol V will be used to describe a vector space defined over the field $\mathbb{K}$, where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$ and is of finite dimension $n < +\infty$. Field $\mathbb{K}$ contains the *scalars* used to construct linear combinations between vectors in V. Note that two finite dimensional vector spaces are isomorphic if and only if they are of the same dimension. Furthermore, if we establish a basis $B = (b_1, \ldots, b_n)$ for V, an *isomorphism* between V and $\mathbb{K}^n$ can be constructed as follows:

$$\begin{array}{llll} I: & V & \longrightarrow & \mathbb{K}^n \\ & v = [v]_B = \sum_{i=1}^{n} v_i b_i & \longmapsto & \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \end{array}$$

that is, I associates each $v \in V$ with the vector of $\mathbb{K}^n$ given by the scalar components of v in relation to the established basis B. Since I is an isomorphism, it follows that $\mathbb{K}^n$ *is the prototype of all vector spaces of dimension* n *over a field* $\mathbb{K}$.

DEFINITION 1.1.– *Let* V *be a vector space defined over a field* $\mathbb{K}$.

A $\mathbb{K}$-form over V is an application defined over $V \times V$ with values in $\mathbb{K}$, that is:

$$\begin{aligned}\phi : V \times V &\longrightarrow \mathbb{K}\\ (v, w) &\longmapsto \phi(v, w).\end{aligned}$$

DEFINITION 1.2.– *Let V be a real vector space. A couple $(V, \langle , \rangle)$ is said to be a real inner product space (or a real pre-Hilbert space) if the form $\langle , \rangle$ is:*

1) bilinear, i.e.[1] linear in relation to each argument (the other being fixed):

$$\langle v_1 + v_2, w_1 + w_2\rangle = \langle v_1, w_1\rangle + \langle v_1, w_2\rangle + \langle v_2, w_1\rangle + \langle v_2, w_2\rangle,\\ \forall\, v_1, v_2, w_1, w_2 \in V$$

and:

$$\langle \alpha v, \beta w\rangle = \alpha\langle v, \beta w\rangle = \beta\langle \alpha v, w\rangle = \alpha\beta\langle v, w\rangle, \quad \forall \alpha, \beta \in \mathbb{R},\ v, w \in V$$

2) symmetrical: $\langle v, w\rangle = \langle w, v\rangle$, $\forall v, w \in V$;

3) defined: $\langle v, v\rangle = 0 \iff v = 0_V$, the null vector of the vector space V;

4) positive: $\langle v, v\rangle > 0\ \forall v \in V,\ v \neq 0_V$.

Upon reflection, we see that, for a real form over V, *the symmetry and bilinearity requirements are equivalent to requiring symmetry and linearity on the left-hand side*, that is:

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle, \quad \langle \alpha v, w\rangle = \alpha\langle v, w\rangle, \quad \forall v, w \in V,\ \forall \alpha \in \mathbb{R}$$

The simplest and most important example of a real inner product is the *canonical inner product*, defined as follows: let $v = (v_1, v_2, \ldots, v_n)$, $w = (w_1, w_2, \ldots, w_n)$ be two vectors in $\mathbb{R}^n$ written with their components in relation to any given, but fixed, basis $(b_i)_{i=1}^n$ in $\mathbb{R}^n$. The canonical inner product of v and w is:

$$\langle v, w\rangle \equiv \sum_{i=1}^{n} v_i w_i = v^t \cdot w = v \cdot w^t,$$

where v^t and w^t in the final equations are the transposed vectors of v and w, giving us the matrix product of a line vector (treated as a $1 \times n$ matrix) and a column vector (treated as an $n \times 1$ matrix).

The extension of these definitions to complex vector spaces is not particularly straightforward. First, note that if V is a *complex* vector space, then *there is no bilinear and definite-positive transformation* over $V \times V$. In this case, any vector $v \in V$ would give the following:

$$\langle iv, iv\rangle = i^2\langle v, v\rangle = -\langle v, v\rangle \leqslant 0 \ \text{ since } \langle v, v\rangle \geqslant 0 \ \text{ by positivity.}$$

1 i.e. is the abbreviation of the Latin expression “id est”, meaning “that is”. This term is often used in mathematical literature.

As we shall see, the property of positivity is essential in order to define a norm (and thus a distance, and by extension, a topology) from a complex inner product. To obtain an algebraic structure for complex scalar products which remains compatible with a topological structure, we are therefore forced to abandon the notion of bilinearity, and to search for an alternative.

We could consider antilinearity[2], i.e.

$\langle \alpha v, \beta w\rangle = \bar{\alpha}\bar{\beta}\langle v, w\rangle$

But it has the same problem as bilinearity, $\langle iv, iv\rangle = (-i)(-i)\langle v, v\rangle = i^2\langle v, v\rangle = -\langle v, v\rangle^2 \leqslant 0$.

A simple analysis shows that, in order to avoid losing the positivity, it is sufficient to request the linearity with respect to one variable and the antilinearity with respect to the other. This property is called *sesquilinearity*[3].

The choice of the linear and antilinear variable is entirely arbitrary.

By convention, the antilinear component is placed on the right-hand side in mathematics, but on the left-hand side in physics.

We have chosen to adopt the mathematical convention here, i.e. $\langle \alpha v, \beta w\rangle = \alpha\bar{\beta}\langle v, w\rangle$.

Next, it is important to note that *sesquilinearity and symmetry are incompatible*: if both properties were verified, then $\langle v, \alpha w\rangle = \bar{\alpha}\langle v, w\rangle$, and also $\langle v, \alpha w\rangle = \langle \alpha w, v\rangle = \alpha\langle w, v\rangle = \alpha\langle v, w\rangle$. Thus, $\langle v, \alpha w\rangle = \bar{\alpha}\langle v, w\rangle = \alpha\langle v, w\rangle$ which can only be verified if $\alpha \in \mathbb{R}$.

Thus $\langle , \rangle$ cannot be both sesquilinear and symmetrical when working with vectors belonging to a complex vector space.

The example shown above demonstrates that, instead of symmetry, the property which must be verified for every vector pair v, w is $\langle v, w\rangle = \overline{\langle w, v\rangle}$, that is, changing the order of the vectors in $\langle , \rangle$ must be equivalent to complex conjugation.

A transform which verifies this property is said to be *Hermitian*[4].

---

2 The symbols z^* and $\bar{z}$ represent the *complex conjugation*, i.e. if $z \in \mathbb{C}$, $z = a + ib$, $a, b \in \mathbb{R}$, then $z^* = \bar{z} = a - ib$. We recall that $\sum_{k=1}^{n} \overline{z_k} = \overline{\sum_{k=1}^{n} z_k}$, $\prod_{k=1}^{n} \overline{z_k} = \overline{\prod_{k=1}^{n} z_k}$, $|z|^2 = z\bar{z}$ and $z = \bar{z}$ if and only if $z \in \mathbb{R}$.

3 Sesqui comes from the Latin *semisque*, meaning *one and a half times*. This term is used to highlight the fact that there are not two instances of linearity, but one "and a half", due to the presence of the complex conjugation.

4 For the French mathematician Charles Hermite (1822, Dieuze-1901, Paris).

These observations provide full justification for Definition 1.3.

DEFINITION 1.3.– *Let V be a complex vector space. The pair $(V, \langle , \rangle)$ is said to be a complex inner product space (or a complex pre-Hilbert space) if $\langle , \rangle$ is a complex form which is:*

1) sesquilinear:

$$\langle v_1 + v_2, w_1 + w_2 \rangle = \langle v_1, w_1 \rangle + \langle v_1, w_2 \rangle + \langle v_2, w_1 \rangle + \langle v_2, w_2 \rangle$$

$\forall\, v_1, v_2, w_1, w_2 \in V$*, and:*

$$\xrightarrow[\text{Linearity on the left}]{\text{Antilinearity on the right}} \langle \alpha v, \beta w \rangle = \alpha \langle v, \beta w \rangle = \bar{\beta} \langle \alpha v, w \rangle = \alpha \bar{\beta} \langle v, w \rangle$$

$\forall\, \alpha, \beta \in \mathbb{C}, \forall\, v, w \in V$ *;*

2) Hermitian: $\langle v, w \rangle = \overline{\langle w, v \rangle}$, $\forall v, w \in V$*;*

3) definite: $\langle v, v \rangle = 0 \iff v = 0_V$*, the null vector of the vector space* V*;*

4) positive: $\langle v, v \rangle > 0\ \forall v \in V$*,* $v \neq 0_V$*.*

As in the case of the canonical inner product, for a complex form over V, *the symmetry and sesquilinearity requirement is equivalent to requiring the Hermitian property and linearity on the left-hand side*; if these properties are verified, then:

$$\langle v, \alpha w \rangle = \overline{\langle \alpha w, v \rangle} = \overline{\alpha \langle w, v \rangle} = \bar{\alpha} \overline{\langle w, v \rangle} = \bar{\alpha} \overline{\overline{\langle v, w \rangle}} = \bar{\alpha} \langle v, w \rangle, \quad \forall \alpha \in \mathbb{C}.$$

Considering the sum of n, rather than two, vectors, sesquilinearity is represented by the following formulae:

$$\boxed{\langle \sum_{i=1}^{n} \alpha_i v_i\,,\ w \rangle = \sum_{i=1}^{n} \alpha_i \langle v_i, w \rangle} \qquad [1.1]$$

$$\boxed{\langle v\,,\ \sum_{i=1}^{n} \alpha_i w_i \rangle = \sum_{i=1}^{n} \overline{\alpha_i} \langle v, w_i \rangle} \qquad [1.2]$$

In $\mathbb{C}^n$, the *complex Euclidean inner product* is defined by:

$$\langle v, w \rangle \equiv \sum_{i=1}^{n} v_i \overline{w_i} = v \cdot (\overline{w})^t = v^t \cdot \overline{w}$$

where $v = (v_1, v_2, \ldots, v_n), w = (w_1, w_2, \ldots, w_n) \in \mathbb{C}^n$ are written with their components in relation to any given, but fixed, basis $(b_i)_{i=1}^n$ in $\mathbb{C}^n$.

