EEE 550 Class Notes

Day 1 - Linear Spaces

**Definition.** A complex linear space (vector space) is a set $X$ together with functions $+: X \times X \to X$ and $\cdot : X \times X \to X$ such that [denoting $+(x, y) = x + y$ and $\cdot (a, x) = ax$, for all $x, y, z$ in $X$ and all complex numbers $a$ and $b$,]

1. (Commulative property of $+$) $x + y = y + x$
2. (Associative property of $+$) $(x + y) + z = x + (y + z)$
3. (Additive identity element) There is an element $0 \in X$ such that $0 + x = x + 0 = x$
4. (Associative property of $\cdot$) $(ab)x = a(bx)$
5. (First distributive property) $a(x + y) = ax + ay$
6. (Second distributive property) $(a + b)x = ax + bx$
7. (Identity element for $\cdot$) $1x = x$

**Note.** A linear space is defined similarly over the real numbers $\mathbb{R}$ or any other field of scalars.

**Example.** $\mathbb{C}^n = \{(x_1, \ldots, x_n)^T \mid x_k \in \mathbb{C}, k = 1, \ldots, n\}$ with addition defined by $(x_1, \ldots, x_n)^T + (y_1, \ldots, y_n)^T = (x_1 + y_1, \ldots, x_n + y_n)^T$ and scalar multiplication by $a(x_1, \ldots, x_n)^T = (ax_1, \ldots, ax_n)^T$.

1. (Commulative property of $+$)

$(x_1, \ldots, x_n)^T + (y_1, \ldots, y_n)^T = (x_1 + y_1, \ldots, x_n + y_n)^T$
$= (y_1 + x_1, \ldots, y_n + x_n)^T = (y_1, \ldots, y_n)^T + (x_1, \ldots, x_n)^T$

2. (Associative property of $+$)

$((x_1, \ldots, x_n)^T + (y_1, \ldots, y_n)^T) + (z_1, \ldots, z_n)^T$
$= (x_1 + y_1, \ldots, x_n + y_n)^T + (z_1, \ldots, z_n)^T$
$= (x_1 + y_1 + z_1, \ldots, x_n + y_n + z_n)^T$
$= (x_1, \ldots, x_n)^T + ((y_1, \ldots, y_n)^T + (z_1, \ldots, z_n)^T)$
3. (Additive identity element) Take $0 = (0, ..., 0)^T \in \mathbb{C}^n$. Then
\[
((x_1, ..., x_n)^T + 0 = (x_1 + 0, ..., x_n + 0)^T = (x_1, ..., x_n)^T
\]
and similarly $0 + (x_1, ..., x_n)^T = (x_1, ..., x_n)^T$.

4. (Associative property of $\cdot$)
\[
(ab)(x_1, ..., x_n)^T = (abx_1, ..., abx_n)^T = a(b(x_1, ..., x_n)^T)
\]

5. (First distributive property)
\[
a((x_1, ..., x_n)^T + (y_1, ..., y_n)^T) = a(x_1 + y_1, ..., x_n + y_n)^T = a(x_1, ..., x_n)^T + a(y_1, ..., y_n)^T
\]

6. (Second distributive property)
\[
(a + b)(x_1, ..., x_n)^T = ((a + b)x_1, ..., (a + b)x_n)^T = (ax_1 + bx_1, ..., ax_n + bx_n)^T = a(x_1, ..., x_n)^T + b(x_1, ..., x_n)^T
\]

7. (Identity element for $\cdot$) $1(x_1, ..., x_n)^T = (1x_1, ..., 1x_n)^T = (x_1, ..., x_n)^T$

**Example.** The space $C_0([0, 1])$ of continuous functions $f : [0, 1] \to \mathbb{C}$ is a complex linear space with addition defined by $(f + g)(t) = f(t) + g(t)$ and scalar multiplication by $(af)(t) = af(t)$. The additive identity element is the function $0(t) = 0$ for all $t \in [0, 1]$.

**Example.** With $1 \leq p < \infty$, the space $\ell^p(\mathbb{Z})$ of complex sequences $(..., x_{-1}, x_0, x_1, x_2, ...)$ indexed by the integers $\mathbb{Z}$ and with $\sum_{k=-\infty}^{\infty} |x_k|^p < \infty$ is a linear space with addition and scalar multiplication defined componentwise.

**Exercise.** Show that $\ell^1(\mathbb{Z}) \neq \ell^2(\mathbb{Z})$ by finding a sequence that is in $\ell^2$ but not in $\ell^1$. 

2
**Definition.** A norm on a linear space \( X \) is a non-negative function \( \varphi : X \to \mathbb{R} \) such that (denoting \( ||x|| = \varphi(x) \)), for all \( x \) and \( y \) in \( X \) and all \( a \in \mathbb{C} \),

1. *(Homogeneity)* \( ||ax|| = |a||x|| \)
2. *(Norm of zero)* \( ||x|| = 0 \iff x = 0 \)
3. *(Triangle inequality)* \( ||x+y|| \leq ||x|| + ||y|| \)

**Definition.** A normed linear space is a linear space \( X \) together with a particular norm on \( X \).

**Example.** \( \mathbb{C}^n \) with the Euclidean norm \( ||x|| = \sqrt{|x_1|^2 + \cdots + |x_n|^2} \) is a normed linear space. Properties 1 and 2 are obvious; property 3 will be verified on Day 3.

**Example.** \( \mathbb{C}^n \) with the supremum norm \( ||x|| = \max\{|x_1|, \ldots, |x_n|\} \) is a normed linear space.

**Example.** The set of continuous functions \( f : [0, 2\pi] \to \mathbb{C} \) with \( \int_0^{2\pi} |f(t)|^2 \, dt < \infty \) (i.e., continuous finite-energy signals on the interval \([0, 2\pi]\)) is a normed linear space with the \( L^2 \) norm \( ||f|| = \sqrt{\int_0^{2\pi} |f(t)|^2 \, dt} \).

**Example.** With \( 1 \leq p < \infty \), the space \( \ell^p(\mathbb{Z}) \) is a normed linear space with the \( p \)-norm

\[
||x|| = \left[ \sum_{k=-\infty}^{\infty} |x_k|^p \right]^{1/p}
\]

**Note.** A norm on a linear space provides a sense of distance between two elements, given by \( ||x - y|| \). Precisely, defining \( \rho(x, y) = ||x - y|| \), it is evident that (i) \( \rho(y, x) = \rho(x, y) \), (ii) \( \rho(x, y) = 0 \iff x = y \), and \( \rho(x, z) \leq \rho(x, y) + \rho(y, z) \) so that \( \rho \) is a metric on \( X \) and \((X, \rho)\) is a metric space.

**Definition.** A sequence \( \{x_1, x_2, \ldots\} \) of elements in a normed linear space \( X \) is a Cauchy sequence if, for any \( \varepsilon > 0 \), there is an integer \( N_\varepsilon \) such that \( ||x_j - x_k|| < \varepsilon \) whenever both \( j \) and \( k \) are greater than \( N_\varepsilon \).

**Definition.** A sequence \( \{x_1, x_2, \ldots\} \) of elements in a normed linear space \( X \) converges to \( x \in X \) if, for any \( \varepsilon > 0 \), there is an integer \( N_\varepsilon \) such that \( ||x_j - x|| < \varepsilon \) for all \( j > N_\varepsilon \). In this case, one writes \( \lim_{j \to \infty} x_j = x \).
**Definition.** A normed linear space $X$ is a Banach space if every Cauchy sequence in $X$ converges to some $x \in X$.

**Example.** The spaces $\mathbb{C}^n$, with either the Euclidean or supremum norms, and $\ell^p$ described above are Banach spaces.

**Example.** As an example of a normed linear space that is not a Banach space, consider a complex linear space $X$ consisting of infinite sequences $\{x_k\} = (x_1, x_2, \ldots)$ with the condition that only finitely many of the elements in a sequence may be non-zero. Thus, for example,

$$(1, 0, 0, \ldots) \in X$$

but $(1, 1, 1, \ldots) \notin X$ and $(1, 1/2, 1/4, \ldots) \notin X$. Defining addition and scalar multiplication componentwise clearly makes $X$ into a linear space. Further,

$$\|(x_1, x_2, \ldots)\| = \sum_{k=1}^{\infty} |x_k|$$

is a norm on $X$. Define a sequence in $X$ by

$$x_1 = (1, 0, 0, \ldots)$$
$$x_2 = (1, 1/2, 0, 0, \ldots)$$
$$\vdots$$
$$x_n = (1, 2^{-1}, \ldots, 2^{-(n-1)}, 0, \ldots)$$

Let $\varepsilon > 0$ and choose $N_\varepsilon$ so that $2^{-N_\varepsilon} < \varepsilon/2$. Suppose $j > k > N_\varepsilon$. Then

$$\|x_j - x_k\| = \left(\frac{1}{2}\right)^k + \left(\frac{1}{2}\right)^{(k+1)} + \cdots + \left(\frac{1}{2}\right)^{(j-1)}$$

$$= \left(\frac{1}{2}\right)^k \left[1 + \frac{1}{2} + \cdots + \left(\frac{1}{2}\right)^{j-k-1}\right] < \frac{\varepsilon}{2}(2) = \varepsilon$$

Thus $x_1, x_2, \ldots$ is a Cauchy sequence of elements in $X$, but $\lim_n x_n = (1, 1/2, 1/4, \ldots) \notin X$. Thus $X$ is not a Banach space.

**Exercise.** Show that the space $X$ of continuous finite-energy signals on $[0, 2\pi]$ described above is not a Banach space by finding a Cauchy sequence in $X$ whose limit is not continuous (and hence not in $X$).
Definition. An inner product on a complex linear space $X$ is a function $p : X \times X \to \mathbb{C}$ such that (denoting $\langle x, y \rangle = p(x, y)$), for all $x$ and $y$ in $X$ and all $a \in \mathbb{C}$,

1. $\langle x, y \rangle = \langle y, x \rangle^*$
2. $\langle ax, y \rangle = a \langle x, y \rangle$
3. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
4. $\langle x, x \rangle \geq 0$
5. $\langle x, x \rangle = 0 \Leftrightarrow x = 0$

Definition. An inner product space is a linear space $X$ together with an inner product on $X$.

Theorem. $\|x\| = \sqrt{\langle x, x \rangle}$ is a norm on $X$.

Proof. That $\|x\| = 0$ if and only if $x = 0$ follows immediately from properties 4 and 5 of the inner product. Also, $\|ax\| = \sqrt{\langle ax, ax \rangle} = \sqrt{a^2 \langle x, x \rangle} = |a| \|x\|$. It remains only to show that the triangle inequality is satisfied. First, note that if $\langle v, w \rangle = 0$ for any $v$ and $w$ in $X$, then

$$||v + w||^2 = \langle v + w, v + w \rangle = ||v||^2 + \langle v, w \rangle + \langle w, v \rangle + ||w||^2 = ||v||^2 + ||w||^2 \quad (1)$$

Now, define

$$u = x - \frac{\langle x, y \rangle}{||y||^2} y$$

By direct calculation, $\langle y, u \rangle = 0$ and

$$x = \frac{\langle x, y \rangle}{||y||^2} y + u$$

Thus (1) gives

$$||x||^2 = \left| \frac{\langle x, y \rangle}{||y||^2} \right|^2 ||y||^2 + ||u||^2$$

Since $||u||^2 \geq 0$, this implies

$$||x||^2||y||^2 \geq |\langle x, y \rangle|^2 \quad (2)$$

1
Now,

$$
||x + y||^2 = (x + y, x + y) = ||x||^2 + ||y||^2 + (x, y) + (y, x)
\leq ||x||^2 + ||y||^2 + 2| (x, y) | \leq ||x||^2 + ||y||^2 + 2||x|| ||y||
= (||x|| + ||y||)^2
$$

So $||x + y|| \leq ||x|| + ||y||$ and the triangle inequality holds. \(\square\)

Note. Equation (1) and inequality (2) both follow from the definition of an inner product. They are known as the Pythagorean theorem and the Cauchy-Schwarz inequality, respectively.

Example. The set of continuous functions $f : [0, 2\pi] \to \mathbb{C}$ with $\int_0^{2\pi} |f(t)|^2 \, dt < \infty$ (i.e., continuous finite-energy signals on the interval $[0, 2\pi]$) is an inner product space with $\langle f, g \rangle = \int_0^{2\pi} f(t)g^*(t) \, dt$.

Definition. A Hilbert space is an inner product space in which every Cauchy sequence converges in the norm $||x|| = \sqrt{(x, x)}$.

Example. $\mathbb{C}^n$ is a Hilbert space with inner product

$$
\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{k=1}^{n} x_k y_k^*
$$

Example. The space $\ell^2(\mathbb{Z})$ is a Hilbert space with inner product

$$
(x, y) = \sum_{k=-\infty}^{\infty} x_k y_k^*
$$

Note. For $p \neq 2$, $\ell^p(\mathbb{Z})$ is not a Hilbert space. These Banach spaces do not admit an inner product that is compatible with their norms.

Definition. Non-zero elements $x$ and $y$ in an inner product space $X$ are orthogonal if $\langle x, y \rangle = 0$. They are orthonormal if, in addition to being orthogonal, $||x|| = ||y|| = 1$.

Definition. A set $\{x_k\}$ in an inner product space is an orthonormal set if any two elements in the set are orthonormal.

Example. In the Hilbert space $H = \mathbb{C}^n$, the set

$$
\left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}
$$

is an orthonormal set.
Example. In the inner product space of continuous finite-energy signals on $[0, 2\pi]$, the set of functions

$$\left\{ x_k(t) = \frac{e^{ikt}}{\sqrt{2\pi}} \mid k \in \mathbb{Z} \right\}$$

is an orthonormal set.
Definition. If \( Y \) is a subset of a linear space \( X \) that is itself a linear space, then \( Y \) is a subspace of \( X \).

Example. With \( X = \mathbb{C}^5 \), the subset
\[
Y = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_5 \\ \end{pmatrix} \mid x_3 = 0 \right\}
\]
is a subspace of \( X \).

Definition. If \( X \) is a Banach space and \( S \subseteq X \), the (finite linear) span of \( S \) is
\[
\text{span } S = \left\{ \sum_{k=1}^{n} a_k s_k \mid s_k \in S \text{ and } a_k > 0 \right\}
\]
(i.e., the collection of all finite weighted sums of elements of \( S \)).

Note. 1. The closed span of \( S \), denoted \( \overline{\text{span } S} \), consists of \( \text{span } S \) together with the limits of every Cauchy sequence in \( \text{span } S \).

2. If \( \text{span } S = X \), \( S \) is called complete in \( X \).

3. \( \text{span } S \) and \( \overline{\text{span } S} \) are subspaces of \( X \); \( \overline{\text{span } S} \) is closed.

4. If \( S \) is complete in \( X \), then each \( x \in X \) is the limit of some sequence in \( \text{span } S \).

Definition. Elements \( x \) and \( y \) in a Hilbert space are orthogonal if \( \langle x, y \rangle = 0 \). They are orthonormal if, in addition, \( ||x|| = ||y|| = 1 \).

Example. In \( \mathbb{C}^4 \), consider discrete sinusoids
\[
x = \frac{1}{2} \begin{pmatrix} 1 \\ i \\ -1 \\ -i \\ \end{pmatrix} \quad \text{and} \quad y = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ \end{pmatrix}
\]
Then \( ||x||^2 = \frac{1}{4}(1 + 1 + 1 + 1) = 1 = ||y||^2 \) and \( \langle x, y \rangle = 1 - i - 1 + i = 0 \). So \( x \) and \( y \) are orthonormal.
\begin{align*}
\text{Analysis} & \quad \text{"Signal"} \quad \xrightarrow{\Phi} \quad \text{"Coefficients"} \\
\text{Synthesis} & \quad \text{"Signal"} \quad \xleftarrow{\Phi} \quad \text{"Signal"}
\end{align*}
\begin{align*}
x & \quad \mapsto \quad a_k = \langle x, \phi_k \rangle \\
\sum_k a_k \phi_k & \quad \mapsto \quad x = \sum_k \langle x, \phi_k \rangle \phi_k
\end{align*}

Note. If \( \Phi = \{ \varphi_1, \varphi_2, \ldots \} \) with \( \varphi_i \) and \( \varphi_j \) orthonormal for \( i \neq j \), \( \Phi \) is called an orthonormal set.

