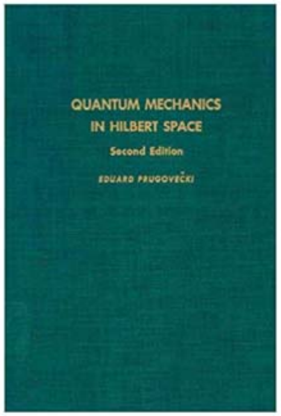
# Modern Algebra

**Chapter I. Basic Ideas of Hilbert Space Theory**

I.2. Euclidean (pre-Hilbert) Spaces—Proofs of Theorems








## Theorem I.2.1

**Theorem I.2.1.** In a Euclidean space $\mathcal{E}$, the inner product $\langle f \mid g \rangle$ satisfies the relations:

(a) $\langle af \mid g \rangle = a^* \langle f \mid g \rangle$, and

(b) $\langle f + g \mid h \rangle = \langle f \mid h \rangle + \langle g \mid h \rangle$

for all $f, g, h \in \mathcal{E}$ and for every scalar $a$.

**Proof.** We have

$$\begin{aligned} \langle af \mid g \rangle &= \langle g \mid af \rangle^* \text{ by Definition I.2.1(2)} \\ &= (a \langle g \mid f \rangle)^* \text{ by Definition I.2.1(3)} \\ &= a^* \langle g \mid f \rangle^* \text{ since } (z_1 z_2)* = z_1^* z_2^* \\ &= a^* \langle f \mid g \rangle \text{ by Definition I.2.1(2),} \end{aligned}$$

and. . .





## Theorem I.2.1 (continued)

**Theorem I.2.1.** In a Euclidean space $\mathcal{E}$, the inner product $\langle f \mid g \rangle$ satisfies the relations:

(a) $\langle af \mid g \rangle = a^* \langle f \mid g \rangle$, and

(b) $\langle f + g \mid h \rangle = \langle f \mid h \rangle + \langle g \mid h \rangle$

for all $f, g, h \in \mathcal{E}$ and for every scalar $a$.

**Proof.** . . .

$$\begin{aligned} \langle f + g \mid h \rangle &= \langle h \mid f + g \rangle^* \text{ by Definition I.2.1(2)} \\ &= (\langle h \mid f \rangle + \langle h \mid g \rangle)^* \text{ by Definition I.2.1(4)} \\ &= \langle h \mid f \rangle^* + \langle h \mid g \rangle^* \text{ since } (z_1 + z_2)* = z_1^* + z_2^* \\ &= \langle f \mid g \rangle + \langle g \mid h \rangle \text{ by Definition I.2.1(2).} \end{aligned}$$

□





## Theorem I.2.2

**Theorem I.2.2. Schwarz-Cauchy Inequality.**
Any two elements $f, g$ of a Euclidean space $\mathcal{E}$ satisfies

$$|\langle f \mid g \rangle|^2 \leq \langle f \mid f \rangle \langle g \mid g \rangle.$$

**Proof.** For any $f, g \in \mathcal{E}$ and any $a \in \mathbb{C}$ we have $\langle f + ag \mid f + ag \rangle \geq 0$ by Definition I.2.1(1). If $\langle f \mid g \rangle = 0$ then the result holds (again, by Definition I.2.1(1)), so we can assume without loss of generality that $\langle f \mid g \rangle \neq 0$. Let $a = \lambda \langle f \mid g \rangle^* / |\langle f \mid g \rangle|$ where $\lambda \in \mathbb{R}$. Then $\langle f + ag \mid f + ag \rangle \geq 0$ implies

$$\begin{aligned} \langle f + ag \mid f + ag \rangle &= \langle f + ag \mid f \rangle_{(} f + ag \mid ag \rangle \text{ by Definition I.2.1(4)} \\ &= \langle f \mid f \rangle + \langle ag \mid f \rangle + \langle f \mid ag \rangle + \langle ag \mid ag \rangle \\ &\quad \text{by Theorem I.2.1(b)} \\ &= a^* a \langle g \mid g \rangle + \langle f \mid ag \rangle^* + \langle f \mid ag \rangle + \langle f \mid f \rangle \\ &\quad \text{by Definition I.2.1(2)} \ldots \end{aligned}$$

