ECE 4213/5213

MODULE 3

Index

DISTRIBUTION THEORY AND THE DIRAC DELTA

- For continuous-time signals $x(t)$, the natural basis was the translates of the Dirac delta $\{\delta(t-\tau)\}_{\tau\in\mathbb{R}}$.
- Writing $x(t)$ as a sum of this basis made it easy to understand why the output of a continuous-time LTI system is convolution.
- In Chapter 3, our main goal is to represent signals $x(t)$ and $x[n]$ in the "frequency domain", which will simply be a change of basis.

  → But, the Dirac delta will show up again.

- So we begin by spending some time to understand the Dirac delta better.

Plan

- Review the "usual" presentation of the delta "function."
- Understand why the "usual" presentation is mathematically incorrect.
- Understand why many books use the incorrect presentation.
- Understand the mathematically correct way to treat the Dirac delta.

- The "usual" definition of $\delta(t)$:
  - It is a unit area pulse that is "infinitely thin" and "infinitely tall":

$$\delta(t) = 0, \quad t \neq 0 \qquad (*)$$

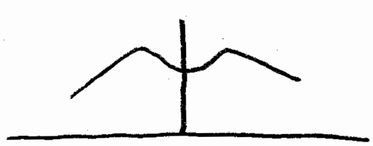
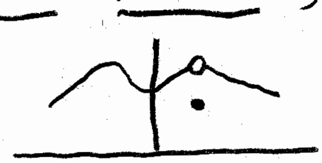
$$\int_{\mathbb{R}} \delta(t)\,dt = 1 \qquad (**)$$

  - The "sifting property":

$$\int_{\mathbb{R}} \delta(t)\,x(t)\,dt = x(0) \qquad (***)$$

$\Rightarrow$ How you know that something is <u>wrong</u> about this:

  - according to $(*)$, the integrand in $(**)$ differs from the function $x(t) = 0$ only at a single point.
  - In calculus, you learned that the value of a (Riemann) integral cannot be affected by the value of the integrand at <u>single</u> <u>points</u>; e.g.,


and

have <u>the</u> <u>same</u> <u>area</u> under the curve.

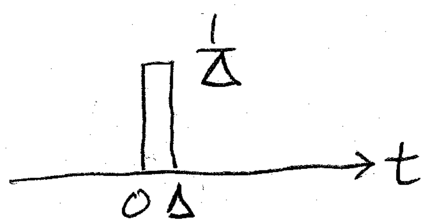
  $\rightarrow$ There can't be any area under a point.

- So, in the theory of the Riemann integral, the integral $(**)$ <u>must</u> be the <u>same</u> as

$$\int_{-\infty}^{\infty} 0\,dt = 0. \qquad ??$$

- Some books present it this way:

let $\delta_\Delta(t) = \begin{cases} \frac{1}{\Delta}, & 0 \le t \le \Delta \\ 0, & \text{otherwise} \end{cases}$

$=$ 

Then they attempt to define

$$\delta(t) = \lim_{\Delta \to 0} \delta_\Delta(t).$$

But this doesn't really fix the problems with (**) on the previous page...

we have

$$(**) = \int_{-\infty}^{\infty} \delta(t)\,dt = \int_{-\infty}^{\infty} \lim_{\Delta \to 0} \delta_\Delta(t)\,dt$$

→ Now they want to switch the order of the limit and the integral (which is also a limit, recall) and write

$$= \lim_{\Delta \to 0} \int_{-\infty}^{\infty} \delta_\Delta(t)\,dt = \lim_{\Delta \to 0} 1 = 1 \quad XX$$

→ But this is <u>INCORRECT</u>.

- to switch the order, the limit in the integrand must converge uniformly.

- But here the integrand is

$$\lim_{\Delta \to 0} \delta_\Delta(t)$$

which converges only pointwise, not uniformly. So it's not okay to swith the order of the limit and the integral in this case.

- Then why do so many books teach it this wrong way?

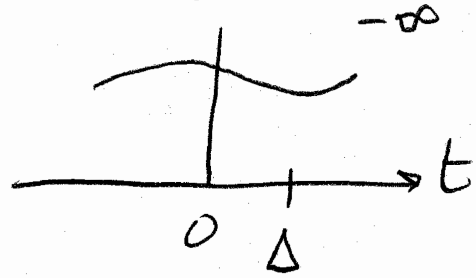
- Let's consider the $\delta_\Delta(t)$ model and take a look at it with respect to (***) on page 3-1:

$$\int_{-\infty}^{\infty} \delta(t)\, x(t)\, dt = x(0)$$

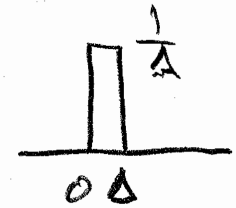
- Suppose $x(t)$ is extremely smooth or "slowly varying" with respect to tiny time intervals Δ.

- Then consider the integral $\int_{-\infty}^{\infty} \delta_\Delta(t)\, x(t)\, dt$. (*)

if $x(t) =$

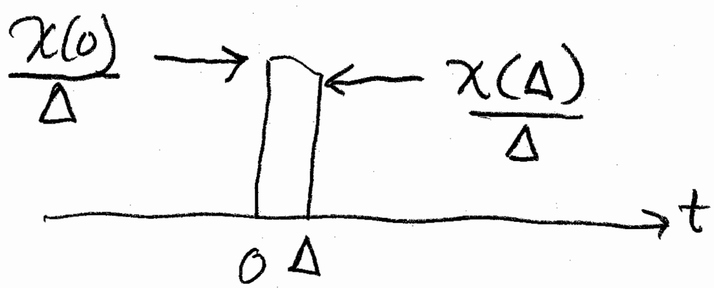


Then the integrand is $x(t) \cdot$



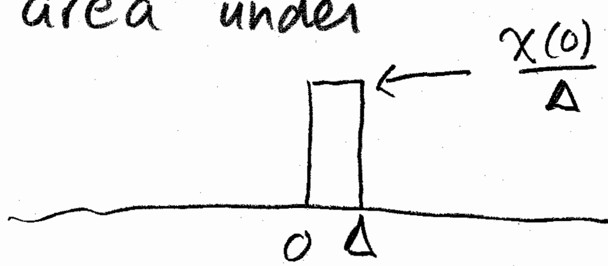
which look like this:

$x(t)\delta_\Delta(t) =$



(**)

Now, as Δ becomes very small, the area under (**) becomes only neglibly different from the area under



, wich is $x(0)$.

- So it seems like it should be okay to say that, for very small Δ, the area under (**) on page 3-4 is approximately equal to $x(0)$.

  → However, because the integrand again converges only pointwise in the limit as Δ vanishes,

  ⇒ The theory of Riemann calculus is not powerful enough to handle this situation.

We would <u>like</u> to write

$$\int_{\mathbb{R}} \delta(t)\, x(t)\, dt = \int_{\mathbb{R}} \lim_{\Delta \to 0} \delta_\Delta(t)\, x(t)\, dt$$

$$\overset{?}{\underset{\times}{=}} \lim_{\Delta \to 0} \int_{\mathbb{R}} \delta_\Delta(t)\, x(t)\, dt = x(0)$$

⇒ But it's not legal to do this because the integrand converges only pointwise.

- So what is the correct way to think about this??

- First, understand the reason that we want to have $\delta(t)$ in the first place:

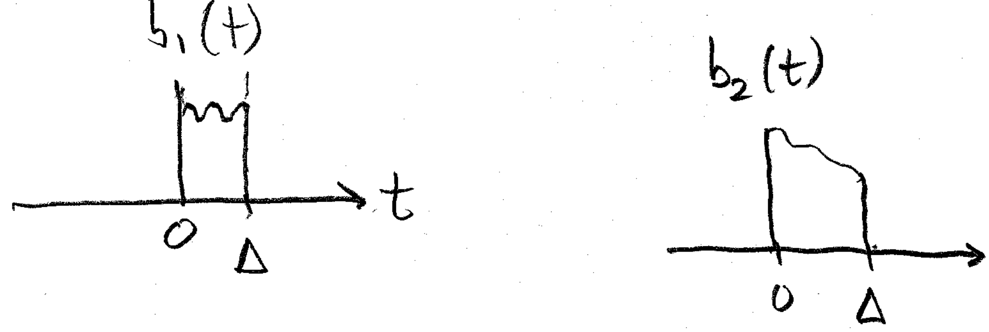
  - To analyze & design systems, we need to be able to write signals $x(t)$ as a weighted sum of a natural basis.

  - Our main concerns about the basis functions (aka "basis signals") are how they interact with other signals $x(t)$ in a dot product and how they interact with LTI systems (convolution... also a dot product).

    → So we don't so much care about the exact values that any basis signal takes at every $t \in \mathbb{R}$,

    so much as we <u>do care</u> about how each basis signal behaves in a dot product.

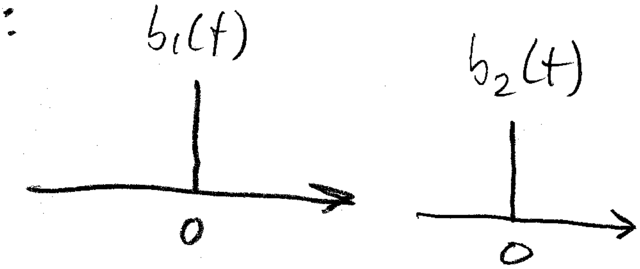
- Now consider two very narrow unit area signals $b_1(t)$ and $b_2(t)$ that look like:



- we have $\int_{-\infty}^{\infty} b_1(t)\,dt = \int_{-\infty}^{\infty} b_2(t)\,dt = 1$.

- Now suppose that Δ is tiny, like 10^{-43} sec.

⇒ Then on the oscilloscope, $b_1(t)$ and $b_2(t)$ both look the same:



→ Moreover, if we put $b_1(t)$ and $b_2(t)$ into an LTI system, the two outputs $y_1(t)$ and $y_2(t)$ will look identical on the scope... the differences will be too small to measure.

- Also, for any reasonable (e.g. smooth) $x(t)$, the difference between the two dot products

$$\langle x(t), b_1(t) \rangle = \int_{\mathbb{R}} x(t)\, b_1^*(t)\, dt \approx x(0)$$

and

$$\langle x(t), b_2(t) \rangle = \int_{\mathbb{R}} x(t)\, b_2^*(t)\, dt \approx x(0)$$

will be <u>too</u> <u>small</u> <u>to</u> <u>measure</u>.

$\Rightarrow$ As far as we can measure things,

so as far as we care,

$\Rightarrow$ There is no <u>practical</u> difference between $b_1(t)$ and $b_2(t)$ when Δ is so small.