The symbol $\mathbb{K}$ will be used throughout to represent either $\mathbb{R}$ or $\mathbb{C}$ in the context of properties which are valid independently of the reality or complexity of the inner product.

THEOREM 1.1.– Let $(V, \langle\ ,\ \rangle)$ be an inner product space. We have:

1) $\langle v, 0_V\rangle = 0\ \forall v \in V$;

2) if $\langle u, w\rangle = \langle v, w\rangle\ \forall w \in V$, then u and v must coincide;

3) $\langle v, w\rangle = 0\ \forall v \in V \iff w = 0_V$, i.e. the null vector is the only vector which is orthogonal to all of the other vectors.

PROOF.–

1) $\langle v, 0_V\rangle = \langle v, 0_V + 0_V\rangle = \langle v, 0_V\rangle + \langle v, 0_V\rangle$ by linearity, i.e. $\langle v, 0_V\rangle - \langle v, 0_V\rangle = 0 = \langle v, 0_V\rangle$.

2) $\langle u, w\rangle = \langle v, w\rangle\ \forall w \in V$ implies, by linearity, that $\langle u - v, w\rangle = 0\ \forall w \in V$ and thus, notably, considering $w = u - v$, we obtain $\langle u - v, u - v\rangle = 0$, implying, due to the definite positiveness of the inner product, that $u - v = 0_V$, i.e. $u = v$.

3) If $w = 0_V$, then $\langle v, w\rangle = 0\ \forall v \in V$ using property (1). Inversely, by hypothesis, it holds that $\langle v, w\rangle = 0 = \langle v, 0_V\rangle\ \forall v \in V$, but then property (2) implies that $w = 0_V$. □

Finally, let us consider a typical property of the complex inner product, which results directly from a property of complex numbers.

THEOREM 1.2.– Let $(V, \langle\ ,\ \rangle)$ be a complex inner product space. Thus:

$$\Im(\langle v, w\rangle) = \Re(\langle v, iw\rangle) \qquad \forall v, w \in V$$

PROOF.– Consider any complex number $z = a + ib$, so $-iz = b - ia$, hence $b = \Im(z) = \Re(-iz)$. Taking $z = \langle v, w\rangle$, we obtain $\Im(\langle v, w\rangle) = \Re(-i\langle v, w\rangle) = \Re(\langle v, iw\rangle)$ by sesquilinearity. □

1.2. The norm associated with an inner product and normed vector spaces

If $(V, \langle, \rangle)$ is an inner product space over $\mathbb{K}$, then a norm on V can be defined as follows:

$$\begin{aligned} \|\ \|:\ V &\to \mathbb{R}_0^+ = [0, +\infty) \\ v &\to \|v\| = \sqrt{\langle v, v\rangle} \end{aligned}$$

Note that $\|v\|$ is well defined since $\langle v, v\rangle \geqslant 0\ \forall v \in V$. Once a norm has been established, it is always possible to define a *distance* between two vectors v, w in V: $d(v, w) = \|v - w\|$.

The vector $v \in V$ such that $\|v\| = 1$ is known as a *unit vector*. Every vector $v \in V$ can be *normalized* to produce a unit vector, simply by dividing it by its norm.

NOTABLE EXAMPLES.–

$$(\mathbb{R}^n, \langle, \rangle) : \|v\| = \sqrt{\sum_{i=1}^{n} v_i^2}$$

$$(\mathbb{C}^n, \langle, \rangle) : \|v\| = \sqrt{\sum_{i=1}^{n} v_i \overline{v}_i} = \sqrt{\sum_{i=1}^{n} |v_i|^2}$$

Three properties of the norm, which should already be known, are listed below. Taking any $v, w \in V$, and any $\alpha \in \mathbb{K}$:

1) $\|v\| \geqslant 0, \ \|v\| = 0 \iff v = 0_V$;
2) $\|\alpha v\| = |\alpha|\|v\|$ (homogeneity);
3) $\|v + w\| \leqslant \|v\| + \|w\|$ (triangle inequality).

DEFINITION 1.4 (normed vector space).– *A normed vector space is a pair* $(V, \| \ \|)$ *given by a vector space* V *and a function, called a norm,* $\| \ \| : V \to \mathbb{R}_0^+$*, satisfying the three properties listed above.*

A norm $\| \ \|$ *is Hilbertian if there exists an inner product* $\langle \ , \ \rangle$ *on* V *such that* $\|v\| = \sqrt{\langle v, v\rangle}\ \forall v \in V$.

Canonically, an inner product space is therefore a normed vector space. Counter-examples can be used to show that the reverse is not generally true.

Note that, by definition, $\langle v, v\rangle = \|v\| \, \|v\|$, but, in general, the magnitude of the inner product between two different vectors is dominated by the product of their norms. This is the result of the well-known inequality shown below.

THEOREM 1.3 (Cauchy-Schwarz inequality).– For all $v, w \in (V, \langle \ , \ \rangle)$ we have:

$$\boxed{|\langle v, w\rangle| \leqslant \|v\|\|w\|}$$

PROOF.– Dozens of proofs of the Cauchy-Schwarz inequality have been produced. One of the most elegant proofs is shown below, followed by the simplest one:

– *first proof*: if $w = 0_V$, then the inequality is verified trivially with $0 = 0$. If $w \neq 0_V$, then we can define $z = v - \frac{\langle v,w\rangle}{\|w\|^2}w$, i.e. $v = \frac{\langle v,w\rangle}{\|w\|^2}w + z$, and we note that:

$$\langle z, w\rangle = \langle v - \frac{\langle v, w\rangle}{\|w\|^2}w, w\rangle = \langle v, w\rangle - \frac{\langle v, w\rangle}{\cancel{\|w\|^2}}\cancel{\langle w, w\rangle} = 0$$

thus:

$$\begin{aligned}\|v\|^2 &= \langle v, v\rangle = \left\langle \frac{\langle v, w\rangle}{\|w\|^2}w + z, \frac{\langle v, w\rangle}{\|w\|^2}w + z \right\rangle \\ &= \frac{\langle v, w\rangle}{\|w\|^2}\left\langle w, \frac{\langle v, w\rangle}{\|w\|^2}w + z \right\rangle + \left\langle z, \frac{\langle v, w\rangle}{\|w\|^2}w + z \right\rangle \\ &= \frac{\langle v, w\rangle}{\|w\|^2}\frac{\overline{\langle v, w\rangle}}{\cancel{\|w\|^2}}\cancel{\langle w, w\rangle} + \frac{\langle v, w\rangle}{\|w\|^2}\langle w, z\rangle + \frac{\overline{\langle v, w\rangle}}{\|w\|^2}\langle z, w\rangle + \langle z, z\rangle \\ &= \frac{|\langle v, w\rangle|^2}{\|w\|^2} + \|z\|^2\end{aligned}$$

as the two intermediate terms in the penultimate step are zero, since $\langle z, w\rangle = \langle w, z\rangle = 0$.

As $\|z\|^2 \geqslant 0$, we have seen that:

$$\|v\|^2 = \frac{|\langle v, w\rangle|^2}{\|w\|^2} + \|z\|^2 \geqslant \frac{|\langle v, w\rangle|^2}{\|w\|^2}$$

i.e. $|\langle v, w\rangle|^2 \leqslant \|v\|^2\|w\|^2$, hence $|\langle v, w\rangle| \leqslant \|v\|\|w\|$;

– *second proof* (in one line!): $\forall t \in \mathbb{R}$ we have:

$$\begin{aligned}&0 \leqslant \|tv - w\|^2 = \langle tv - w, tv - w\rangle = t^2\|v\|^2 - 2t\langle v, w\rangle + \|w\|^2 \\ &\iff \Delta = 4\langle v, w\rangle^2 - 4\|v\|^2\|w\|^2 \leqslant 0\end{aligned}$$

□

The Cauchy-Schwarz inequality allows the concept of the angle between two vectors to be generalized for abstract vector spaces. In fact, it implies the existence of a coefficient k between -1 and $+1$ such that $\langle v, w\rangle = \|v\|\|w\|k$, but, given that the restriction of cos to $[0, \pi]$ creates a bijection with $[-1, 1]$, this means that there is only one $\vartheta \in [0, \pi]$ such that $\langle v, w\rangle = \|v\|\|w\| \cos\vartheta$. $\vartheta \in [0, \pi]$ is known as the angle between the two vectors v and w.