If \( \Phi = \{ \varphi_1, \varphi_2, \ldots \} \) is an orthonormal set in \( \mathcal{H} \) and \( x \in \text{span} \Phi \), then
\[ x = \sum_k a_k \varphi_k \]
for scalars \( a_1, a_2, \ldots \). In this case
\[ \langle x, \varphi_j \rangle = \left( \sum_k a_k \varphi_k, \varphi_j \right) = \sum_k a_k \langle \varphi_k, \varphi_j \rangle = a_j \]
Hence, for each \( x \in \text{span} \Phi \),
\[ x = \sum_k \langle x, \varphi_k \rangle \varphi_k \]

Example. If \( U \) is an \( n \times n \) unitary matrix with columns \( u_1, u_2, \ldots, u_n \), then \( \{ u_1, \ldots, u_n \} \) is an orthonormal set of vectors in \( \mathbb{C}^n \). In particular, consider the \( 4 \times 4 \) discrete Fourier transform matrix
\[ D = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\
1 & i & -1 & -i \\
1 & -1 & 1 & -1 \\
1 & -i & -1 & i \end{bmatrix} \]
The discrete sinusoids comprising the columns of \( D \) are an orthonormal set in \( \mathbb{C}^4 \).

Example. With \( \mathcal{H} = L^2([0, 2\pi]) \), define \( \varphi_k(t) = \frac{1}{\sqrt{2\pi}} e^{ikt} \) for \( k \in \mathbb{Z} \). Then
\[ \langle \varphi_j, \varphi_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{i(j-k)t} \, dt = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \]
Thus \( \Phi = \{ \varphi_k | k \in \mathbb{Z} \} \) is an orthonormal set in \( \mathcal{H} \). If
\[ x(t) = \sum_k a_k \varphi_k(t) = \frac{1}{\sqrt{2\pi}} \sum_k a_k e^{ikt} \]
(i.e., \( x \) has a pointwise Fourier series representation), then its coefficients in this expansion (i.e., its Fourier coefficients) are given by
\[ a_k = \langle x, \varphi_j \rangle = \frac{1}{2\pi} \int_0^{2\pi} x(t) e^{-ikt} \, dt \]
EEE 550 Class Notes

Day 5 – Unitarity of Orthonormal Analysis

**Theorem.** Let \( \Phi = \{ \varphi_k \mid k \in K \} \) with \( K \subset \mathbb{Z} \) an index set be an orthonormal set in a Hilbert space \( \mathcal{H} \). If

\[
x = \sum_{k \in K} a_k \varphi_k \quad \text{and} \quad y = \sum_{j \in K} b_j \varphi_j
\]

are elements of span \( \Phi \), then

\[
\langle x, y \rangle = \sum_{k \in K} a_k^* b_k
\]

**Proof.**

\[
\langle x, y \rangle = \left\langle \sum_{k \in K} a_k \varphi_k, \sum_{j \in K} b_j \varphi_j \right\rangle = \sum_k \sum_j a_k^* b_j \langle \varphi_k, \varphi_j \rangle = \sum_{k \in K} a_k^* b_k
\]

**Corollary** (Identity resolution formula). \( \langle x, y \rangle = \sum_k \langle x, \varphi_k \rangle \langle \varphi_k, y \rangle \)

**Proof.** Note that \( a_k = \langle x, \varphi_k \rangle \) and \( b_k^* = \langle \varphi_k, y \rangle \).

**Corollary** (Parseval’s theorem). \( ||x||^2 = \sum_k |\langle x, \varphi_k \rangle|^2 \)

**Proof.** Taking \( x = y \) in the theorem gives \( a_k = b_k = \langle x, \varphi_k \rangle \).

**Operator Theoretic Interpretation**

Consider \( \{a_k\} = \{\langle x, \varphi_k \rangle\} \) and \( \{b_k\} = \{\langle y, \varphi_k \rangle\} \) as sequences indexed by \( K \subset \mathbb{Z} \). Parseval’s theorem shows that these sequences are elements of the Hilbert space \( \ell^2(K) \). Further, the linear transform \( T : \operatorname{span} \Phi \to \ell^2(K) \) by \( T(x) = \{\langle x, \varphi_k \rangle\} \) has the following properties:

1. It is a unitary transformation; i.e.,

\[
\langle x, y \rangle_{\mathcal{H}} = \langle T(x), T(y) \rangle_{\ell^2(K)}
\]

2. It is an isometric transformation, also called an isometry; i.e., \( ||x||_{\mathcal{H}} = ||T(x)||_{\ell^2} \)
Figure 1: The linear transformation $T : \text{span } \Phi \rightarrow \ell^2(K)$ preserves both lengths and angles.
As discussed earlier, the span of a collection $S$ of elements in a Hilbert space $\mathcal{H}$ is a subspace of $\mathcal{H}$. If $\mathcal{H}$ is finite-dimensional, every subspace of $\mathcal{H}$ is closed. When $\mathcal{H}$ is infinite-dimensional, the span of $S$ may or may not be a closed subspace. But its closure $\text{span} S$ is, of course, closed.

**Definition.** Subspaces $U$ and $V$ of $\mathcal{H}$ are orthogonal if $\langle u, v \rangle = 0$ for every $u \in U$ and $v \in V$.

**Note.** If $U$ and $V$ are orthogonal subspaces of $\mathcal{H}$, then $U \cap V = \{0\}$.

**Definition.** The orthogonal complement of a subspace $U$ of $\mathcal{H}$ is the set $V$ of all $v \in \mathcal{H}$ with $\langle v, u \rangle = 0$ for all $u \in U$.

**Note.** The orthogonal complement of $U$ in $\mathcal{H}$ is often denoted by $U^\perp$.

**Theorem.** The orthogonal complement $U^\perp$ of a subspace $U$ is a subspace of $\mathcal{H}$.

**Proof.** If $v \in U^\perp$, then $\langle av, u \rangle = a \langle v, u \rangle = 0$ for all $a \in \mathbb{C}$ and all $u \in U$. Also, if $v' \in U^\perp$, then $\langle v + v', u \rangle = \langle v, u \rangle + \langle v', u \rangle = 0$ for all $u \in U$. Thus $U^\perp$ is a subspace of $\mathcal{H}$.

**Theorem.** If $x$ and $y$ are elements of $\overline{\text{span} \Phi} \subset \mathcal{H}$ and $\langle x, \varphi_k \rangle = \langle y, \varphi_k \rangle$ for all $\varphi_k \in \Phi$, then $x = y$.

**Proof.**

$$x = \sum_k \langle x, \varphi_k \rangle \varphi_k = \sum_k \langle y, \varphi_k \rangle \varphi_k = y$$

**Theorem.** If $\Phi = \{\varphi_1, \varphi_2, \ldots\}$ is an orthonormal set in $\mathcal{H}$, there is exactly one way to express each $x \in \mathcal{H}$ as

$$x = y + e$$

with $y \in \overline{\text{span} \Phi}$ and $\langle e, \varphi_k \rangle = 0$ for all $k$.

**Proof.** Define

$$y = \sum \varphi \in \Phi \langle x, \varphi_k \rangle \varphi_k$$

(1)
and note that $y \in \overline{\text{span } \Phi}$ by construction. Take $e = x - y$. Then $x = y + e$ and for any $z = \sum a_j \varphi_j \in \text{span } \Phi$,

$$
\langle e, z \rangle = \left( x - \sum_k \langle x, \varphi_k \rangle \varphi_k, \sum_j a_j \varphi_j \right) = \sum_j a_j^2 \langle x, \varphi_j \rangle - \sum_j \left[ \sum_k a_k \langle x, \varphi_k \rangle \langle \varphi_k, \varphi_j \rangle \right] = 0
$$

To show uniqueness, suppose $x = y + e = y' + e'$ are two expressions for $x$ in the desired form. Then $y' = y + (e - e')$, and because $e - e' \in (\text{span } \Phi)^\perp$,

$$
\langle y', \varphi_k \rangle = \langle y + (e - e'), \varphi_k \rangle = \langle y, \varphi_k \rangle
$$

for all $\varphi_k \in \Phi$. Hence $y = y'$ and $e = e'$.

**Theorem (Bessel's inequality).** If $\Phi = \{ \varphi_k | k \in K \}$ is an orthonormal set in $\mathcal{H}$, then

$$
\|x\|^2 \geq \sum_{k \in K} |\langle x, \varphi_k \rangle|^2
$$

for any $x \in \mathcal{H}$.

**Proof.** Write $x = y + e$ with $y \in \overline{\text{span } \Phi}$ and $e \in (\overline{\text{span } \Phi})^\perp$. Then the Pythagorean theorem gives $\|x\|^2 = \|y\|^2 + \|e\|^2$ so that $\|x\|^2 \geq \|y\|^2$. But $y$ is given by (1), so Parseval's theorem gives

$$
\|y\|^2 = \sum_{k \in K} |\langle x, \varphi_k \rangle|^2
$$

**Definition.** An orthonormal set $\Phi = \{ \varphi_k | k \in K \}$ in $\mathcal{H}$ is complete in $\mathcal{H}$ if $\overline{\text{span } \Phi} = \mathcal{H}$.

**Theorem.** If an orthonormal set is complete in $\mathcal{H}$, then $\langle x, \varphi_k \rangle = 0$ for all $k$ if and only if $x = 0$.

**Proof.** If $x = 0$, $\langle x, \varphi_k \rangle = 0$ for all $k$. Conversely, suppose $\langle x, \varphi_k \rangle = 0$ for all $k$ and write

$$
x = \sum_k a_k \varphi_k
$$

Then $a_k = \langle x, \varphi_k \rangle = 0$ for all $k$, and hence $x = 0$.

**Corollary.** If $\Phi$ is complete, there is no larger orthonormal set containing $\Phi$.

**Definition.** A complete orthonormal set in $\mathcal{H}$ is called an orthonormal basis (ONB) for $\mathcal{H}$.

**Notes.** 1. If $\Phi = \{ \varphi_k | k \in K \}$ is an ONB for $\mathcal{H}$, the linear transform $T: \mathcal{H} \to \ell^2(K)$ taking $x$ to $\{ \langle x, \varphi_k \rangle \}$ is unitary (hence isometric).
2. If $\Phi$ is an ONB,

$$\sum_{k \in K} a_k \varphi_k = 0$$

if and only if $a_k = 0$ for all $k$ (i.e., orthonormal signals are linearly independent).

3. If $\langle x, z \rangle = \langle y, z \rangle$ for all $z \in \mathcal{H}$ then $x = y$. This follows from writing $z$ in terms of an ONB and considering $\langle x - y, z \rangle$. 
EEE 550 Class Notes

Day 7 – Orthonormalization

Starting with a linearly independent set $\Psi$ in a Hilbert space $\mathcal{H}$, there are procedures to produce an orthonormal set $\Phi$ such that $\text{span } \Phi = \text{span } \Psi$. One such procedure is described here.

Suppose $\Psi = \{\psi_k \mid k \in K\} \subset \mathcal{H}$ has the property that

$$\sum_{k \in K} a_k \psi_k = 0$$

if and only if $a_k = 0$ for all $k$ (i.e., $\Psi$ is a linearly independent set). In particular, this implies that no $\psi_k = 0$.

1. Define

$$\varphi_1 = \frac{\psi_1}{||\psi_1||}$$

Then $||\varphi_1|| = 1$.

2. Next, set

$$\varphi_2 = \frac{\psi_2 - (\psi_2, \varphi_1) \varphi_1}{||\psi_2 - (\psi_2, \varphi_1) \varphi_1||}$$

Then $||\varphi_2|| = 1$ and

$$\langle \varphi_2, \varphi_1 \rangle = \frac{(\psi_2, \varphi_1) - (\psi_2, \varphi_1) ||\varphi_1||^2}{||\psi_2 - (\psi_2, \varphi_1) \varphi_1||} = 0$$

so that $\{\varphi_1, \varphi_2\}$ is an orthonormal set.

3. Proceeding iteratively for $k > 2$ by defining

$$\varphi_k = \frac{\psi_k - (\psi_k, \varphi_1) \varphi_1 - \cdots - (\psi_k, \varphi_{k-1}) \varphi_{k-1}}{||\psi_k - (\psi_k, \varphi_1) \varphi_1 - \cdots - (\psi_k, \varphi_{k-1}) \varphi_{k-1}||}$$

yields an orthonormal set $\Phi = \{\varphi_k \mid k \in K\}$.

This algorithm is known as the Gram-Schmidt procedure.

Example. Consider polynomials $\psi_k(t)$ defined for $k = 0, 1, 2, \ldots$ on the interval $[0,1]$ by $\psi_k(t) = t^k$. Note that this is a linearly independent set, since the polynomial

$$a_0 + a_1 t + \cdots + a_k t^k$$

is not identically zero unless $a_0 = a_1 = \cdots = a_k$. Following the Gram-Schmidt procedure gives
1. \( \varphi_0(t) = 1, \) so
\[
||\varphi_0||^2 = \int_0^1 |\varphi_0(t)|^2 \, dt = 1
\]

2. \[
\varphi_1(t) = \frac{\psi_1(t) - \langle \psi_1, \varphi_0 \rangle \varphi_0(t)}{||\psi_1(t) - \langle \psi_1, \varphi_0 \rangle \varphi_0(t)||} = \frac{t - 1/2}{(1/\sqrt{12})} = \sqrt{12} \left( t - \frac{1}{2} \right)
\]

3. \[
\varphi_2 = \frac{\psi_2 - \langle \psi_2, \varphi_1 \rangle \varphi_1 - \langle \psi_2, \varphi_1 \rangle \varphi_1}{||\psi_2 - \langle \psi_2, \varphi_1 \rangle \varphi_1 - \langle \psi_2, \varphi_1 \rangle \varphi_1||}
\]

So \( \varphi_2(t) = \sqrt{180} (t^2 - t + \frac{1}{2}) \).

Proceeding in this fashion, an arbitrarily large set of orthonormal polynomials can be constructed.

**Note.** Re-indexing the elements in \( \Psi \) before the orthonormalization process will lead to a different orthonormal set \( \Phi' \). But \( \text{span} \Phi' = \text{span} \Phi = \text{span} \Psi \).
Day 8 – Properties of Linear Maps

If $X$ and $Y$ are complex linear spaces, the following terminology is common (though not completely standard) for linear functions on $X$:

- $T : X \to \mathbb{C}$ is called a "linear functional"
- $T : X \to X$ is called a "linear operator"
- $T : X \to Y$ is called a "linear transform"

**Definition.** The adjoint of a linear transform $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a linear transform $T^* : \mathcal{H}_2 \to \mathcal{H}_1$ with the property that

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

for all $x \in \mathcal{H}_1$ and all $u \in \mathcal{H}_2$.

**Example.** Let $\mathcal{H}_1 = \mathbb{C}^n$, $\mathcal{H}_2 = \mathbb{C}^m$, and $A$ be an $m \times n$ matrix. With $x \in \mathbb{C}^n$ and $u \in \mathbb{C}^m$, $\langle u, Ax \rangle = (Ax)^H u = x^H A^H u$. If $B$ is an $n \times m$ matrix with $\langle Bu, x \rangle = x^H Bu = x^H A^H u$, $B$ must equal $A^H$. Hence the adjoint of the transform defined by $A$ is the one defined by $A^H$; i.e., $A^* = A^H$.

**Definition.** A linear operator $T$ on a Hilbert space $\mathcal{H}$ is unitary if $T^* T = I$, where $I$ denotes the identity operator $I(x) = x$ for all $x \in \mathcal{H}$.

**Note.** If $T$ is unitary, $T^* = T^{-1}$. So $TT^* = I$.

**Theorem.** If $T$ is unitary, $\langle T(x), T(y) \rangle = \langle x, y \rangle$ for all $x$ and $y$ in $\mathcal{H}$. In particular, $\|T(x)\| = \|x\|$.

**Proof.**

$$\langle T(x), T(y) \rangle = \langle T^* T(x), y \rangle = \langle I(x), y \rangle = \langle x, y \rangle$$

**Example.** If $\mathcal{H} = \mathbb{C}^n$ and $A$ is an $n \times n$ matrix,

$$\|Ax\|^2 = \langle Ax, Ax \rangle = (Ax)^H Ax = x^H (A^H A) x$$

So, to have $\|Ax\| = \|x\|$ for every $x$, $A^H A = I$; i.e., $A$ must be a unitary matrix.