- Thus, we would like to have <u>one</u> mathematical object to model $b_1(t)$, $b_2(t)$, and <u>all</u> the other little guys who behave exactly the same as $b_1(t)$ and $b_2(t)$ in a dot product or a system (at least as far as our ability to measure any practical difference).

- Having such a math model (object) would be nice because it would free us from having to worry about all the exact details of the values $b_1(t)$ for every t, and these details make no practical difference whatsoever to anything we can measure.

⇒ But this has two important consequences:

① The object that will model all of these indistinguishable little pulses <u>cannot</u> be a <u>function</u> ⟶

- The reason is: a function is a rule that assigns to each member of the domain exactly one "buddy" who is a member of the range.

  ⇒ So, by definition, to model $b_1(t)$ with a <u>function</u> you have to specify all the buddy assignments... which is exactly what we are trying to avoid doing.

2. A Riemann integral is the limit of the sum of the area under a Riemann partition as the widths of the elements become vanishingly small.

   ⇒ To work out the limit, you have to know all the buddy assignments of the integrand at a microscopic level... e.g., the integrand has to be a function.

<u>So</u>: to get the one math object we want to model all the little indistinguishable unit area pulses,

- we can't make it a function,
- and we won't be able to use Riemann calculus.

⇒ Recall: for any <u>one</u> guy like $b_1(t)$, we <u>can</u> model him as a function and we <u>can</u> use Riemann calculus on him.

⇒ It's only when we demand <u>one</u> math object to model not only $b_1(t)$, but also at the same time <u>all</u> of the little guys who are indistinguishable from $b_1(t)$, that we run into these problems (can't use a function, can't use Riemann integral).

Consider an LTI system H:

$$x(t) \longrightarrow \boxed{\begin{matrix}\text{LTI}\\ H\end{matrix}} \longrightarrow y(t)$$

- In any practical system, there will be a limit on the time resolution with which we can measure the input and output signals.
- There will be <u>many</u> very narrow unit-area pulses that we can't distinguish from one another.
- They will all make outputs $y(t)$ that we also can't distinguish from one another.

$\longrightarrow$ So it would be <u>nice</u> to have <u>one</u> math model to handle all of these narrow pulses, since we can't tell them apart anyway... and we can't detect any differences in how they behave in an LTI system.

☆ The Dirac delta $\delta(t)$ is a <u>single</u> math model for all of the members of this class of very narrow pulses.

- We can't measure the fine-scale structure of these signals.
- The exact "buddy matches" between the domain and range are <u>not</u> important to us.

$\Rightarrow$ What <u>is</u> important, what we <u>can</u> measure, and what we <u>do</u> care about is:

- How these narrow pulses interact with systems,
- How they behave in dot products with other signals,

$\Rightarrow$ In other words, we <u>do</u> care how they behave under an integral.

Recall: a function is a rule that matches each member of one set (the domain) to a unique member of a second set (the range).

→ This is exactly what we are trying to avoid with $\delta(t)$.

→ So the math model we seek for $\delta(t)$ cannot be a function. ☆☆

- Two more powerful ways to treat $\delta(t)$ were developed in 20th century mathematics:

1. Treat it as a measure. This is an important part of the modern theory of integration that is taught to graduate students in math. See "Lebesgue Integration" on Wikipedia to get an idea about this. For DSP, we will not do it this way, because the 2nd way is simpler.

2. Treat it as a special kind of functional known as a "generalized function" or "distribution." This is how we will treat $\delta(t)$ is DSP class.

☆ NOTE: This is a completely different use of the word "distribution" from what you are used to. It has nothing to do with "probability distributions."

A Good Book for the details:

A.H. Zemanian, "Distribution Theory and Transform Analysis: An Introduction to Generalized Functions, with Applications," Dover, 2010.

→ \$14.81 on amazon.com!

DEF: a functional is a function from a vector space to the underlying field.

- It assigns scalar buddies to vectors.

→ For our purposes, the domain will be a space of vectors or of functions... i.e., a space of signals.

→ The range will be numbers.

Recall: you have already seen an example of a functional: the norm. It is a function that maps each signal to a number, which we think of as "length."

$$\|x(t)\|_{L^p} = \left[\int_{\mathbb{R}} |x(t)|^p \, dt\right]^{1/p}$$

DEF: a functional f is linear if the action of the functional commutes with linear combinations:

$$f[a x_1(t) + b x_2(t)] = a f[x_1(t)] + b f[x_2(t)].$$

DEF: a functional f is continuous if the action of the functional commutes with uniformly convergent limits of sequences of functions (signals):

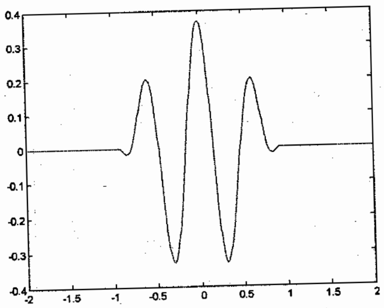
$$\lim_{k\to\infty} f[x_k(t)] = f\left[\lim_{k\to\infty} x_k(t)\right].$$

EX: Suppose that $\{x_k(t)\}_{k\in\mathbb{Z}}$ is a sequence of signals converging uniformly to the signal $x(t)=0$. For example,

$$x_k(t) = \begin{cases} \frac{1}{k} e^{\frac{1}{t^2-1}} \cos 10t, & |t|<1 \\ 0, & \text{other} \end{cases}$$

→

$x_1(t)$



- Then $\lim_{k\to\infty} x_k(t) = 0$ (a signal that's everywhere zero).

- Let f be a functional.

- If f is linear, then it follows that $f[0] = 0$ ... f must map the "zero signal" to the number zero.

- If f is also continuous, then

$$\lim_{k\to\infty} f[x_k(t)] = f\left[\lim_{k\to\infty} x_k(t)\right] = 0.$$

DEF: A distribution is a continuous linear functional.

DEF: $x(t)$ is locally integrable if $\forall a, b \in \mathbb{R}$ s.t. $-\infty < a < b < \infty$, $\int_a^b |x(t)|\,dt < \infty$.

$\Rightarrow$ In other words, if $|x(t)|$ has a finite integral on every finite interval.

- One easy way to construct distributions is by inner product with a locally integrable function.

EX: $u(t)$ is locally integrable. For a signal $x(t) \in L^1(\mathbb{R})$, the inner product with $u(t)$ is given by

$$\langle x(t), u(t)\rangle = \int_{-\infty}^{\infty} x(t)u^*(t)\,dt = \int_0^{\infty} x(t)\,dt = \text{a number}$$

$\longrightarrow$

- The mapping $x(t) \xmapsto{f} \langle x(t), u(t) \rangle$ is a continuous linear functional on $L^1(\mathbb{R})$ ... a <u>distribution</u>.

- The distributions that can be defined this way... by dot product with a locally integrable function... are called <u>regular</u> <u>distributions</u>.

- There are also perfectly good continuous linear functionals (distributions) that <u>cannot</u> be expressed as a dot product with a locally integrable function.

  ☆ These are called the <u>singular</u> <u>distributions</u>.

<u>EX</u>: $x(t) \xmapsto{f} x(0)$

$\uparrow$ a signal $\quad$ $\uparrow$ a number

- Although this is a perfectly good continuous linear functional, there is <u>no</u> locally integrable <u>function</u> $f(t)$ that can satisfy

$$\langle x(t), f(t) \rangle = \int_{-\infty}^{\infty} x(t) f^*(t)\, dt = x(0)$$

in general (i.e., for all the $x(t)$).

- It's easy to prove that such a function $f(t)$ <u>can</u> <u>not</u> exist.

- But this singular distribution <u>is</u> the Dirac delta $\delta(t)$.

- You should not think of it as an ordinary function.

- It is a model for how all the tiny unit-area pulses behave in a dot product.

- Mathematically, it is a functional that maps $x(t)$ to what the dot product should be.

- So, in terms of rigorous mathematics, the symbol $\delta(t)$ represents a tiny unit-area pulse, but it has meaning <u>only</u> when it's in a dot product with a signal like $x(t)$ or an impulse response like $h(t)$.

- Because it models the dot product of $x(t)$ with <u>any</u> tiny unit-area pulse, you <u>can't</u> think of

$$\int_{-\infty}^{\infty} x(t)\,\delta(t)\,dt = \langle x(t), \delta(t) \rangle = x(0) \qquad (***)$$

as a Riemann integral... it's <u>not</u>.

- What $\langle \cdot, \delta(t) \rangle$ is: it is a mapping that takes $x(t)$ to the number $x(0)$. There is no calculus involved.

- So why do we write an integral sign in (***) ??

$\Rightarrow$ It is for convenience. It does <u>not</u> mean integral and it does <u>not</u> imply integration.

- But in working with signals and distributions, we will need to be able to define <u>operations</u> on distributions, such as: scalar multiplication, addition, time shifting, time scaling, differentiation, etc...

- The theory of distributions provides a mathematically rigorous way to do this.

- The "symbolic" integral sign in (***) will help us to get the right answer when we do this, even though it is <u>not</u> an integral.

EX: Your undergrad signals & systems book said that $\frac{d}{dt}u(t) = \delta(t)$.

- It is true.
- But what it really means is: $\delta(t)$ and $\frac{d}{dt}u(t)$ are two singular distributions that both map every $x(t)$ to the same number, i.e.,

$$\langle x(t), \delta(t)\rangle = \langle x(t), \tfrac{d}{dt}u(t)\rangle = x(0).$$

$\Rightarrow$ We say that $\delta(t)$ and $\frac{d}{dt}u(t)$ are "equal in the sense of distributions."

- To define an operation on a distribution, the general strategy is to move the operation onto the signal.
  - For example, $\delta(t-t_0)$ means the mapping $\langle x(t), \delta(t-t_0)\rangle$.
  - In order to figure out what number $\delta(t-t_0)$ maps $x(t)$ to, we move the operation (time shift in this case) onto the signal.
  - Then we can figure it out from the definition of the unshifted distribution $\delta(t)$.

Recall: a regular distribution is defined by inner product with a locally integrable function:

$$\langle x(t), f(t) \rangle = \int_{-\infty}^{\infty} x(t) f^*(t)\, dt = \text{a number.}$$

- So, to define an operation on distributions:
  - First, use ordinary calculus to figure out how it works for the regular distributions.
  - Then, define it to be the same for the singular distributions.

$\Longrightarrow$ Since the operation will work the same for both the regular and singular distributions,
  - it will be helpful to write a fake integral sign for the singular distributions, just to help us keep track of the manipulations,
  - but always remember that a singular distribution is not really an integral... it's nothing more than a simple rule for assigning numbers to signals.