## Theorem I.2.2 (continued 1)

**Proof (continued).** ...

$$\begin{aligned} \langle f + ag \mid f + ag \rangle &= |a|^2 \langle g \mid g \rangle + a^* \langle f \mid g \rangle^* + a \langle f \mid g \rangle + \langle f \mid f \rangle \\ &\quad \text{by Definition I.2.1(3)} \\ &= \left| \lambda \frac{\langle f \mid g \rangle^*}{|\langle f \mid g \rangle|} \right|^2 \langle g \mid g \rangle + \lambda \frac{\langle f \mid g \rangle}{|\langle f \mid g \rangle|} \langle f \mid g \rangle^* \\ &\quad + \lambda \frac{\langle f \mid g \rangle^*}{|\langle f \mid g \rangle|} \langle f \mid g \rangle + \langle f \mid \rangle \\ &= \lambda^2 \langle g \mid g \rangle + 2\lambda |\langle f \mid g \rangle| + \langle f \mid \rangle \\ &\quad \text{since } |z^*| = |z| \text{ and } z^* z = |z|^2 \\ &\geq 0. \end{aligned}$$

Define polynomial of real variable

$$p(\lambda) = \lambda^2 + \langle g \mid g \rangle + 2\lambda |\lambda f \mid g \rangle| + \langle f \mid f \rangle.$$





## Theorem I.2.2 (continued 2)

**Theorem I.2.2. Schwarz-Cauchy Inequality.**
Any two elements $f, g$ of a Euclidean space $\mathcal{E}$ satisfies

$$|\langle f \mid g \rangle|^2 \leq \langle f \mid f \rangle \langle g \mid g \rangle.$$

**Proof (continued).** Then $p(\lambda)$ is a second degree nonnegative concave up polynomial and so it must have at most one root. This means the discriminant from the quadratic equation must no non-positive. So we need $(2|\langle f \mid g \rangle|)^2 - 4(\langle g \mid g \rangle)(\langle f \mid f \rangle) \leq 0$ or $|\langle f \mid g \rangle|^2 \leq \langle f \mid f \rangle \langle g \mid g \rangle$, as claimed. □





## Theorem I.2.3

**Theorem I.2.3.** In a Euclidean space $\mathcal{E}$ with inner product $\langle f \mid g \rangle$, the real-valued function $\|f\| = \sqrt{\langle f \mid f \rangle}$ is a norm.

**Proof.** We check the four parts of Definition I.2.2.

**(1)** If $f \neq \mathbf{0}$ then $\langle f \mid f \rangle > 0$ by Definition I.2.1(1), so $\|f\| = \sqrt{\langle f \mid f \rangle} > 0$ for $f \neq \mathbf{0}$.

**(2)** $\langle \mathbf{0} \mid \mathbf{0} \rangle = \langle \mathbf{0} \mid \mathbf{0} + \mathbf{0} \rangle = \langle \mathbf{0} \mid \mathbf{0} + \langle \mathbf{0} \mid \mathbf{0} \rangle$ by Definition I.2.1(4), so that $\langle \mathbf{0} \mid \mathbf{0} \rangle = 0$ and $\|\mathbf{0}\| = \sqrt{\langle \mathbf{0} \mid \mathbf{0} \rangle} = 0$.

**(3)**

$$\begin{aligned} \|af\| &= \sqrt{\langle af \mid af \rangle} = \sqrt{a^* a \langle f \mid f \rangle} \text{ by Definition I.2.1(3)} \\ &\quad \text{and Theorem I.2.1(a)} \\ &= \sqrt{a^* a} \sqrt{\langle f \mid f \rangle} = |a| \|f\| \text{ since } |z|^2 = z^* z. \end{aligned}$$