**Theorem.** $(T^*)^* = T$. 

1
Proof. For all \( x \) and \( u \), \( \langle u, Tx \rangle = \langle T^*u, x \rangle = \langle x, (T^*)^*u \rangle = \langle x, (T^*)^*x \rangle \). Thus \((T^*)^* = T\).

\[ \square \]

**Theorem.** If \( T \) is unitary, so is \( T^* \).

**Proof.** \( ||T^*x||^2 = \langle T^*x, T^*x \rangle = \langle (T^*)^*T^*x, x \rangle = \langle TT^*x, x \rangle = \langle x, x \rangle = ||x||^2 \).

\[ \square \]

**Positive Operators**

**Definition.** A linear operator \( T : \mathcal{H} \to \mathcal{H} \) is positive definite if \( \langle Tx, x \rangle > 0 \) for all \( x \in \mathcal{H} \). \( T \) is positive semidefinite if \( \langle Tx, x \rangle \geq 0 \) for all \( x \in \mathcal{H} \).

**Notes.**
1. For operators, the notation \( T > 0 \) means \( T \) is positive definite and \( T > S \) means \( T - S > 0 \). Similarly, \( T \geq 0 \) means \( T \) is positive semidefinite and \( T \geq S \) means \( T - S \geq 0 \).
2. If \( \mathcal{H} = \mathbb{C}^n \), then a linear operator may be represented by an \( n \times n \) matrix \( A \). The operator is positive definite precisely if \( x^HAx > 0 \) for all \( x \); i.e., when \( A \) is a positive definite matrix.
3. If \( T = T^*T \), then \( \langle Tx, x \rangle = \langle T^*T x, x \rangle = \langle Tx, Tx \rangle = ||Tx||^2 \geq 0 \). Thus \( T \) is positive semidefinite.
4. If \( x_0 \) is an eigenvector of a positive definite operator \( T \) with eigenvalue \( \lambda \), then \( Tx_0 = \lambda x_0 \). Thus \( \langle T x_0, x_0 \rangle = \lambda ||x_0||^2 > 0 \) so that every eigenvalue of a positive definite operator is positive.

**Bounded Transforms**

**Definition.** A linear transform \( T : X \to Y \) from a normed linear space \( X \) to a normed linear space \( Y \) is bounded if \( ||Tx|| < B||x|| \) for some \( B \in \mathbb{R} \) and all \( x \in X \).

**Note.** \( T \) is bounded if and only if \( ||Tx|| < B \) for all \( x \in X \) with \( ||x|| = 1 \). To assess boundedness, it is sufficient to consider the image of the unit sphere in \( X \) under \( T \).

**Theorem.** A linear transform between normed linear spaces is continuous if and only if it is bounded.

**Proof.** If \( T \) is bounded, let \( \varepsilon > 0 \) and choose \( \delta = \varepsilon / B \). Then \( ||x - x_0|| < \delta \) implies \( ||T(x - x_0)|| = ||Tx - Tx_0|| < B\delta < \varepsilon \). Hence \( T \) is continuous at each \( x_0 \in X \).

Conversely, if \( T \) is continuous, it is (in particular) continuous at \( x_0 = 0 \). Denote \( N = \{ x \in X \mid ||x|| = 1 \} \) and observe that \( N \) is a compact set in the metric space \( X \). Hence \( T(N) \), the image of \( N \) under a continuous function, is compact in \( Y \). So \( ||Tx|| < B \) for all \( x \in N \) and \( T \) is hence bounded. \( \square \)
Theorem (Riesz representation theorem). If $T$ is a continuous linear functional on $\mathcal{H}$, then there is an element $\varphi \in \mathcal{H}$ such that $Tx = \langle x, \varphi \rangle$ for all $x \in \mathcal{H}$.

Proof. Define $M = \{ x \in \mathcal{H} \mid Tx = 0 \}$. Since $T$ is linear, $M$ is a subspace. If $M = \mathcal{H}$, then taking $\varphi = 0$ gives the desired result. Otherwise, $M$ has a non-trivial orthogonal complement $M^\perp$. Take any non-zero $y \in M^\perp$ and define

$$\varphi = \frac{(Ty)^*}{||y||^2}$$

Because $\varphi \in M^\perp$, $T\varphi \neq 0$. Further,

$$T \left( x - \frac{Tx}{T\varphi} \varphi \right) = Tx - \frac{Tx}{T\varphi} T\varphi = 0$$

so that

$$x - \frac{Tx}{T\varphi} \varphi \in M$$

Recalling that $\varphi \in M^\perp$,

$$0 = \left< x - \frac{Tx}{T\varphi} \varphi, \varphi \right> = \langle x, \varphi \rangle - \left< \frac{Tx}{T\varphi} \varphi, \varphi \right>$$

Thus,

$$\langle x, \varphi \rangle = \left< \frac{Tx}{T\varphi} \varphi, \varphi \right> = \frac{Tx}{T\varphi} ||\varphi||^2 = \frac{Tx}{||Ty||^2} ||y||^2 = Tx$$

Note. It is easy to verify that the “representative” $\varphi$ corresponding to a linear functional is unique. Thus there is a one-to-one correspondence between linear functionals on $\mathcal{H}$ and elements of $\mathcal{H}$. 

\[ \square \]
Definition. A collection \( \Phi = \{ \varphi_k \mid k \in K \} \) of elements in a Hilbert space \( \mathcal{H} \) is a frame for \( \mathcal{H} \) if there are real numbers \( A > 0 \) and \( B < \infty \) such that

\[
A\|x\|^2 \leq \sum_{k \in K} |\langle x, \varphi_k \rangle|^2 \leq B\|x\|^2
\]

(1)

for all \( x \in \mathcal{H} \).

Example. If \( \Phi \) is an orthonormal basis for \( \mathcal{H} \), Parseval’s theorem shows that it is a frame with \( A = B = 1 \).

Notes. 1. The right-hand inequality in (1) ensures that the sequence of “coefficients” \( \{ \langle x, \varphi_k \rangle \mid k \in K \} \) is in \( \ell^2(K) \) for every \( x \in \mathcal{H} \).

2. The left-hand inequality in (1) guarantees that the linear transform \( T : \mathcal{H} \to \ell^2(K) \) taking \( x \) to \( \{ \langle x, \varphi_k \rangle \mid k \in K \} \) is one-to-one. To see this, suppose \( x \) and \( y \) are distinct elements of \( \mathcal{H} \) with \( T x = T y \). Define \( z = x - y \) and note that \( \|z\| \neq 0 \). But \( T z = T x - T y = 0 \); i.e., \( \langle z, \varphi_k \rangle = 0 \) for all \( k \in K \). This would violate the left-hand inequality, so \( T \) must be one-to-one.

3. The linear transform \( T : \mathcal{H} \to \ell^2(K) \) taking a signal \( x \) to the coefficient sequence \( \{ \langle x, \varphi_k \rangle \mid k \in K \} \) is known as the Bessel map for the frame \( \Phi \).

Theorem. The map \( T^* : \ell^2(K) \to \mathcal{H} \) by \( T^*(\{c_k\}) = \sum c_k \varphi_k \) is the adjoint of the Bessel map \( T \).

Proof. \( \langle \{c_k\}, T x \rangle_{\ell^2} = \sum c_k \langle \varphi_k, x \rangle_{\mathcal{H}} = \langle \sum c_k \varphi_k, x \rangle_{\mathcal{H}} = \langle T^*(\{c_k\}), x \rangle_{\mathcal{H}} \). \( \square \)

Definition. The frame operator associated with \( \Phi \) is \( T = T^* T \).

Notes. 1. \( T \) takes signals in \( \mathcal{H} \) to signals in \( \mathcal{H} \) by first analyzing them with the frame elements to obtain a coefficient sequence and then synthesizing a new signal as a weighted sum of these coefficients with the frame elements, as pictured below.

2. Because it has the form \( T^* T \), \( T \) is positive semidefinite.

3. \( \langle T x, y \rangle = \langle T x, T y \rangle = \langle x, T y \rangle \) so \( T^* = T \); i.e., \( T \) is self-adjoint.

4. If \( \Phi \) is an orthonormal basis for \( \mathcal{H} \), then \( T x = x \) for all \( x \) and thus \( T = I \); i.e., the identity map on \( \mathcal{H} \).
Theorem. If $T$ is the frame operator associated with a frame $\Phi$ having frame bounds $A$ and $B$, then $A I \leq T \leq B I$.

Proof. $\langle (B I - T)x, x \rangle = B \langle Ix, x \rangle - \langle Tx, x \rangle = B \langle x, x \rangle - \langle Tx, Tx \rangle = B \|x\|^2 - \sum |\langle x, \varphi_k \rangle|^2 \geq 0$ for all $x$ from the definition of an upper frame bound. The proof is identical for the lower bound. □

Corollary. The frame operator is continuous (bounded) and invertible.

Corollary. If $\lambda$ is an eigenvalue of $T$, then $A \leq \lambda \leq B$.

Proof. Suppose $x$ is an eigenvector of $T$ with corresponding eigenvalue $\lambda$. Then

$\langle (B I - T)x, x \rangle = B \|x\|^2 - \lambda \|x\|^2 \geq 0$

Thus $\lambda \leq B$. Similarly $\lambda \geq A$. □
Referring to the frame condition

\[ A||x||^2 \leq \sum |\langle x, \varphi_k \rangle|^2 \leq B||x||^2 \quad \forall x \in \mathcal{H} \]

recall that a frame for a Hilbert space \( \mathcal{H} \) is tight if \( A = B \); i.e., if

\[ ||x||^2 = \frac{1}{A} \sum |\langle x, \varphi_k \rangle|^2 \]

for some \( A > 0 \) and all \( x \in \mathcal{H} \). In this case, the polarization identity gives, for all \( x \) and \( y \) in \( \mathcal{H} \),

\[
\langle x, y \rangle = \frac{1}{4} \left[ ||x + y||^2 + ||x - y||^2 + i||x + iy||^2 - i||x - iy||^2 \right]
= \frac{1}{4A} \left[ \sum |\langle x + y, \varphi_k \rangle|^2 + \cdots \right]
= \frac{1}{4A} \left[ \sum \langle x + y, \varphi_k \rangle \langle \varphi_k, x + y \rangle + \cdots \right]
= \frac{1}{4A} \left[ \sum (\langle x, \varphi_k \rangle + \langle y, \varphi_k \rangle)(\langle \varphi_k, x \rangle + \langle \varphi_k, y \rangle) + \cdots \right]
= \frac{1}{A} \sum \langle x, \varphi_k \rangle \langle \varphi_k, y \rangle = \left\langle \frac{1}{A} \sum \langle x, \varphi_k \rangle \varphi_k, y \right\rangle
\]

Thus

\[ x = \frac{1}{A} \sum \langle x, \varphi_k \rangle \varphi_k \]  \hspace{1cm} (1)

for all \( x \in \mathcal{H} \). In terms of the frame operator \( T \), this result is summarized in the following theorem.

**Theorem.** If \( \Phi \) is a tight frame for \( \mathcal{H} \) with frame bound \( A \), the corresponding frame operator is \( T = A I \).

**Proof.** By (1), \( Tx = \sum \langle x, \varphi_k \rangle \varphi_k = Ax \). \( \square \)

**Corollary.** \( T^{-1} = \frac{1}{A} I \).

**Tight Frames and Orthonormal Bases**

As remarked earlier, it follows from Parseval’s theorem that an orthonormal basis is a tight frame with \( A = 1 \). Further, if \( \Phi \) is a tight frame with frame bound
$A = 1$, the corresponding reconstruction formula (1) is identical to what it would be if $\Phi$ were an orthonormal basis. The relationship between orthonormal bases and frames is further elucidated by the following result and subsequent example.

**Theorem.** Let $\Phi = \{\varphi_k \mid k \in K\}$ be a tight frame for $\mathcal{H}$ with frame bound $A = 1$. If $\|\varphi_k\| = 1$ for all $k \in K$, then $\Phi$ is an orthonormal basis for $\mathcal{H}$.

**Proof.** It follows from (1) that $\Phi$ spans $\mathcal{H}$. Also, $\|\varphi_k\| = 1$ for all $k$ by assumption. So it remains only to show that $\Phi$ is an orthogonal set. For any fixed $j \in K$,

$$1 = \|\varphi_j\|^2 = \sum_{k \in K} |\langle \varphi_j, \varphi_k \rangle|^2 = 1 + \sum_{k \neq j} |\langle \varphi_j, \varphi_k \rangle|^2$$

So $\langle \varphi_j, \varphi_k \rangle = 0$ for $j \neq k$. \hfill \Box

**Example.** With $\mathcal{H} = \mathbb{R}^2$, let

$$\varphi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \varphi_2 = \frac{1}{2} \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}, \quad \varphi_3 = \frac{1}{2} \begin{pmatrix} -1 \\ -\sqrt{3} \end{pmatrix}$$

![Figure 1: A tight frame in $\mathbb{R}^2$.](image)

$$T \left( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right)$$

$$= \left\{ \langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle, \langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \frac{1}{2} \left( \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix} \right) \rangle, \langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \frac{1}{2} \left( \begin{pmatrix} -1 \\ -\sqrt{3} \end{pmatrix} \right) \rangle \right\}$$

$$= \left\{ x_1 - \frac{x_1}{2} + \frac{\sqrt{3}}{2} x_2, \frac{-x_1}{2} - \frac{\sqrt{3}}{2} x_2 \right\}$$

2
So

$$\left\| \mathbf{T} \left( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) \right\|^2 = \frac{3}{2} x_1^2 + \frac{3}{2} x_2^2 = \frac{3}{2} \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\|^2$$

and hence $\Phi$ is a tight frame with frame bound $A = B = 3/2$.

Note that

$$\mathbf{T} \left( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right)$$

$$= x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \left[ -\frac{x_1}{2} + \frac{\sqrt{3}}{2} x_2 \right] \frac{1}{2} \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix} + \left[ -\frac{x_1}{2} - \frac{\sqrt{3}}{2} x_2 \right] \frac{1}{2} \begin{pmatrix} -1 \\ -\sqrt{3} \end{pmatrix}$$

$$= \frac{3}{2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

as expected from (1).

In this example, the frame bound $3/2$ reflects the redundancy of the frame; i.e., it contains three unit vectors in a two-dimensional space. Re-normalizing to $\psi_k = \frac{\sqrt{3}}{3} \varphi_k$ for $k = 1, 2, 3$ gives a tight frame (of non-unit vectors) with bound $A = 1$ which is clearly not an orthonormal basis.
Consider three unit vectors in $\mathbb{R}^2$ defined as follows:

$$\varphi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \varphi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \varphi_3 = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

![Figure 1: A frame of three unit vectors in $\mathbb{R}^2$.](image)

The set $\Phi = \{\varphi_1, \varphi_2, \varphi_3\}$ is a frame for $\mathbb{R}^2$ with Bessel map $T : \mathbb{R}^2 \to l^2(\{1, 2, 3\})$ given by the matrix

$$T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

Its adjoint is

$$T^* = \begin{bmatrix} 1 & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & \frac{\sqrt{2}}{2} \end{bmatrix}$$

So the frame operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ is

$$T = T^*T = \frac{1}{2} \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$
To examine the eigenstructure of $T$, note that the characteristic polynomial is $\det(\lambda I - T) = \lambda^2 - 3\lambda + 2$ which has roots $\lambda_1 = 1$ and $\lambda_2 = 2$. Hence, frame bounds for $\Phi$ are given by $A = 1$ and $B = 2$ (i.e., the smallest and largest eigenvalues of $T$). The eigenvector corresponding to $\lambda = 2$ is determined by the equation $Tx = 2x$ which gives the relationship $x_1 = x_2$ for the components of $x$. Since the eigenvalues of $T$ are distinct, the second eigenvector is orthogonal to the first one. Weighting each eigenvector by its corresponding eigenvalue defines the solid ellipse pictured in Figure 2.