EX: time shifting.

The problem: given a distribution $f(t)$, we want to define the distribution $f(t-t_0)$.

- Start with the regular distributions:
  - Let $f(t)$ be a locally integrable function.
  - Then $\langle x(t), f(t)\rangle$ is a regular distribution.
  - So, using ordinary calculus, we have

$$\langle x(t), f(t-t_0)\rangle = \int_{-\infty}^{\infty} x(t) f^*(t-t_0)\, dt$$

$$\left\{\begin{array}{ll} \text{Let } \theta = t-t_0 & t = \theta + t_0 \\ d\theta = dt & dt = d\theta \\ \text{when } t\to\infty, & \theta\to\infty \\ \text{when } t\to-\infty, & \theta\to-\infty \end{array}\right.$$

$$= \int_{-\infty}^{\infty} x(\theta+t_0) f^*(\theta)\, d\theta = \underline{\underline{\langle x(t+t_0), f(t)\rangle}}.$$

- Note that this defines the new distribution $f(t-t_0)$ in terms of the known distribution $f(t)$.

- Now, for all distributions, including both regular and singular, we define the shifted distribution $f(t-t_0)$ by:

$$\langle x(t), f(t-t_0)\rangle \equiv \langle x(t+t_0), f(t)\rangle .$$

→ Note that this moves the operation from the distribution to the signal.

- Applying this definition to the Dirac delta, we have

$$\langle x(t), \delta(t-t_0)\rangle = \langle x(t+t_0), \delta(t)\rangle$$
$$= x(t+t_0)\Big|_{t=0}$$
$$= x(t_0).$$

<u>Note</u>: this agrees with the "sifting property":

$$\int_{-\infty}^{\infty} x(t)\,\delta(t-t_0)\,dt = x(t_0)$$

→ But this is <u>not</u> really an integral!!

<u>Note</u>: given the definition $\langle x(t), \delta(t)\rangle = x(0)$, we can get this same answer by manipulating a <u>fake</u> integral:

$$\langle x(t), \delta(t-t_0)\rangle = \underbrace{\int_{-\infty}^{\infty} x(t)\,\delta(t-t_0)\,dt}_{\text{fake integral}}$$

$$= \int_{-\infty}^{\infty} x(\theta+t_0)\,\delta(\theta)\,d\theta \quad \leftarrow \left(\begin{array}{c}\text{change of}\\ \text{variable}\end{array}\right)$$

$$= \langle x(t+t_0), \delta(t)\rangle$$

$$= x(t_0) \quad \leftarrow \left(\begin{array}{c}\underline{\text{real}}\text{ math... but}\\ \text{not calculus}\end{array}\right)$$

Product of $\delta(t)$ with a signal

- if $f(t)$ is a regular distribution and $x(t)$ is a signal,
  - Then $x(t)f(t)$ is also a regular distribution (provided that $x(t)f(t)$ is still locally integrable).
- So for a signal $y(t)$, we have

$$\langle y(t), x(t)f(t)\rangle = \int_{-\infty}^{\infty} y(t)\,[x(t)f(t)]^* \, dt$$

$$= \int_{-\infty}^{\infty} [y(t)\,x^*(t)]\, f^*(t)\, dt$$

$$= \langle y(t)\,x^*(t), f(t)\rangle .$$

- Applying this to the product $x(t)\delta(t)$, we get

$$\langle y(t), x(t)\delta(t)\rangle = \langle y(t)\,x^*(t), \delta(t)\rangle$$

$$= y(0)\,x^*(0)$$

$$= \langle y(t), x(0)\,\delta(t)\rangle .$$

- So, for a signal $x(t)$,

$$x(t)\,\delta(t) = x(0)\,\delta(t)$$

(in the sense of distributions)

- since $\delta(t)$ is a model for the class of very narrow unit-area pulses concentrated at $t=0$,
- we have that for a bounded signal $x(t)$,

$$\lim_{t\to\infty} x(t)\delta(t) = \lim_{t\to-\infty} x(t)\delta(t) = 0$$

Derivative of a Distribution

- if $f(t)$ is a locally integrable function that is differentiable, then

$$\langle x(t), f'(t)\rangle = \int_{-\infty}^{\infty} x(t)\left[f'(t)\right]^{*} dt \qquad \begin{cases} dv = f'(t)^{*}\,dt \\ v = f^{*}(t) \\ u = x(t) \\ du = x'(t)\,dt \end{cases}$$

$$= \int_{-\infty}^{\infty} u\, dv$$

$$= uv\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} v\, du$$

$$= x(t)f^{*}(t)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f^{*}(t)\, x'(t)\, dt$$

$$= \lim_{t\to\infty} x(t)f^{*}(t) - \lim_{t\to-\infty} x(t)f^{*}(t) - \int_{-\infty}^{\infty} x'(t) f^{*}(t)\, dt$$

$$= \lim_{t\to\infty} x(t)f^{*}(t) - \lim_{t\to-\infty} x(t)f^{*}(t) - \langle x'(t), f(t)\rangle$$

- This <u>defines</u> the distributional derivative.

- Note that the differentiation is moved onto the signal $x(t)$.

- Applying this definition to the singular distribution $\delta(t)$, we get

$$\langle x(t), \delta'(t)\rangle = \lim_{t\to\infty} x(t)\delta(t) - \lim_{t\to -\infty} x(t)\delta(t) - \langle x'(t), \delta(t)\rangle$$

$$= 0 - 0 - x'(0)$$

$$= -x'(0)$$

- This follows from the top of p. 3-23.
  → So it's not valid for unbounded signals like e^{2t}.

- $\delta'(t)$ is called the "unit doublet".
- Higher-order derivatives of $\delta(t)$:

$$\langle x(t), \delta^{(k)}(t)\rangle = (-1)^k x^{(k)}(0).$$

- More properties of $\delta(t)$:

$$\delta(at) = \frac{1}{|a|}\delta(t)$$

$$x(t) * \delta(t-t_0) = x(t-t_0)$$

$$x(t) * \delta'(t-t_0) = x'(t-t_0)$$

- The theory of distributions is set up to make it easy to compute the mappings that are made by a distribution that has an operation applied to it (like a time shift).
  - For a singular distribution, you do this by writing a "fake" integral for the inner product and then manipulating it by the ordinary rules of calculus.

- Then why am I wasting your time by telling you all of this ????

⇒ Because: if you <u>forget</u>, and you <u>incorrectly</u> think that these fake integral signs really mean integration,

☆ Then you can get into <u>serious</u> <u>trouble</u>, get <u>seriously</u> <u>confused</u>, and get <u>wrong</u> <u>answers</u>.

→

Here is an example:

- your undergrad signals & systems book said: $\mathcal{F}\{\cos \Omega_0 t\} = \pi[\delta(\Omega-\Omega_0)+\delta(\Omega+\Omega_0)]$.
- Your book probably "proved" this as follows:

$$\mathcal{F}^{-1}\{\pi[\delta(\Omega-\Omega_0)+\delta(\Omega+\Omega_0)]\}$$
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\pi[\delta(\Omega-\Omega_0)+\delta(\Omega+\Omega_0)]e^{j\Omega t}\,d\Omega$$
$$= \frac{1}{2}\int_{-\infty}^{\infty} e^{j\Omega t}\delta(\Omega-\Omega_0)\,d\Omega + \frac{1}{2}\int_{-\infty}^{\infty} e^{j\Omega t}\delta(\Omega+\Omega_0)\,d\Omega$$
$$\left(= \frac{1}{2}\langle e^{j\Omega t}, \delta(\Omega-\Omega_0)\rangle + \frac{1}{2}\langle e^{j\Omega t}, \delta(\Omega+\Omega_0)\rangle \right)$$

This part was not in your book

$$= \frac{1}{2}e^{j\Omega_0 t} + \frac{1}{2}e^{-j\Omega_0 t} = \cos\Omega_0 t.$$

- But those are <u>all</u> "fake" integrals. In reality, there is no calculus at all in the above.
- If you incorrectly believe that they <u>are</u> really integrals, then you will conclude that it must be possible to use Riemann integration to work out this integral as well:

$$\mathcal{F}\{\cos\Omega_0 t\} = \int_{-\infty}^{\infty}\cos\Omega_0 t\, e^{-j\Omega t}\,dt = ?$$

→ But it is <u>not</u> possible... and if you try you will become hopelessly confused and very likely get wrong answers.

- The theory of Riemann integration that you learned in calculus is not powerful enough to handle integrals such as $\mathcal{F}\{\cos(\Omega_0 t)\} = \int_{\mathbb{R}} \cos(\Omega_0 t) e^{-j\Omega t} dt$

  → As a Riemann integral, this is <u>divergent</u>.

- But it can be interpreted and evaluated meaningfully by carefully applying 20th century mathematics.

- We will use distribution theory and interpret it as a distributional integral ... that converges not as a function, but as a distribution.

- We will need the following important result:

<u>Riemann-Lebesgue Lemma (RLL)</u>

- The RLL states that, in the sense of distributions,

$$\lim_{\Omega\to\infty} e^{j\Omega t} = 0.$$

- It means that, in dot products and in systems, the <u>distribution</u> $\lim_{\Omega\to\infty} e^{j\Omega t}$ behaves exactly like the regular distrubution $s(t) = 0$.

- In other words, for any $x(t) \in L^p(\mathbb{R})$,

$$\langle x(t), \lim_{\Omega\to\infty} e^{j\Omega t}\rangle = \langle x(t), 0\rangle$$

$$= \int_{\mathbb{R}} x(t)\cdot 0 \, dt = 0.$$

- This also implies that $\lim_{\Omega\to\infty} \cos\Omega t = 0$ and $\lim_{\Omega\to 0} \sin\Omega t = 0$ in the sense of distributions.

- For our purposes, the most important consequence of the Riemann-Lebesgue Lemma is:

FACT: $$\lim_{A\to\infty} \frac{\sin At}{\pi t} = \delta(t) \qquad \star$$

in the sense of distributions.

- It means that, for any $x(t) \in L^p(\mathbb{R})$,

$$\left\langle x(t), \lim_{A\to\infty} \frac{\sin At}{\pi t} \right\rangle = \langle x(t), \delta(t) \rangle = x(0).$$

- We will also refer to this important result as the RLL.

