Lecture 1: Hilbert Space

Tianpei Xie

Nov. 18th., 2022

Contents

1 Metric Space

1.1 Basics

- **Definition** A ***metric space*** is a set M and a real-valued function $d(\cdot,\cdot) : M \times M \to \mathbb{R}$ which satisfies:

  1. (***Non-Negativity***) $d(x,y) \geq 0$
  2. (***Definiteness***) $d(x,y) = 0$ if and only if $x = y$
  3. (***Symmetric***) $d(x,y) = d(y,x)$
  4. (***Triangle Inequality***) $d(x,z) \leq d(x,y) + d(y,z)$

  The function d is called a ***metric*** on M. The metric space M equipped with metric d is denoted as (M,d).

- **Definition** (***Cauchy Sequence***)
  A sequence of elements $\{x_n\}$ of a metric space (M,d) is called a ***Cauchy sequence*** if $\forall \epsilon > 0$, there exists $N \in \mathbb{N}$, for all $n,m \geq N$, $d(x_n, x_m) < \epsilon$.

- **Proposition 1.1** *Any convergent sequence is a Cauchy sequence.*

  Note that this is the direct result of *triangle inequality property of a metric.*

- **Definition** (***Complete Metric Space***)
  A metric space in which ***all Cauchy sequences converge*** is called ***complete***.

- **Remark** In complete metric space, one can prove convergence *without knowing what point the sequence converges to.*

- **Example** The space of *all absolutely integrable functions* $\mathcal{L}^1(X,\mu)$ is ***complete***.

- **Definition** (***Denseness***)
  A set B in a *metric space* M is called ***dense*** if *every* $m \in M$ *is a limit of elements in* B.

- **Definition** (***Continuity***)
  A function $f : (X,d) \to (Y,p)$ is called ***continuous*** at x if $f(x_n) \xrightarrow{p} f(x)$ whenever $x_n \xrightarrow{d} x$.

- **Definition** (***Isometry***)
  A ***bijection*** $h : (X,d) \to (Y,p)$ which ***preserves*** *the metric*, that is,

  $$p(h(x), h(y)) = d(x,y)$$

  is called an ***isometry***. It is automatically *continuous*. (X,d) and (Y,p) are said to be ***isometric*** if such an isometry exists.

1.2 Equicontinuity

- **Definition** (***Equicontinuity***) [Reed and Simon, 1980]
  Let $\mathscr{F}$ be a family of functions from a metric space (X,p) to another metric space (Y,d). We say $\mathscr{F}$ is an ***equicontinuous family*** if and only if for all $\epsilon > 0$ and all $x \in X$, there exists $\delta > 0$ such that $d(f(x), f(x')) < \epsilon$ whenever $p(x,x') < \delta$ ***for every*** $f \in \mathscr{F}$ and all $x' \in X$.

We say $\mathscr{F}$ is a ***uniformly equicontinuous family*** if and only if for all $\epsilon > 0$, there exists $\delta > 0$ such that $d(f(x), f(x')) < \epsilon$ whenever $p(x, x') < \delta$ for all $x, x' \in X$ and ***every*** $f \in \mathscr{F}$.

- **Remark** An *equicontinuous family* of functions is *a family of continuous functions.*
- **Remark** The concept of ***equicontinuity*** is with respect to ***a family of functions***, while the concept of *continuity* is *for one fixed function.* In other word, for continuous function f, the radius of input $\delta := \delta(\epsilon, x, f)$ depends on threshold ϵ, the point of continuity x and the function of concern f. But for an ***equicontinuous family***, $\delta := \delta(\epsilon, x)$ does not depends on which function $f \in \mathscr{F}$. For a ***uniform equicontinuous family***, $\delta := \delta(\epsilon)$ does not depends on which function $f \in \mathscr{F}$ and which point x for continuity.
- **Remark** We can control the behavior of $\lim_{n\to\infty} f_n(x)$ in two ways [Reed and Simon, 1980]:
  1. ***Control its dependence on*** x: If the *convergence* of $\{f_n(x)\}$ *does not depends on the* ***choice of*** x, we have ***uniform convergence***. ***Otherwise***, we have ***pointwise convergence***.
  2. ***Control its dependence on*** n: If the *convergence* of $\{f_n(x)\}$ *does not depends on* ***choice of function*** f_n, we have an ***equicontinuous family*** $\{f_n\}$. This time it reveals the behavior of x in the limit. What we will see is that one can obtain not only information about the x behavior of the limit but that one can also turn weak information about the approach to the limit into stronger information.
- **Proposition 1.2** *Let f_n be a sequence of functions from one metric space to another with the property that the family $\{f_n\}$ is* ***equicontinuous****. Suppose that $f_n(x) \to f(x)$* ***pointwise*** *for each x. Then f is* ***continuous****.*
- We see that ***pointwise convergence*** on a ***dense set*** combined with ***equicontinuity*** implies ***pointwise convergence everywhere***.

  Proposition 1.3 *[Reed and Simon, 1980]*
  Let $\{f_n\}$ be an ***equicontinuous family*** *of functions from one metric space (X, p) to another (Y, d) with Y complete. Suppose that for a* ***dense*** *set $D \subseteq X$, we know $f_n(x)$ converges for all $x \in D$. Then $f_n(x)$ converges for all $x \in X$.*
- The following shows that uniformly equicontinuous combined with pointwise convergence implies uniform convergence.