![Figure 2: The solid ellipse has its major and minor axes are given by the eigenvectors of $T$ weighted by their corresponding eigenvalues. The dotted ellipse corresponds to $T^{-1}$ in a similar way.](image)

Finally, note that $T$ is invertible with inverse

$$T^{-1} = \frac{1}{4} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

and that $T^{-1}$ has eigenstructure depicted by the dotted ellipse in Figure 2.
EEE 550 Class Notes

Day 12 – Inversion in General Frames

It remains to establish two key facts regarding the analysis of signals with frames. Both of these are associated with, and obtainable by analysis of, the positive definite operator \( T : \mathcal{H} \to \mathcal{H} \) by

\[
Tx = \sum_k \langle x, \varphi_k \rangle \varphi_k
\]

First, taking \( \tilde{\varphi}_k = T^{-1}(\varphi_k) \) for all \( k \), the set \( \tilde{\Phi} = \{ \tilde{\varphi}_k \} \) is a frame with lower frame bound \( 1/B \) and upper frame bound \( 1/A \). From this, the second result will follow; i.e.,

\[
x = \sum_k \langle x, \varphi_k \rangle \tilde{\varphi}_k = \sum_k \langle x, \tilde{\varphi}_k \rangle \varphi_k
\]

**Lemma.** \( T \) is invertible and \( T^{-1} \) is bounded by \( 1/A \).

**Proof.** That \( T \) is a one-to-one mapping follows from the lower frame bound condition. If \( Tx_1 = Tx_2 \), then \( T(x_1 - x_2) = 0 \). But

\[
A ||x_1 - x_2||^2 \leq \langle T(x_1 - x_2), x_1 - x_2 \rangle = 0
\]

so \( x_1 = x_2 \). To show \( T \) maps \( \mathcal{H} \) onto \( \mathcal{H} \), denote \( M = \text{Range} T = \{ y \in \mathcal{H} | y = Tx \text{ for some } x \in \mathcal{H} \} \). If \( \psi \in M^\perp \), then \( \langle \psi, Tx \rangle = 0 \) for all \( x \in \mathcal{H} \); in particular, \( \langle \psi, Ty \rangle = 0 \). But \( T > 0 \), so \( \psi = 0 \). Hence \( M = \mathcal{H} \) and \( T \) maps \( \mathcal{H} \) onto \( \mathcal{H} \) and thus \( T \) is invertible.

Finally, denote \( y = Tx \) so that \( x = T^{-1}y \). Then

\[
A ||x||^2 = A ||T^{-1}y||^2 = A \langle T^{-1}y, T^{-1}y \rangle \\
\leq \langle TT^{-1}y, T^{-1}y \rangle = \langle y, T^{-1}y \rangle \leq ||y|| ||T^{-1}y||
\]

Thus

\[
||T^{-1}y|| \leq \frac{1}{A} ||y||
\]

(1)

**Note.** A similar argument shows \( \frac{1}{A} ||y|| \leq ||T^{-1}y|| \).

**Lemma.** \( T \) and \( T^{-1} \) are self-adjoint.
Proof. This was already proven for $T$, since it has the form $T^*T$. For $T^{-1}$, let $x_1 = T^{-1}y_1$ and $x_2 = T^{-1}y_2$. Then

$$
\langle T^{-1}y_1, y_2 \rangle = \langle x_1, Tx_2 \rangle = \langle Tx_1, x_2 \rangle = \langle TT^{-1}y_1, T^{-1}y_2 \rangle = \langle y_1, T^{-1}y_2 \rangle
$$

So $(T^{-1})^* = T^{-1}$.

Definition. The dual frame associated with a frame $\Phi = \{\varphi_k | k \in K\}$ is $\tilde{\Phi} = \{\tilde{\varphi}_k | k \in K\}$ where $\tilde{\varphi}_k = T^{-1}\varphi_k$.

Theorem. $\tilde{\Phi}$ is a frame with lower frame bound $1/B$ and upper frame bound $1/A$.

Proof. For a fixed $k$, note that $\langle x, \tilde{\varphi}_k \rangle = \langle x, T^{-1}\varphi_k \rangle = \langle T^{-1}x, \varphi_k \rangle$. So,

$$
\sum_k |\langle x, \tilde{\varphi}_k \rangle|^2 = \sum_k |\langle T^{-1}x, \varphi_k \rangle|^2 = ||TT^{-1}x||^2
$$

$$
= \langle TT^{-1}x, TT^{-1}x \rangle = \langle T^{-1}x, TTT^{-1}x \rangle = \langle T^{-1}x, x \rangle
$$

Expression (1) thus implies

$$
\sum_k |\langle x, \tilde{\varphi}_k \rangle|^2 = \langle T^{-1}x, x \rangle \leq ||x|| \cdot ||T^{-1}x|| \leq \frac{1}{A} ||x||^2
$$

Similarly,

$$
\sum_k |\langle x, \tilde{\varphi}_k \rangle|^2 \geq \frac{1}{B} ||x||^2
$$

Note. Define $\tilde{T} = \{(x, \tilde{\varphi}_k)\}$. Then $\tilde{T}^*\{a_k\} = \sum a_k \tilde{\varphi}_k$ and $\tilde{T}^*\tilde{T} = T^{-1}$; i.e.,

$$
TT^*\tilde{T}x = T\left(\sum_k \langle x, \tilde{\varphi}_k \rangle \tilde{\varphi}_k\right) = \sum_k \langle x, \varphi_k \rangle \varphi_k = x
$$

where the last step is shown below.

Theorem. $x = \sum_k \langle x, \varphi_k \rangle \tilde{\varphi}_k = \sum_k \langle x, \tilde{\varphi}_k \rangle \varphi_k$

Proof. For a fixed $k$,

$$
(\tilde{T}x)_k = \langle x, \tilde{\varphi}_k \rangle = \langle x, T^{-1}\varphi_k \rangle = \langle T^{-1}x, \varphi_k \rangle = \langle (T(T^{-1}x))_k
$$

So $\tilde{T} = TT^{-1}$. Thus $T^*\tilde{T} = T^*TT^{-1} = TT^{-1} = I$. So $x = \sum_k \langle x, \tilde{\varphi}_k \rangle \varphi_k$.

Similarly, $\tilde{T}^*T = (TT^{-1})^*T = T^{-1}T^*T = I$ and $x = \sum_k \langle x, \varphi_k \rangle \tilde{\varphi}_k$. 

$\square$
Definition. Let $S \subset \mathbb{R}$ be non-empty. The real number $u$ is an upper bound for $S$ if $u \geq s$ for all $s \in S$. The real number $\ell$ is a lower bound for $S$ if $\ell \leq s$ for all $s \in S$.

Example. With $S = (0, 1]$,
- $\pi$, 2, and 1 are upper bounds for $S$
- $-e$, 0, and $-2/7$ are lower bounds for $S$
- $1/2$ is neither an upper bound nor a lower bound for $S$

Some subsets of $\mathbb{R}$ have no upper bound. Some have no lower bound. And some, such as $S = \mathbb{Z}$, have neither upper nor lower bounds, in which case they are called unbounded.

Definition. If $\xi$ is an upper bound for $S \subset \mathbb{R}$ with the property that $\xi \leq u$ for all upper bounds $u$ of $S$, then $\xi$ is the supremum of $S$ and one writes $\xi = \sup S$ (or $\xi = \text{l.u.b. } S$).

Definition. If $\lambda$ is a lower bound for $S \subset \mathbb{R}$ with the property that $\lambda \geq \ell$ for all lower bounds $\ell$ of $S$, then $\lambda$ is the infimum of $S$ and one writes $\lambda = \inf S$ (or $\lambda = \text{g.l.b. } S$).

Examples. 1. For $S = (0, 1]$, $\sup S = 1$ and $\inf S = 0$. Note that, in this case, $\sup S \in S$ and $\inf S \notin S$.
2. With

$$S = \left\{ (-1)^n \left( 1 - \frac{1}{n} \right) \mid n = 1, 2, 3, \ldots \right\}$$

$\inf S = -1$ and $\sup S = 1$.

Definition. A non-empty subset $S \subset \mathbb{R}$ is bounded if $-\infty < \inf S \leq \sup S < \infty$.

Construction of the Lebesgue Measure on $\mathbb{R}$

Definition. 1. The length of an open interval $I = (a, b) \subset \mathbb{R}$ is $\ell(I) = b - a$.
2. If $S \subset \mathbb{R}$, a covering of $S$ by open intervals is a countable collection $C = \{I_1, I_2, \ldots\}$ of open intervals with $S \subset I_1 \cup I_2 \cup \cdots$. 

1
3. The length of $C$ is $\ell(C) = \sum_{k=1}^{\infty} \ell(I_k)$.

**Definition.** The Lebesgue measure of $S \subset \mathbb{R}$ is

$$\mu(S) = \inf \{ \ell(C) \mid C \text{ is a covering of } S \text{ by open intervals} \}$$

**Notes.**
1. $\mu(\emptyset) = 0$
2. $\mu(S) \geq 0$ for all $S \subset \mathbb{R}$
3. $\mu((a,b)) = b - a$
4. $A \subset B \Rightarrow \mu(A) \leq \mu(B)$
5. If $A_i \cap A_j = \emptyset$ for all $i \neq j$, then $\mu(A_1 \cup A_2 \cup \cdots) = \mu(A_1) + \mu(A_2) + \cdots$

**Theorem.** If $S \subset \mathbb{R}$ is countable, then $\mu(S) = 0$.

**Proof.** Write $S = \{s_1, s_2, \ldots\}$. Let $\varepsilon > 0$ and define

$$I_k = \left(s_k - \frac{\varepsilon}{2k+1}, s_k + \frac{\varepsilon}{2k+1}\right)$$

Then $C = \{I_1, I_2, \ldots\}$ is a covering for $S$ and

$$\ell(C) = \sum_{k=1}^{\infty} \ell(I_k) = \sum_{k=1}^{\infty} \varepsilon \left(\frac{1}{2}\right)^k \leq \varepsilon.$$

Hence $0 \leq \mu(S) \leq \varepsilon$ for any $\varepsilon > 0$, and so $\mu(S)$ must be zero. \qed

**Corollary.** $\mu(\mathbb{Q}) = \mu(\mathbb{Z}) = \mu(\{s\}) = 0$

**Abstract Measures**

The Lebesgue measure on $\mathbb{R}$ is a special case of the more general concept of a measure on a set $U$.

**Definition.** Let $U$ be a set and denote by $PU$ the collection of all subsets of $U$. A measure on $U$ is a function $\nu : PU \to \mathbb{R}$ with the properties

1. $\nu(S) \geq 0$ for all $S \subset U$
2. $\nu(\emptyset) = 0$
3. $A \subset B \Rightarrow \nu(A) \leq \nu(B)$
4. If $S_i \cap S_j = \emptyset$ for all $i \neq j$, then $\nu(S_1 \cup S_2 \cup \cdots) = \nu(S_1) + \nu(S_2) + \cdots$

**Examples (Measures).**

1. The Lebesgue measure on $\mathbb{R}$.

2. Define $\eta(S)$ to be the number of integers in $S$. Then $\eta$ is a measure on $\mathbb{R}$ called the counting measure.

3. If $\nu$ is a measure on $S$ that satisfies $\nu(S) = 1$, then $\nu$ is a probability measure on $S$.  

2
Definition. Let $\nu$ be a measure on $S$. Then $A \subset S$ is measurable with respect to $\nu$ if

$$\nu(B) = \nu(B \cap A) + \nu(B \cap (S - A))$$

for all $B \subset S$ with $\nu(B) < \infty$.

Notes.

1. Measurability is a technical condition that ensures "countable additivity" of integrals; i.e., that

$$\int_{A_1 \cup A_2 \cup \ldots} f(s) \, d\nu = \int_{A_1} f(s) \, d\nu + \int_{A_2} f(s) \, d\nu + \ldots$$

when the $A_k$ are all disjoint.

2. If $\nu$ is a probability measure, the $\nu$-measurable sets are called "events."

3. The following subsets of $\mathbb{R}$ are all measurable with respect to Lebesgue measure:

   - All intervals $[a, b], [a, b), (a, b), (a, b]$.
   - All unions of sequences of intervals $I_1 \cup I_2 \cup \ldots$.
   - All intersections of sequences of intervals $I_1 \cap I_2 \cap \ldots$.
   - Set complements of any of the above.
   - Unions or intersections of any sequence of the above.
   - Any set of measure zero.

Despite this rich class of Lebesgue measurable sets, there do exist subsets of $\mathbb{R}$ that are not Lebesgue measurable. This will be discussed further later in the class.

4. Sets that are measurable with respect to a measure form a sigma algebra; i.e., $\emptyset$ is a member and all complements and countable unions are members.

Definition. If $f : A \to B$ and $C \subset B$, the image set of $f$ is

$$f(A) = \{b \in B \mid b = f(a) \text{ for some } a \in A\} \subset B$$

The inverse image of $C$ under $f$ is

$$f^{-1}(C) = \{a \in A \mid f(a) \in C\} \subset A$$
Definition. A function $f : S \to \mathbb{R}$ is

- Positive if $f(s) > 0$ for all $s \in S$,
- Non-negative if $f(s) \geq 0$ for all $s \in S$,
- Negative if $f(s) < 0$ for all $s \in S$.

Definition. A function $f : A \to B$ is simple if its image set is countable.

Examples.
- $f : \mathbb{R} \to \mathbb{R}$ by $f(t) = 1$ for all $t$ is simple because its image set $\{1\}$ is countable.
- $f : \mathbb{R} \to \mathbb{R}$ by $f(t) = \sin(t)$ is not simple because its image set $[-1,1]$ is not countable.

Definition. If $\nu_1$ is a measure on $S_1$ and $\nu_2$ is a measure on $S_2$, a function $f : S_1 \to S_2$ is measurable if $f^{-1}(A)$ is $\nu_1$-measurable in $S_1$ whenever $A$ is non-empty and $\nu_2$ measurable in $S_2$.

Notes. 1. If $S_2 = \mathbb{R}$ and $\nu_2 = \mu$, $f$ is called Lebesgue measurable.

2. If $S_2 = \mathbb{C}$, $f$ is called Lebesgue measurable if its real and imaginary parts are Lebesgue measurable.

3. The following functions $f : \mathbb{R} \to \mathbb{C}$ are Lebesgue measurable:

- Continuous functions
- Piecewise-continuous functions
- $f = g + h$ with $g$ and $h$ Lebesgue measurable
- $f = gh$ with $g$ and $h$ Lebesgue measurable
- $f = g/h$ with $g$ and $h$ Lebesgue measurable, provided $h(t)$ is never zero
- $f = |g|$ with $g$ Lebesgue measurable
- The pointwise limit of a sequence of Lebesgue measurable functions
EEE 550 Class Notes

Day 15 – Lebesgue Integration

Definition. Let $\nu_1$ be a measure on $S$ and let $\nu_2$ be a measure on $\mathbb{R}$. If $f : S \to \mathbb{R}$ is non-negative, measurable, and simple with image set $\{x_1, x_2, \ldots\} \subseteq \mathbb{R}$, then

$$\int_S f(s) \, d\nu_1 = \sum_{k=1}^{\infty} x_k \nu_1[f^{-1}(x_k)]$$

Note.

$$\int_S f(s) \, d\nu_1 \geq 0$$

and it may be infinite.

Definition. Let $\nu$ be a measure on $S$ and let $f : S \to \mathbb{R}$ be non-negative and measurable. Then

$$\int_S f(s) \, d\nu = \sup \left\{ \int_S g(s) \, d\nu \middle| 0 \leq g(s) \leq f(s) \text{ for all } s \in S \text{ and } g \text{ measurable and simple} \right\}$$

(1)

Definition. A measurable function $f : S \to \mathbb{R}$ is summable if

$$\int_S |f(s)| \, d\nu < \infty$$

Note. In the above definition, $|f|$ is non-negative and measurable, so its integral is defined by (1).

Theorem. If $f : S \to \mathbb{R}$ is summable, then $f$ can be written as $g - h$ where $g$ and $h$ are non-negative summable functions.

Proof. Define $g = |f|$ and $h = |f| - f$. Then $g$ and $h$ are measurable and non-negative. Also $f = g - h$. Moreover,

$$\int_S g(s) \, d\nu = \int_S |f(s)| \, d\nu < \infty$$

and

$$\int_S h(s) \, d\nu = \int_S |f(s)| - f(s) \, d\nu \leq 2 \int_S |f(s)| \, d\nu < \infty$$

so that $g$ and $h$ are both summable. \qed
**Theorem.** If $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ are non-negative measurable functions, then

$$\int_S [f(s) + g(s)] \, d\nu = \int_S f(s) \, d\nu + \int_S g(s) \, d\nu$$

**Proof.** (Sketch) If $\varphi : S \to \mathbb{R}$ and $\psi : S \to \mathbb{R}$ are simple functions with $0 \leq \varphi(s) \leq f(s)$ and $0 \leq \psi(s) \leq g(s)$ for all $s \in S$, then $[\varphi + \psi] : S \to \mathbb{R}$ is non-negative and simple with $0 \leq [\varphi + \psi](s) \leq [f + g](s)$ for all $s$. Thus

$$\int_S [f(s) + g(s)] \, d\nu \geq \int_S f(s) \, d\nu + \int_S g(s) \, d\nu$$

Conversely, if $\theta : S \to \mathbb{R}$ is simple with $0 \leq \theta(s) \leq f(s) + g(s)$ for all $s$, then $	heta$ can be written as $\theta = \varphi + \psi$ with $\varphi$ and $\psi$ simple and $0 \leq \varphi(s) \leq f(s)$ and $0 \leq \psi(s) \leq g(s)$ for all $s$. Thus

$$\int_S [f(s) + g(s)] \, d\nu \leq \int_S f(s) \, d\nu + \int_S g(s) \, d\nu$$

and the equality is established. $\square$

**Theorem (Uniqueness).** Let $f : S \to \mathbb{R}$ be summable. If $f = g_1 - h_1 = g_2 - h_2$ with $g_1, g_2, h_1, h_2$ all non-negative and summable, then

$$\int_S g_1(s) \, d\nu - \int_S h_1(s) \, d\nu = \int_S g_2(s) \, d\nu - \int_S h_2(s) \, d\nu$$  \hspace{1cm} (2)

**Proof.** Since $g_1 + h_2 = g_2 + h_1$,

$$\int_S g_1(s) \, d\nu + \int_S h_2(s) \, d\nu = \int_S g_2(s) \, d\nu + \int_S h_1(s) \, d\nu$$

All these integrals are finite, so the terms may be re-arranged to get (2). $\square$

**Definition.** Let $f : S \to \mathbb{R}$ be summable. Then

$$\int_S f(s) \, d\nu = \int_S g(s) \, d\nu - \int_S h(s) \, d\nu$$

where $g$ and $h$ are any two non-negative summable functions with $f = g - h$.