EX: calculate $\mathcal{F}\{\cos \Omega_0 t\}$ as a distributional Fourier transform:

$$\mathcal{F}\{\cos \Omega_0 t\} = \int_{\mathbb{R}} \cos(\Omega_0 t)\, e^{-j\Omega t}\, dt$$

$$= \int_{\mathbb{R}} \left[\tfrac{1}{2} e^{j\Omega_0 t} + \tfrac{1}{2} e^{-j\Omega_0 t} \right] e^{-j\Omega t}\, dt$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} e^{-j(\Omega-\Omega_0)t}\, dt + \frac{1}{2} \int_{-\infty}^{\infty} e^{-j(\Omega+\Omega_0)t}\, dt$$

$$= \lim_{A\to\infty} \frac{1}{2} \int_{-A}^{A} e^{-j(\Omega-\Omega_0)t}\, dt + \lim_{A\to\infty} \frac{1}{2} \int_{-A}^{A} e^{-j(\Omega+\Omega_0)t}\, dt$$

$$= \lim_{A\to\infty} \frac{1}{2} \left[\frac{1}{-j(\Omega-\Omega_0)} e^{-j(\Omega-\Omega_0)t} \right]_{t=-A}^{A} + \lim_{A\to\infty} \frac{1}{2} \left[\frac{1}{-j(\Omega+\Omega_0)} e^{-j(\Omega+\Omega_0)t} \right]_{t=-A}^{A}$$

$$= \lim_{A\to\infty} \frac{1}{2} \left[\frac{e^{-j(\Omega-\Omega_0)A} - e^{+j(\Omega-\Omega_0)A}}{-j(\Omega-\Omega_0)} \right] + \lim_{A\to\infty} \frac{1}{2} \left[\frac{e^{-j(\Omega+\Omega_0)A} - e^{-j(\Omega+\Omega_0)A}}{-j(\Omega+\Omega_0)} \right]$$

$$= \lim_{A\to\infty} \left[\frac{e^{j(\Omega-\Omega_0)A} - e^{-j(\Omega-\Omega_0)A}}{2j(\Omega-\Omega_0)} \right] + \lim_{A\to\infty} \left[\frac{e^{j(\Omega+\Omega_0)A} - e^{-j(\Omega+\Omega_0)A}}{2j(\Omega+\Omega_0)} \right]$$

$$\cdots = \lim_{A\to\infty} \frac{\sin[(\Omega-\Omega_0)A]}{\Omega-\Omega_0} + \lim_{A\to\infty} \frac{\sin[(\Omega+\Omega_0)A]}{\Omega+\Omega_0}$$

$$= \pi \left\{ \lim_{A\to\infty} \frac{\sin[A(\Omega-\Omega_0)]}{\pi(\Omega-\Omega_0)} + \lim_{A\to\infty} \frac{\sin[A(\Omega+\Omega_0)]}{\pi(\Omega+\Omega_0)} \right\}$$

$$= \pi\left[\delta(\Omega-\Omega_0) + \delta(\Omega+\Omega_0)\right]$$

"in the sense of distributions."

→ The last step follows from the RLL on page 3-25 by substituting $(\Omega-\Omega_0)$ and $(\Omega+\Omega_0)$ for "t".

Now back to the main point of our story:

- In chapter 2, we wrote our signals $x(t)$ as linear combinations of the natural basis $\{\delta(t-t_0)\}_{t_0\in\mathbb{R}}$.
- This is called the <u>time domain</u> representation of $x(t)$.
- This representation made it easy to see that the output of an LTI system is given by the convolution of the input with the impulse response.

- The translates of $\delta(t)$ form a basis for a very large space of signals that includes all of the L^p spaces as subspaces.
    - So even though $\delta(t)$ is not a function and does not itself belong to any of the L^p spaces,
        - we can use this basis to represent any $x(t)$ that is in any of the L^p spaces.

- Now we will show that, in a distributional sense, the natural basis $\{\delta(t-t_0)\}_{t_0 \in \mathbb{R}}$ is an <u>orthonormal</u> <u>basis</u>.

- Let $t_1, t_2 \in \mathbb{R}$ be real constants and consider the basis vectors $\delta(t-t_1)$ and $\delta(t-t_2)$.
- The inner product between them is given by

$$\langle \delta(t-t_1), \delta(t-t_2) \rangle = \int_{\mathbb{R}} \delta(t-t_1)\, \delta^*(t-t_2)\, dt$$

$$= \int_{\mathbb{R}} \delta(t-t_1)\, \delta(t-t_2)\, dt$$

$$= \delta(t_1 - t_2). \qquad (*)$$

$\Rightarrow$ If $t_1 \neq t_2$, then $\delta(t-t_1)$ and $\delta(t-t_2)$ are two <u>different</u> basis vectors and

$$(*) = \delta(t_1 - t_2) = \delta(\text{not zero}) = 0$$

→ The dot product between any two <u>distinct</u> basis vectors is <u>zero</u>. This shows that the basis is orthogonal.

$\Rightarrow$ If $t_1 = t_2$, then $\delta(t-t_1)$ and $\delta(t-t_2)$ are two different names for the same basis vector.

→ Then $(*)$ above is the dot product of a basis vector with itself, which <u>must</u> be the norm squared.

→ We have $(*) = \delta(t_1 - t_2) = 1\,\delta(0)$.

↑ norm squared... in a distributional sense

- This shows that the set of basis vectors (basis "signals") $\{\delta(t-t_0)\}_{t_0 \in \mathbb{R}}$ are all mutually orthogonal and, in a distributional sense, all have unit norm.

$\Rightarrow$ Thus, it is an orthonormal basis in the sense of distributions.

<u>FACT</u>: the set of signals $\{e^{j\Omega t}\}_{\Omega \in \mathbb{R}}$ is an orthogonal basis for a very large space of signals that also includes all of the L^p spaces and most of the $x(t)$ that we care about.

$\rightarrow$ But it is <u>not</u> an orthonormal basis.

$\rightarrow$ In the sense of distributions, each vector (signal) in this basis has norm $\sqrt{2\pi}$, not one.

- Let's show this:

$$\langle e^{j\Omega_1 t}, e^{j\Omega_2 t}\rangle = \int_{\mathbb{R}} e^{j\Omega_1 t} e^{-j\Omega_2 t}\, dt$$

$$= \int_{-\infty}^{\infty} e^{j(\Omega_1-\Omega_2)t}\, dt = \lim_{A\to\infty} \int_{-A}^{A} e^{j(\Omega_1-\Omega_2)t}\, dt$$

$$= \lim_{A\to\infty} \frac{1}{j(\Omega_1-\Omega_2)} \left[e^{j(\Omega_1-\Omega_2)t} \right]_{t=-A}^{A}$$

$$= \lim_{A\to\infty} \frac{2}{\Omega_1-\Omega_2} \; \frac{e^{j(\Omega_1-\Omega_2)A} - e^{-j(\Omega_1-\Omega_2)A}}{2j}$$

$\longrightarrow$

$$\ldots = \lim_{A\to\infty} \frac{2\sin[A(\Omega_1-\Omega_2)]}{\Omega_1-\Omega_2}$$

$$= 2\pi \lim_{A\to\infty} \frac{\sin[A(\Omega_1-\Omega_2)]}{\pi(\Omega_1-\Omega_2)} = 2\pi\delta(\Omega_1-\Omega_2). \quad \text{by the RLL}$$

- If $\Omega_1 \neq \Omega_2$, then the dot product is zero, which shows that the basis is orthogonal.
- If $\Omega_1 = \Omega_2$, then the dot product is $\underbrace{2\pi\delta(0)}_{\text{norm squared}}$, which shows that, in a distributional sense, each basis vector has norm $\sqrt{2\pi}$.
- This basis is called the "spectral basis" or the "Fourier basis".
- Writing $x(t)$ as a linear combination of this basis is called the "frequency domain" representation of $x(t)$.

NOTE: if we used Hertzian frequency "$2\pi f$" instead of radian frequency Ω, then the basis would be $\{e^{j2\pi f}\}_{f\in\mathbb{R}}$, which IS orthonormal.

- But the book uses radian frequency, so we will too.

- The big question now is: why would you want to write $x(t)$ as a sum of this spectral basis (instead of the natural basis) ???

- The answer is: just like a change of coordinates (basis) sometimes makes a calculus problem easier, a change to the spectral basis sometimes makes it easier to analyze a signal or its behavior in an LTI system.

<u>Recall</u>:

- if A is a matrix, $\vec{u}$ is a vector, and $\lambda \in \mathbb{C}$ is a constant,
- and if $A\vec{u} = \lambda \vec{u}$,

→ Then $\vec{u}$ is called an <u>eigenvector</u> of A with associated <u>eigenvalue</u> λ.

<u>EX</u>:

$$A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix} \qquad \vec{u} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

$$A\vec{u} = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 8 \end{bmatrix} = 4\vec{u}.$$

→ $\vec{u}$ is an eigenvector of the matrix A with associated eigenvalue $\lambda = 4$.

→ Multiplication by A doesn't change the direction of $\vec{u}$. It just "stretches" the length by $\lambda = 4$.

- For a continuous-time system H, there may be certain inputs $x(t)$ such that

$$x(t) \longrightarrow \boxed{H} \longrightarrow y(t) = \lambda x(t)$$

→ In this case, we call $x(t)$ an <u>eigenfunction</u> of the system H with associated eigenvalue λ.

<u>FACT</u>: the signal $e^{j\Omega t}$ is an eigenfunction of <u>any</u> continuous-time LTI system.

- The eigenvalue depends on the specific value Ω and on the specific system (i.e., on $h(t)$).

<u>Proof</u>; Let $\Omega_0 \in \mathbb{R}$ and let $x(t) = e^{j\Omega_0 t}$. Let H be a continuous-time LTI system with impulse response $h(t)$:

$$x(t) \longrightarrow \boxed{\begin{matrix}\text{LTI}\\ H\end{matrix}} \longrightarrow y(t) = x(t) * h(t).$$

Then
$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(t-\tau)\, h(\tau)\, d\tau$$

$$= \int_{-\infty}^{\infty} e^{j\Omega_0 (t-\tau)}\, h(\tau)\, d\tau$$

$$= \underbrace{e^{j\Omega_0 t}}_{x(t)} \underbrace{\int_{-\infty}^{\infty} h(\tau) e^{-j\Omega_0 \tau}\, d\tau}_{\text{a number... there's no "}t\text{" in it... call it } \lambda.}$$

$$= \lambda x(t)$$

- This shows that the spectral basis is a basis of eigenfunctions of any LTI system.

- If we write the system input $x(t)$ as a sum of these basis vectors,
  - → then every term in the sum is an eigenfunction.
  - → the action of the system is simply to multiply each term by the appropriate eigenvalue.

⇒ Making a change of basis from the natural basis to the spectral basis turns convolution into multiplication.