  Proposition 1.4 *[Reed and Simon, 1980]*
  Let $\{f_n\}$ be a ***uniformly equicontinuous family*** *of functions on $[0, 1]$. Suppose that $f_n(x) \to f(x)$ for each x in $[0, 1]$. Then $f_n(x) \to f(x)$* ***uniformly*** *in x.*
- **Remark** For functions on $[0, 1]$, *every* ***equicontinuous family*** is ***uniformly equicontinuous***.
- **Theorem 1.5** ***(Ascoli's Theorem)*** *[Reed and Simon, 1980]*
  Let $\{f_n\}$ be a family of ***uniformly bounded equicontinuous functions*** *on $[0, 1]$. Then* ***some subsequence*** *$\{f_{n,m}\}$ converges* ***uniformly*** *on $[0, 1]$.*

1.3 Proof of Completeness

- **Remark** To prove completeness, we take an arbitrary ***Cauchy sequence*** (x_n) in X and show that *it converges in* X. For different spaces, such proofs may vary in complexity, but they have approximately the same general pattern:

  1. ***Construct an element*** x (to be used as a ***limit***).
  2. ***Prove that*** x ***is in the space considered.***
  3. ***Prove convergence*** $x_n \to x$ *(in the sense of the* ***metric****)*

- **Proposition 1.6** ***(Completeness of*** ℓ^∞***)***
  The space ℓ^∞ *is complete.*

  Proof: Let (x_m) be any Cauchy sequence in the space ℓ^∞, where $x_m = (\xi_1^{(m)}, \xi_2^{(m)}, \ldots)$. Since the metric on ℓ^∞ is given by

  $$d(x, y) = \sup_j |\xi_j - \eta_j|$$

  where $x = (\xi_j)$ and $y = (\eta_j)$ and (x_m) is Cauchy, for any $\epsilon > 0$ there is an N such that for all $m, n > N$,

  $$d(x_m, x_n) = \sup_j \left|\xi_j^{(m)} - \xi_j^{(n)}\right| < \epsilon.$$

  A fortiori, for every fixed j,

  $$\left|\xi_j^{(m)} - \xi_j^{(n)}\right| < \epsilon, \quad (m, n > N). \tag{1}$$

  Hence for every fixed j, the sequence $(\xi_j^{(1)}, \xi_j^{(2)}, \ldots)$ is a *Cauchy sequence* of numbers. It converges, say, $\xi_j^{(m)} \to \xi_j$ as $m \to \infty$. Using these infinitely many limits $\xi_1, \xi_2, \ldots$, we define $x = (\xi_1, \xi_2, \ldots)$ and show that $x \in \ell^\infty$ and $x_m \to x$. From (1) with $n \to \infty$ we have

  $$\left|\xi_j^{(m)} - \xi_j\right| \le \epsilon, \quad (m > N). \tag{2}$$

  Since $x_m = (\xi_j^{(m)}) \in \ell^\infty$, there is a real number k_m such that $\left|\xi_j^{(m)}\right| \le k_m$ for all j. Hence by the triangle inequality

  $$|\xi_j| \le \left|\xi_j - \xi_j^{(m)}\right| + \left|\xi_j^{(m)}\right| \le \epsilon + k_m \quad (m > N).$$

  This inequality holds for every j, and the right-hand side does not involve j. Hence (ξ_j) is a *bounded* sequence of numbers. This implies that $x = (\xi_j) \in \ell^\infty$. Also, from (2) we obtain

  $$d(x_m, x) = \sup_j \left|\xi_j^{(m)} - \xi_j\right| \le \epsilon, \quad (m, n > N).$$

  This shows that $x_m \to x$. Since (x_m) was an arbitrary Cauchy sequence, ℓ^∞ is complete. ∎

- **Proposition 1.7** ***(Completeness of*** $\mathcal{C}^0[a, b]$***)***
  The space of all continuous function under supremum norm on $[a, b]$ *(, denoted as* $\mathcal{C}^0[a, b]$*) is complete.*

Proof: Let (x_m) be any Cauchy sequence in $\mathcal{C}^0[a,b]$. Then, given any $\epsilon > 0$, there is an N such that for all $m, n > N$ we have

$$d(x_m, x_n) = \max_{t \in J} |x_m(t) - x_n(t)| \le \epsilon \tag{3}$$

where $J = [a,b]$. Hence for any fixed $t = t_0 \in J$,

$$|x_m(t_0) - x_n(t_0)| \le \epsilon, \quad (m, n > N).$$

This shows that $(x_1(t_0), x_2(t_0), \ldots)$ is a Cauchy sequence of real numbers. Since $\mathbb{R}$ is complete, the sequence converges, say, $x_m(t_0) \to x(t_0)$ as $m \to \infty$. In this way we can associate with each $t \in J$ a unique real number $x(t)$. This defines (*pointwise*) a function x on J, and we show that $x \in \mathcal{C}^0[a,b]$ and $x_m \to x$. From (3) with $n \to \infty$ we have

$$\max_{t \in J} |x_m(t) - x(t)| \le \epsilon, \quad (m > N).$$

Hence for every $t \in J$,

$$|x_m(t) - x(t)| \le \epsilon, \quad (m > N).$$

This shows that $(x_m(t))$ converges to $x(t)$ *uniformly* on J. Since the x_m's are *continuous* on J and the convergence is *uniform*, the limit function x is *continuous* on J, as is well known from calculus. Hence $x \in \mathcal{C}^0[a,b]$. Also $x_m \to x$. This proves completeness of $\mathcal{C}^0[a,b]$. ∎

- **Theorem 1.8** ***(Riesz-Fisher Theorem)*** *[Reed and Simon, 1980]*
  Let $\mathcal{L}^1(X, \mu)$ be the space of all absolutely integrable functions on measure space $(X, \mathscr{B}, \mu)$ with $\mathcal{L}^1$ norm. $\mathcal{L}^1(X, \mu)$ is complete.