Another very important property of the norm is as follows.

THEOREM 1.4.– Let $(V, \| \ \|)$ be an arbitrary normed vector space and $v, w \in V$. We have:

$$\boxed{|\|v\| - \|w\|| \leqslant \|v - w\|} \qquad [1.3]$$

PROOF.– On one side:

$$\|v\| = \|v - w + w\| = \|(v - w) + w\| \leqslant \|v - w\| + \|w\|$$

by the triangle inequality, thus $\|v\| - \|w\| \leqslant \|v - w\|$. On the other side:

$$\|w\| = \|w - v + v\| = \|(w - v) + v\| \leqslant \|w - v\| + \|w\|$$

thus $\|w\| - \|v\| \leqslant \|v - w\|$, i.e. $\|v\| - \|w\| \geqslant -\|v - w\|$.

Hence, $-\|v - w\| \leqslant \|v\| - \|w\| \leqslant \|v - w\|$, i.e. $|\|v\| - \|w\|| \leqslant \|v - w\|$. □

The following formula is also extremely useful.

THEOREM 1.5 (Carnot's theorem).– Taking $v, w \in (V, \langle\ ,\ \rangle)$:

$$\boxed{\|v \pm w\|^2 = \|v\|^2 + \|w\|^2 \pm 2\langle v, w\rangle, \quad (\mathbb{K} = \mathbb{R})} \qquad [1.4]$$

and

$$\boxed{\|v \pm w\|^2 = \|v\|^2 + \|w\|^2 \pm \langle v, w\rangle \pm \langle w, v\rangle, \quad (\mathbb{K} = \mathbb{C})} \qquad [1.5]$$

PROOF.– Direct calculation:

$$\begin{aligned}\|v \pm w\|^2 &= \langle v \pm w, v \pm w\rangle = \langle v, v\rangle \pm \langle v, w\rangle \pm \langle w, v\rangle + \langle w, w\rangle \\ &= \|v\|^2 + \|w\|^2 \pm \langle v, w\rangle \pm \langle w, v\rangle\end{aligned}$$ □

If $\mathbb{K} = \mathbb{C}$, then $\langle w, v\rangle = \overline{\langle v, w\rangle}$, and since, if $z = a + ib = \Re(z) + i\Im(z)$, then $z + \bar{z} = 2a = 2\Re(z)$, we can rewrite [1.5] as:

$$\boxed{\|v \pm w\|^2 = \|v\|^2 + \|w\|^2 \pm 2\Re(\langle v, w\rangle)} \qquad [1.6]$$

The laws presented in this section have immediate consequences which will be highlighted in section 1.2.1.

1.2.1. *The parallelogram law and the polarization formula*

The parallelogram law in $\mathbb{R}^2$ is shown in Figure 1.1. This law can be generalized on a vector space with an arbitrary inner product.

THEOREM 1.6 (Parallelogram law).– Let $(V, \langle , \rangle)$ be an inner product space on $\mathbb{K}$. Thus, $\forall v, w \in V$:

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2)$$

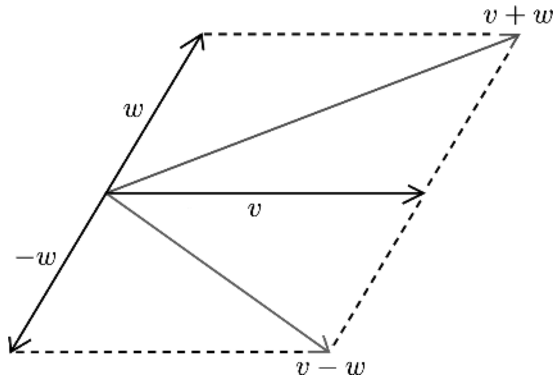

Figure 1.1. *Parallelogram law in $\mathbb{R}^2$: The sum of the squares of the two diagonal lines is equal to two times the sum of the squares of the edges v and w. For a color version of this figure, see www.iste.co.uk/provenzi/spaces.zip*

PROOF.– A direct consequence of law [1.4] or law [1.5] taking $\|v+w\|^2$ then $\|v-w\|^2$. □

As we have seen, an inner product induces a norm. The polarization formula can be used to "reverse" roles and write the inner product using the norm.

THEOREM 1.7 (Polarization formula).– Let $(V, \langle , \rangle)$ be an inner product space on $\mathbb{K}$. In this case, $\forall v, w \in V$:

$$\langle v, w \rangle = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right), \quad (\mathbb{K} = \mathbb{R})$$

and:

$$\langle v, w \rangle = \frac{1}{4}\left[\|v+w\|^2 - \|v-w\|^2 + i\left(\|v+iw\|^2 - \|v-iw\|^2\right)\right], \quad (\mathbb{K} = \mathbb{C})$$

PROOF.– This law is a direct consequence of law [1.4], in the real case. For the complex case, w is replaced by iw in law [1.5], and by sesquilinearity, we obtain:

$$\|v \pm iw\|^2 = \|v\|^2 + \|w\|^2 \mp i\langle v, w \rangle \pm i\langle w, v \rangle$$

By direct calculation, we can then verify that $\|v+w\|^2 - \|v-w\|^2 + i\,\|v+iw\|^2 - i\,\|v-iw\|^2 = 4\langle v, w \rangle$. □

It may seem surprising that something as simple as the parallelogram law may be used to establish a necessary and sufficient condition to guarantee that a norm over a vector space will be induced by an inner product, that is, the norm is Hilbertian. This notion will be formalized in Chapter 4.

1.3. Orthogonal and orthonormal families in inner product spaces

The "geometric" definition of an inner product in $\mathbb{R}^2$ and $\mathbb{R}^3$ indicates that this product is zero if and only if ϑ, the angle between the vectors, is $\pi/2$, which implies $\cos(\vartheta) = 0$.

In more complicated vector spaces (e.g. polynomial spaces), or even Euclidean vector spaces of more than three dimensions, it is no longer possible to visualize vectors; their orthogonality must therefore be "axiomatized" *via* the nullity of their scalar product.

DEFINITION 1.5.– *Let* $(V, \langle,\rangle)$ *be a real or complex inner product space of finite dimension* n*. Let* $F = \{v_1, \cdots, v_n\}$ *be a family of vectors in* V*. Thus:*

– F *is an* orthogonal family *of vectors if each different vector pair has an inner product of 0:* $\langle v_i, v_j\rangle = 0$*;*

– F *is an* orthonormal family *if it is orthogonal and, furthermore,* $\|v_i\| = 1\ \forall i$*. Thus, if* $\{v_i\}_{i=1}^n$ *is an orthogonal family,* $\{u_i = \|v_i\|^{-1} v_i\}_{i=1}^n$ *is an orthonormal family.*

An orthonormal family (unit and orthogonal vectors) may be characterized as follows:

$$\boxed{\langle v_i, v_j\rangle = \delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \qquad \textit{Orthonormal family}}$$

$\delta_{i,j}$ is the *Kronecker delta*[5].

1.4. Generalized Pythagorean theorem

The Pythagorean theorem can be generalized to abstract inner product spaces. The general formulation of this theorem is obtained using a lemma.

LEMMA 1.1.– Let $(V, \langle,\rangle)$ be a real or complex inner product space. Let $u \in V$ be orthogonal to all vectors $v_1, \ldots, v_n \in V$. Hence, u is also orthogonal to all vectors in V obtained as a linear combination of $v_1, \ldots, v_n$.