- On page 3-31, we saw that the needed eigenvalues are given by
$$\lambda = \int_{-\infty}^{\infty} h(\tau) e^{-j\Omega_0 \tau} d\tau$$
  - → From your undergrad signals and systems class, you can see that this is the Fourier transform of $h(t)$ evaluated at Ω_0, the frequency of the input signal.
  - → In other words, the eigenvalues are given precisely by the system "frequency response"... more on this a little later.

- So, let's take a signal $x(t) \in L^2(\mathbb{R})$ and write him as a weighted sum of the spectral basis $\{e^{j\Omega t}\}_{\Omega \in \mathbb{R}}$.

<u>Step 1</u>: Compute the dot product of $x(t)$ with the Ω^{th} basis signal:

$$\langle x(t), e^{j\Omega t} \rangle = \int_{\mathbb{R}} x(t)\, e^{-j\Omega t}\, dt$$

= a complex number (for each basis signal)

= a complex number for each $\Omega \in \mathbb{R}$

= a complex-valued function of Ω.

$\Rightarrow$ This is called the <u>Fourier</u> <u>Transform</u>

- You get a dot product (complex number) for each basis vector,

  e.g. you get one for each $\Omega \in \mathbb{R}$.

- So you can think of the Fourier Transform as a function of Ω.

  - We usually write $X(\Omega)$ or sometimes $X(j\Omega)$.

    - The second notation arises from the fact that one often uses a partial fraction expansion to invert the Fourier transform; you do the PFE in terms of "$j\Omega$".

- To show $X(\Omega)$ graphically, you have to use separate graphs for the real and imaginary parts,

  or, more comonly,

  for the magnitude and phase.

<u>Step 2</u>: add up the dot products times the basis vectors to get $x(t)$.

$\Rightarrow$ Don't forget to divide by 2π to account for the fact that the basis is orthogonal but not orthonormal:

$$x(t) = \frac{1}{2\pi}\int_{\mathbb{R}} X(\Omega)\, e^{j\Omega t}\, d\Omega .$$

$\Rightarrow$ This is called the <u>Inverse Fourier Transform</u>.

We write:

$$x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$$

$$X(\Omega) = \mathcal{F}\{x(t)\}$$

$$x(t) = \mathcal{F}^{-1}\{X(\Omega)\} \qquad (*)$$

<u>Note</u>: the inverse F.T. $(*)$ converges to the <u>midpoints</u> of discontinuities in $x(t)$.

- if $x(t) \in L^1(\mathbb{R})$, then $X(\Omega)$ is convergent as a Riemann integral.

- if $x(t) \in L^2(\mathbb{R})$, then $X(\Omega)$ is convergent as a Riemann integral or as a Cauchy Principle value.

$$x(t) \longrightarrow \overset{LTI}{\boxed{H}} \longrightarrow y(t)$$

- So our plan is to write the system input $x(t)$ as a weighted sum of the spectral basis

  - All the basis functions are eigenfunctions of the system.
  - Each term in the sum comes through the system times a complex eigenvalue

- Look back at page 3-34:
  $\Rightarrow$ The eigenvalues are given by

$$H(\Omega) = \int_{\mathbb{R}} h(t) e^{-j\Omega t} dt = \mathcal{F}\{h(t)\}.$$

$H(\Omega)$ is called the <u>Frequency Response</u> of the LTI system $\mathcal{H}$.

- you can think of the graph of $H(\Omega)$ as a tabulation of the eigenvalues that are associated with the eigenfunctions that make up the spectral basis.

- This whole idea is made precise by the "convolution property" of the Fourier Transform:

if $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$, $h(t) \overset{\mathcal{F}}{\longleftrightarrow} H(\Omega)$,

and $y(t) = x(t) * h(t)$,

then $y(t) \overset{\mathcal{F}}{\longleftrightarrow} \underline{\underline{Y(\Omega) = X(\Omega)H(\Omega)}}$.

$\Rightarrow$ You should already know how to prove this from your undergrad signals and systems class.

so

$$x(t) \longrightarrow \boxed{\begin{matrix}\text{LTI}\\ H\end{matrix}} \longrightarrow y(t)$$

$$y(t) = x(t) * h(t)$$

$$Y(j\Omega) = X(j\Omega)H(j\Omega)$$

- $Y(j\Omega)$ expresses the output $y(t)$ as a linear composition of the spectral basis:

$$y(t) = \frac{1}{2\pi}\int_{\mathbb{R}} Y(j\Omega)e^{j\Omega t}\,d\Omega$$

- $X(j\Omega)$ expresses the input $x(t)$ as a linear composition of the spectral basis:

$$x(t) = \frac{1}{2\pi}\int_{\mathbb{R}} X(j\Omega)\,e^{j\Omega t}\,d\Omega$$

- The eigenvalues $H(j\Omega)$ are given precisely by the coordinates of $h(t)$ with respect to the spectral basis:

$$H(j\Omega) = \int_{\mathbb{R}} h(t)e^{-j\Omega t}\,dt$$

$$h(t) = \frac{1}{2\pi}\int_{\mathbb{R}} H(j\Omega)e^{j\Omega t}\,d\Omega$$

- We think of $x(t)$ as a weighted sum of complex sinusoids: (3-42)

$$x(t) = \frac{1}{2\pi}\int_{\mathbb{R}} \underbrace{X(\Omega)}_{\text{weights}} \underbrace{e^{j\Omega t}}_{\text{sinusoids}} \, d\Omega$$

- When $x(t)$ is input to an LTI system H, each one of these terms gets multiplied by an eigenvalue when it goes through the system:

Input term in $x(t)$	$\longmapsto$ Makes $\longmapsto$	Output term in $y(t)$
$X(\Omega_1)e^{j\Omega_1 t}$	$\longmapsto$	$X(\Omega_1)H(\Omega_1)e^{j\Omega_1 t}$
$X(\Omega_2)e^{j\Omega_2 t}$	$\longmapsto$	$X(\Omega_2)H(\Omega_2)e^{j\Omega_2 t}$
$X(\Omega_3)e^{j\Omega_3 t}$	$\longmapsto$	$X(\Omega_3)H(\Omega_3)e^{j\Omega_3 t}$
$\vdots$		$\vdots$

- and we think of $y(t)$ as being a sum of all the terms of this form:

$$y(t) = \frac{1}{2\pi}\int_{\mathbb{R}} Y(\Omega)e^{j\Omega t}\,d\Omega = \frac{1}{2\pi}\int_{\mathbb{R}} X(\Omega)H(\Omega)e^{j\Omega t}\,d\Omega$$

In Pictures:

$$x(t) = \frac{1}{2\pi}\int_{\mathbb{R}} X(\Omega)\, e^{j\Omega t}\, d\Omega \longrightarrow \boxed{\begin{matrix}\text{LTI} \\ H\end{matrix}} \longrightarrow y(t) = \frac{1}{2\pi}\int_{\mathbb{R}} Y(\Omega)\, e^{j\Omega t}\, d\Omega$$

$$\begin{matrix} \vdots & & & & \vdots \\ X(\Omega_1)e^{j\Omega_1 t} & \longrightarrow & & \longrightarrow & X(\Omega_1)H(\Omega_1)e^{j\Omega_1 t} \\ X(\Omega_2)e^{j\Omega_2 t} & \longrightarrow & \boxed{\begin{matrix}\text{LTI} \\ H \\ h(t) \\ H(\Omega)\end{matrix}} & \longrightarrow & X(\Omega_2)H(\Omega_2)e^{j\Omega_2 t} \\ X(\Omega_3)e^{j\Omega_3 t} & \longrightarrow & & \longrightarrow & X(\Omega_3)H(\Omega_3)e^{j\Omega_3 t} \\ \vdots & & & & \vdots \end{matrix}$$

$$y(t) = \frac{1}{2\pi}\int_{\mathbb{R}} X(\Omega)H(\Omega)\, e^{j\Omega t}\, d\Omega$$

i.e., the coordinates of $y(t)$ with respect to the spectral basis are given by

$$Y(\Omega) = X(\Omega)\,H(\Omega).$$

- What is the effect of multiplying a sinusoidal input term by a complex eigenvalue?

  - suppose the input term is

$$e^{j\Omega_0 t} = \cos\Omega_0 t + j\sin\Omega_0 t$$

  - suppose the eigenvalue is

$$H(\Omega_0) = Ae^{j\theta}, \text{ where } A, \theta \in \mathbb{R}.$$

  - then the corresponding output term is

$$H(\Omega_0)e^{j\Omega_0 t} = Ae^{j\theta}e^{j\Omega_0 t}$$

$$= Ae^{j(\Omega_0 t + \theta)}$$

$$= A\cos(\Omega_0 t + \theta) + jA\sin(\Omega_0 t + \theta)$$

e.g., the magnitudes of the real and imaginary parts get scaled by the magnitude of the eigenvalue, and the phase gets shifted by the angle of the eigenvalue.

- if the impulse response is <u>real</u>, then the eigenvalues (frequency response) are conjugate symmetric, and there is a corresponding result for real sinusoids:

$$h(t) \in \mathbb{R} \longrightarrow H(\Omega) = H^*(-\Omega).$$

- if the input is $\cos\Omega_0 t$, then the output is

$$\begin{aligned} y(t) &= h(t) * \cos\Omega_0 t \\ &= h(t) * \left[\tfrac{1}{2}e^{j\Omega_0 t} + \tfrac{1}{2}e^{-j\Omega_0 t}\right] \\ &= \tfrac{1}{2}h(t) * e^{j\Omega_0 t} + \tfrac{1}{2}h(t) * e^{-j\Omega_0 t} \\ &= \tfrac{1}{2}H(\Omega_0)e^{j\Omega_0 t} + \tfrac{1}{2}H(-\Omega_0)e^{-j\Omega_0 t} \\ &= \tfrac{1}{2}|H(\Omega_0)|\, e^{j[\Omega_0 t + \arg H(\Omega_0)]} \\ &\quad + \tfrac{1}{2}|H(-\Omega_0)|\, e^{j[-\Omega_0 t + \arg H(-\Omega_0)]} \end{aligned}$$

$\longrightarrow$

$$\ldots y(t) = \tfrac{1}{2}|H(\Omega_0)|e^{j[\Omega_0 t + \arg H(\Omega_0)]}$$

$$+ \tfrac{1}{2}\underbrace{|H(\Omega_0)|}\, e^{j[-\Omega_0 t \underbrace{- \arg H(\Omega_0)}]}$$

because $h(t)$ is real, $H(\Omega)$ is conjugate symmetric, so the magnitude is even and the phase is odd.

$$= |H(\Omega_0)| \left\{ \tfrac{1}{2}e^{j[\Omega_0 t + \arg H(\Omega_0)]} + \tfrac{1}{2}e^{-j[\Omega_0 t + \arg H(\Omega_0)]} \right\}$$

$$= |H(\Omega_0)| \cos\left[\Omega_0 t + \arg H(\Omega_0)\right]$$

→ i.e., the magnitude is scaled by the magnitude of the eigenvalue and the phase is shifted by the angle of the eigenvalue.