  Proof: Let (f_n) be Cauchy sequence in $\mathcal{L}^1$. It is enough to prove some ***subsequence converges*** (i.e. "*Cauchy sequence in a metric space is convergent if and only if it has convergent sub-sequence*".) so pass to a subsequence (also labeled f_n with $\|f_n - f_{n+1}\|_{\mathcal{L}^1} \le 2^{-n}$). Let

  $$g_m(x) = \sum_{n=1}^{m} |f_n(x) - f_{n+1}(x)|$$

  Let g_∞ be the *infinite sum* (which may be ∞). Then $g_m \le g_{m+1} \nearrow g_\infty$ and

  $$\int_X |g_m(x)|\, d\mu = \int_X \left| \sum_{n=1}^{m} |f_n(x) - f_{n+1}(x)| \right| d\mu \le \int_X \sum_{n=1}^{m} |f_n(x) - f_{n+1}(x)|\, d\mu = \sum_{n=1}^{m} \|f_n - f_{n+1}\|_{\mathcal{L}^1} \le 1,$$

  so by *the monotone convergence theorem*, $g_\infty \in \mathcal{L}^1(X, \mu)$. Thus $|g_\infty(x)| < \infty$ a.e. As a result

  $$f_m(x) = f_1(x) - \sum_{n=1}^{m-1} (f_n(x) - f_{n+1}(x))$$

  converges pointwise a.e. to a function $f(x)$. Moreover,

  $$|f_m(x)| \le |f_1(x)| + |g_\infty(x)| < \infty$$

  i.e. $f_m(x) \in \mathcal{L}^1(X, \mu)$, so $f_n \to f$ in $\mathcal{L}^1$ norm by *the dominated convergence theorem.* ∎

- **Proposition 1.9** *$\mathcal{C}^0[a,b]$ is* ***dense*** *(w.r.t. $\|\|_1$) in $\mathcal{L}^1([a,b])$. Thus $\mathcal{L}^1([a,b])$ is the* ***completion*** *of $\mathcal{C}^0[a,b]$ w.r.t. $\mathcal{L}^1$ norm.*

2 Hilbert Space

- **Remark** (***Hilbert Space vs. Banach Space***)
  Hilbert space is a special Banach space equipped with inner product. Historically, Hilbert space appears eariler. The theory of inner product and Hilbert spaces is richer than that of general normed and Banach spaces. ***Distinguishing features*** are

  1. ***representations*** of $\mathcal{H}$ as a ***direct sum*** of a ***closed subspace*** and its ***orthogonal complement*** (section 2.3),
  2. ***orthonormal sets*** and sequences and corresponding ***representations*** of elements of $\mathcal{H}$ (section 2.5),
  3. ***the Riesz representation*** of ***bounded linear functionals*** by inner products, (section 2.4)
  4. ***the Hilbert-adjoint operator*** T^* of a bounded linear operator T (section 2.10).

2.1 Inner Product Space

- **Remark** Finite-dimensional vector spaces have *three kinds of properties* whose generalizations we will study in the next four chapters:

  1. ***linear*** *properties*,
  2. ***metric*** *properties*,
  3. and ***geometric*** *properties*.

  A ***Hilbert space*** generalizes the ***geometric*** property of a finite-dimensional vector space to *infinite-dimensional* via definition of inner product.

- **Definition** A *complex vector space* V is called ***an inner product space*** if there is a *complex-valued function* $\langle \cdot \, , \, \cdot \rangle : V \times V \to \mathbb{C}$ that satisfies the following four conditions for an $x, y, z \in V$ and $a, b \in \mathbb{C}$:

  1. (***Positive Definiteness***): $\langle x \, , \, x \rangle \geq 0$ and $\langle x \, , \, x \rangle = 0$ if and only if $x = 0$
  2. (***Linearity***): $\langle a\, x + b\, y \, , \, z \rangle = a \, \langle x \, , \, z \rangle + b \, \langle y \, , \, z \rangle$
  3. (***Hermitian***): $\langle x \, , \, y \rangle = \overline{\langle y \, , \, x \rangle}$

  The function $\langle \cdot \, , \, \cdot \rangle$ is called ***an inner product***.

- **Remark** Without "condition $\langle x \, , \, x \rangle = 0$ if and only if $x = 0$", we have ***semi-inner product*** [Conway, 2019].
- **Remark** From *Hermitian property*, we have $\langle x \, , \, a\, y + b\, z \rangle = \overline{a} \, \langle x \, , \, y \rangle + \overline{b} \, \langle x \, , \, z \rangle$.
- **Remark** For *real vector space*, an inner product is a ***symmetric covariant*** 2***-tensor***, or a ***symmetric bilinear form***.
- **Remark** Some books [Reed and Simon, 1980] define inner product via ***linearity in second argument***; while others [Kreyszig, 1989, Luenberger, 1997, Conway, 2019] defines it in terms

of ***linearity in first argument***. The difference is the position of conjugate.

- **Proposition 2.1** *Every **inner product space** V is a **normed linear space** with the norm $\|x\| = \sqrt{\langle x\,,\,x\rangle}$.*

- **Remark** We denote $\|x\| = \sqrt{\langle x\,,\,x\rangle}$ as *the* ***length*** of a vector. With the definition of length, we can define the ***distance*** d as

$$d(x,y) := \|x-y\| = \sqrt{\langle x-y\,,\,x-y\rangle}.$$

  As a consequence of *the Pythagorean Theorem*, d satisfies the triangle inequality so it is a *metric*. Thus ***every inner product space is a metric space***.

- **Proposition 2.2** ***(Parallelogram Law)***
  For any $x,y \in (V, \langle\cdot\,,\,\cdot\rangle)$,

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2$$

- **Remark** The followings are other versions of ***Parallelogram Law***:

$$\begin{aligned}
\Re\langle x\,,\,y\rangle &= \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right)\\
\Re\langle x\,,\,y\rangle &= \frac{1}{2}\left(\|x\|^2 + \|y\|^2 - \|x-y\|^2\right)\\
\Re\langle x\,,\,y\rangle &= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)\\
\langle x\,,\,y\rangle &= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2\right)\\
&= \Re\langle x\,,\,y\rangle + i\Re\langle x\,,\,iy\rangle
\end{aligned}$$

- The converse holds true as well.

  Proposition 2.3 *In a **normed space** $(V, \|\cdot\|)$, if **the parallelogram law***

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2$$

  *holds, then there exists a **unique inner product** $\langle\cdot,\ \cdot\rangle$ on V such that $\|x\| = \sqrt{\langle x\,,\,x\rangle}$ for all $x \in V$.*

- **Remark** The inner product defines the concept of ***angle*** (and *orthorgonality*), and ***distance***. Hence it allows the ***geometric property*** of Euclidean space to be generalized.