PROOF.– Let $w = \sum_{i=1}^{n} \alpha_i v_i$, $\alpha_i \in \mathbb{K}\ \forall i = 1, \ldots, n$, be an arbitrary linear combination of vectors $v_1, \ldots, v_n$. By direct calculation:

$$\langle u, w\rangle = \langle u, \sum_{i=1}^{n} \alpha_i v_i\rangle \underset{\text{(sesquilinearity)}}{=} \sum_{i=1}^{n} \overline{\alpha_i}\langle u, v_i\rangle \underset{u \perp v_i}{=} \sum_{i=1}^{n} \overline{\alpha_i} 0 = 0 \qquad \square$$

5 Leopold Kronecker (1823, Liegnitz-1891, Berlin).

THEOREM 1.8 (Generalized Pythagorean theorem).– Let $(V, \langle , \rangle)$ be an inner product space on $\mathbb{K}$. Let $u, v \in V$ be orthogonal to each other. Hence:

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

More generally, if the vectors $v_1, \ldots, v_n \in V$ are orthogonal, then:

$$\left\| \sum_{i=1}^{n} v_i \right\|^2 = \sum_{i=1}^{n} \|v_i\|^2 \qquad \Longleftrightarrow \qquad \|v_1 + \ldots + v_n\|^2 = \|v_1\|^2 + \ldots + \|v_n\|^2$$

PROOF.– The two-vector case can be proven thanks to Carnot's formula:

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \cancelto{0}{\langle u, v \rangle} + \cancelto{0}{\langle v, u \rangle} + \langle v, v \rangle \\ &= \|u\|^2 + \|v\|^2 \end{aligned}$$

Proof for cases with n vectors is obtained by recursion:

– the case where $n = 2$ is demonstrated above;

– we suppose that $\left\| \sum\limits_{i=1}^{n-1} v_i \right\|^2 = \sum\limits_{i=1}^{n-1} \|v_i\|^2$ (recursion hypothesis);

– now, we write $u = v_n$ and $z = \sum\limits_{i=1}^{n-1} v_i$, so $u \perp z$ using Lemma 1.1. Hence, using case $n = 2$: $\|u + z\|^2 = \|u\|^2 + \|z\|^2$, but:

$$u + z = v_n + \sum_{i=1}^{n-1} v_i = \sum_{i=1}^{n} v_i$$

so:

$$\|u + z\|^2 = \left\| \sum_{i=1}^{n} v_i \right\|^2$$

and:

$$\|u\|^2 + \|z\|^2 = \|v_n\|^2 + \left\| \sum_{i=1}^{n-1} v_i \right\|^2 \underset{\text{(Recursion hypothesis)}}{=} \|v_n\|^2 + \sum_{i=1}^{n-1} \|v_i\|^2 = \sum_{i=1}^{n} \|v_i\|^2$$

giving us the desired thesis. □

Note that *the Pythagorean theorem thesis is a double implication if and only if V is real*, in fact, using law [1.6] we have that $\|u + v\|^2 = \|u\|^2 + \|v\|^2$ holds true if and only if $\Re(\langle u, v \rangle) = 0$, which is equivalent to orthogonality if and only if V is real.

The following result gives information concerning the distance between any two vectors within an orthonormal family.

THEOREM 1.9.– Let $(V, \langle , \rangle)$ be an inner product space on $\mathbb{K}$ and let F be an orthonormal family in V. The distance between any two elements of F is constant and equal to $\sqrt{2}$.

PROOF.– Using the Pythagorean theorem: $\|u + (-v)\|^2 = \|u\|^2 + \|v\|^2 = 2$, from the fact that $u \perp v$. □

1.5. Orthogonality and linear independence

The orthogonality condition is more restrictive than that of linear independence: *all orthogonal families are free*.

THEOREM 1.10.– Let F be an orthogonal family in $(V, \langle , \rangle)$, $F = \{v_1, \cdots, v_n\}$, $v_i \neq 0 \; \forall i$, then F is free.

PROOF.– We need to prove the linear independence of the elements v_i, that is, $\sum_{i=1}^{n} a_i v_i = 0 \implies a_i = 0 \; \forall i$. To this end, we calculate the inner product of the linear combination $\sum_{i=1}^{n} a_i v_i$ and an arbitrary vector v_j with $j \in \{1, \ldots, n\}$:

$$\langle \sum_{i=1}^{n} a_i v_i, v_j \rangle \underset{[1.1]}{=} \sum_{i=1}^{n} a_i \langle v_i, v_j \rangle \underset{(\langle v_i, v_j \rangle \neq 0 \Leftrightarrow i=j)}{=} a_j \langle v_j, v_j \rangle = a_j \|v_j\|^2$$

By hypothesis, none of the vectors in F are zero; the hypothesis that $\sum_{i=1}^{n} a_i v_i = 0$ therefore implies that:

$$\underbrace{\langle 0, v_j \rangle}_{\substack{\shortparallel \\ 0}} = a_j \underbrace{\|v_j\|^2}_{\substack{\nparallel \\ 0}} \Rightarrow a_j = 0.$$

This holds for any $j \in \{1, \ldots, n\}$, so the orthogonal family F is free. □

Using the general theory of vector spaces in finite dimensions, an immediate corollary can be derived from Theorem 1.10.

COROLLARY 1.1.– An orthogonal family of n non-null vectors in a space $(V, \langle , \rangle)$ of dimension n is a basis of V.

DEFINITION 1.6.– *A family of n non-null orthogonal vectors in a vector space $(V, \langle , \rangle)$ of dimension n is said to be an* orthogonal basis *of V. If this family is also orthonormal, it is said to be an* orthonormal basis *of V.*

The extension of the orthogonal basis concept to inner product spaces of infinite dimensions will be discussed in Chapter 5. For the moment, it is important to note that *an orthogonal basis is made up of the maximum number of mutually orthogonal vectors in a vector space*. Taking n to represent the dimension of the space V and proceeding by reductio ad absurdum, imagine the existence of another vector $u^* \in V$, $u \neq 0$, orthogonal to all of the vectors in an orthogonal basis $(u_i)_{i=1}^n$; in this case, the set $(u^*, u_i)_{i=1}^n$ would be free as orthogonal vectors are linearly independent, and the dimension of V would be $n+1$ instead of n! This property is usually expressed by saying that an orthogonal family is a basis if it is not a subset of another orthogonal family of vectors in V.

Note that in order to determine the components of a vector in relation to an arbitrary basis, we must solve a linear system of n equations with n unknown variables. In fact, if $v \in V$ is any vector and (u_i) $i = 1, \ldots, n$ is a basis of V, then the components of v in (u_i) are the scalars $\alpha_1, \ldots, \alpha_n$ such that:

$$v = \sum_{i=1}^{n} \alpha_i u_i \qquad \Longleftrightarrow \qquad \begin{cases} v_1 = \sum_{i=1}^{n} \alpha_i u_{i,1} \\ \vdots \\ v_n = \sum_{i=1}^{n} \alpha_i u_{i,n}, \end{cases}$$

where $u_{i,j}$ is the j-th component of vector u_i.

However, in the presence of an orthogonal or orthonormal basis, components are determined by inner products, as seen in Theorem 1.11.

Note, too, that solving a linear system of n equations with n unknown variables generally involves far more operations than the calculation of inner products; this highlights one advantage of having an orthogonal basis for a vector space.

THEOREM 1.11.– Let $B = \{u_1, \ldots, u_n\}$ be an *orthogonal* basis of $(V, \langle , \rangle)$. Then:

$$\boxed{v = \sum_{i=1}^{n} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i}$$

Notably, if B is an *orthonormal* basis, then:

$$\boxed{v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i}$$

PROOF.– B is a basis, so there exists a set of scalars $\alpha_1, \dots, \alpha_n$ such that $v = \sum_{j=1}^{n} \alpha_j u_j$. Consider the inner product of this expression of v with a fixed vector u_i, $i \in \{1, \dots, n\}$:

$$\langle v, u_i \rangle = \langle \sum_{j=1}^{n} \alpha_j u_j, u_i \rangle = \sum_{j=1}^{n} \alpha_j \langle u_j, u_i \rangle \underset{(u_i \perp u_j \ \forall i \neq j)}{=} \alpha_i \langle u_i, u_i \rangle = \alpha_i \|u_i\|^2$$

so $\alpha_i = \frac{\langle v, u_i \rangle}{\|u_i\|^2}$ $\forall i = 1, \cdots, n$, and thus $v = \sum_{i=1}^{n} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i$. If B is an orthonormal basis, $\|u_i\| = 1$ giving the second law in the theorem. □

Geometric interpretation of the theorem: The theorem that we are about to demonstrate is the generalization of the decomposition theorem of a vector in plane $\mathbb{R}^2$ or in space $\mathbb{R}^3$ on a canonical basis of unit vectors on axes. To simplify this, consider the case of $\mathbb{R}^2$.

If $\hat{\imath}$ and $\hat{\jmath}$ are, respectively, the unit vectors of axes x and y, then the decomposition theorem says that:

$$v = \underbrace{\|v\| \cos \alpha}_{\langle v, \hat{\imath} \rangle} \hat{\imath} + \underbrace{\|v\| \cos \beta}_{\langle v, \hat{\jmath} \rangle} \hat{\jmath} = \langle v, \hat{\imath} \rangle \hat{\imath} + \langle v, \hat{\jmath} \rangle \hat{\jmath}$$

which is a particular case of the theorem above.

We will see that the *Fourier series* can be viewed as a further generalization of the decomposition theorem on an orthogonal or orthonormal basis.

1.6. Orthogonal projection in inner product spaces

The definition of orthogonal projection can be extended by examining the geometric and algebraic properties of this operation in $\mathbb{R}^2$ and $\mathbb{R}^3$. Let us begin with $\mathbb{R}^2$.

In the Euclidean space $\mathbb{R}^2$, the inner product of a vector v and a *unit* vector evidently gives us the orthogonal projection of v in the direction defined by this vector, as shown in Figure 1.2 with an orthogonal projection along the x axis.