**Note.** The uniqueness theorem above ensures that $\int_S f(s) \, d\nu$ is well-defined.
EEE 550 Class Notes

Day 16 – Almost Everywhere

Definition. Two functions $f$ and $g$ on a measure space $S$ are equal almost everywhere with respect to a measure $\nu$ on $S$ if the set $A = \{ s \in S \mid f(s) \neq g(s) \}$ is contained in a $\nu$-measurable set $B$ with $\nu(B) = 0$.

Notes. 1. When $f = g$ almost everywhere with respect to $\nu$, the abbreviation "$f = g$ a.e. ($\nu$)" is often used.

2. For Lebesgue measure, this condition is the same as $\mu(A) = 0$.

3. If $S$ is a probability space (i.e., $\nu(S) = 1$) and $f$ and $g$ are random variables (i.e., measurable functions) on $S$ with $f = g$ a.e., one says "$f = g$ almost surely." 

4. If $f$ and $g$ are summable with $f = g$ a.e., then

$$\int_S f(s) \, d\nu = \int_S g(s) \, d\nu$$

Examples. 1. $S = \mathbb{R}$, $\nu = \mu$. Define

$$f(t) = \begin{cases} 
1 & t \in [-1, 1] \\
0 & \text{otherwise} 
\end{cases} \quad g(t) = \begin{cases} 
1 & t \in (-1, 1) \\
0 & \text{otherwise} 
\end{cases}$$

Then

$$\int_{\mathbb{R}} f(t)e^{-i\omega t} \, d\mu = \int_{\mathbb{R}} g(t)e^{-i\omega t} \, d\mu$$

because $\mu(\{t \mid f(t) \neq g(t)\}) = \mu(\{-1, 1\}) = 0$ so that $f(t)e^{-i\omega t} = g(t)e^{-i\omega t}$ a.e. ($\mu$).

2. With $f$ and $g$ as above, but $\nu = \eta$ (counting measure), $\eta(\{t \mid f(t) \neq g(t)\}) = 2$ and

$$\int_{\mathbb{R}} f(t)e^{-i\omega t} \, d\eta \neq \int_{\mathbb{R}} g(t)e^{-i\omega t} \, d\eta$$

Notes. 1. If $f : \mathbb{R} \to \mathbb{C}$,

$$\int_{\mathbb{R}} f(t) \, d\eta = \sum_{k=-\infty}^{\infty} f(k)$$

Hence, for example, the discrete-time Fourier transform (DTFT) of $f$ is

$$\int_{\mathbb{R}} f(t)e^{-i\omega t} \, d\eta$$. 

1
2. If the Riemann integral of $f : \mathbb{R} \to \mathbb{C}$ exists,

$$\int_{\mathbb{R}} f(t) \, dt = \int_{\mathbb{R}} f(t) \, d\mu$$

i.e., the Lebesgue integral exists and is equal to the Riemann integral. In general, the Lebesgue integral will exist for many $f$ that are not Riemann integrable.

**Theorem.** Let $x \in \mathbb{R}$ and define a function $\delta_x$ on the subsets of $\mathbb{R}$ by

$$\delta_x(S) = \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}$$

Then $\delta_x$ is a measure on $\mathbb{R}$.

**Proof.** Straightforward verification the the properties of a measure hold. \qed

**Note.**

$$\int_{\mathbb{R}} f(t) \, d\delta_x = f(x)$$

**Definition.** Let $S \subset \mathbb{R}$ and $f : S \to \mathbb{C}$. Then

$$\int_{S} f(t) \delta(t - x) \, d\mu = \int_{S} f(t) \, d\delta_x = \begin{cases} f(x) & x \in S \\ 0 & x \notin S \end{cases}$$

**Notes.**
1. The symbol "$\delta(t)$" is only meaningful inside an integral. In particular, $\delta$ is not a function from $\mathbb{R}$ to $\mathbb{R}$ and is not in $L^p(\mathbb{R})$ for any $p$.

2. The symbol may be manipulated formally using the properties of Lebesgue integration; e.g.,

$$\int_{\mathbb{R}} f(t)[\delta(t - x_1) - \delta(t - x_2)] \, d\mu = \int_{\mathbb{R}} f(t) \delta(t - x_1) \, d\mu - \int_{\mathbb{R}} f(t) \delta(t - x_2) \, d\mu = f(x_1) - f(x_2)$$

**Example.** Consider the inverse Fourier transform of $\pi \delta(\omega + \Omega) + \pi \delta(\omega - \Omega)$ (which is an integral).

$$\frac{1}{2\pi} \int_{\mathbb{R}} (\pi \delta(\omega + \Omega) + \pi \delta(\omega - \Omega)) e^{i\omega t} \, d\mu(\omega) = \frac{1}{2}(e^{-i\Omega t} + e^{i\Omega t}) = \cos(\Omega t)$$

For this reason, one sometimes sees the "Fourier transform" of $\cos(\Omega t)$ expressed in this way.

**Theorem.** If $f$ is non-negative and measurable with $\int_S f(s) \, dv = 0$, then $f(s) = 0$ a.e. on $S$. 

2
Proof. Denoting \( E = \{ s \in S \mid f(s) \neq 0 \} \), the goal is to show \( \nu(E) = 0 \). Suppose \( \nu(E) > 0 \) and define \( A_0 = \{ s \in S \mid f(s) > 1 \} \) and, for \( k = 1, 2, \ldots \), define \( A_k = \{ s \in S \mid \frac{1}{k+1} < f(s) \leq \frac{1}{k} \} \). Note that

\[
E = \bigcup_{k=0}^{\infty} A_k
\]

and, since \( A_i \) and \( A_j \) are disjoint for \( i \neq j \), \( \nu(E) = \nu(A_0) + \nu(A_1) + \cdots \). Thus \( \nu(A_{k_0}) > 0 \) for some \( k_0 \). Take

\[
g(s) = \begin{cases} 
1/k_0 & s \in A_{k_0} \\
0 & \text{otherwise}
\end{cases}
\]

Then \( g \) is simple and \( g(s) \leq f(s) \) for all \( s \in S \). But

\[
\int_S g(s) \, d\nu = \frac{1}{k_0} \nu(A_{k_0}) > 0
\]

and \( \int_S f(s) \, d\nu \geq \int_S g(s) \, d\nu \). This contradiction means \( \nu(E) = 0 \) and hence \( f = 0 \) a.e. \( \square \)

Corollary. For any \( p \geq 1 \), \( \int_S |f(s) - g(s)|^p \, d\nu = 0 \) if and only if \( f = g \) a.e.

Notes. 1. Functions \( f \) and \( g \) on \( \mathbb{R} \) are called “Lebesgue equivalent” if \( f = g \) a.e. (\( \mu \)).

2. Formally, for \( p \geq 1 \), \( L^p(\mathbb{R}) \) consists of Lebesgue equivalence classes of functions \( f : \mathbb{R} \to \mathbb{C} \) with \( \int_{\mathbb{R}} |f(t)|^p \, dt < \infty \); i.e., two such functions that are Lebesgue equivalent are regarded as the same element of \( L^p(\mathbb{R}) \), corresponding to the fact that \( ||f - g||_p = 0 \).
EEE 550 Class Notes

Day 17 – Fourier Transform on \( \mathbb{R} \)

**Definition.** For \( f \in L^1(\mathbb{R}) \), its Fourier transform is the function \( \hat{f} : \mathbb{R} \to \mathbb{C} \) defined by

\[
\hat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{-i\omega t} \, d\mu(t)
\]

**Notes.**

1. Several variations on this basic definition are used in the mathematics and engineering literature.

2. This Fourier integral is well defined for each \( \omega \in \mathbb{R} \). Note that

\[
|\hat{f}(\omega)| = \left| \int_{\mathbb{R}} f(t) e^{-i\omega t} \, d\mu(t) \right| \\
\leq \int_{\mathbb{R}} |f(t) e^{-i\omega t}| \, d\mu(t) = \int_{\mathbb{R}} |f(t)| \, d\mu(t) = \|f\|_1 < \infty
\]

3. \( \hat{f} \) is not necessarily in \( L^1(\mathbb{R}) \). Example:

\[
f(t) = \begin{cases} 
1 & t \in [-\pi, \pi] \\
0 & \text{otherwise}
\end{cases}
\]

With \( f \in L^1(\mathbb{R}) \), \( \hat{f} \) is guaranteed to be continuous.

4. If \( f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \), then (1) is valid and Parseval's theorem gives

\[
\|f\|_2^2 = \frac{1}{2\pi} \|\hat{f}\|_2^2
\]

**Plancherel's Extension**

Suppose \( \varphi_k \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \) for \( k = 1, 2, \ldots \) and \( \lim \varphi_k = f \) in \( L^2(\mathbb{R}) \). Because \( L^2(\mathbb{R}) \) is a closed space, \( f \in L^2(\mathbb{R}) \). Also, Parseval's theorem shows that \( \varphi_k \to g \in L^2(\mathbb{R}) \).

Specifically, starting with \( f \in L^2(\mathbb{R}) \), define

\[
\varphi_k(t) = \begin{cases} 
f(t) & |t| \leq k \\
0 & \text{otherwise}
\end{cases}
\]
Then
\[
\int_{-k}^{k} |\varphi_k(t)|^2 \, d\mu \leq \int_{\mathbb{R}} |f(t)|^2 \, d\mu = ||f||_2^2 < \infty
\]
so that \( \varphi_k \in L^2(\mathbb{R}) \) and \( \varphi_k \in L^2([-k,k]) \) for all \( k \). Further, \( L^2([-k,k]) \subset L^1([-k,k]) \) so that
\[
\int_{-k}^{k} |\varphi_k(t)| \, d\mu \leq \int_{\mathbb{R}} |\varphi_k(t)| \, d\mu < \infty
\]
and \( \varphi_k \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \) for all \( k \). Since \( \varphi_k \to f \) in \( L^2(\mathbb{R}) \), \( \varphi_k \to g \) in \( L^2(\mathbb{R}) \). Defining \( f = g \) extends the definition of the Fourier transform from \( L^1(\mathbb{R}) \) to an isometry (up to the multiplicative factor of 2\( \pi \)) on \( L^2(\mathbb{R}) \).

**Notes.**

1. If \( f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \), \( g = \hat{f} \); i.e.,
\[
g(\omega) = \int_{\mathbb{R}} f(t)e^{-i\omega t} \, d\mu(t) \text{ a.e.}(\omega)
\]

2. When \( f \in L^2(\mathbb{R}) \) but \( f \notin L^1(\mathbb{R}) \), the Fourier transform of \( f \) is only defined in the sense of \( L^2 \). The integral (1) may fail to be defined for some \( \omega \in \mathbb{R} \), but it will be defined for almost all \( \omega \). In this case, \( \hat{f} \) might not be continuous (for example, \( f(t) = \text{sinc}(t) \)).
Definition. Let $f \in L^2(\mathbb{R})$ have positive norm. The time uncertainty of $f$ is

$$U_t(f) = \frac{1}{||f||^2} \int_{\mathbb{R}} |t f(t)|^2 \, dt$$

The frequency uncertainty of $f$ is

$$U_\omega(f) = \frac{1}{||f||^2} \int_{\mathbb{R}} |\omega \hat{f}(\omega)|^2 \, d\omega$$

Notes. 1. $U_t(f)$ and $U_\omega(f)$ may be infinite.

2. The time uncertainty measures the "spread" of a signal in the following familiar sense. Define

$$\rho_f(t) = \frac{|f(t)|^2}{||f||^2}$$

Then $\rho_f$ is non-negative and

$$\int_{\mathbb{R}} \rho_f(t) \, dt = \frac{1}{||f||^2} \int_{\mathbb{R}} |f(t)|^2 \, dt = 1$$

so that $\rho_f$ may be regarded as a probability density function. Moreover,

$$U_t(f) = \int_{\mathbb{R}} t^2 \rho_f(t) \, dt$$

With this perspective, $U_t$ is the sum of the variance and the squared mean of the density $\rho_f$. The frequency uncertainty has a similar interpretation.

Definition. With the above notation, the time mean of $f$ is

$$M_t(f) = \int_{\mathbb{R}} t \rho_f(t) \, dt$$

The frequency mean $M_\omega(f)$ of $f$ is defined analogously.

Definition. The time variance of $f$ is

$$\sigma_t(f) = \int_{\mathbb{R}} (t - M_t(f))^2 \rho_f(t) \, dt$$

The frequency variance $\sigma_\omega(f)$ of $f$ is defined analogously.
Note. The time and frequency variances of \( f \) are equal to its time and frequency uncertainties, respectively, when \( M_t(f) = M_\omega(f) = 0 \). The variances are unaffected by time and frequency shifts of the signal and represent the minimal uncertainties under these shifts.

Lemma. Suppose \( f : \mathbb{R} \to \mathbb{C} \) is differentiable. Then

\[
2|f^*(t)f'(t)| \geq \left| \frac{d}{dt}|f^2(t)| \right| \geq \left| \frac{d}{dt}|f^2(t)| \right|
\]

Proof.

\[
\frac{d}{dt}|f^2(t)| \leq \left| \frac{d}{dt}|f^2(t)| \right| = \left| \frac{d}{dt}|f(t)|^2 \right| = \left| \frac{d}{dt}f(t)f^*(t) \right|
= |f(t)[f^*]'(t) + f'(t)f^*(t)| \leq |f(t)[f^*]'(t)| + |f'(t)f^*(t)|
= 2|f^*(t)f'(t)|
\]

\[\square\]

Theorem (Heisenberg Principle). If \( f \in L^2(\mathbb{R}) \) with \( f \neq 0 \),

\[
U_t(f)U_f(f) \geq \frac{1}{4}
\]

Proof. If either \( U_t(f) \) or \( U_\omega(f) \) is infinite, there is nothing to prove. So assume both are finite. Denote by \( A \subset L^2(\mathbb{R}) \) the set of all functions that are continuously differentiable with\(^1\)

\[
\lim_{t \to \pm \infty} tf(t) = 0
\]

Suppose \( f \in A \). Then the derivative formula for the Fourier transform holds, implying \( f' \in L^2(\mathbb{R}) \) and

\[
\hat{f}'(\omega) = i\omega \hat{f}(\omega)
\]

\(^1\) One often sees this proof for the Schwarz class of infinitely differentiable functions with \( \lim_{t \to \pm \infty} |t^j f^{(k)}(t)| = 0 \) for all \( i, j \in \mathbb{Z}^+ \).
Applying Parseval's relation and the Cauchy-Schwarz inequality,

\[
U_t(f) U_\omega(f) = \frac{1}{\|f\|^2 \|\hat{f}\|^2} \int_\mathbb{R} |tf(t)|^2 \, dt \int_\mathbb{R} |\omega \hat{f}(\omega)|^2 \, d\omega \\
= \frac{1}{2\pi \|f\|^4} \int_\mathbb{R} |tf(t)|^2 \, dt \int_\mathbb{R} |\hat{f}'(\omega)|^2 \, d\omega \\
= \frac{1}{\|f\|^4} \int_\mathbb{R} |tf^*(t)|^2 \, dt \int_\mathbb{R} |f'(t)|^2 \, dt \\
\geq \frac{1}{\|f\|^4} \left[ \int_\mathbb{R} |tf^*(t)f'(t)| \, dt \right]^2 \\
\geq \frac{1}{4\|f\|^4} \left[ \int_\mathbb{R} t \frac{d}{dt} \{ |f(t)|^2 \} \, dt \right]^2 \\
\geq \frac{1}{4\|f\|^4} \left[ \int_\mathbb{R} t \frac{d}{dt} \{ |f(t)|^2 \} \, dt \right]^2 \\
= \frac{1}{4\|f\|^4} \|f\|^4 = \frac{1}{4}
\]

The last step uses integration by parts and assumption (1).