→ The same is true for $\sin \Omega_0 t$ when $h(t)$ is real.

Distributional Fourier Transforms

- A general $x(t)$ can be written as

$$x(t) = x_a(t) + x_p(t)$$

where $x_a(t)$ is <u>aperiodic</u> and $x_p(t)$ is <u>periodic</u>.

→ $x_a(t)$ can generally be represented as a Fourier integral using Riemann calculus.

→ $x_p(t)$ can generally be represented as a <u>Fourier Series</u>.

→ To represent $x_p(t)$ with a Fourier integral, one must use 20th century math... measure theoretic integration or distribution theory.

→ here, we shall use distribution theory.

EX:

$$\mathcal{F}\{\delta(t)\} = \int_{-\infty}^{\infty} \delta(t) e^{-j\Omega t}\, dt = e^{-j\Omega t}\Big|_{t=0} = 1$$

- This you have seen before.

- However, great pitfalls await you if you try to interpret $\mathcal{F}^{-1}\{1\}$ as a Riemann integral.

- Using distribution theory, we have

$$\mathcal{F}^{-1}\{1\} = \frac{1}{2\pi}\int_{\mathbb{R}} e^{j\Omega t}\, d\Omega$$

$$= \frac{1}{2\pi}\int_{\mathbb{R}} \cos\Omega t\, d\Omega + \underbrace{\frac{j}{2\pi}\int_{\mathbb{R}} \sin\Omega t\, d\Omega}_{\text{zero: odd integrand on symmetric limits.}}$$

$$= \frac{1}{2\pi}\lim_{A\to\infty}\int_{-A}^{A} \cos\Omega t\, d\Omega$$

$$= \lim_{A\to\infty} \frac{2\sin At}{2\pi t} = \lim_{A\to\infty}\frac{\sin At}{\pi t}$$

$$= \delta(t) \quad \text{(in the sense of distributions)}$$

which establishes

$$\delta(t) \overset{\mathcal{F}}{\longleftrightarrow} 1 \quad \text{rigorously.}$$

- Applying the F.T. duality property to this pair, we get

$$1 \overset{\mathcal{F}}{\longleftrightarrow} 2\pi\,\delta(\Omega)$$

(the properties will be reviewed later in this module of the notes).

- Applying the frequency shift property to the pair above, we get

$$e^{j\Omega_0 t} \overset{\mathcal{F}}{\longleftrightarrow} 2\pi\,\delta(\Omega-\Omega_0)$$

- Linearity then gives us:

$$\mathcal{F}\{\cos\Omega_0 t\} = \mathcal{F}\{\tfrac{1}{2}e^{j\Omega_0 t} + \tfrac{1}{2}e^{-j\Omega_0 t}\}$$

$$= \pi\left[\delta(\Omega-\Omega_0) + \delta(\Omega+\Omega_0)\right]$$

and

$$\mathcal{F}\{\sin\Omega_0 t\} = j\pi\left[\delta(\Omega+\Omega_0) - \delta(\Omega-\Omega_0)\right]$$

- using the fact that

$$\int_0^\infty \sin \Omega t \, dt = \frac{1}{\Omega}$$ in the sense of distributions,

you can show easily that

$$\operatorname{sgn} t = \frac{t}{|t|} \overset{\mathcal{F}}{\longleftrightarrow} \frac{2}{j\Omega}$$

- with the linearity property, we then have

$$\mathcal{F}\{u(t)\} = \mathcal{F}\{ \tfrac{1}{2} + \tfrac{1}{2} \operatorname{sgn} t \}$$

$$= \tfrac{1}{2} \mathcal{F}\{1\} + \tfrac{1}{2} \mathcal{F}\{ \operatorname{sgn} t \}$$

$$= \pi \delta(\Omega) - j \frac{1}{\Omega}$$

- combining this with the convolution property, we can easily derive the time integration property:

if $x_1(t) \overset{\mathcal{F}}{\longleftrightarrow} X_1(\Omega)$ and

$$x_2(t) = \int_{-\infty}^{t} x_1(\tau) \, d\tau ,$$

Then

$$x_2(t) = \int_{-\infty}^{t} x_1(\tau)\,d\tau = \int_{-\infty}^{\infty} x_1(\tau)\,u(t-\tau)\,d\tau$$

$$= x_1(t) * u(t),$$

so

$$X_2(\Omega) = \mathcal{F}\{x_1(t)\}\,\mathcal{F}\{u(t)\}$$

$$= X_1(\Omega)\left[\pi\delta(\Omega) - j\frac{1}{\Omega}\right]$$

- By applying the frequency shift property to the transform of $u(t)$, we get

$$\mathcal{F}\{u(t)\cos\Omega_0 t\} = \mathcal{F}\{u(t)\tfrac{1}{2}e^{j\Omega_0 t} + u(t)\tfrac{1}{2}e^{-j\Omega_0 t}\}$$

$$= \frac{\pi}{2}\left[\delta(\Omega-\Omega_0) + \delta(\Omega+\Omega_0)\right] - j\frac{1}{2(\Omega-\Omega_0)} - j\frac{1}{2(\Omega+\Omega_0)}$$

$$= \frac{\pi}{2}\left[\delta(\Omega-\Omega_0) + \delta(\Omega+\Omega_0)\right] + \frac{j\Omega}{\Omega_0^2-\Omega^2}$$

- Likewise,

$$u(t)\sin\Omega_0 t \overset{\mathcal{F}}{\longleftrightarrow} \frac{\pi}{2j}\left[\delta(\Omega-\Omega_0) - \delta(\Omega+\Omega_0)\right] + \frac{\Omega_0}{\Omega_0^2-\Omega^2}$$

More useful transform pairs that take some work to prove: (3-52)

$$e^{-t^2/2} \overset{\mathcal{F}}{\longleftrightarrow} \sqrt{2\pi}\, e^{-\Omega^2/2}$$

$$\frac{d^n}{dt^n}\delta(t) \overset{\mathcal{F}}{\longleftrightarrow} (j\Omega)^n$$

$$t^n \overset{\mathcal{F}}{\longleftrightarrow} 2\pi j^n \frac{d^n}{d\Omega^n}\delta(\Omega)$$

$$|t| \overset{\mathcal{F}}{\longleftrightarrow} -\frac{2}{\Omega^2}$$

$$t\,u(t) \overset{\mathcal{F}}{\longleftrightarrow} j\pi\delta'(\Omega) - \frac{1}{\Omega^2}$$

Symmetry Properties of the Fourier Transform

- You should know how to prove these already from your undergrad signals & systems course.

→ if $x(t)$ is real, then $X(\Omega)$ is conjugate symmetric: $X(\Omega) = X^*(-\Omega)$.

This implies:

$\text{Re}\, X(\Omega)$ is even

$\text{Im}\, X(\Omega)$ is odd

$|X(\Omega)|$ is even

$\arg X(\Omega)$ is odd

→ if $x(t)$ is pure imaginary, then $X(\Omega)$ is conjugate antisymmetric: $X(\Omega) = -X^*(-\Omega)$.

This implies:

$\text{Re}\, X(\Omega)$ is odd

$\text{Im}\, X(\Omega)$ is even

$|X(\Omega)|$ is odd

$\arg X(\Omega)$ is even

→ if $x(t)$ is real and even, then $X(\Omega)$ is real and even.

→ if $x(t)$ is real and odd, then $X(\Omega)$ is pure imaginary and odd.

→ Any real $x(t)$ can be decomposed into a unique even part and odd part:

$$x(t) = x_{ev}(t) + x_{od}(t).$$

If $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$, then

$$x_{ev}(t) \overset{\mathcal{F}}{\longleftrightarrow} \operatorname{Re} X(\Omega)$$

$$x_{od}(t) \overset{\mathcal{F}}{\longleftrightarrow} j \operatorname{Im} X(\Omega)$$

→ Any (complex) $x(t)$ can be decomposed into a unique conjugate symmetric part and conjugate antisymmetric part:

$$x(t) = x_{cs}(t) + x_{ca}(t)$$

If $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$, then

$$x_{cs}(t) \overset{\mathcal{F}}{\longleftrightarrow} \operatorname{Re} X(\Omega)$$

$$x_{ca}(t) \overset{\mathcal{F}}{\longleftrightarrow} j \operatorname{Im} X(\Omega).$$

Fourier Transform Properties

- one normally sets out to evaluate a Fourier Transform by integration only as a last resort.
- Your first hope is to find the transform pair you need in a table.
- If it's not in the table, your second strategy is to try to use one or more "properties" to make the one you've got look like one that's in your table.
- Here we will only summarize the main properties.

  → You should already know how to prove all of these from your undergraduate Signals & Systems course.

Linearity

if $x_1(t) \overset{\mathcal{F}}{\longleftrightarrow} X_1(\Omega)$, $x_2(t) \overset{\mathcal{F}}{\longleftrightarrow} X_2(\Omega)$, and $a, b \in \mathbb{C}$ are constants, then

$$a x_1(t) + b x_2(t) \overset{\mathcal{F}}{\longleftrightarrow} a X_1(\Omega) + b X_2(\Omega).$$

Time Shift

if $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$ and $t_0 \in \mathbb{R}$, then

$$x(t - t_0) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega) e^{-j\Omega t_0}$$

Frequency Shift

if $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$ and $\Omega_0 \in \mathbb{R}$, then

$$x(t) e^{j\Omega_0 t} \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega - \Omega_0)$$

Time Scaling

if $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$, then $x(at) \overset{\mathcal{F}}{\longleftrightarrow} \frac{1}{|a|} X\left(\frac{\Omega}{a}\right)$

Frequency Scaling

if $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$, then $\frac{1}{|a|} x\left(\frac{t}{a}\right) \overset{\mathcal{F}}{\longleftrightarrow} X(a\Omega)$

Duality

if $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$, then

$$X(t) \overset{\mathcal{F}}{\longleftrightarrow} 2\pi x(-\Omega)$$

Time Differentiation

if $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$, then $\frac{d}{dt}x(t) \overset{\mathcal{F}}{\longleftrightarrow} j\Omega X(\Omega)$.