The properties verified by this projection are as follows:

1) projecting onto the x axis a second time, vector $P_x v$ obviously remains unchanged given that it is already on the x axis, i.e. $P_x^2(v) := P_x(P_x v) = P_x v$ $\forall v \in V$. Put differently, the operator P_x bound to the x axis is the identity of this axis;

2) the difference vector between v and its projection $v - P_x v$ is orthogonal to the x axis, as we see from Figure 1.3;

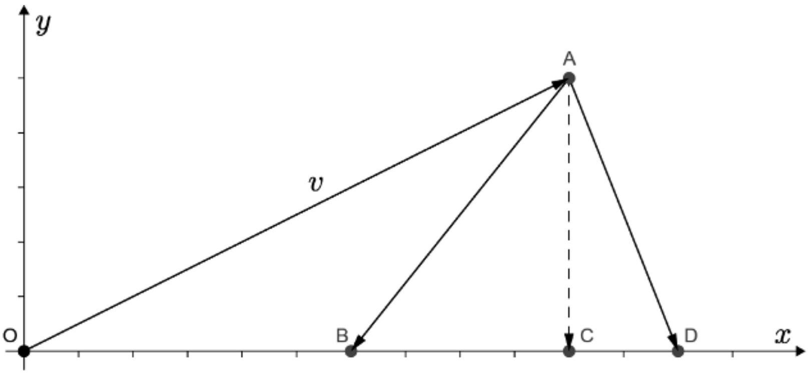

Figure 1.2. *Orthogonal projection $P_x v = \vec{OC}$ and diagonal projections $\vec{OB}$ and $\vec{OD}$ of a vector in $v \in \mathbb{R}^2$ onto the x axis. For a color version of this figure, see www.iste.co.uk/provenzi/spaces.zip*

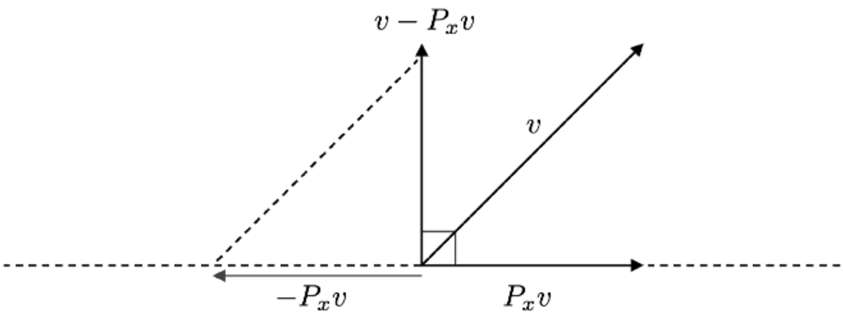

Figure 1.3. *Visualization of property 2 in $\mathbb{R}^2$. For a color version of this figure, see www.iste.co.uk/provenzi/spaces.zip*

3) $P_x v$ minimizes the distance between the terminal point of v and the x axis. In Figure 1.2, $\vec{AB}$ and $\vec{AD}$ are, in fact, the hypotenuses of right-angled triangles ABC and ACD; on the other hand, $\vec{AC}$ is another side of these triangles, and is therefore smaller than $\vec{AB}$ and $\vec{AD}$. $\vec{AC}$ is the distance between the terminal point of v and the terminal point of $P_x v$, while $\vec{AB}$ and $\vec{AD}$ are the distances between the terminal point of v and the diagonal projections of v onto x rooted at B and D, respectively.

We wish to define an orthogonal projection operation for an abstract inner product space of dimension n which retains these same geometric properties.

Analyzing orthogonal projections in $\mathbb{R}^3$ helps us to establish an idea of the algebraic definition of this operation. Figure 1.4 shows a vector $v \in \mathbb{R}^3$ and the plane produced by the *orthogonal* vectors u_1 and u_2. We see that the projection p of v onto this plane is the vector sum of the orthogonal projections $p_1 = \frac{\langle v,u_1\rangle}{\|u_1\|^2}u_1$ and $p_2 = \frac{\langle v,u_2\rangle}{\|u_2\|^2}u_2$ onto the two vectors u_1 and u_2 taken separately, i.e. $p = p_1 + p_2 = \sum_{i=1}^{2} \frac{\langle v,u_i\rangle}{\|u_i\|^2} u_i$.

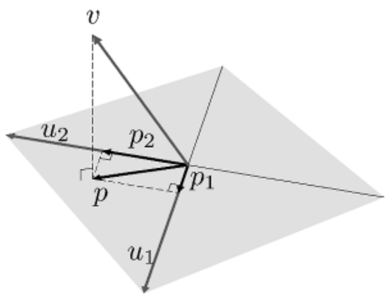

Figure 1.4. *Orthogonal projection p of a vector in $\mathbb{R}^3$ onto the plane produced by two unit vectors. For a color version of this figure, see www.iste.co.uk/provenzi/spaces.zip*

Generalization should now be straightforward: consider an inner product space $(V, \langle,\rangle)$ of dimension n and an *orthogonal family* of non-zero vectors $F = \{u_1, \ldots, u_m\}$, $m \leqslant n$, $u_i \neq 0_V$ $\forall i = 1, \ldots, m$.

The vector subspace of V produced by all linear combinations of the vectors of F shall be written Span(F):

$$\text{span}(F) \equiv S = \left\{ s \in V \ : \ \exists \alpha_1, \ldots, \alpha_m \in \mathbb{K} \text{ such that } s = \sum_{j=1}^{m} \alpha_j u_j \right\}$$

The *orthogonal projection operator* or *orthogonal projector* of a vector $v \in V$ onto S is defined as the following application, which is obviously linear:

$$\begin{aligned} P_S : V &\longrightarrow S \subseteq V \\ v &\longmapsto \boxed{P_S(v) = \sum_{i=1}^{m} \frac{\langle v, u_i\rangle}{\|u_i\|^2} u_i} \end{aligned}$$

Theorem 1.12 shows that the orthogonal projection defined above retains all of the properties of the orthogonal projection demonstrated for $\mathbb{R}^2$.

THEOREM 1.12.– Using the same notation as before, we have:

1) if $s \in S$ then $P_S(s) = s$, i.e. the action of P_S on the vectors in S is the identity;

2) $\forall v \in V$ and $s \in S$, the *residual* vector of the projection, i.e. $v - P_S(v)$, is $\perp$ to S:

$$\langle v - P_S(v), s\rangle = 0 \quad \iff \quad v - P_S(v) \perp s$$

3) $\forall v \in V$ et $s \in S$: $\|v - P_S(v)\| \leqslant \|v - s\|$ and the equality holds if and only if $s = P_S(v)$. We write:

$$\boxed{P_S(v) = \underset{s \in S}{\textit{argmin}} \, \|v - s\|}$$

PROOF.–

1) Let $s \in S$, i.e. $s = \sum_{j=1}^{m} \alpha_j u_j$, then:

$$P_S(s) = \sum_{i=1}^{m} \frac{\langle \sum_{j=1}^{m} \alpha_j u_j, u_i\rangle}{\|u_i\|^2} u_i = \sum_{i=1}^{m} \frac{\sum_{j=1}^{m} \alpha_j \langle u_j, u_i\rangle}{\|u_i\|^2} u_i$$

$$\underset{(u_i \perp u_j \ \forall i \neq j)}{=} \sum_{i=1}^{m} \frac{\alpha_i \langle u_i, u_i\rangle}{\|u_i\|^2} u_i = \sum_{i=1}^{m} \alpha_i u_i = s$$

2) Consider the inner product of $P_S(v)$ and a fixed vector u_j, $j \in \{1, \dots, m\}$:

$$\langle P_S(v), u_j\rangle = \langle \sum_{i=1}^{m} \frac{\langle v, u_i\rangle}{\|u_i\|^2} u_i, u_j\rangle \underset{\text{(linearity)}}{=} \sum_{i=1}^{m} \frac{\langle v, u_i\rangle}{\|u_i\|^2} \langle u_i, u_j\rangle$$

$$\underset{(u_i \perp u_j \ \forall i \neq j)}{=} \frac{\langle v, u_j\rangle}{\|u_j\|^2} \langle u_j, u_j\rangle = \langle v, u_j\rangle$$

hence:

$$\langle v, u_j\rangle - \langle P_S(v), u_j\rangle = 0 \underset{\text{linearity of } \langle\ ,\ \rangle}{\iff} \langle v - P_S(v), u_j\rangle = 0 \quad \forall j \in \{1, ..., m\}$$

Lemma 1.1 guarantees that $\langle v - P_S(v), s\rangle = 0 \ \forall s = \sum_{j=1}^{m} \alpha_j u_j$.

3) It is helpful to rewrite the difference $v - s$ as $v - P_S(v) + P_S(v) - s$. From property 2, $v - P_S(v) \perp S$, however $P_S(v), s \in S$ so $P_S(v) - s \in S$. Hence $(v - P_S(v)) \perp (P_S(v) - s)$. The generalized Pythagorean theorem implies that:

$$\|v - s\|^2 = \|v - P_S(v) + P_S(v) - s\|^2 = \|v - P_S(v)\|^2 + \underbrace{\|P_S(v) - s\|^2}_{\geqslant 0} \geqslant \|v - P_S(v)\|^2$$

hence $\|v - s\| \geqslant \|v - P_S(v)\| \ \forall v \in V, \ s \in S$.