The result may be extended from \(A\) to all of \(L^2(\mathbb{R})\) by density: given \(f \in L^2(\mathbb{R})\), there is a sequence \(\{f_k\} \subset A\) with \(f = \lim f_k\) and

\[
\lim \int_\mathbb{R} \left|t(f_k(t) - f(t))\right| \, dt = 0
\]

\(\square\)

**Corollary.** \(\sigma_t(f) \sigma_\omega(f) \geq 1/4\).

**Note.** With \(f(t) = e^{-t^2}\), \(U_t(f) U_\omega(f) = 1/4\). Hence the inequality is sharp.
Definition. A group is a set $G$ with a binary operation $\odot : G \times G \rightarrow G$ satisfying

1. Associative law: $a \odot (b \odot c) = (a \odot b) \odot c$ for all $a$, $b$, and $c$ in $G$. 

2. Identity element: There is an element $e \in G$ such that $a \odot e = e \odot a = a$ for all $a \in G$. 

3. Inverses: For every $a \in G$, there exists an element $\hat{a} \in G$ such that $a \odot \hat{a} = \hat{a} \odot a = e$. 

Definition. A group $G$ is commutative (abelian) if $a \odot b = b \odot a$ for all $a$ and $b$ in $G$. 

Examples. 1. Additive integers: $G = \mathbb{Z}; \odot = +$ (i.e., usual addition). Clearly $e = 0$ and $\hat{a} = -a$. Associativity follows from the usual associative law of addition. Further, this group is abelian. 

2. Additive reals: $G = \mathbb{R}; \odot = +$ 

3. General Linear Group $GL_n(\mathbb{R})$: With $n \in \mathbb{Z}$, $n \geq 1$, take 

$$G = \{\text{Invertible } n \times n \text{ matrices with real entries}\}$$ 

and $\odot$ to be matrix multiplication. Then $e = I_n$ (the identity matrix) and $\hat{A} = A^{-1}$. This group is not abelian for $n > 1$. 

4. Affine Group $A$: $G = \{(a, b) \mid a, b \in \mathbb{R}, a \neq 0\};$ 

$$(a, b) \odot (a', b') = (aa', ab' + b)$$ 

Associativity is by direct calculation 

$$[(a, b) \odot (a'b')] \odot (\hat{a}, \hat{b}) = (aa', ab' + b) \odot (\hat{a}, \hat{b}) = (aa'\hat{a}, aa'\hat{b} + ab' + b)$$ 

$$= (a(a'\hat{a}), a(a'\hat{b} + b') + b) = (a, b) \odot (a'\hat{a}, a'\hat{b} + b')$$ 

$$= (a, b) \odot [(a', b') \odot (\hat{a}, \hat{b})]$$ 

Also, clearly $e = (1, 0)$ and the inverse of $(a, b)$ is $(1/a, -b/a)$. This is called the affine group because of its relationship with affine functions. If $f(t) = at + b$ and $g(t) = a't + b'$ with $a \neq 0$ and $a' \neq 0$, then $h(t) = f(g(t)) = (aa')t + (ab' + b)$. This group is non-abelian.
5. Weyl-Heisenberg Group: \( G = \{(z, q, p) \mid p, q \in \mathbb{R} \text{ and } z \in \mathbb{C} \text{ with } |z| = 1\}. \)

\[(z, q, p) \odot (z', q', p') = (e^{i(pq' - qp')} z z', q + q', p + p')\]

Since \( |e^{i(pq' - qp')}| = 1 \), the operation \( \odot \) clearly takes \( G \times G \) into \( G \). For this group, which is not abelian, \( e = (1, 0, 0) \) and the inverse of \((z, q, p)\) is \((z', -q, -p)\). Associativity may be verified by direct (somewhat laborious) calculation.

**Action of a Group on a Set**

**Definition.** Let \( G \) be a group with operation \( \odot \) and let \( S \) be a set. A group action of \( G \) on \( S \) is a function \( \Gamma : G \times S \to S \) such that

\[\Gamma(a \odot b, s) = \Gamma(a, \Gamma(b, s))\]

for all \( a \) and \( b \) in \( G \) and all \( s \) in \( S \).

**Examples.**

1. Time shift on \( \ell^2(\mathbb{Z}) \): \( G = \mathbb{Z} \) (additive integers); \( S = \ell^2(\mathbb{Z}) \);
\[
\Gamma(n, \{c_k\}) = \{c_{k-n}\}. \text{ Note } e = 0 \text{ and } \Gamma(0, \{c_k\}) = \{c_k\}; \text{ i.e., } \Gamma(e, \cdot) \text{ is the identity map on } S. \text{ Also,}
\[
\Gamma(m+n, \{c_k\}) = \Gamma(m, \{c_{k-n}\}) = \Gamma(m, \Gamma(n, \{c_k\})) = \{c_{k-n-m}\}
\]

2. Weyl-Heisenberg action on \( L^2(\mathbb{R}) \): \( G = \text{Weyl-Heisenberg group}; \ S = L^2(\mathbb{R}); \)
\[
\Gamma_W \left( (e^{i\theta}, q, p), f \right)(t) = e^{i\theta} e^{-ipq/2} e^{ipt} f(t - q)
\]

Since
\[
\int_{\mathbb{R}} \left| \Gamma_W \left( (e^{i\theta}, q, p), f \right)(t) \right|^2 dt = \int_{\mathbb{R}} \left| e^{i\theta} e^{-ipq/2} e^{ipt} f(t - q) \right|^2 dt = \|f\|^2 < \infty
\]

this action maps \( L^2 \) into \( L^2 \). It is straightforward to verify
\[
\Gamma_W \left( (e^{i\theta}, q, p) \odot (e^{i\theta'}, q', p'), f \right) = \Gamma_W \left( (e^{i\theta}, q, p), \Gamma_W \left( (e^{i\theta'}, q', p'), f \right) \right)
\]

so that this is a proper group action. Note that \( p \) causes a frequency shift and \( q \) a time shift in \( f \), while \( e^{i\theta} \) contributes only to a bulk phase shift.

3. Affine group action on \( L^2(\mathbb{R}) \): \( G = A; \ S = L^2(\mathbb{R}); \)
\[
\Gamma_A \left( (a, b), f \right)(t) = \frac{1}{|a|} f \left( \frac{t - b}{a} \right)
\]

Note that
\[
\|\Gamma_A ((a, b), f)\|^2 = \frac{1}{|a|} \int_{\mathbb{R}} \left| f \left( \frac{t - b}{a} \right) \right|^2 dt = \frac{1}{|a|} \int_{\mathbb{R}} |a| |f(u)|^2 du = \|f\|^2
\]

The parameter \( a \) dilates the signal, while \( b \) time shifts it.
Definition. The Weyl operator $\mathcal{W}(p,q)$ on $L^2(\mathbb{R})$ is defined in terms of the Weyl-Heisenberg group action as

$$\mathcal{W}(p,q)g(t) = g^{(p,q)}(t) = \Gamma_w ((1,q,p),g)) (t) = e^{-ipq/2} e^{iqt} g(t-q)$$

Notes. 1. The main effects of $\mathcal{W}(p,q)$ on a signal $g$ are to time shift it by $q$ and frequency shift it by $p$. It also introduces a phase shift that depends on both $p$ and $q$.

2. Some literature uses the alternative definition

$$g^{(p,q)}(t) = e^{ipt} g(t - q)$$

so that $g^{(p,q)}$ is a time and frequency shifted replicate of $g$ with no ancillary phase shift. This convention provides some convenience, but it is easily seen that it creates asymmetry in the factor distinguishing $g^{(p,q)}g^{(p',q')}$ from $g^{(p',q')}g^{(p,q)}$.

3. Assume, only to simplify the upcoming calculation, $\|g\| = 1$. The time mean of $g^{(p,q)}$ is

$$M_t \left(g^{(p,q)}\right) = \int_{\mathbb{R}} t|g^{(p,q)}(t)|^2 \, dt = \int_{\mathbb{R}} t|e^{-ipq/2} e^{ipt} g(t-q)|^2 \, dt$$

$$= \int_{\mathbb{R}} t|g(t-q)|^2 \, dt = \int_{\mathbb{R}} (q+u)|g(u)|^2 \, du$$

$$= M_t(g) + q$$

Similarly, $M_\omega \left(g^{(p,q)}\right) = M_\omega(g) + p$. Direct calculation also verifies that $\sigma_t \left(g^{(p,q)}\right) = \sigma_t(g)$ and $\sigma_\omega \left(g^{(p,q)}\right) = \sigma_\omega(g)$. In this sense, the signal is "re-centered" by $q$ in time and $p$ in frequency without a change in its "spread" in either time or frequency.

Resolution of the Identity

Lemma. Let $\hat{h} \in L^2(\mathbb{R})$. Then

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \hat{h}(y) e^{-ip(y-u)} \, dy \, dp = 2\pi \hat{h}(x)$$
Proof. Denote by $h$ the inverse Fourier transform of $\hat{h}$. Then
\[
\int_{\mathbb{R}} \int_{\mathbb{R}} \hat{h}(y) e^{-ip(x-y)} \, dy \, dp = \int_{\mathbb{R}} \left[ \int_{\mathbb{R}} \hat{h}(y) e^{ipy} \, dy \right] e^{-ipx} \, dp
\]
\[
= \int_{\mathbb{R}} \left[ 2\pi \hat{h}(p) \right] e^{-ipx} \, dp = 2\pi \hat{h}(x)
\]
\[\square\]

**Theorem** (Identity resolution). Let $f_1$, $f_2$, $g_1$, and $g_2$ be functions in $L^2(\mathbb{R})$. Then
\[
\int_{\mathbb{R}} \int_{\mathbb{R}} \left< f_1, g_1^{(p,q)} \right> \left< g_2^{(p,q)}, f_2 \right> \, dp \, dq = 2\pi \left< f_1, f_2 \right> \left< g_2, g_1 \right>
\]

**Proof.**
\[
\int_{\mathbb{R}} \int_{\mathbb{R}} \left< f_1, g_1^{(p,q)} \right> \left< g_2^{(p,q)}, f_2 \right> \, dp \, dq
\]
\[
= \int_{\mathbb{R}} \int_{\mathbb{R}} \left[ \int_{\mathbb{R}} f_1(x) e^{ipx/2} e^{-ip(x-q)} \, dx \right] \left[ \int_{\mathbb{R}} e^{-ipx/2} e^{ipy} g_2(y-q) \, dy \right] \, dp \, dq
\]
\[
= \int_{\mathbb{R}} \int_{\mathbb{R}} f_1(x) g_1^*(x-q) \left[ \int_{\mathbb{R}} g_2(y-q) f_2^*(y) e^{-ip(x-y)} \, dy \right] \, dx \, dq
\]

Using the lemma, this becomes
\[
2\pi \int_{\mathbb{R}} \int_{\mathbb{R}} f_1(x) f_2^*(x) g_1^*(x-q) g_2(x-q) \, dx \, dq
\]
\[
= 2\pi \int_{\mathbb{R}} f_1(x) f_2^*(x) \left[ \int_{\mathbb{R}} g_1^*(x-q) g_2(x-q) \, dq \right] \, dx
\]
\[
= 2\pi \left< g_2, g_1 \right> \int_{\mathbb{R}} f_1(x) f_2^*(x) \, dx = 2\pi \left< f_1, f_2 \right> \left< g_2, g_1 \right>
\]
\[\square\]

**Notes.**
1. Letting $g_1 = g_2 = g$, $f_1 = f$, and $f_2(x) = \delta(x-t)$, the above result gives
\[
f = \frac{1}{2\pi ||g||^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \left< f, g^{(p,q)} \right> g^{(p,q)} \, dp \, dq \tag{1}
\]
This can be shown without the use of the $\delta$ symbol.

2. Regarding $f$ as the "signal" and $g$ as an "analysis function," (1) shows that $f$ can be reconstructed from all its projections onto time-shifted, frequency shifted replicates of $g$. 


The Weyl-Heisenberg Isometry

Setting $f_1 = f_2 = f$ and $g_1 = g_2 = g$ in the identity resolution formula gives

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \left| \langle f, g^{(p,q)} \rangle \right|^2 dp \, dq = 2\pi \|g\|^2 \|f\|^2$$

Thus the "coefficient function" $c_f(p, q) = \langle f, g^{(p,q)} \rangle$ is in $L^2(\mathbb{R}^2)$. If the window $g$ has unit energy, the mapping $L^2(\mathbb{R}) \to L^2(\mathbb{R}^2)$ by $f \mapsto c_f$ is isometric up to the constant factor $\sqrt{2\pi}$. 
The Weyl-Heisenberg identity resolution theorem shows that \( f \in L^2(\mathbb{R}) \) is recoverable from
\[
\langle f, g^{(p,q)} \rangle
\]
for any non-zero "window" \( g \in L^2(\mathbb{R}) \) provided these values are available for all \( p \) and \( q \) in \( \mathbb{R} \). Taking
\[
g(t) = \begin{cases} 
\frac{1}{\sqrt{2\pi}} & t \in [0, 2\pi) \\
0 & \text{otherwise}
\end{cases}
\]
standard facts about Fourier series show that \( \{g^{(2\pi n, m)} \mid n, m \in \mathbb{Z}\} \) is an orthonormal basis for \( L^2(\mathbb{R}) \). Hence, with this \( g \), samples of \( c_f(p,q) \) taken on a rectangular grid are sufficient to recover \( f \):

Wigh \( g \) a rectangular function of length \( q_0 > 0 \), it is clear that this holds for \( \{g^{(n p_0, m q_0)} \mid n, m \in \mathbb{Z}\} \) provided that \( p_0 q_0 = 2\pi \). If \( p_0 q_0 = a < 2\pi \), this set is a tight frame for \( L^2(\mathbb{R}) \) with bound \( A = B = 2\pi / a \); in particular, \( f \) is reconstructible from the samples of \( c_f \).

Notes. (cf. Ten Lectures on Wavelets by I. Daubechies)

1. If \( p_0 q_0 > 2\pi \), \( G = \{g^{(n p_0, m q_0)} \mid n, m \in \mathbb{Z}\} \) is not complete in \( L^2(\mathbb{R}) \) for any \( g \in L^2(\mathbb{R}) \).

2. If \( p_0 q_0 = 2\pi \), \( G \) may or may not be complete. But if it is a frame:

Theorem. (Balian-Low) Either \( \sigma_t(g) \) or \( \sigma_\omega(g) \) is infinite.
3. If \( p_0 q_0 < 2\pi \), \( G \) may be a frame for \( L^2(\mathbb{R}) \) and \( \sigma_z(g)\sigma_\omega(g) \) can still be finite - even near the Heisenberg bound.

**Ambiguity Functions**

**Definition.** The (auto) ambiguity function of \( f \in L^2(\mathbb{R}) \) is \( A_f : \mathbb{R} \times \mathbb{R} \to \mathbb{C} \) by

\[
A_f(p,q) = \langle f^{(p,q)}, f \rangle
\]

**Theorem.** (*Moyal’s identity*)

\[
\|f\|_{L^2(\mathbb{R})}^2 = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} |A_f(p,q)|^2 \, dp \, dq = \frac{1}{2\pi} \|A_f\|_{L^2(\mathbb{R}^2)}^2
\]

**Proof.** Take \( f_1 = f_2 = g_1 = g_2 = f \) in the Weyl-Heisenberg identity resolution theorem. \( \square \)

**Note.** Cross ambiguity is defined analogously.

**Ambiguity in Radar**

Figure 1: A simplified monostatic radar system.