- **Definition** Two vectors, x and y, in an inner product space V are said to be ***orthogonal*** if $\langle x\,,\,y\rangle = 0$. A collection $\{x_n\}$ of vectors in V is called ***an orthonormal set*** if $\langle x_i\,,\,x_i\rangle = 1$ for all i, and $\langle x_i\,,\,x_j\rangle = 0$ if $i \neq j$.

- **Theorem 2.4** ***(Pythagorean Theorem)***
  *Let $\{x_i\}_{i=1}^n$ be an **orthonormal** set in an inner product space V. Then for all $x \in V$,*

$$\|x\|^2 = \sum_{i=1}^n |\langle x_i\,,\,x\rangle|^2 + \left\|x - \sum_{i=1}^n \langle x_i\,,\,x\rangle\, x_i\right\|^2$$

- **Corollary 2.5** ***(Bessel's inequality)***
  *Let $\{x_i\}_{i=1}^n$ be an **orthonormal** set in an inner product space V. Then for all $x \in V$,*
  $$\|x\|^2 \geq \sum_{i=1}^{n} |\langle x_i\,,\, x\rangle|^2$$

- **Corollary 2.6** ***(Cauchy-Schwartz's inequality)***
  Let V be an inner product space. For $x, y \in V$,
  $$|\langle x\,,\, y\rangle| \leq \|x\|\ \|y\|\,.$$

2.2 Hilbert Space

- **Definition** A ***complete*** *inner product space* is called ***a Hilbert space***.

  Inner product spaces are sometimes called ***pre-Hilbert spaces***.

- **Definition** Two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ are said to be ***isomorphic*** if there is a ***surjective linear*** *operator* $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that
  $$\langle Ux\,,\, Uy\rangle_{\mathcal{H}_2} = \langle x\,,\, y\rangle_{\mathcal{H}_1}$$
  for all $x, y \in \mathcal{H}_1$. Such an operator is called ***unitary***.

- **Remark** (***Isomorphism***)
  For vector space, an ***(linear) isomorphism*** is a ***bijective linear mapping*** from *one vector spaces* to *another vector space* that ***preserve*** the ***structure*** of that vector space. However, depending on definition of specific structure, we can have various different definition of isomorphisms:

  1. For ***metric space***, an *isomorphism* is a *bijective linear operator* that ***preserves the metric***. It is often called an ***isometry***.

  2. For ***inner product space***, an *isomorphism* is a *surjective linear operator* that ***preserves the inner product***. It is often called an ***surjective isometry***.

  3. For ***linear algebra***, an *isomorphism* is a ***bijective linear mapping*** that ***preserves all algebraic operations*** (i.e. *the vector addition* and *scalar multiplication*).

  In general, ***isomorphism*** is a ***structure-preserving mapping*** between two structures of the same type that *can be **reversed*** by ***an inverse mapping***. It means that "***two spaces are essentially of the same form***". For instance, the followings are also called ***isomorphism*** depending on the context:

  1. ***homemorphism*** between *topological spaces*,

  2. ***diffeomorphism*** between *smooth manifolds*,

  3. ***bijective homomorphism*** between *algebraic groups / rings / fields*,

  4. ***graph isomorphism*** between *graphs* that preseves the edge structure,

  Also an isomorphism is called a ***transformation*** in ***geometry***, e.g. *rigid transformation*, *affine transformation* etc.

- **Example** ($\mathcal{L}^2[a,b]$)
  Define $\mathcal{L}^2([a,b])$ to be the set of complex-valued measurable functions on $[a,b]$, a finite interval, that satisfy $\int_{[a,b]} |f(x)|^2\, dx < \infty$. We define an inner product by

$$\langle f\,,\, g\rangle = \int_a^b f(x)\overline{g(x)}dx$$

  $\mathcal{L}^2([a,b])$ is a complete metric space. Actually, $\mathcal{L}^2([a,b])$ is a completion of $\mathcal{C}^0([a,b])$ with finite $\mathcal{L}^2$ norm

$$\|f\|_{\mathcal{L}^2} = \left(\int_a^b |f(x)|^2\, dx\right)^{\frac{1}{2}}$$

  Thus $\mathcal{L}^2([a,b])$ is a *Hilbert space.*

- **Example** (ℓ^2)
  Define ℓ^2 to be *the set of sequences* $(x_n)_{n=1}^{\infty}$ *of complex numbers* which satisfy $\sum_{n=1}^{\infty} |x_n|^2 < \infty$ with the inner product

$$\langle (x_n)_{n=1}^{\infty}\,,\, (y_n)_{n=1}^{\infty}\rangle = \sum_{n=1}^{\infty} \overline{x_n}\, y_n.$$

  ℓ^2 is a complete metric space with ℓ^2 norm

$$\|(x_n)_{n=1}^{\infty}\|_2 = \left(\sum_{n=1}^{\infty} |x_n|^2\right)^{\frac{1}{2}}.$$

  So ℓ^2 is a *Hilbert space.*

  We will see that any Hilbert space that has a ***countable dense set*** and is ***not finite dimensional*** is ***isomorphic*** to ℓ^2 In this sense, ℓ^2 is *the canonical example* of a Hilbert space.

- **Example** ($\mathcal{L}^2(\mathbb{R}^n, \mu)$)
  Define μ to be a *Borel measure* on $\mathbb{R}^n$ and $\mathcal{L}^2(\mathbb{R}^n, \mu)$ to be the set of complex-valued measurable functions on $\mathbb{R}^n$ that satisfy $\int_{\mathbb{R}^n} |f(x)|^2\, d\mu < \infty$. We define an inner product by

$$\langle f\,,\, g\rangle = \int_{\mathbb{R}^n} f(x)\overline{g(x)}d\mu$$

  $\mathcal{L}^2(\mathbb{R}^n, \mu)$ is a *Hilbert space.*

2.3 The Projection Theorem

- **Remark** ***Orthogonality*** is the central concept of Hilbert space. In the presence of closed subspaces, the orthogonality allows us to decompose the Hilbert space into the direct sum of the *subspace* and its *orthogonal complement.*