## Theorem I.2.3 (continued)

**Proof (continued). (4)** For the Triangle Inequality, we have

$$\begin{aligned} \|f + g\|^2 &= \langle f + g \mid f + g \rangle \\ &= \langle f \mid f \rangle + \langle f \mid g \rangle + \langle g \mid f \rangle + \langle g \mid g \rangle \\ &\quad \text{by Definition I.2.1(4) and Theorem I.2.1(b)} \\ &= \|f\|^2 + \langle f \mid g \rangle + \langle f \mid g \rangle^* + \langle g \mid g \rangle \text{ by Definition I.2.1(2)} \\ &= \|f\|^2 + 2\mathrm{Re}(\langle f \mid g \rangle) + \|g\|^2 \text{ since } z + z^* = 2\mathrm{Re}(z) \\ &\leq \|f\|^2 + r|\mathrm{Re}(\langle f \mid g \rangle)| + \|g\|^2 \\ &\leq \|f\|^2 + 2|\langle f \mid g \rangle| + \|g\|^2 \text{ since } |\mathrm{Re}(z)| \leq |z| \\ &\leq \|f\|^2 + 2\sqrt{\langle f \mid f \rangle \langle g \mid g \rangle} + \|g\|^2 \\ &\quad \text{by Schwarz-Cauchy Inequality (Theorem I.2.2)} \\ &= \|f\|^2 + 2\|f\|\|g\| + \|g\|^2 = (\|f\| + \|g\|)^2, \end{aligned}$$

and so $\|f + g\| \leq \|f\| + \|g\|$. □

## Theorem I.2.4

**Theorem I.2.4.** If $S$ is a finite of countably infinite set of vectors in a Euclidean space $\mathcal{E}$ and $\mathcal{V}$ is the vector subspace of $\mathcal{E}$ spanned by $S$, then there is an orthonormal system $T$ of vectors which spans $\mathcal{V}$; that is, for which $\text{span}(T) = \mathcal{V}$ (that is, the set of all linear combinations of elements of $T$; Prugovečki denotes the space of $T$ as $(T)$). $T$ is a finite set when $S$ is a finite set.

**Proof.** Let $S = \{f_1 f_2, \ldots\}$. Define $g_1$ to be the first nonzero vector in $S$. Then recursively define $g_n$ for $n \geq 2$ as $g_n = f_m$ where $m$ is the minimum index such that $\{g_1, g_2, \ldots, g_{n-1}, f_n\}$ is linearly independent; if no such $m$ exists, take $S_0 = \{g_1, g_2, \ldots, g_{n-1}\}$ (so if $S_0$ is finite then $T$finite). Then $S_0 = \{g_1, g_2, \ldots\}$. Notice that $S_0$ is linearly independent and $\text{span}(S_0) = \text{span}(S)$. We now create orthonormal set $T$ from $S_0$ using the Gram-Schmidt process (or the "Schmidt orthonormalization procedure").





## Theorem I.2.4 (continued 1)

**Proof (continued).** Since $g \neq \mathbf{0}$, let $e_1 = g_1/\|g_1\|$. Then recursively define $e_n$ for $n \geq 2$ in terms of $g_n$ and $e_1, e_2, \ldots, e_{n-1}$ as

$$e_n = \frac{g_n - \langle e_{n-1} \mid g_n\rangle e_{n-1} - \langle e_{n-2} \mid g_n\rangle e_{n-2} - \cdots - \langle e_1 \mid g_n\rangle e_1}{\|g_n - \langle e_{n-1} \mid g_n\rangle e_{n-1} - \langle e_{n-2} \mid g_n\rangle e_{n-2} - \cdots - \langle e_1 \mid g_n\rangle e_1\|}.$$

We now show the denominator here is nonzero. ASSUME

$$g_n - \langle e_{n-1} \mid g_n\rangle e_{n-1} - \langle e_{n-2} \mid g_n\rangle e_{n-2} - \cdots - \langle e_1 \mid g_n\rangle e_1 = \mathbf{0}.$$

We have such $g_n$ expanded as a linear combination of $e_1, e_2, \ldots, e_k$ (so $\text{span}(g_1, g_2, \ldots, g_k) \subset \text{span}(e_1, e_2, \ldots, e_k)$) and so we can solve for $g_k$ as follows (where $c_\ell$, are scalars):

$$\begin{aligned} g_1 &= c_{11} e_1 \\ g_2 &= c_{21} e_1 + c_{22} e_2 \\ &\vdots \\ g_{n-1} &= c_{n-1,1} e_1 + c_{n-1,2} e_2 + \cdots + c_{n-1,n-1} e_{n-1}. \end{aligned}$$