Frequency Differentiation

if $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$, then $-jt\,x(t) \overset{\mathcal{F}}{\longleftrightarrow} \frac{d}{d\Omega}X(\Omega)$

Conjugation

if $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$, then $x^*(t) \overset{\mathcal{F}}{\longleftrightarrow} X^*(-\Omega)$

Time Convolution

if $x(t) \overset{\mathcal{F}}{\longleftrightarrow} X(\Omega)$, $h(t) \overset{\mathcal{F}}{\longleftrightarrow} H(\Omega)$, and if $y(t) = x(t) * h(t)$, then

$$y(t) \overset{\mathcal{F}}{\longleftrightarrow} Y(\Omega) = X(\Omega)H(\Omega).$$

Frequency Convolution

if $x_1(t) \xleftrightarrow{\mathcal{F}} X_1(\Omega)$ and $x_2(t) \xleftrightarrow{\mathcal{F}} X_2(\Omega)$,

then $x_1(t)x_2(t) \xleftrightarrow{\mathcal{F}} \frac{1}{2\pi}\left[X_1(\Omega) * X_2(\Omega)\right]$

$$= \frac{1}{2\pi}\int_{\mathbb{R}} X_1(\theta)\, X_2(\Omega-\theta)\, d\theta$$

Time Integration

if $x_1(t) \xleftrightarrow{\mathcal{F}} X_1(\Omega)$ and if $x_2(t) = \int_{-\infty}^{t} x_1(\tau)\, d\tau$,

then $x_2(t) \xleftrightarrow{\mathcal{F}} X_1(\Omega)\left[\pi\delta(\Omega) + \frac{1}{j\Omega}\right]$

Parseval Relation (Plancharell's Theorem)

$$\int_{\mathbb{R}} |x(t)|^2\, dt = \frac{1}{2\pi}\int_{\mathbb{R}} |X(\Omega)|^2\, d\Omega$$

→ in other words, up to the "irritant" factor $\frac{1}{2\pi}$, the <u>norm</u> of $x(t)$ and the <u>norm</u> of $X(\Omega)$ are the same (in the $L^2(\mathbb{R})$ ... Hilbert space) sense...

- Intuitively, this means that $x(t)$ and $X(\Omega)$ can be interpreted as two vectors with equal length

- In this interpretation, the Fourier transform may be viewed as a rotation. It takes a vector and changes it's direction, while keeping the length the same...

i.e., a "Hilbert space rotation."

i.e., a "Hilbert space Isometry."

The Uncertainty Principle

- In modern physics, the Heisenberg uncertainty principle states that the <u>position</u> and <u>momentum</u> of a particle cannot both be known simultaneously with arbitrary accuracy.

→ in modern formulation, Heisenberg's statement is

$$\Delta_x \Delta_p \geq \hbar/2$$

uncertainty in position

uncertainty in momentum

normalized Planck constant.

- In Heisenberg's formulation, the momentum is an operator and the relationship between position and momentum is a matrix equation.
- The formulation is based on the interpretation that the position and the momentum are a

<u>Fourier Transform Pair</u>

- Thus, the Heisenberg Uncertainty Principle is really just a restatement of what engineers call the "reciprocal spreading property"!

Reciprocal Spreading

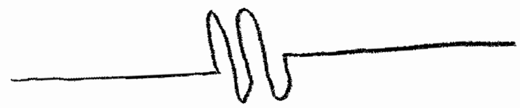
- if $x(t)$ is <u>skinny</u>, then $X(\Omega)$ is <u>fat</u>
- if $x(t)$ is <u>fat</u>, then $X(\Omega)$ is <u>skinny</u>.

→ Now, the Heisenberg inequality may be interpreted as a limit on the <u>rate</u> at which $x(t)$ may become fat (skinny) in relation to the rate at which $X(\Omega)$ will become skinny (fat).

What does it mean?

- suppose we want to build a detector for police "pulse" radar

- The signal we are trying to detect is a short burst of a sinusoid at a specific frequency:



- To detect the radar frequency without the nuisance of frequent false alarms, we need a detection filter that is localized in frequency... i.e., we need the frequency response to be skinny, so that we don't get false alarms from energy at other frequencies.
- But we also want the detection filter to produce an alarm <u>immediately</u> when the pulsed radar signal is present--- without any significant delay.

  → This implies that the filter impulse response must be skinny (localized) in time.

- So we need the filter impulse response to be <u>jointly</u> <u>localized</u> in time and frequency... we need it to be skinny in time and in frequency simultaneously.

- The functional form of the uncertainty principle is known as the Heisenberg-Weyl inequality.

  → This inequality provides a quantitative limit on the degree to which a function can be simultaneously localized in time and in frequency.

- Let $x(t) \in L^2(\mathbb{R})$. The norm of $x(t)$ is a number given by

$$\|x(t)\|_{L^2} = \left[\int_{\mathbb{R}} |x(t)|^2 \, dt \right]^{1/2}.$$

- So the function

$$\frac{1}{\|x(t)\|_{L^2}^2} |x(t)|^2$$

is $\geq 0 \ \forall t$ and has <u>unit</u> <u>area</u>.

- Thus, $\frac{1}{\|x(t)\|_{L^2}^2}|x(t)|^2$ may be rigorously interpreted as a probability density that quantifies the distribution of the signal's energy in time.
- Since $x(t) \in L^2(\mathbb{R})$, we have also that $X(\Omega) \in L^2(\mathbb{R})$.
  - The norm of $X(\Omega)$ is a number given by

$$\|X(\Omega)\|_{L^2}^2 = \int_{\mathbb{R}} |X(\Omega)|^2 \, d\Omega$$

- So the function

$$\frac{1}{\|X(\Omega)\|_{L^2}^2}|X(\Omega)|^2$$

is ≥ 0, has <u>unit area</u>, and may be rigorously interpreted as a probability density function that quantifies the distribution of the signal's energy in frequency.
- The <u>mean time</u> of $x(t)$ is the mean of the pdf $\frac{1}{\|x(t)\|_{L^2}^2}|x(t)|^2$:

$$\langle t \rangle = E[t] = \frac{1}{\|x(t)\|_{L^2}^2}\int_{\mathbb{R}} t\,|x(t)|^2 \, dt$$

- The mean frequency of $x(t)$ is the mean value of the pdf $\frac{1}{\|X(\Omega)\|_{L^2}^2} |X(\Omega)|^2$:

$$\langle \Omega \rangle = E[\Omega] = \frac{1}{\|X(\Omega)\|_{L^2}^2} \int_{\mathbb{R}} \Omega \, |X(\Omega)|^2 \, d\Omega .$$

- The time localization, or time uncertainty of $x(t)$ is quantified by the <u>variance</u> of the pdf $\frac{1}{\|x(t)\|_{L^2}^2} |x(t)|^2$:

Sometimes called "duration"

$$\sigma_t^2(x) = E\left[(t - \langle t \rangle)^2\right]$$

$$= \frac{1}{\|x(t)\|_{L^2}^2} \int_{\mathbb{R}} (t - \langle t \rangle)^2 |x(t)|^2 \, dt$$

- The frequency localization, or frequency uncertainty of $x(t)$ is quantified by the variance of the pdf $\frac{1}{\|X(\Omega)\|_{L^2}^2} |X(\Omega)|^2$:

$$= E\left[(\Omega - \langle \Omega \rangle)^2\right]$$

$$\sigma_\Omega^2(x) = \frac{1}{\|X(\Omega)\|_{L^2}^2} \int_{\mathbb{R}} (\Omega - \langle \Omega \rangle)^2 |X(\Omega)|^2 \, d\Omega .$$

- The Heisenberg-Weyl inequality, also known as the Heisenberg-Pauli-Weyl inequality:

$$\sigma_t^2(x)\,\sigma_\Omega^2(x) \geq \frac{1}{4}.$$

- The lower bound $\sigma_t^2(x)\sigma_\Omega^2(x) = \frac{1}{4}$ is achieved iff $x(t)$ is a <u>Gabor function</u> ... the product of a Gaussian with a complex exponential, e.g. iff

$$x(t) = K e^{j(\Omega_0 t + \phi)} e^{-(t-\mu)^2/2\sigma^2}$$

for constants $K, \Omega_0, \phi, \mu, \sigma \in \mathbb{R}$.

→ The Fourier transform of this Gabor function is a Gaussian (in frequency) centered at $\Omega = \Omega_0$.

- if $a \in \mathbb{R}$ and $y(t) = x(at)$, then

$$\sigma_t^2(x)\,\sigma_\Omega^2(x) = \sigma_t^2(y)\,\sigma_\Omega^2(y).$$

⇒ The time-bandwidth product is unchanged, although the individual variances generally do change.

Formal Definition of "Inner Product Space"

Def (Rudin):

A complex vector space $\mathcal{X}$ is called an inner product space if to each ordered pair of vectors $x, y \in \mathcal{X}$ there is associated a complex number $\langle x, y \rangle$ such that:

(a) $\langle y, x \rangle = \langle x, y \rangle^* \quad \forall\, x, y \in \mathcal{X}$.

(b) $\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \quad \forall\, x, y, z \in \mathcal{X}$.

(c) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle \quad \forall\, x, y \in \mathcal{X}$ and all scalars α.

(d) $\langle x, x \rangle \geq 0 \quad \forall\, x \in \mathcal{X}$.

(e) $\langle x, x \rangle = 0$ iff $x = 0$.

The number $\langle x, y \rangle$ is called the "inner product" or "scalar product" of x and y.

Immediate consequences:

- from (c), $\langle 0, x \rangle = 0 \quad \forall\, x \in \mathcal{X}$.

- from (b) and (c), $\forall\, y \in \mathcal{X}$, the mapping $x \mapsto \langle x, y \rangle$ is a linear functional on $\mathcal{X}$.

- from (a) and (c), $\langle x, \alpha y \rangle = \alpha^* \langle x, y \rangle \quad \forall\, x, y \in \mathcal{X}$ and all scalars α.

- from (a) and (b), $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$
$\forall\, x, y, z \in \mathcal{X}$.

Note: Any valid inner product on $\mathcal{X}$ naturally induces a "2-norm" on $\mathcal{X}$ that is given by

$$\|x\|^2 = \langle x, x \rangle.$$

i.e., $\|x\|$ is the non-negative square root of $\langle x, x \rangle$.