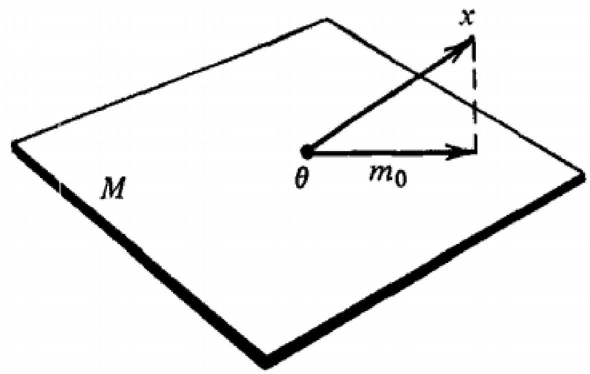

Figure 1: The projection theorem in Hilbert space [Luenberger, 1997]

- **Definition** (***Direct Sum***)
  Suppose that $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces. Then the set of *pairs* (x, y) with $x \in \mathcal{H}_1, y \in \mathcal{H}_2$ is a *Hilbert space* with *inner product*

  $$\langle (x_1, y_1)\,,\,(x_2, y_2)\rangle := \langle x_1\,,\,x_2\rangle_{\mathcal{H}_1} + \langle y_1\,,\,y_2\rangle_{\mathcal{H}_2}$$

  This space is called ***the direct sum*** *of the spaces* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ and is denoted by $\mathcal{H}_1 \oplus \mathcal{H}_2$.

- **Definition** (***Orthogonal Complement***)
  Let $\mathcal{M} \subseteq \mathcal{H}$ is a ***closed*** *linear subspace* of Hilbert space $\mathcal{H}$ with *induced inner product* $\langle\,,\,\rangle$ (i.e. $\langle x\,,\,y\rangle_{\mathcal{M}} = \langle x\,,\,y\rangle_{\mathcal{H}}$ for all $x, y \in \mathcal{M}$). $\mathcal{M}$ is also a *Hilbert space.*

  We denote by $\mathcal{M}^{\perp}$ the set of vectors in $\mathcal{H}$ which are *orthogonal* to $\mathcal{M}$; $\mathcal{M}^{\perp}$ is called ***the orthogonal complement of*** $\mathcal{M}$. It follows from the linearity of the inner product that $\mathcal{M}^{\perp}$ is a *linear subspace* of $\mathcal{H}$ and an elementary argument shows that $\mathcal{M}^{\perp}$ is *closed.* So $\mathcal{M}^{\perp}$ is also a *Hilbert space.*

- **Remark** The following theorem is going to show that

  $$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp} = \left\{x + y : x \in \mathcal{M}, y \in \mathcal{M}^{\perp},\ \text{i.e.}\ \langle x\,,\,y\rangle = 0\right\}.$$

  This important geometric property is one of the main reasons that Hilbert spaces are ***easier*** to handle than Banach spaces.

- **Lemma 2.7** *Let $\mathcal{H}$ be a Hilbert space, $\mathcal{M}$ a closed subspace of $\mathcal{H}$, and suppose $x \in \mathcal{H}$. Then there exists in $\mathcal{M}$ a* ***unique*** *element z* ***closest*** *to x.*

- **Theorem 2.8** ***(The Projection Theorem)***
  Let $\mathcal{H}$ be a Hilbert space, $\mathcal{M}$ a closed subspace. Then every $x \in \mathcal{H}$ can be ***uniquely*** *written $x = z + w$ where $z \in \mathcal{M}$ and $w \in \mathcal{M}^{\perp}$.*

- **Remark** The projection theorem sets up a natural *isomorphism* $\mathcal{M} \oplus \mathcal{M}^{\perp} \to \mathcal{H}$ given by

  $$(z, w) \mapsto z + w$$

  We will often suppress the isomorphism and simply write $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$.

2.4 The Riesz Representation Theorem

- **Definition** (***Bounded Linear Operator***)
  A ***bounded linear transformation*** (or ***bounded operator***) is a mapping $T : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ from a normed linear space X to a normed linear space Y that satisfies

  1. (***Linearity***) $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ for all $x, y \in X$, $\alpha, \beta \in \mathbb{R}$ or $\mathbb{C}$
  2. (***Boundedness***) $\|Tx\|_Y \le C \, \|x\|_X$ for small $C \ge 0$.

  The smallest such C is called ***the norm of*** T, written $\|T\|$ or $\|T\|_{X,Y}$. Thus

$$\|T\| := \sup_{\|x\|_X = 1} \|Tx\|_Y$$

- **Remark** Denote the space of ***all bounded linear operator*** between Hilbert space $\mathcal{H}_1$ and $\mathcal{H}_2$ as $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. The space $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is linear space with norm

$$\|T\| := \sup_{\|x\|_{\mathcal{H}_1} = 1} \|Tx\|_{\mathcal{H}_2}, \quad \forall T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2).$$

  It can be shown that $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is a *complete normed space (i.e. a Banach space).*

- **Definition** (***Dual Space***)
  The space $\mathcal{L}(\mathcal{H}, \mathbb{C})$ is called the ***dual space*** of $\mathcal{H}$ and is denoted by $\mathcal{H}^*$. The elements of $\mathcal{H}^*$ are called ***continuous linear functionals***. That is, the dual space $\mathcal{H}^*$ is the space of *continuous linear functionals* on $\mathcal{H}$.

- **Remark** The *dual space* $\mathcal{H}^*$ is also called ***covector space*** with respect to a vector space $\mathcal{H}$ and the linear functionals are called ***covectors***. This terms are mostly used in *differential geometry* when the vector space is *the tangent space*.

- **Theorem 2.9** *(**The Riesz Representation Theorem**) [Reed and Simon, 1980, Kreyszig, 1989, Conway, 2019]*
  *For each $T \in \mathcal{H}^*$, there is a **unique** $y_T \in \mathcal{H}$ such that*

$$T(x) = \langle x \, , \, y_T \rangle$$

  for all $x \in \mathcal{H}$. In addition $\|y_T\|_{\mathcal{H}} = \|T\|_{\mathcal{H}^*}$.