## Theorem I.2.4 (continued 2)

**Proof (continued).** We can inductively solve this system of equations for $e_1$ in terms of $g_1$, solve for $e_2$ in terms of $g_1$ and $g_2$, ..., solve for $e_{n-1}$ in terms of $g_1, g_2, \ldots, g_{n-1}$ (this is sometimes called "back substitution"). This implies $\text{span}(e_1, e_2, \ldots, e_k) \subset \text{span}(g_1, g_2, \ldots, g_k)$ and hence $\text{span}(e_1, e_2, \ldots, e_k) = \text{span}(g_1, g_2, \ldots, g_k)$ and $\text{span}(T) = \text{span}(S_0) = \text{span}(S)$. If the assumed equality holds then $g_n$ is a linear combination of $e_1, e_2, \ldots, e_{n-1}$, but then we can substitute for $e_i$ its expression in terms of $g_1, g_2, \ldots, g_i$ (for $1 \leq i \leq n-1$) and then we can express $g_n$ as a linear combination of $g_1, g_2, \ldots, g_{n-1}$, CONTRADICTING the choice of $g_n$. So the denominator in the formula for $e_n$ is not 0 and $e_n$ is defined.
The vectors in $I = \{e_1, e_2, \ldots\}$ are unit vectors (i.e., normalized). We now show orthogonality. Let $\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$. We give an induction argument.





## Theorem I.2.4 (continued 3)

**Proof (continued).** Suppose we have shown that $\langle e_i \mid e_j\rangle = \delta_{ij}$ for $i, j = 1, 2, \ldots, n-1$. Then for $m < n$

$$\begin{aligned} \langle e_m \mid e_n\rangle &= \left\langle e_m \,\middle|\, \frac{g_n - \langle e_{n-1} \mid g_n\rangle e_{n-1} - \cdots - \langle e_1 \mid g_n\rangle e_1}{\|g_n - \langle e_{n-1} \mid g_n\rangle e_{n-1} - \cdots - \langle e_1 \mid g_n\rangle e_1\|} \right\rangle \\ &= \frac{1}{\|g_n - \langle e_{n-1} \mid g_n\rangle e_{n-1} - \cdots - \langle e_1 \mid g_n\rangle e_1\|} \times \\ &\quad \left( \langle e_m \mid g_n\rangle - \sum_{k=1}^{n-1} \langle e_k \mid g_n\rangle \langle e_m \mid e_k\rangle \right) \\ &= \frac{1}{\|g_n - \langle e_{n-1} \mid g_n\rangle e_{n-1} - \cdots - \langle e_1 \mid g_n\rangle e_1\|} \times \\ &\quad \left( \langle e_m \mid g_n\rangle - \sum_{k=1}^{n-1} \langle e_k \mid g_n\rangle \delta_{mk} \right) = 0. \end{aligned}$$

## Theorem I.2.4 (continued 4)

**Theorem I.2.4.** If $S$ is a finite of countably infinite set of vectors in a Euclidean space $\mathcal{E}$ and $\mathcal{V}$ is the vector subspace of $\mathcal{E}$ spanned by $S$, then there is an orthonormal system $T$ of vectors which spans $\mathcal{V}$; that is, for which $\text{span}(T) = \mathcal{V}$ (that is, the set of all linear combinations of elements of $T$; Prugovečki denotes the space of $T$ as $(T)$). $T$ is a finite set when $S$ is a finite set.

**Proof (continued).** So $\langle e_i \mid e_j \rangle = \delta_{ij}$ for $i, j = 1, 2, \ldots, n-1, n$ and by induction $T$ is orthonormal. □





## Theorem I.2.5

**Theorem I.2.5.** All complex Euclidean $n$-dimensional spaces are isomorphic to $\ell^2(n)$ and consequently mutually isomorphic.

**Proof.** If $\mathcal{E}$ is an $n$-dimensional Euclidean space then there is, by Theorem I.1.2, a set of $n$ vectors $f_1, f_2, \ldots f_n$ spanning $\mathcal{E}$. By Theorem I.2.4, there is an orthonormal system of $n$ vectors $e_1, e_2, \ldots, e_n$ which also spaces $\mathcal{E}$. Consider the mapping $f \mapsto [\langle e_1 \mid r\rangle, \langle e_2 \mid f\rangle, \ldots, \langle e_n \mid f\rangle]^T$. It is to be shown that this mapping is an isomorphism between $\mathcal{E}$ and $\ell^2(n)$ in Exercise I.2.7. □