Evidently, $\|P_S(v) - s\|^2 = 0$ if and only if $s = P_S(v)$, and in this case $\|v - s\|^2 = \|v - P_S(v)\|^2$. □

The theorem demonstrated above tells us that the vector in the vector subspace $S \subseteq V$ which is the most "similar" to $v \in V$ (in the sense of the norm induced by the inner product) is given by the orthogonal projection. The generalization of this result to infinite-dimensional Hilbert spaces will be discussed in Chapter 5.

As already seen for the projection operator in $\mathbb{R}^2$ and $\mathbb{R}^3$, the non-negative scalar quantity $\frac{|\langle v,u_i\rangle|}{\|u_i\|}$ gives a measure of the importance of $\frac{u_i}{\|u_i\|}$ in the reconstruction of the best approximation of v in S *via* the formula $P_S(v) = \sum\limits_{i=1}^{m} \frac{\langle v,u_i\rangle}{\|u_i\|^2} u_i$: if this quantity is large, then $\frac{u_i}{\|u_i\|}$ is very important to reconstruct $P_S(v)$, otherwise, in some circumstances, it may be ignored. In the applications to signal compression, a usual strategy consists of reordering the summation that defines $P_S(v)$ in descent order of the quantities $\frac{|\langle v,u_i\rangle|}{\|u_i\|}$ and trying to eliminate as many small terms as possible without degrading the signal quality.

This observation is crucial to understanding the significance of the Fourier decomposition, which will be examined in both discrete and continuous contexts in the following chapters.

Finally, note that the seemingly trivial equation $v = v - s + s$ is, in fact, far more meaningful than it first appears when we know that $s \in S$: in this case, we know that $v - s$ and s are orthogonal.

The decomposition of a vector as the sum of a component belonging to a subspace S and a component belonging to its orthogonal is known as the *orthogonal projection theorem*.

This decomposition is unique, and its generalization for infinite dimensions, alongside its consequences for the geometric structure of Hilbert spaces, will be examine in detail in Chapter 5.

1.7. Existence of an orthonormal basis: the Gram-Schmidt process

As we have seen, projection and decomposition laws are much simpler when an orthonormal basis is available.

Theorem 1.13 states that in a finite-dimensional inner product space, an orthonormal basis can always be constructed from a free family of generators.

THEOREM 1.13.– (The iterative Gram-Schmidt process[6]) If $(v_1, \ldots, v_n)$, $n \leqslant \infty$ is a basis of $(V, \langle , \rangle)$, then an orthonormal basis of $(V, \langle , \rangle)$ can be obtained from $(v_1, \ldots, v_n)$.

PROOF.– This proof is constructive in that it provides the method used to construct an orthonormal basis from any arbitrary basis.

– Step 1: normalization of v_1:

$$u_1 = \frac{v_1}{\|v_1\|}$$

– Step 2, illustrated in Figure 1.5: v_2 is projected in the direction of u_1, that is, we consider $\langle v_2, u_1 \rangle u_1$. We know from Theorem 1.12 that the vector difference $v_2 - \langle v_2, u_1 \rangle u_1$ is orthogonal to u_1. The result is then normalized:

$$u_2 = \frac{v_2 - \langle v_2, u_1 \rangle u_1}{\|v_2 - \langle v_2, u_1 \rangle u_1\|}$$

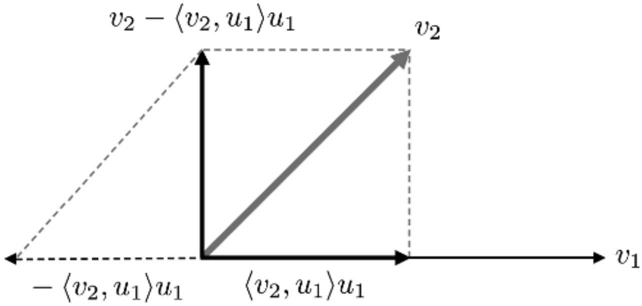

Figure 1.5. *Illustration of the second step in the Gram-Schmidt orthonormalization process. For a color version of this figure, see www.iste.co.uk/provenzi/spaces.zip*

– Step n, by iteration:

$$u_n = \frac{v_n - (\langle v_n, u_{n-1} \rangle u_{n-1} + \ldots + \langle v_n, u_1 \rangle u_1)}{\|v_n - (\langle v_n, u_{n-1} \rangle u_{n-1} + \ldots + \langle v_n, u_1 \rangle u_1)\|} \qquad \square$$

1.8. Fundamental properties of orthonormal and orthogonal bases

The most important properties of an orthonormal basis are listed in Theorem 1.14.

6 Jørgen Pedersen Gram (1850, Nustrup-1916, Copenhagen), Erhard Schmidt (1876, Tatu-1959, Berlin).

THEOREM 1.14.– *Let* $(u_1, \ldots, u_n)$ *be an orthonormal basis of* $(V, \langle,\rangle)$, *dim*$(V) = n$. *Then,* $\forall v, w \in V$*:*

1) Decomposition theorem on an orthonormal basis*:*

$$\boxed{v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i} \tag{1.7}$$

2) Parseval's identity[7]*:*

$$\boxed{\langle v, w \rangle = \sum_{i=1}^{n} \langle v, u_i \rangle \langle u_i, w \rangle} \tag{1.8}$$

3) Plancherel's theorem[8]*:*

$$\boxed{\|v\|^2 = \sum_{i=1}^{n} |\langle v, u_i \rangle|^2} \tag{1.9}$$

Proof of 1: an immediate consequence of Theorem 1.12. Given that $(u_1, \ldots, u_n)$ is a basis, $v \in \text{span}(u_1, \ldots, u_n)$; furthermore, $(u_1, \ldots, u_n)$ is orthonormal, so $v = P_S(v) = \sum_{i=1}^{n} \langle v, u_i \rangle u_i$. It is not necessary to divide by $\|u_i\|^2$ when summing since $\|u_i\| = 1 \; \forall i$.

Proof of 2: using point 1 it is possible to write $v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i$, and calculating the inner product of v, written in this way, and w, using equation [1.1], we obtain:

$$\langle v, w \rangle = \langle \sum_{i=1}^{n} \langle v, u_i \rangle u_i, w \rangle = \sum_{i=1}^{n} \langle v, u_i \rangle \langle u_i, w \rangle$$

Proof of 3: writing $w = v$ on the left-hand side of Parseval's identity gives us $\langle v, v \rangle = \|v\|^2$. On the right-hand side, we have:

$$\sum_{i=1}^{n} \langle v, u_i \rangle \langle u_i, v \rangle = \sum_{i=1}^{n} \langle v, u_i \rangle \overline{\langle v, u_i \rangle} = \sum_{i=1}^{n} |\langle v, u_i \rangle|^2$$

hence $\|v\|^2 = \sum_{i=1}^{n} |\langle v, u_i \rangle|^2$. □

7 Marc-Antoine de Parseval des Chêsnes (1755, Rosières-aux-Salines-1836, Paris).

8 Michel Plancherel (1885, Bussy-1967, Zurich).

NOTE.–

1) The *physical interpretation of Plancherel's theorem* is as follows: the energy of v, measured as the square of the norm, can be decomposed using the sum of the squared moduli of each projection of v on the n directions of the orthonormal basis $(u_1, ..., u_n)$.

In Fourier theory, the directions of the orthonormal basis are fundamental harmonics (sines and cosines with defined frequencies): this is why Fourier analysis may be referred to as *harmonic analysis*.

2) If $(u_1, \dots, u_n)$ is an orthogonal, rather than an orthonormal, basis, then using the projector formula and Theorem 1.12, the results of Theorem 1.14 can be written as:

a) *decomposition of $v \in V$ on an orthogonal basis*:

$$v = \sum_{i=1}^{n} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i \qquad [1.10]$$

b) *Parseval's identity for an orthogonal basis*:

$$\langle v, w \rangle = \sum_{i=1}^{n} \frac{\langle v, u_i \rangle \langle u_i, w \rangle}{\|u_i\|^2} \qquad [1.11]$$

c) *Plancherel's theorem for an orthogonal basis*:

$$\|v\|^2 = \sum_{i=1}^{n} \frac{|\langle v, u_i \rangle|^2}{\|u_i\|^2} \qquad [1.12]$$

The following exercise is designed to test the reader's knowledge of the theory of finite-dimensional inner product spaces. The two subsequent exercises explicitly include inner products which are non-Euclidean.

Exercise 1.1

Consider the complex Euclidean inner product space $\mathbb{C}^3$ and the following three vectors:

$$u = (0, i, 2i), \quad v = (2i, 0, -i), \quad w = \left(0, i, \frac{1}{2} e^{-\frac{\pi}{2} i}\right)$$

1) Determine the orthogonality relationships between vectors u, v, w.

2) Calculate the norm of u, v, w and the Euclidean distances between them.

3) Verify that (u, v, w) is a (non-orthogonal) basis of $\mathbb{C}^3$.