  Proof: Let $\mathcal{N} = \mathrm{Ker}(T) = \{x \in \mathcal{H} : T(x) = 0\}$. By continuity of T, $\mathcal{N}$ is a closed subspace of $\mathcal{H}$. If $\mathcal{H} = \mathcal{N}$, then $T(x) = 0$ for all x so for $y_T = 0$, $T(x) = \langle x \, , \, 0 \rangle$. If $\mathcal{N} \subset \mathcal{H}$, then there exists $x_0 \notin \mathcal{N}$. Define $y_T = \overline{T(x_0)} \frac{x_0}{\|x_0\|^2}$ so for all $x = \alpha x_0$ for any $\alpha \neq 0$

$$T(x) = T(\alpha x_0) = \alpha T(x_0) = \left\langle \alpha x_0 \, , \, \overline{T(x_0)} \frac{x_0}{\|x_0\|^2} \right\rangle = \langle x \, , \, y_T \rangle \, .$$

  Note that $\mathcal{H} = \mathrm{span}\,\{x_0\} \oplus \mathcal{N}$ since for any $x \in \mathcal{H}$

$$x = \left(x - \frac{T(x)}{T(x_0)} x_0 \right) + \frac{T(x)}{T(x_0)} x_0 \in \mathcal{N} \oplus \mathrm{span}\,\{x_0\} \, .$$

  Also T and $\langle \cdot \, , \, y_T \rangle$ agree on both $\mathcal{N}$ and $\mathrm{span}\,\{x_0\}$, so they must agree on entire $\mathcal{H}$.

To prove $\|y_T\|_{\mathcal{H}} = \|T\|_{\mathcal{H}^*}$, we see that

$$\|T\|_{\mathcal{H}^*} = \sup_{\|x\| \le 1} |Tx| = \sup_{\|x\| \le 1} |\langle x\,,\, y_T \rangle| \le \sup_{\|x\| \le 1} \|x\| \, \|y_T\| = \|y_T\|\,,$$

and

$$\|T\|_{\mathcal{H}^*} = \sup_{\|x\| \le 1} |Tx| \ge \left| T\left(\frac{y_T}{\|y_T\|}\right)\right| = \|y_T\|^{-1} |\langle y_T\,,\, y_T \rangle| = \|y_T\|\,. \quad \blacksquare$$

- **Remark** *The Riesz Representation Theorem* [Conway, 2019, Kreyszig, 1989] is also called ***The Riesz Lemma*** [Reed and Simon, 1980].

- **Remark** We note that *the Cauchy-Schwarz inequality* shows that the ***converse*** of *the Riesz Representation Theorem* is ***true***. Namely, each $y \in \mathcal{H}$ defines *a continuous linear functional* T_y on $\mathcal{H}^*$ by

$$T_y(x) = \langle x\,,\, y \rangle\,.$$

Thus *the Riesz Representation Theorem* together with *the Cauchy-Schwarz inequality* defines an ***isomorphism*** $\mathcal{H}^* \to \mathcal{H}$ between a Hilbert space $\mathcal{H}$ and its dual $\mathcal{H}^*$. In other words, unlike the case in Banach space, the bounded linear functional on Hilbert space has a simple form.

- **Corollary 2.10** ***(The Riesz Representation for Sesquilinear Form)***
Let $B(\cdot, \cdot)$ be a function from $\mathcal{H} \times \mathcal{H}$ to $\mathbb{C}$ which satisfies:

  1. *(**Linearity**) $B(\alpha x + \beta y, z) = \alpha B(x, z) + \beta B(y, z)$*
  2. *(**Conjugate Linearity**) $B(x, \alpha y + \beta z) = \overline{\alpha} B(x, y) + \overline{\beta} B(x, z)$*
  3. *(**Boundedness**) $|B(x, y)| \le C\, \|x\|_{\mathcal{H}}\, \|y\|_{\mathcal{H}}$*

*for all $x, y, z \in \mathcal{H}$, $\alpha, \beta \in \mathbb{C}$. Then there is a **unique bounded linear transformation** $A : \mathcal{H} \to \mathcal{H}$ so that*

$$B(x, y) = \langle x\,,\, Ay \rangle$$

*for all $x, y \in \mathcal{H}$. The **norm** of A is the smallest constant C such that (3) holds.*

Proof: Fix z, (1), (3) shows that $B(\cdot, z)$ is a continuous linear functional on $\mathcal{H}$. Thus by *the Riesz Representation theorem*, there exists some $y_{B,z} \in \mathcal{H}$,

$$B(x, z) = \langle x\,,\, y_{B,z} \rangle\,, \quad \forall x \in \mathcal{H}$$

Define $Az = y_{B,z}$. It is not difficult to show that A is a continuous linear operator with right property. $\blacksquare$

- **Remark** A bilinear function on $\mathcal{H}$ obeying (1) and (2) is called a ***sesquilinear form*** (as a generalization of ***blinear form*** in complex vector space).

In terms of this, an inner product in complex vector space is ***a complex Hermitian form*** (also called a ***symmetric sesquilinear form***).

2.5 Orthonormal Bases

- **Definition** (***Complete Orthonormal Basis***)
  If S is *an orthonormal set* in a Hilbert space $\mathcal{H}$ and *no other orthonormal set* contains S as a proper subset, then S is called *an* ***orthonormal basis*** (or a ***complete orthonormal system***) for $\mathcal{H}$.