Transmitted waveform: \( x(t) \in L^2(\mathbb{R}) \)

Received signal model: \( r(t) = ax(t - \tau)e^{j2\pi \alpha t} + n(t) \)

- \( a \) models attenuation in the propagation channel
- \( \tau \) accounts for the travel time of the transmitted waveform to the scatterer and back to the receiver
- \( \alpha \) is a frequency shift approximating the Doppler effect induced by transmitter, receiver, or target motion
- \( n \) is additive noise
If $x, \tau,$ and $\alpha$ are known and $n(\cdot)$ is white Gaussian, the optimal (Neyman-Pearson) detection statistic is a function of

$$
\rho = \langle x^{(\alpha, \tau)}, x^* \rangle = \int_{\mathbb{R}} x^{(\alpha, \tau)}(t) x^*(t) \, dt
$$

$$
= \int_{\mathbb{R}} x(t - \tau)e^{i\alpha t} [a^* x^*(t - \tau)e^{-i\alpha t} + n(t)] \, dt
$$

$$
= a^* |x|^2 + m
$$

But $\alpha$ and $\tau$ are generally not known, so $\rho$ is computed for a range of $\alpha$ and $\tau$ values, yielding statistics

$$
a^* A_x(\alpha', \tau') + m(\alpha', \tau')
$$

Note that this is a scaled and noisy replicate of the auto-ambiguity function of the transmitted waveform $x$. Consequently, the design of waveforms for radar (and sonar) is often characterized in terms of the ambiguity function.
Definition. A function \( f : \mathbb{C} \rightarrow \mathbb{C} \) is differentiable at a point \( a \in \mathbb{C} \) if its derivative at \( a \)

\[
f'(a) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}
\]

exists and is in \( \mathbb{C} \).

Note. This is deeper than differentiation of functions of real variables because \( h \) is complex. There are many directions from which \( h \) can approach zero; all must yield the same value of the limit.

Definition. A function \( f : \mathbb{C} \rightarrow \mathbb{C} \) is analytic (holomorphic) on a subset \( S \subset \mathbb{C} \) if it is differentiable at all \( a \in S \). It is entire if it is differentiable at all \( a \in \mathbb{C} \).

Operations on Analytic Functions

If \( \alpha \in \mathbb{C} \) and \( f : \mathbb{C} \rightarrow \mathbb{C} \) and \( g : \mathbb{C} \rightarrow \mathbb{C} \) are differentiable at \( z \in \mathbb{C} \), then so are

1. \( \alpha f \) with \( (\alpha f)'(z) = \alpha f'(z) \)
2. \( f + g \) with \( (f + g)'(z) = f'(z) + g'(z) \)
3. \( fg \) with \( (fg)'(z) = f(z)g'(z) + f'(z)g(z) \)
4. \( f/g \) provided \( g(z) \neq 0 \) with

\[
\left(\frac{f}{g}\right)'(z) = \frac{f'(z)g(z) - f(z)g'(z)}{[g(z)]^2}
\]

Also, the composition of differentiable functions is differentiable in the sense of the following chain rule. If \( g \) is differentiable at the point \( f(z) \in \mathbb{C} \), then \( g \circ f \) is differentiable at \( z \) with \( (g \circ f)' = g'(f(z))f'(z) \).

Examples of Analytic Functions

1. Constant functions: Define \( f : \mathbb{C} \rightarrow \mathbb{C} \) by \( f(z) = \alpha \) for all \( z \). Then

\[
f'(z) = \lim_{h \to 0} \frac{\alpha - \alpha}{h} = 0
\]

for all \( z \in \mathbb{C} \).
2. Polynomials: Define \( f : \mathbb{C} \rightarrow \mathbb{C} \) by \( f(z) = z \) for all \( z \). Then
\[
f'(z) = \lim_{h \to 0} \frac{(z + h) - z}{h} = 1
\]
for all \( z \in \mathbb{C} \). The properties above imply \( p(z) = \alpha_0 + \cdots + \alpha_n z^n \) is entire.

3. Exponential functions: Define \( f : \mathbb{C} \rightarrow \mathbb{C} \) by \( f(z) = e^z \) for all \( z \). Then
\[
f'(z) = \lim_{h \to 0} \frac{e^{(z+h)} - e^z}{h} = e^z \lim_{h \to 0} \frac{e^h - 1}{h} = e^z
\]
for all \( z \in \mathbb{C} \). So \( \cos(z) \), \( \sin(z) \), et cetera are entire.

### Cauchy-Riemann Equations

If \( f : \mathbb{C} \rightarrow \mathbb{C} \), write \( f = u + iv \) with \( u : \mathbb{C} \rightarrow \mathbb{R} \) and \( v : \mathbb{C} \rightarrow \mathbb{R} \); i.e., \( u \) is the real part of \( f \) and \( v \) is the imaginary part of \( f \). Denoting \( z = x + iy \), these may be regarded as functions \( \mathbb{R}^2 \rightarrow \mathbb{R} \) by setting
\[
u(x, y) = u(x + iy) \quad v(x, y) = v(x + iy)
\]

**Theorem.** Let \( f = u + iv \) and suppose \( f \) is differentiable at \( a = \alpha + i\beta \in \mathbb{C} \). Then \( u \) and \( v \) satisfy the Cauchy-Riemann partial differential equations
\[
\frac{\partial u}{\partial x}(a) = \frac{\partial v}{\partial y}(a) \quad \frac{\partial u}{\partial y}(a) = -\frac{\partial v}{\partial x}(a)
\]

**Proof.** Let \( \Delta > 0 \) be real. Then
\[
f'(a) = \lim_{\Delta \to 0} \frac{f(a + \Delta) - f(a)}{\Delta}
\]
\[
= \lim_{\Delta \to 0} \frac{u(a + \Delta) - u(a) + iv(a + \Delta) - iv(a)}{\Delta}
\]
\[
= \lim_{\Delta \to 0} \frac{u(a + \Delta) - u(a)}{\Delta} + i \lim_{\Delta \to 0} \frac{v(a + \Delta) - v(a)}{\Delta}
\]
\[
= \frac{\partial u}{\partial x}(a) + i \frac{\partial v}{\partial x}(a)
\]

But \( f'(a) \) is also given by
\[
f'(a) = \lim_{\Delta \to 0} \frac{f(a + i\Delta) - f(a)}{i\Delta} = -i \frac{\partial u}{\partial y}(a) + \frac{\partial v}{\partial y}(a)
\]
Equating real and imaginary parts of (1) and (2) completes the proof. \( \square \)

**Example.** Let \( f : \mathbb{C} \rightarrow \mathbb{C} \) by \( f(z) = e^z = e^{x[\cos(y) + i\sin(y)]} \). Then
\[
\frac{\partial u}{\partial x} = e^x \cos(y) = \frac{\partial v}{\partial y}
\]
and
\[
\frac{\partial u}{\partial y} = -e^x \sin(y) = -\frac{\partial v}{\partial x}
\]
**EEE 550 Class Notes**

**Day 23 – Integration on Paths**

**Definition.** A path in $\mathbb{C}$ is a continuous function $\gamma : [0,1] \rightarrow \mathbb{C}$. Its reverse path $-\gamma$ is the path with $(-\gamma)(t) = \gamma(1 - t)$.

**Example.** With $\gamma : [0,1] \rightarrow \mathbb{C}$ by $\gamma(t) = \cos(\pi t) + 2i \sin(\pi t)$, $-\gamma$ is defined by $(-\gamma)(t) = \cos(\pi(1 - t)) + 2i \sin(\pi(1 - t)) = -\cos(\pi t) + 2i \sin(\pi t)$.

![Diagram of paths $\gamma$ and $-\gamma$](image)

**Definition.** A piecewise differentiable path in $\mathbb{C}$ is a function $\gamma : [0,1] \rightarrow \mathbb{C}$ that is differentiable except possibly at finitely many points in $[0,1]$.

**Definition.** If $\gamma$ is a piecewise differentiable path in $\mathbb{C}$ and $f : \mathbb{C} \rightarrow \mathbb{C}$ is such that $f \circ \gamma$ is Lebesgue measurable,

$$\int_{\gamma} f(z) \, dz = \int_0^1 f(\gamma(t)) \gamma'(t) \, dt$$

**Note.** If $\xi : [0,1] \rightarrow [0,1]$ is differentiable with $\xi(0) = 0$ and $\xi(1) = 1$, let $p(t) = \gamma(\xi(t))$ (i.e., $p$ is a “reparameterization of $\gamma$”). Then

$$\int_{p} f(z) \, dz = \int_0^1 f(p(t))p'(t) \, dt = \int_0^1 f(\gamma(\xi(t))) \gamma'(\xi(t)) \xi'(t) \, dt$$

$$= \int_0^1 f(\gamma(u)) \gamma'(u) \, du = \int_{\gamma} f(z) \, dz$$

In other words, the integral does not depend on the parameterization chosen for the path $\gamma$. 

1
Example. With $f(z) = z$ and $\gamma(t) = \cos(2\pi t) + 2i \sin(2\pi t)$,
\[
\int_{\gamma} f(z) \, dz = \int_{0}^{1} \left( \cos(2\pi t) + 2i \sin(2\pi t) \right) \left[ -2\pi \sin(2\pi t) + 4\pi i \cos(2\pi t) \right] \, dt = 0
\]
by ordinary calculus.

Notes. 1.
\[
\int_{-\gamma} f(z) \, dz = - \int_{\gamma} f(z) \, dz
\]

2. Denoting by $\gamma_1 + \gamma_2 + \cdots + \gamma_n$ the union of the oriented paths $\gamma_k$,
\[
\int_{\gamma_1 + \cdots + \gamma_n} f(z) \, dz = \int_{\gamma_1} f(z) \, dz + \cdots + \int_{\gamma_n} f(z) \, dz
\]

Path Length and Estimation

Definition. The length of a piecewise differentiable path $\gamma$ is
\[
\ell(\gamma) = \int_{0}^{1} |\gamma'(t)| \, dt
\]

Theorem. (Estimation)
\[
\left| \int_{\gamma} f(z) \, dz \right| \leq \ell(\gamma) \sup_{z=\gamma(t)} |f(z)|
\]

Proof.
\[
\left| \int_{\gamma} f(z) \, dz \right| = \left| \int_{0}^{1} f(\gamma(t)) \gamma'(t) \, dt \right| \leq \int_{0}^{1} |f(\gamma(t)) \gamma'(t)| \, dt
\]
\[
\leq \sup_{z=\gamma(t)} |f(z)| \int_{0}^{1} |\gamma'(t)| \, dt = \ell(\gamma) \sup_{z=\gamma(t)} |f(z)|
\]

Examples. 1. With $\gamma(t) = e^{\pi it}$,
\[
\ell(\gamma) = \int_{0}^{1} |\pi e^{\pi it}| \, dt = \int_{0}^{1} \pi \, dt = \pi
\]
This is as expected, since $\gamma$ is a semicircle of radius one.

2. Taking $\gamma_1(t) = e^{\pi it^2}$, $\gamma_1$ is a reparameterization of $\gamma$.
\[
\ell(\gamma_1) = \int_{0}^{1} |2\pi ie^{\pi it^2}| \, dt = \int_{0}^{1} 2\pi t \, dt = \pi = \ell(\gamma)
\]
3. With \( f(z) = \frac{i}{z} \),

\[
\left| \int_{\gamma} f(z) \, dz \right| \leq \ell(\gamma) \sup_{z=\gamma(t)} |f(z)| = \pi
\]

This can be confirmed by direct calculation of \( \int_{\gamma} f(z) \, dz \).

Jordan Paths in \( \mathbb{C} \)

**Definition.** A subset \( S \subset \mathbb{C} \) is path connected if, given any two points \( x \) and \( y \) in \( S \), there is a path \( \gamma : [0,1] \to \mathbb{C} \) such that \( \gamma(0) = x \), \( \gamma(1) = y \), and \( \gamma(t) \in S \) for all \( t \in (0,1) \).

**Definition.** A piecewise differentiable path \( \gamma \) in \( \mathbb{C} \) is a Jordan path if it is injective on \((0,1)\), \( \gamma(0) = \gamma(1) \), and \( \gamma(t) \neq \gamma(0) \) for any \( t \in (0,1) \).

![Figure 1](image)

**Figure 1:** Sets (a)-(d) are path connected. Sets (e) and (f) are not path connected.

**Theorem** (Jordan curve theorem). Let \( \gamma \) be a Jordan path in \( \mathbb{C} \) with image set \( \Gamma \). \( \mathbb{C} - \Gamma \) is not path connected, but \( \mathbb{C} - \Gamma = S_1 \cup S_2 \) with \( S_1 \) and \( S_2 \) both path connected, \( S_1 \) bounded (the “interior of \( \gamma \)”) and \( S_2 \) unbounded.
EEE 550 Class Notes

Day 24 – Cauchy’s Theorem

Notation. If \( \gamma \) is a Jordan path in \( \mathbb{C} \), one often denotes

\[
\oint_{\gamma} f(z) \, dz = \int_{\gamma} f(z) \, dz
\]

Theorem (Cauchy). If \( f \) is analytic in a region containing the image set and the interior of a Jordan path \( \gamma \), then

\[
\oint_{\gamma} f(z) \, dz = 0
\]

Proof. Step 1: \( f \) a polynomial; Step 2: \( \gamma \) a rectangular path; Step 3: \( \gamma \) a Jordan path.

1. If \( f(z) = a_n z^n + \cdots + a_1 z + a_0 \), then

\[
\oint_{\gamma} f(z) \, dz = \int_{0}^{1} [a_n z^n + \cdots + a_0] \gamma'(t) \, dt = a_n \int_{0}^{1} \gamma^n(t) \gamma'(t) \, dt + \cdots + a_0 \int_{0}^{1} \gamma'(t) \, dt
\]

\[
= a_n \frac{\gamma^{n+1}(t)}{n+1} \bigg|_{0}^{1} + \cdots + a_0 \gamma(t) \bigg|_{0}^{1} = 0
\]

since \( \gamma(1) = \gamma(0) \).

2. Suppose the interior of \( \gamma \) is a rectangle \( R_0 \) and denote its boundary (the image set of \( \gamma \)) by \( \partial R_0 \).

![Diagram of a rectangle and its boundary]

Writing

\[
\int_{\partial R_0} f(z) \, dz = \int_{\gamma} f(z) \, dz
\]
the goal is to show this integral is zero. To accomplish this, subdivide \( R_0 \)
into four rectangles \( S_1, S_2, S_3, \) and \( S_4 \) of equal area.

Since the integrals on the inner edges cancel out,

\[
\int_{\partial S_1} f(z) \, dz + \int_{\partial S_2} f(z) \, dz + \int_{\partial S_3} f(z) \, dz + \int_{\partial S_4} f(z) \, dz = \int_{\partial R_0} f(z) \, dz
\]

For this to hold, at least one \( S_k \) must satisfy

\[
\left| \int_{\partial S_k} f(z) \, dz \right| \geq \frac{1}{4} \left| \int_{\partial R_0} f(z) \, dz \right|
\]

Let \( R_1 \) be the first \( S_k \) for which this holds. Repeating this process yields
a nested sequence of rectangles \( R_0 \supset R_1 \supset \cdots \supset R_n \supset \cdots \) with

\[
\left| \int_{\partial R_n} f(z) \, dz \right| \geq \frac{1}{4^n} \left| \int_{\partial R_0} f(z) \, dz \right|
\]

or

\[
\left| \int_{\partial R_n} f(z) \, dz \right| \leq 4^n \left| \int_{\partial R_0} f(z) \, dz \right|
\]

Now, note that \( \ell(\partial R_n) = \frac{1}{2^n} \ell(\partial R_{n-1}) = \cdots = \frac{1}{2^n} \ell(\partial R_0) \) and

\[
\sup_{z_1, z_2 \in \partial R_n} |z_1 - z_2| < \frac{1}{2} \ell(\partial R_n) = \frac{1}{2^{n+1}} \ell(\partial R_0)
\]

The rectangles \( R_0, R_1, \ldots \) converge to some \( z_0 \in R_0 \) in the sense that it
is possible, given any \( \delta > 0 \), to choose \( n \) so large that \( |z - z_0| < \delta \) for all
\( z \in R_n \). By differentiability of \( f \), for any \( \varepsilon > 0 \) one can choose \( \delta \) small
enough that

\[
\left| \frac{f(z) - f(z_0)}{z - z_0} - f'(z_0) \right| < \varepsilon
\]

or

\[
|f(z) - f(z_0) - (z - z_0)f'(z_0)| < \varepsilon |z - z_0|
\]

for all \( z \in R_n \). Note that

\[
\int_{\partial R_n} f(z) \, dz = \int_{\partial R_n} f(z) \, dz - \int_{\partial R_n} f(z_0) \, dz - \int_{\partial R_n} (z - z_0)f'(z_0) \, dz
\]

\[
= \int_{\partial R_n} [f(z) - f(z_0) - (z - z_0)f'(z_0)] \, dz
\]

2
So,

\[
\left| \int_{\partial R_n} f(z) \, dz \right| \leq 4^n \left| \int_{\partial R_n} f(z) \, dz \right| \leq 4^n \varepsilon \int_{\partial R_n} |z - z_0| \, dz \\
\leq 4^n \varepsilon \ell(\partial R_n) \sup_{z \in \partial R_n} |z - z_0| < \varepsilon 4^n \left[ \frac{1}{2^n} \ell(\partial R_0) \right] \left[ \frac{1}{2^{n+1}} \ell(\partial R_0) \right] \\
= \frac{\varepsilon}{2} \ell(\partial R_0)^2
\]

Since \( \ell(\partial R_0) \) is fixed and \( \varepsilon \) is arbitrary,

\[
\int_{\gamma} f(z) \, dz = 0
\]

3. (Sketch of Proof) Let \( \gamma \) be a Jordan path in \( \mathbb{C} \) and let \( S \) be a rectangular complex contained in the interior of \( \gamma \) as depicted in Figure 3.