4) Let S be the vector subspace of $\mathbb{C}^3$ generated by u and w. Calculate $P_S v$, the orthogonal projection of v onto S. Calculate $d(v, P_S v)$, that is, the Euclidean distance between v and its projection onto S, and verify that this minimizes the distance between v and the vectors of S (*hint*: look at the square of the distance).

5) Using the results of the previous questions, determine an orthogonal basis and an orthonormal basis for $\mathbb{C}^3$ without using the Gram-Schmidt orthonormalization process (*hint*: remember the geometric relationship between the residual vector r and the subspace S).

6) Given a vector $a = (2i, -1, 0)$, write the decomposition of a and Plancherel's theorem in relation to the orthonormal basis identified in point 5. Use these results to identify the vector from the orthonormal basis which has the heaviest weight in the decomposition of a (and which gives the best "rough approximation" of a). Use a graphics program to draw the progressive vector sum of a, beginning with the rough approximation and adding finer details supplied by the other vectors.

Solution to Exercise 1.1

1) Evidently, $e^{-\frac{\pi}{2}i} = -i$, so by directly calculating the inner products: $\langle u, v\rangle = -2$, $\langle u, w\rangle = 0$ et $\langle v, w\rangle = \frac{1}{2}$.

2) By direct calculation: $\|u\|^2 = 5$, $\|v\|^2 = 5$, $\|w\|^2 = \frac{5}{4}$. After calculating the difference vectors, we obtain: $d(u, v) = \|u - v\| = \sqrt{14}$, $d(u, w) = \|u - w\| = \frac{5}{2}$, $d(v, w) = \|v - w\| = \frac{\sqrt{21}}{2}$.

3) The three vectors u, v, w are linearly independent, so they form a basis in $\mathbb{C}^3$. This basis is not orthogonal since only vectors u and w are orthogonal.

4) $S = \text{span}(u, w)$. Since (u, w) is an orthogonal basis in S, we can write:

$$P_S(v) = \frac{\langle v, u\rangle}{\|u\|^2} u + \frac{\langle v, w\rangle}{\|w\|^2} w = (0, 0, -i)$$

The residual vector of the projection of v on S is $r = v - P_S v = (2i, 0, 0)$ and thus $d(v, P_S v)^2 = \|r\|^2 = 4$. The most general vector in S is $s = \alpha u + \beta w = (0, (\alpha + \beta)i, (2\alpha - \frac{\beta}{2})i)$ and $d(v, s)^2 = \|v - s\|^2 = 4 + (\alpha + \beta)^2 + (2\alpha - \frac{\beta}{2} + 1)^2 \geqslant 4 = d(v, P_S v)^2$. This confirms that $P_S v$ is the vector in S with the minimum distance from v in relation to the Euclidean norm.

5) r is orthogonal to S, which is generated by u and w, hence (u, w, r) is a set of orthogonal vectors in $\mathbb{C}^3$, that is, an orthogonal basis of $\mathbb{C}^3$. To obtain an orthonormal basis, we then simply divide each vector by its norm:

$$(\hat{u}, \hat{w}, \hat{r}) \equiv \left(\frac{u}{\|u\|}, \frac{w}{\|w\|}, \frac{r}{\|r\|}\right) = \left(\left(0, \frac{i}{\sqrt{5}}, \frac{2i}{\sqrt{5}}\right), \left(0, \frac{2i}{\sqrt{5}}, -\frac{i}{\sqrt{5}}\right), (i, 0, 0)\right)$$

6) Decomposition: $a = \langle a, \hat{u}\rangle\hat{u} + \langle a, \hat{w}\rangle\hat{w} + \langle a, \hat{r}\rangle\hat{r} = \frac{i}{\sqrt{5}}\hat{u} + \frac{2i}{\sqrt{5}}\hat{w} + 2\hat{r}$.

Plancherel's theorem: $\|a\|^2 = 5 = |\langle a, \hat{u}\rangle|^2 + |\langle a, \hat{w}\rangle|^2 + |\langle a, \hat{r}\rangle|^2 (= \frac{1}{5} + \frac{4}{5} + 4 = 5)$.

The vector with the heaviest weight in the reconstruction of a is thus $\hat{r}$: this vector gives the best rough approximation of a. By calculating the vector sum of this rough representation and the other two vectors, we can reconstruct the "fine details" of a, first with $\hat{w}$ and then with $\hat{u}$. □

Exercise 1.2

Let $M(n, \mathbb{C})$ be the space of $n \times n$ complex matrices. The application $\phi : M(n, \mathbb{C}) \times M(n, \mathbb{C}) \to \mathbb{C}$ is defined by:

$$\phi(A, B) = \text{tr}(B^\dagger A)$$

where $B^\dagger := \overline{B}^t$ denotes the adjoint matrix of B and tr is the matrix trace. Prove that ϕ is an inner product.

Solution to Exercise 1.2

The distributive property of matrix multiplication for addition and the linearity of the trace establishes the linearity of ϕ in relation to the first variable.

Now, let us prove that ϕ is Hermitian. Let $A = (a_{i,j})_{1\leqslant i,j\leqslant n}$ and $B = (b_{i,j})_{1\leqslant i,j\leqslant n}$ be two matrices in $M(n, \mathbb{C})$. Let $(c_{i,j})_{1\leqslant i,j\leqslant n} = (\overline{b_{j,i}})_{1\leqslant i,j\leqslant n}$ be the coefficients of the matrix $B^\dagger$ and let $(d_{i,j})_{1\leqslant i,j\leqslant n} = (\overline{a_{j,i}})_{1\leqslant i,j\leqslant n}$ be the coefficients of $A^\dagger$.

This gives us:

$$\overline{\phi(A, B)} = \overline{\text{tr}(B^\dagger A)} = \overline{\text{tr}\left[\left(\sum_{k=1}^{n} c_{i,k} a_{k,j}\right)_{i,j}\right]} = \overline{\text{tr}\left[\left(\sum_{k=1}^{n} \overline{b_{k,i}} a_{k,j}\right)_{i,j}\right]}$$

$$= \sum_{i,k=0}^{n} b_{k,i} \overline{a_{k,i}} = \sum_{i,k=0}^{n} d_{i,k} b_{k,i} = \text{tr}(A^\dagger B)$$

$$= \phi(B, A)$$

Thus, ϕ is a sesquilinear Hermitian form. Furthermore, ϕ is positive:

$$\forall A \in M(n, \mathbb{C}),\ \phi(A, A) = \sum_{i,k=0}^{n} |a_{k,i}|^2 \geqslant 0$$

It is also definite:

$$\begin{aligned}\phi(A,A)=0 &\iff \sum_{i,k=0}^{n} |a_{k,i}|^2 = 0\\ &\iff \forall 1 \leqslant k,i \leqslant n,\ a_{k,i}=0\\ &\iff A=0\end{aligned}$$

Thus, ϕ is an inner product. □

Exercise 1.3

Let $E = \mathbb{R}[X]$ be the vector space of single variable polynomials with real coefficients. For $P, Q \in E$, take:

$$\Phi(P,Q) = \int_{-1}^{1} \frac{P(t)Q(t)}{\sqrt{1-t^2}} dt$$

1) Remember that $f(t) \underset{t\to t_0}{=} O(g(t))$ means that $\exists\ a, C > 0$ such that $|t-t_0| < a \implies |f(t)| \leqslant C\,|g(t)|$. Prove that for all $P, Q \in E$, this is equal to:

$$\frac{P(t)Q(t)}{\sqrt{1-t^2}} \underset{t\to 1}{=} O\left(\frac{1}{\sqrt{1-t}}\right)$$

and:

$$\frac{P(t)Q(t)}{\sqrt{1-t^2}} \underset{t\to -1}{=} O\left(\frac{1}{\sqrt{1+t}}\right)$$

Use this result to deduce that Φ is definite over $E \times E$.

2) Prove that Φ is an inner product over E, which we shall note $\langle\ ,\ \rangle$.

3) For $n \in \mathbb{N}$, let T_n be the n-th *Chebyshev polynomial*, that is, the only polynomial such that $\forall\theta \in \mathbb{R}$, $T_n(\cos\theta) = \cos(n\theta)$. Applying the substitution $t = \cos\theta$, show that $(T_n)_{n\in\mathbb{N}}$ is an orthogonal family in E. Hint: use the trigonometric formula [1.13]:

$$\frac{1}{2}(\cos((n+m)\theta)+\cos((n-m)\theta)) = \cos(n\theta)\cos(m\theta) \qquad \forall n, m \in \mathbb{N}. \quad [1.13]$$

4) Prove that for all $n \in \mathbb{N}$, $(T_0, \ldots, T_n)$ is an orthogonal basis of $\mathbb{R}_n[X]$, the vector space of polynomials in $\mathbb{R}[X]$ of degree less than or equal to n. Deduce that $(T_n)_{n\in\mathbb{N}}$ is an orthogonal basis in the algebraic sense: every element in E is a *finite* linear combination of elements in the basis of E.

5) Calculate the norm of T_n for all n and deduce an orthonormal basis (in the algebraic sense) of E using this result.