- **Theorem 2.11** ***(Existence of Orthonormal Basis)***
  Every Hilbert space $\mathcal{H}$ has an ***orthonormal basis****.*

- **Proposition 2.12** ***(Orthogonal Representation of Element in Hilbert Space)***
  Let $\mathcal{H}$ be a Hilbert space and $S = (x_\alpha)_{\alpha \in A}$ an ***orthonormal basis****. Then for each $y \in \mathcal{H}$,*

$$y = \sum_{\alpha \in A} \langle y \,,\, x_\alpha \rangle \, x_\alpha \tag{4}$$

  and

$$\|y\|_{\mathcal{H}} = \sum_{\alpha \in A} |\langle y \,,\, x_\alpha \rangle|^2 \tag{5}$$

  The equality in (4) *means that the sum on the right-hand side converges (independent of order) to y in $\mathcal{H}$.* ***Conversely,*** *if $\sum_{\alpha \in A} |c_\alpha|^2 < \infty$, $c_\alpha \in \mathbb{C}$, then $\sum_{\alpha \in A} c_\alpha x_\alpha$ converges to an element of $\mathcal{H}$.*

- **Remark** From Bessel's inequality, we already seen that for any finite collection A' of x_α, we have $\sum_{\alpha \in A'} |\langle y \,,\, x_\alpha \rangle|^2 \le \|y\|_{\mathcal{H}}$. The main difficulty is on how to prove convergence of $\sum_{n=1}^{N} |\langle y \,,\, x_n \rangle|^2$ as $N \to \infty$. Similarly we need to prove that $y - \sum_{n=1}^{m} \langle y \,,\, x_{\alpha_n} \rangle \, x_{\alpha_n}$ is still orthogonal to x_α as $m \to \infty$.

- **Remark** The unique coefficients $(\langle y \,,\, x_\alpha \rangle)$ is called ***the Fourier coefficients of*** *y with respect to basis (x_α).*

- **Remark** (***Gram-Schmidt Orthogonalization***)
  Given *any set of independent vectors* $(v_1, v_2, \ldots)$. we can construct *an orthonormal basis* $(b_1, b_2, \ldots)$ via

$$\begin{aligned} b_1 &= \frac{v_1}{\|v_1\|} \\ b_j &= \frac{v_j - \sum_{i=1}^{j-1} \langle v_j \,,\, b_i \rangle \, b_i}{\left\| v_j - \sum_{i=1}^{j-1} \langle v_j \,,\, b_i \rangle \, b_i \right\|}, \quad j \ge 2 \end{aligned}$$

  Thus $\text{span}\,\{v_1, , \ldots, v_m\} = \text{span}\,\{b_1, , \ldots, b_m\}$ for all $m \ge 1$.

2.6 Separability

- **Definition** (***Separability***)
  A *metric space* which has a ***countable dense subset*** is said to be ***separable***.

- **Remark** Most Hilbert space we have seen is separable.

- **Proposition 2.13** *(**Canonical Hilbert Space**)*
  *A Hilbert space $\mathcal{H}$ is **separable** if and only if it has a **countable orthonormal basis** S. If there are $N < \infty$ elements in S, then $\mathcal{H}$ is **isomorphic** to $\mathbb{C}^N$, If there are **countably many** elements in S, then $\mathcal{H}$ is **isomorphic** to ℓ^2.*

- **Remark** Consider the map $v \mapsto (\langle v\,,\, x_n\rangle)_{n=1}^{\infty}$ for orthonormal basis $(x_n)_{n=1}^{\infty}$ as the isomorphism $\mathcal{H} \to \ell^2$.

- **Remark** Notice that in the separable case, *the Gram-Schmidt process* anows us to construct an orthonormal basis without using *Zorn's lemma*.

2.7 Fourier Series

- **Definition** If $f(x)$ is integrable in $[0, 2\pi]$, define ***the Fourier series*** as

$$F(x) := \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} c_n e^{inx}$$
$$\text{where } c_n = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{-inx} f(x) dx.$$

- **Proposition 2.14** *Suppose that $f(x)$ is periodic of period 2π and is continuously differentiable. Then the functions $\sum_{n=-M}^{M} c_n e^{inx}$ converge uniformly to $f(x)$ as $M \to \infty$.*

- **Proposition 2.15** *If $f \in \mathcal{L}^2[0, 2\pi]$, then $\sum_{n=-M}^{M} c_n e^{inx}$ converges to f in the $\mathcal{L}^2$ norm as $M \to \infty$.*

- **Remark** The collection of functions, $(\frac{1}{\sqrt{2\pi}} e^{inx})_{n=-\infty}^{\infty}$ is an ***orthonormal set*** in $\mathcal{L}^2[0, 2\pi]$. The above proposition states that it is a ***complete*** *orthonormal set*, that is, for every $f \in \mathcal{L}^2[0, 2\pi]$

$$f = \lim_{M\to\infty} \frac{1}{\sqrt{2\pi}} \sum_{n=M}^{M} c_n e^{inx}$$

2.8 Legendre, Hermite and Laguerre Polynomials

2.9 Hilbert-Adjoint Operator

- **Definition** (***Hilbert Space Adjoint***)
  Let $T : \mathcal{H}_1 \to \mathcal{H}_2$ be a *bounded linear operator*, where $\mathcal{H}_1$ and $\mathcal{H}_2$ are *Hilbert spaces*. Then ***the Hilbert-adjoint operator*** T^* of T is the operator

$$T^* : \mathcal{H}_2 \to \mathcal{H}_1$$

  such that for all $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$,

$$\langle Tx\,,\, y\rangle_{\mathcal{H}_2} = \langle x\,,\, T^*y\rangle_{\mathcal{H}_1} \tag{6}$$

- **Proposition 2.16** *(**Existence of Adjoint Operator**) [Kreyszig, 1989]*
  *The Hilbert-adjoint operator T^* of T **exists**, is **unique** and is a **bounded linear operator** with norm*

$$\|T^*\| = \|T\| .$$

- **Lemma 2.17** *(**Zero operator**). [Kreyszig, 1989] Let X and Y be inner product spaces and $Q : X \to Y$ a bounded linear operator. Then:*

  1. *$Q = 0$ if and only if $\langle Qx\,,\,y\rangle = 0$ for all $x \in X$ and $y \in Y$.*
  2. *If $Q : X \to X$, where X is complex, and $\langle Qx\,,\,x\rangle = 0$ for all $x \in X$, then $Q = 0$.*

- **Proposition 2.18** *(**Properties of Hilbert-adjoint operators**). [Reed and Simon, 1980, Kreyszig, 1989]*
  Let $\mathcal{H}_1$, $\mathcal{H}_2$ be Hilbert spaces, $S : \mathcal{H}_1 \to \mathcal{H}_2$ and $T : \mathcal{H}_1 \to \mathcal{H}_2$ bounded linear operators and α any scalar. Then we have