![Figure 1: The Jordan path \( \gamma \) is approximated by the boundary of a rectangular complex.](image)

Since each \( S_k \) is a rectangle,

\[
\int_{\partial S} f(z) \, dz = \sum_{k=1}^{K} \int_{\partial S_k} f(z) \, dz = 0
\]

The Jordan path \( \gamma \) can be approximated with arbitrary accuracy by \( \partial S \) for a rectangular complex \( S \) contained in its interior. Thus

\[
\int_{\gamma} f(z) \, dz = 0
\]
Integration on Homotopic Paths

**Definition.** Let $\gamma_1$ and $\gamma_2$ be paths in $S \subset \mathbb{C}$ with $\gamma_1(0) = \gamma_2(0)$ and $\gamma_1(1) = \gamma_2(1)$. These paths are path homotopic in $S$ if there is a continuous function $\psi : [0, 1] \times [0, 1] \to S$ with $\psi(t, 0) = \gamma_1(t)$ and $\psi(t, 1) = \gamma_2(t)$ for all $t \in [0, 1]$.

**Definition.** A path $\gamma : [0, 1] \to \mathbb{C}$ is closed if $\gamma(0) = \gamma(1)$.

**Definition.** A subset $S \subset \mathbb{C}$ is simply connected if every closed path in $S$ is path homotopic to a constant path in $S$.

**Theorem.** Suppose $f$ is analytic in a simply connected $S \subset \mathbb{C}$. If $\gamma_1$ and $\gamma_2$ are path homotopic in $S$,

$$\int_{\gamma_1} f(z) \, dz = \int_{\gamma_2} f(z) \, dz$$
Proof. (Sketch) By Cauchy’s theorem,
\[ \int_{\gamma_1 - \gamma_2} f(z) \, dz = 0 \]

So,
\[ \int_{\gamma_1} f(z) \, dz - \int_{\gamma_2} f(z) \, dz = 0 \quad \text{and} \quad \int_{\gamma_1} f(z) \, dz = \int_{\gamma_2} f(z) \, dz \]

\[ \square \]

Corollary. If \( f \) is analytic in a simply connected \( S \subset \mathbb{C} \) and \( \gamma \) is a path in \( S \), then
\[ \int_{\gamma} f(z) \, dz \]

depends only on the endpoints of \( \gamma \).
Suppose $n(\gamma, a)$ is to represent the number of times the closed path $\gamma$ winds around the point $a \in \mathbb{C}$ in the counterclockwise direction. It is generally possible to establish an intuitive idea of what the value of $n(\gamma, a)$ should be by examining the image set of $\gamma$:

This notion can be formalized as follows.

**Definition.** Let $\gamma$ be a piecewise differentiable path in $\mathbb{C}$ and let $a \in \mathbb{C}$ with $a$ not in the image set of $\gamma$. The winding number of $\gamma$ with respect to $a$ is

$$n(\gamma, a) = \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z - a}$$

That $n(\gamma, a)$ is an integer can be proven, but the argument is technical and will be omitted here. A heuristic understanding of the winding number is generally sufficient to support interpretation and application of the following important theorem.

**Theorem (Cauchy Integral Formula).** Suppose $f$ is analytic in a simply connected region $S \subset \mathbb{C}$ and let $\gamma$ be a closed piecewise differentiable path in $S$. Then, for any $a \in S$ not in the image set of $\gamma$,

$$n(\gamma, a)f(a) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - a} \, dz$$

**Note.** This formula gives the value of $f(a)$ from the values $f$ takes on $\gamma$!

**Examples.**

1. To evaluate
$$\oint_{|z|=1} \frac{e^z}{z} \, dz$$
Let $f(z) = e^z$ which is analytic for all $z \in \mathbb{C}$. Cauchy’s formula with $a = 0$ gives

$$n(\gamma, 0)f(0) = \frac{1}{2\pi i} \int_{|z|=1} \frac{f(z)}{z} \, dz = \frac{1}{2\pi i} \int_{|z|=1} \frac{e^z}{z} \, dz$$

Thus,

$$1 = \frac{1}{2\pi i} \int_{|z|=1} \frac{e^z}{z} \, dz \implies \int_{|z|=1} \frac{e^z}{z} \, dz = 2\pi i$$

2.

$$\int_{|z|=2} \frac{dz}{z^2 + 1} = \frac{1}{2i} \left[ \int_{|z|=2} \frac{dz}{z + i} - \int_{|z|=2} \frac{dz}{z - i} \right] = \frac{1}{2i} [2\pi i - 2\pi i] = 0$$

Derivatives of Cauchy’s Formula

Assume, for simplicity, that $n(\gamma, a) = 1$. Then

$$f(a) = \frac{1}{2\pi i} \int_{|z|=1} \frac{f(z)}{z - a} \, dz$$

Differentiating both sides with respect to $a$ gives

$$f'(a) = \frac{1}{2\pi i} \int_{|z|=1} \frac{f(z)}{(z - a)^2} \, dz$$

Differentiating $m$ times yields

$$f^{(m)}(a) = \frac{m!}{2\pi i} \int_{|z|=1} \frac{f(z)}{(z - a)^{m+1}} \, dz$$

Example. To evaluate

$$\int_0^\infty \frac{e^z}{z^n} \, dz$$

take $f(z) = e^z$ and $a = 0$. Then $f^{(n-1)}(z) = e^z$ and

$$f^{(n-1)}(0) = \frac{(n-1)!}{2\pi i} \int_{|z|=1} \frac{e^z}{z^n} \, dz$$

and hence

$$\int_0^\infty \frac{e^z}{z^n} \, dz = \frac{2\pi i}{(n-1)!}$$
Definition. Suppose \( f : \mathbb{C} \to \mathbb{C} \cup \{\infty\} \) is not analytic at \( a \in \mathbb{C} \). The point \( a \) is an isolated pole of \( f \) if

1. There is some \( \delta > 0 \) so that \( f \) is analytic at each \( z \) with \( 0 < |z - a| < \delta \)

2. \((z - a)^n f(z)\) is analytic at \( z = a \) for some (finite) \( n \geq 1 \). The smallest integer \( n \) for which the second condition holds is the order of the pole.

If \( f \) has an isolated pole of order \( n \) at \( a \in \mathbb{C} \), then \( g(z) = (z - a)^n f(z) \) is analytic for \( |z - a| < \delta \). So \( g \) can be expanded in a Taylor series around \( z = a \)

\[
g(z) = A_0 + A_1(z - a) + A_2(z - a)^2 + \cdots
\]

that converges for \( |z - a| < \delta \). If \( z \neq a \),

\[
f(z) = \frac{g(z)}{(z - a)^n} = \frac{A_0}{(z - a)^n} + \frac{A_1}{(z - a)^{n-1}} + \cdots + \frac{A_{n-1}}{z - a} + A_n + \cdots
\]

Denoting \( B_k = A_{n+k} \) for \( k = -n, -n + 1, \ldots \)

\[
f(z) = \sum_{k=-n}^{\infty} B_k(z - a)^k \quad \text{for} \quad 0 < |z - a| < \delta
\]

This representation for \( f(z) \) in a neighborhood of an isolated pole is called the Laurent series for \( f \) around \( a \).

Example.

\[
f(z) = \frac{\cos(z)}{z^2} = \frac{1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \cdots}{z^2} = \frac{1}{z^2} + \frac{0}{z} + \left( \frac{1}{2} \right) + 0z + \frac{1}{4!}z^2 + \cdots
\]

Thus \( B_{-2} = 1, B_{-1} = 0, B_0 = -\frac{1}{2}, B_1 = 0 \), etc.

Residues

Definition. If \( f \) has an isolated pole at \( a \in \mathbb{C} \), the residue of \( f \) at \( a \) is

\[
\text{Res}_a f = B_{-1}
\]

(i.e., the coefficient of \((z - a)^{-1}\) in the Laurent series for \( f \) around \( a \)).
Examples.  1. \( f(z) = 1/(z-i) \). As a Laurent series around \( i \),

\[
f(z) = \frac{1}{(z-i)} + 0 + 0(z-i) + \cdots
\]

So \( \text{Res}_i f(z) = 1 \).

2. 

\[
f(z) = z + \frac{3}{(z-4)^2}
\]

As a Laurent series around \( 4 \),

\[
f(z) = \frac{3}{(z-4)^2} + \frac{0}{(z-4)} + 4 + (z-4) + 0(z-4)^2 + \cdots
\]

So \( \text{Res}_4 f(z) = 0 \).

3. 

\[
f(z) = \frac{e^z}{z^n} = \frac{1}{z^n} + \frac{1}{(n-1)!} + \cdots + \frac{1}{(n-1)!} z + \cdots
\]

Hence \( \text{Res}_0 f(z) = 1/(n-1)! \).

Suppose \( f \) has an isolated pole of order \( n \) at \( a \in \mathbb{C} \). Define \( g(z) = (z-a)^n f(z) \). Then, as a Taylor series around \( a \),

\[
g(z) = g(a) + g'(a)(z-a) + \frac{g''(a)}{2!}(z-a)^2 + \cdots + \frac{g^{(n-1)}(a)}{(n-1)!}(z-a)^{n-1} + \cdots
\]

So, for \( 0 < |z-a| < \delta \),

\[
f(z) = \frac{g(a)}{(z-a)^n} + \frac{g'(a)}{(z-a)^{n-1}} + \cdots + \frac{g^{(n-1)}(a)}{(n-1)!}(z-a)^{-1} + \cdots
\]

Thus

\[
\text{Res}_a f(z) = \frac{g^{(n-1)}(a)}{(n-1)!}
\]

This expression provides a formula for computation of residues.
EEE 550 Class Notes

Day 27 – Calculus of Residues

Many path integrals may be evaluated by computation of residues at poles. Techniques for computation of certain integrals on the real line, including some Fourier transforms, via residue computations are a well developed consequence of the study of analytic functions. The main enabling result, for which the proof is only roughly sketched here, is the following theorem.

**Theorem** (Residue theorem). Suppose $f$ is analytic in a simply connected set $S \subseteq \mathbb{C}$ except at isolated poles $a_1, a_2, \ldots$. If $\gamma$ is a closed path in $S$,

$$
\int_{\gamma} f(z) \, dz = 2\pi i \sum_k n(\gamma, a_k) \text{Res}_{a_k} f
$$

**Proof.** (Sketch) Consider the integral of $f$ on a closed path $\xi$ encircling the isolated pole $a_1$ and whose image set and interior contain no other poles of $f$. Considering $f$ as a Laurent series around $a_1$, the constant term, $(z - a_1)$ term, $(z - a_1)^2$ term and so forth are all analytic in the interior and image set of $\xi$, so these terms contribute nothing to the integral. Also, a simple estimate shows the $(z - a_1)^{-2}$ term, the $(z - a_1)^{-3}$ also integrate to zero around $\xi$. So only the residue term contributes to the integral, exactly in the amount

$$
\frac{1}{2\pi i} \text{Res}_{a_k} f
$$

The original path $\gamma$ is path homotopic to a path constructed by connecting small circles around the poles of $f$ contained in its interior. So the value of the integral of $f$ around $\gamma$ is $2\pi i$ times the sum of the residues at all the poles in the interior of $\gamma$, suitable weighted by winding numbers. \hfill \square

**Examples.**

1. 

$$
\int_{|z| = 1} \frac{e^z}{z^n} \, dz = 2\pi i \text{Res}_0 \frac{e^z}{z^n} = \frac{2\pi i}{(n-1)!}
$$

where the residue was computed previously. Note that this is the same answer obtained via Cauchy’s integral formula.

2. With $\gamma$ as pictured,

$$
\int_{\gamma} \frac{dz}{(z+2)(z-3)} = 2\pi i \left[ (-1) \text{Res}_{-2} \frac{1}{(z+2)(z-3)} + 2 \text{Res}_3 \frac{1}{(z+2)(z-3)} \right]
$$
The residues are
\[
\text{Res}_{z=2} \left( \frac{1}{(z+2)(z-3)} \right) = \frac{z+2}{(z+2)(z-3)} \bigg|_{z=2} = -\frac{1}{5}
\]
and
\[
\text{Res}_{z=3} \left( \frac{1}{(z+2)(z-3)} \right) = \frac{z-3}{(z+2)(z-3)} \bigg|_{z=3} = \frac{1}{5}
\]
So
\[
\int_{\gamma} \frac{dz}{(z+2)(z-3)} = \frac{6\pi i}{5}
\]

Integrals on the Real Line

As noted above, a central use of complex integration in signal analysis is for computation of transform integrals on \( \mathbb{R} \) that are difficult or impossible to evaluate by elementary methods.

**Example (Fourier integral).** Let \( f : \mathbb{R} \to \mathbb{R} \) by
\[
f(t) = \frac{1}{t^2 + a^2} \quad a > 0
\]

Note that \( f \) is even, so \( \hat{f} \) will also be even and it suffices to evaluate
\[
\hat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{-i\omega t} dt
\]
for \( \omega < 0 \).

**Step 1:** Extend the integrand to a function on \( \mathbb{C} \); i.e.,
\[
g(z) = \frac{e^{-i\omega z}}{z^2 + a^2}
\]
Note that \( g \) has poles at \( z = \pm a \).

**Step 2:** Construct a closed path of integration in \( \mathbb{C} \) that runs along the real axis.
Step 3: Write the closed-path integral in terms of elementary path integrals.

\[
\int_{\gamma_R + \Gamma_R} g(z) \, dz = \int_{\gamma_R} g(z) \, dz + \int_{\Gamma_R} g(z) \, dz = \int_{-R}^{R} g(t) \, dt + \int_{\Gamma_R} g(z) \, dz
\]

Step 4: Assume \( R \) is sufficiently large (since the limit \( R \to \infty \) is what is of interest) and apply the residue theorem.

\[
\text{Res}_{z=a} g = \frac{(z - ia)e^{-i\omega z}}{(z + ia)(z - ia)} \bigg|_{z=ia} = \frac{e^{\omega a}}{2ia}
\]

So,

\[
\int_{\gamma_R + \Gamma_R} g(z) \, dz = \frac{e^{\omega a}}{2ia}
\]

and

\[
\int_{-R}^{R} g(t) \, dt = \frac{e^{\omega a}}{2ia} - \int_{\Gamma_R} g(z) \, dz
\]

Step 5: Let \( R \to \infty \).

\[
\int_{-R}^{R} g(t) \, dt = \frac{e^{\omega a}}{2ia} - \lim_{R \to \infty} \int_{\Gamma_R} g(z) \, dz
\]

Step 6: Estimate the off-axis integral.

\[
\left| \int_{\Gamma_R} g(z) \, dz \right| \leq \ell(\Gamma_R) \sup_{z \in \Gamma_R} |g(z)| = \pi R \sup_{\theta \in [0,1]} \left| \frac{e^{-i\omega R \exp(i\pi \theta)}}{R^2 e^{2\pi i \theta} + a^2} \right|
\]

\[
= \pi R \sup_{\theta \in [0,1]} \left| \frac{e^{-i\omega R \cos(\pi \theta)} e^{i\omega R \sin(\pi \theta)}}{R^2 e^{2\pi i \theta} + a^2} \right| \leq \pi R \left| \frac{1}{R^2 - a^2} \right|
\]

The last inequality follows from
1. \( |\exp(-i\omega R \cos(\pi \theta))| = 1 \)

2. For \( 0 \leq \theta \leq 1, \sin(\pi \theta) \in [0, 1] \). Since \( \omega < 0 \) and \( R > 0 \), \( \exp[R \omega \sin(\pi \theta)] \leq 1 \)

3. The smallest possible denominator is \( |R^2 - a^2| \)

Thus
\[
\lim_{R \to \infty} \int_{\Gamma_R} g(z) \, dz = 0
\]

Step 7: Reach the conclusion.

\[
\hat{f}(\omega) = \frac{\pi e^{-|\omega|a}}{a} \quad \omega \neq 0
\]