Solution to Exercise 1.3

1) We write $f(t) = \frac{P(t)Q(t)}{\sqrt{1-t^2}} = \frac{P(t)Q(t)}{\sqrt{1-t}\sqrt{1+t}}$. Since P and Q are polynomials, the function $t \mapsto \frac{P(t)Q(t)}{\sqrt{1+t}}$ is continuous in a neighborhood $V_1(1)$ and thus, according to the Weierstrass theorem, it is bounded in this neighborhood, that is, $\exists\, C_1 > 0$ such that $t \in V_1(1) \implies \left|\frac{P(t)Q(t)}{\sqrt{1+t}}\right| \leqslant C_1$. Similarly, the function $t \mapsto \frac{P(t)Q(t)}{\sqrt{1-t}}$ is continuous in a neighborhood $V_2(-1)$, thus $\exists\, C_2 > 0$ such that $t \in V_2(-1) \implies \left|\frac{P(t)Q(t)}{\sqrt{1-t}}\right| \leqslant C_2$. This gives us:

$$t \in V_1(1) \implies |f(t)| \leqslant C_1 \frac{1}{\sqrt{1-t}} \iff f(t) \underset{t\to 1}{=} O\left(\frac{1}{\sqrt{1-t}}\right)$$

and:

$$t \in V_2(-1) \implies |f(t)| \leqslant C_2 \frac{1}{\sqrt{1+t}} \iff f(t) \underset{t\to -1}{=} O\left(\frac{1}{\sqrt{1+t}}\right)$$

This implies that the integral defining Φ is definite; $f(t)$ is continuous over $(-1,1)$ and therefore can be integrated. The result which we have just proved shows that $f(t)$ is integrable in a right neighborhood of –1 and a left neighborhood of 1, as the integral of its absolute value is incremented by an integrable function in both cases.

2) The bilinearity of Φ is obtained from the linearity of the integral using direct calculation. Its symmetry is a consequence of that of the dot product between functions. The only property which is not immediately evident is definite positiveness. Let us start by proving positiveness:

$$\Phi(P,P) = \int_{-1}^{1} \frac{P^2(t)}{\sqrt{1-t^2}} dt \geqslant 0$$

and[9]:

$$\Phi(P,P) = 0 \iff \frac{P^2(t)}{\sqrt{1-t^2}} dt = 0 \text{ a.e. on } (-1,1) \iff P(t) = 0 \text{ a.e. on } (-1,1)$$

but the only polynomial with an infinite number of roots is the null polynomial $\mathbf{0}(t) \equiv 0$, so $P = \mathbf{0}$. Φ is therefore an inner product on E.

9 a.e.: almost everywhere (see Chapter 3).

3) For all $n, m \in \mathbb{N}$:

$$\langle T_n, T_m\rangle = \int_{-1}^{1} \frac{T_n(t)T_m(t)}{\sqrt{1-t^2}}dt \qquad (t = \cos\theta,\ dt = -\sin\theta d\theta)$$

$$t = \cos\theta = -1 \iff \theta = \pi,\ t = \cos\theta = 1 \iff \theta = 0$$

$$\begin{aligned}
&= \int_{\pi}^{0} -\sin\theta \frac{T_n(\cos\theta)T_m(\cos\theta)}{\sqrt{1-\cos^2(\theta)}}d\theta \\
&= \int_0^{\pi} \sin\theta \frac{\cos(n\theta)\cos(m\theta)}{|\sin\theta|}d\theta \\
&= \int_0^{\pi} \cancel{\sin\theta} \frac{\cos(n\theta)\cos(m\theta)}{\cancel{\sin\theta}}d\theta \quad (\sin\theta \geqslant 0 \text{ on } [0,\pi]) \\
&= \int_0^{\pi} \cos(n\theta)\cos(m\theta)d\theta \\
&= \frac{1}{2}\left(\int_0^{\pi} \cos((n+m)\theta)d\theta + \int_0^{\pi} \cos((n-m)\theta)d\theta\right) \quad (\text{from } [1.13])
\end{aligned}$$

So, for all $n \neq m$, we have:

$$\langle T_n, T_m\rangle = \frac{1}{2}\left(\left[\frac{\sin((n+m)\theta)}{n+m}\right]_0^{\pi} + \left[\frac{\sin((n-m)\theta)}{n-m}\right]_0^{\pi}\right) = 0$$

that is, Chebyshev polynomials form an orthogonal family of polynomials in relation to the inner product defined above.

4) The family $(T_0, T_1, \ldots, T_n)$ is an orthogonal (and thus free) of $n+1$ elements of $\mathbb{R}[X]$, which is of dimension $n+1$, meaning that it is an orthogonal basis of $\mathbb{R}_n[X]$. To show that $(T_n)_{n\in\mathbb{N}}$ is a basis in the algebraic sense of E, consider a polynomial $P \in E$ of an arbitrary degree $d \in \mathbb{N}$, i.e. $P \in \mathbb{R}_d[X]$, and note that $(T_0, T_1, \ldots, T_d)$ is an orthogonal (free) family of generators of $\mathbb{R}_d[X]$, that is, a basis in the algebraic sense of the term.

5) The norm of T_n is calculated using the following equality:

$$\langle T_n, T_m\rangle = \frac{1}{2}\left(\int_0^{\pi} \cos((n+m)\theta)d\theta + \int_0^{\pi} \cos((n-m)\theta)d\theta\right)$$

which was demonstrated in point 3. Taking $n = m$, we have:

$$\|T_n\|^2 = \langle T_n, T_n\rangle = \frac{1}{2}\left(\int_0^{\pi} \cos(2n\theta)d\theta + \int_0^{\pi} d\theta\right) = \frac{1}{2}\left(\left[\frac{\sin 2n\theta}{2n}\right]_0^{\pi} + \pi\right) =$$

$$\|T_n\|^2 = \langle T_n, T_n\rangle = \tfrac{1}{2}\left(\int_0^{\pi} \cos(2n\theta)d\theta + \int_0^{\pi} d\theta\right) =$$

$$= \begin{cases} \frac{1}{2}\left(\int_0^{\pi} d\theta + \int_0^{\pi} d\theta\right) = \pi & \text{if } n = 0 \\ \frac{1}{2}\left(\left[\frac{\sin 2n\theta}{2n}\right]_0^{\pi} + \pi\right) = \frac{\pi}{2} & \text{if } n \geqslant 1, \end{cases}$$

hence $\|T_0\| = \sqrt{\pi}$ and $\|T_n\| = \sqrt{\pi/2}$ for $n \geqslant 1$. Finally, the family:

$$\left\{ \frac{T_0}{\sqrt{\pi}} \right\} \cup \left\{ \sqrt{\frac{2}{\pi}} T_n \right\}_{n \geqslant 1}$$

is an orthonormal basis of the vector space of first-order polynomials with real coefficients E. □

1.9. Summary

In this chapter, we have examined the properties of real and complex inner products, highlighting their differences. We noted that the symmetrical and bilinear properties of the real inner product must be replaced by conjugate symmetry and sesquilinearity in order to obtain a set of properties which are compatible with definite positivity. This final property is essential in order to produce a norm from a scalar product.

We noted that the prototype for all inner product spaces, or pre-Hilbert spaces, of finite dimension n is the Euclidean space $\mathbb{K}^n$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

Using the inner product, the concept of orthogonality between vectors can be extended to any inner product space. Two vectors are orthogonal if their inner product is null. The null vector is the only vector which is orthogonal to all other vectors, and the property of definite positiveness means that it is the only vector to be orthogonal to itself. If two vectors have the same inner product with all other vectors, that is, the same projection in every direction, then these vectors coincide.

A norm on a vector space is said to be a Hilbert norm if an inner product can be defined which generates the norm in a canonical manner. Remarkably, a norm is a Hilbert norm if and only if it satisfies the parallelogram law; this holds true for both finite and infinite dimensions. The polarization law can be used to define an inner product which is compatible with a Hilbert norm.

Vector orthogonality implies linear independence, guaranteeing that a set of n orthogonal vectors in a vector space of dimension n will constitute a basis. The expansion of a vector on an orthonormal basis is trivial: the components in relation to this basis are the inner products of the vector with the basis vectors. It is therefore much simpler to calculate components in such cases because, if the basis is not orthonormal, then a linear system of equations must be solved.

The concept of orthogonal projection on a vector subspace S was also presented. Given an orthogonal basis of this space, the projection can be represented as an expansion over the vectors of the basis, with coefficients given by the inner products (which are normalized if the basis is not orthonormalized). We have seen that the difference between a vector and its orthogonal projection, known as the residual vector, is orthogonal to the projection subspace S. We also demonstrated that the orthogonal projection is the vector in S which minimizes the distance (in relation to the norm of the vector space) between the vector and the vectors of S.

Given an inner product space, of finite or infinite dimensions, an orthonormal basis can always be defined using the Gram-Schmidt orthonormalization algorithm.

Finally, we proved the important Parseval identity and Plancherel's theorem in relation to an orthonormal or orthogonal basis. The extension of these properties to infinite dimensions is presented in Chapter 5.