  1. *$\langle T^*y\,,\,x\rangle = \langle y\,,\,Tx\rangle$, $(x \in H_1, y \in \mathcal{H}_2)$*
  2. *$(S+T)^* = S^* + T^*$*
  3. *$(\alpha T)^* = \alpha T^*$*
  4. *$(T^*)^* = T$*
  5. *$\|T^*T\| = \|TT^*\| = \|T\|^2$*
  6. *$T^*T = 0 \Leftrightarrow T = 0$*
  7. *$(ST)^* = T^*S^*$ (assuming $\mathcal{H}_2 = \mathcal{H}_1$)*
  8. *If T has a **bounded inverse**, T^{-1}, then T^* has a **bounded inverse** and $(T^*)^{-1} = (T^{-1})^*$.*

2.10 Self-Adjoint, Unitary and Normal Operators

- **Definition** A ***bounded linear operator*** $T : \mathcal{H} \to \mathcal{H}$ on a Hilbert space $\mathcal{H}$ is said to be

  1. ***self-adjoint*** or ***Hermitian*** if

$$T^* = T \quad \Leftrightarrow \quad \langle Tx\,,\,y\rangle = \langle x\,,\,Ty\rangle$$

  2. ***unitary*** if T is *bijective* and

$$T^* = T^{-1}$$

  3. ***normal*** if

$$T^*T = TT^*$$

- **Definition** (***Projection Operator***)
  If $P \in \mathcal{L}(\mathcal{H})$ and $P^2 = P$, then P is called a ***projection***. If in addition $P = P^*$, then P is called an ***orthogonal projection***.

- **Remark** If T is ***self-adjoint*** and ***unitary***, then T is ***normal***.
- **Remark** If a basis for $\mathbb{C}^n$ is given and *a* ***linear operator*** on $\mathbb{C}^n$ is represented by a certain ***matrix***, then its ***Hilbert-adjoint operator*** is represented by the ***complex conjugate transpose*** of that matrix. For $\mathbb{R}^n$, then the ***Hilbert-adjoint operator*** is represented by the ***transpose*** of that matrix
- **Remark** Similarly we have

  1. The matrix representation for *self-adjoint operator* is ***Hermitian*** or ***Symmetric***.
  $$T^* = T \quad \Leftrightarrow \quad \boldsymbol{T}^H = \boldsymbol{T} \text{ (or for real vector space } \boldsymbol{T}^T = \boldsymbol{T})$$
  2. The matrix representation for *unitary operator* is ***unitary*** or ***orthogonal***.
  $$T^* = T^{-1} \quad \Leftrightarrow \quad \boldsymbol{T}^H = \boldsymbol{T}^{-1} \text{ (or for real vector space } \boldsymbol{T}^T = \boldsymbol{T}^{-1})$$
  3. The matrix representation for *normal operator* is ***normal***.
  $$T^*T = TT^* \quad \Leftrightarrow \quad \boldsymbol{T}^H\boldsymbol{T} = \boldsymbol{T}\boldsymbol{T}^H \text{ (or for real vector space } \boldsymbol{T}^T\boldsymbol{T} = \boldsymbol{T}\boldsymbol{T}^T)$$

- **Proposition 2.19** ***(Self-adjointness)****. [Kreyszig, 1989]*
  Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator on a Hilbert space $\mathcal{H}$. Then:

  1. *If T is **self-adjoint**, $\langle Tx\,,\, x\rangle$ is **real** for all $x \in \mathcal{H}$.*
  2. *If $\mathcal{H}$ is complex and $\langle Tx\,,\, x\rangle$ is **real** for all $x \in \mathcal{H}$, the operator T is **self-adjoint***

- **Proposition 2.20** ***(Self-adjointness of product)****. [Kreyszig, 1989]*
  *The product of two bounded **self-adjoint** linear operators S and T on a Hilbert space $\mathcal{H}$ is **self-adjoint** if and only if the operators **commute**,*
  $$ST = TS.$$

- **Proposition 2.21** ***(Sequences of self-adjoint operators)****. [Kreyszig, 1989]*
  *Let (T_n) be a sequence of **bounded self-adjoint** linear operators $T_n : \mathcal{H} \to \mathcal{H}$ on a Hilbert space $\mathcal{H}$. Suppose that (T_n) converges, say,*
  $$T_n \to T, \qquad i.e.\ \|T_n - T\| \to 0$$
  *where $\|\cdot\|$ is the norm on the space $\mathcal{L}(\mathcal{H}, \mathcal{H})$. Then the limit operator T is a **bounded self-adjoint** linear operator on H.*

- **Proposition 2.22** ***(Unitary operator)****. [Kreyszig, 1989]*
  *Let the operators $U : \mathcal{H} \to \mathcal{H}$ and $V : \mathcal{H} \to \mathcal{H}$ be **unitary**; here, $\mathcal{H}$ is a Hilbert space. Then:*

  1. *U is **isometric**; thus $\|Ux\| = \|x\|$ for all $x \in \mathcal{H}$;*
  2. *$\|U\| = 1$, provided $\mathcal{H} \neq \{0\}$,*
  3. *$U^{-1} = U^*$ is **unitary**,*
  4. *UV is unitary,*

5. *U is normal.*

6. *A bounded linear operator T on a complex Hilbert space $\mathcal{H}$ is* ***unitary*** *if and only if T is* ***isometric*** *and* ***surjective****.*

- **Remark** Note that an ***isometric operator*** need not be *unitary* since it may fail to be ***surjective***. An example is the *right shift operator* $T : \ell^2 \to \ell^2$ given by

$$(\xi_1, \xi_2, \xi_3, \dots) \mapsto (0, \xi_1, \xi_2, \xi_3, \dots).$$

References

John B Conway. *A course in functional analysis*, volume 96. Springer, 2019.

Erwin Kreyszig. *Introductory functional analysis with applications*, volume 81. wiley New York, 1989.

David G Luenberger. *Optimization by vector space methods*. John Wiley & Sons, 1997.

Michael Reed and Barry Simon. *Methods of modern mathematical physics: Functional analysis*, volume 1. Gulf Professional Publishing, 1980.