EX: in the continuous-time Hilbert space $L^2(\mathbb{R})$, we have

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

$$\|x(t)\|^2 = \langle x(t), x(t) \rangle = \int_{\mathbb{R}} x(t)\, x^*(t)\, dt = \int_{\mathbb{R}} |x(t)|^2\, dt.$$

EX: in the discrete-time Hilbert space $\ell^2(\mathbb{Z})$, we have:

$$\langle x[n], y[n] \rangle = \sum_{n \in \mathbb{Z}} x[n] y^*[n]$$

$$\|x[n]\|^2 = \langle x[n], x[n] \rangle = \sum_{n \in \mathbb{Z}} x[n] x^*[n] = \sum_{n \in \mathbb{Z}} |x[n]|^2.$$

Cauchy-Schwarz Inequality

Let $x(t), y(t) \in L^2(\mathbb{R})$. Then

$$\left| \int_{\mathbb{R}} x(t) y^*(t)\, dt \right|^2 \leq \int_{\mathbb{R}} |x(t)|^2 dt \int_{\mathbb{R}} |y(t)|^2 dt.$$

- More generally, in any abstract Hilbert space $\mathcal{H}$, the inequality is given by

$$|\langle x, y \rangle|^2 \leq \|x\|^2 \|y\|^2 \quad \forall \ x, y \in \mathcal{H}.$$

- And in $\ell^2(\mathbb{Z})$, we have

$$\left| \sum_{n \in \mathbb{Z}} x[n] y^*[n] \right|^2 \leq \sum_{n \in \mathbb{Z}} |x[n]|^2 \sum_{n \in \mathbb{Z}} |y[n]|^2.$$

- The inequality is known variously as "Schwarz's inequality," "the Schwarz inequality," "Bunyakovsky's inequality," "the Cauchy-Schwarz inequality" and "the Cauchy-Schwarz-Bunyakovsky inequality."

Historical Note:

- In 1821, A.L. Cauchy proved that, for any two finite or countably infinite sequences of reals x_k and y_k,

$$\left[\sum_k x_k y_k\right]^2 \le \sum_k x_k^2 \sum_k y_k^2.$$

- For vectors $x, y \in \mathbb{R}^n$, this may be written compactly as

$$\|\langle x, y\rangle\|^2 \le \|x\|^2 \|y\|^2$$, e.g., the C-S inequality.

- In 1859, Bunyakovsky proved the integral version for real-valued functions:

$$\left[\int x(t)y(t)\,dt\right]^2 \le \int x^2(t)\,dt \int y^2(t)\,dt,$$

which holds for any finite or infinite limits of integration.

- Unaware of the work of Bunyakovsky, H.A. Schwarz proved independently in 1884 that, $\forall\, x, y \in L^2(\mathbb{R})$,

$$\left[\int x(t)y^*(t)\,dt\right]^2 \le \int |x(t)|^2\,dt \int |y(t)|^2\,dt,$$

also valid on any finite or infinite limits of integration.

Proof of the Cauchy-Schwarz Inequality:

<u>Thm:</u> Let $\mathcal{X}$ be an inner product space. Then

$$|\langle x,y\rangle|^2 \le \|x\|^2\|y\|^2 \qquad \forall\ x,y\in\mathcal{X}.$$

<u>Pf:</u> Let $x,y\in\mathcal{X}$ and let $\lambda\in\mathbb{C}$.
Let $z = x+\lambda y$. Then $z\in\mathcal{X}$, $\|z\|<\infty$, $\|x\|<\infty$, and $\|y\|<\infty$.

We have

$$0\le \|y\|^2\|z\|^2 = \langle y,y\rangle\langle z,z\rangle$$

$$= \langle y,y\rangle\langle x+\lambda y, x+\lambda y\rangle$$

$$= \langle y,y\rangle\left[\langle x,x\rangle + \lambda^*\langle y,x\rangle + \lambda\langle x,y\rangle + \lambda\lambda^*\langle y,y\rangle\right]$$

$$= \langle y,y\rangle\langle x,x\rangle + \lambda^*\langle y,y\rangle\langle y,x\rangle + \lambda\langle y,y\rangle\langle x,y\rangle + \lambda\lambda^*\langle y,y\rangle\langle y,y\rangle. \qquad ☺$$

Now choose $\lambda = -\dfrac{\langle y,x\rangle}{\langle y,y\rangle}$.

Then ☺ becomes

$$0\le \langle y,y\rangle\langle x,x\rangle - \frac{\langle x,y\rangle}{\langle y,y\rangle}\langle y,y\rangle\langle y,x\rangle - \frac{\langle y,x\rangle}{\langle y,y\rangle}\langle y,y\rangle\langle x,y\rangle$$

$$+ \frac{\langle y,x\rangle}{\langle y,y\rangle}\frac{\langle x,y\rangle}{\langle y,y\rangle}\langle y,y\rangle\langle y,y\rangle \longrightarrow$$

$$= \langle y,y\rangle\langle x,x\rangle - \langle x,y\rangle\langle y,x\rangle - \langle y,x\rangle\langle x,y\rangle + \langle y,x\rangle\langle x,y\rangle$$

$$= \langle y,y\rangle\langle x,x\rangle - \langle x,y\rangle\langle x,y\rangle^* - \langle x,y\rangle^*\langle x,y\rangle + \langle x,y\rangle^*\langle x,y\rangle$$

$$= \langle y,y\rangle\langle x,x\rangle - \langle x,y\rangle\langle x,y\rangle^*$$

$$= \|x\|^2\|y\|^2 - |\langle x,y\rangle|^2 \geq 0.$$

So $|\langle x,y\rangle|^2 \leq \|x\|^2\|y\|^2$.

QED

Corollary: the Triangle inequality:

$$\forall\, x,y \in \mathcal{X}, \quad \|x+y\| \leq \|x\| + \|y\|.$$

Pf: $\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle$

Applying the Cauchy-Schwarz inequality to the two middle terms, we have

$$\|x+y\|^2 \leq \|x\|^2 + |\langle x,y\rangle| + |\langle y,x\rangle| + \|y\|^2$$
$$\leq \|x\|^2 + \|x\|\,\|y\| + \|x\|\,\|y\| + \|y\|^2 \longrightarrow$$

$$\cdots = \left[\|x\| + \|y\|\right]^2.$$

Taking the positive square root of the far left and far right sides,

$$\|x+y\| \leq \|x\| + \|y\|.$$

- Triangle inequality for complex numbers:

$$|x+y| \leq |x| + |y| \quad \forall \; x, y \in \mathbb{C}$$

- Triangle inequality for sums:

$\forall \; x[n] \in \ell^1(\mathbb{Z})$,

$$\left|\sum_k x[k]\right| \leq \sum_k |x[k]|$$

- Triangle inequality for integrals:

$\forall \; x(t) \in L^1(\mathbb{R})$,

$$\left|\int_{\mathbb{R}} x(t)\,dt\right| \leq \int_{\mathbb{R}} |x(t)|\,dt$$

A usefull related inequality: the Hölder inequality

Hölder's inequality for sums:

Let p and q be conjugate exponent. This means $p, q \in \mathbb{R}$, $p > 1$, and $\frac{1}{p} + \frac{1}{q} = 1$. $(p+q = pq)$

Let $x[n] \in \ell^p(\mathbb{Z})$ and $y[n] \in \ell^q(\mathbb{Z})$.

Then

$$\sum_n |x[n]y[n]| \leq \left[\sum_n |x[n]|^p\right]^{\frac{1}{p}} \left[\sum_n |y[n]|^q\right]^{\frac{1}{q}},$$

with equality if $|y[n]| = c|x[n]|^{p-1}$ for some constant $c \in \mathbb{C}$.

Hölder's inequality for integrals:

Let p and q be conjugate exponents and let $x(t) \in L^p(\mathbb{R})$ and $y(t) \in L^q(\mathbb{R})$.

Then

$$\int_{\mathbb{R}} |x(t)y(t)|\,dt \leq \left[\int_{\mathbb{R}} |x(t)|^p\right]^{\frac{1}{p}} \left[\int_{\mathbb{R}} |y(t)|^q\right]^{\frac{1}{q}}$$

with equality if $|y(t)| = c|x(t)|^{p-1}$ for some constant $c \in \mathbb{C}$.

- Next we need to build up some results that will be useful in proving the uncertainty principle (Heisenberg-Weyl Inequality).

- If we divide both sides of ☺ on page 3-71 by $\langle y,y\rangle = \|y\|^2$, we get

$$0 \leq \langle x,x\rangle + \lambda^*\langle y,x\rangle + \lambda\langle x,y\rangle + \lambda\lambda^*\langle y,y\rangle$$

- If we let $x, y \in L^2(\mathbb{R})$, this becomes

$$0 \leq \int_{\mathbb{R}} x(t)x^*(t)dt + \lambda^*\int_{\mathbb{R}} y(t)x^*(t)dt + \lambda\int_{\mathbb{R}} x(t)y^*(t)dt + \lambda\lambda^*\int_{\mathbb{R}} y(t)y^*(t)dt.$$

- Now suppose we restrict our attention to only <u>real</u> λ. This becomes

$$0 \leq \int_{\mathbb{R}} x(t)x^*(t)dt + \lambda\int_{\mathbb{R}}[x^*(t)y(t) + x(t)y^*(t)]dt + \lambda^2\int_{\mathbb{R}} y(t)y^*(t)dt$$

$$\leq \int_{\mathbb{R}} x(t)x^*(t)dt + \lambda\left|\int_{\mathbb{R}} x^*(t)y(t) + x(t)y^*(t)dt\right| + \lambda^2\int_{\mathbb{R}} y(t)y^*(t)dt \quad ☹$$

Let $A = \int_{\mathbb{R}} y(t)y^*(t)\,dt$

$$B = \left| \int_{\mathbb{R}} x^*(t)y(t) + x(t)y^*(t)\,dt \right|$$

$$C = \int_{\mathbb{R}} x(t)x^*(t)\,dt.$$

Then ☹ on page 3-75 becomes

$$A\lambda^2 + B\lambda + C \geq 0,$$

a <u>non-negative</u> quadratic in λ.

Then the discriminant of this quadratic must be non-positive; i.e.

$$B^2 - 4AC \leq 0$$

$$B^2 \leq 4AC$$

$$\left| \int_{\mathbb{R}} x^*(t)y(t) + x(t)y^*(t)\,dt \right|^2 \leq 4 \int_{\mathbb{R}} x(t)x^*(t)\,dt \int_{\mathbb{R}} y(t)y^*(t)\,dt$$

I will refer to this as "Bracewell's form of the Cauchy-Schwarz inequality."

Lemma A:

Let $x(t) \in L^2(\mathbb{R})$, $t_0 \in \mathbb{R}$, and $x_1(t) = x(t-t_0)$.
Then $\sigma_t^2(x_1) = \sigma_t^2(x)$, i.e., the time duration measure σ_t^2 is translation invariant.

Pf:

Let $\langle t \rangle_1 = \frac{1}{\|x_1(t)\|_{L^2}^2} \int_{\mathbb{R}} t\, |x_1(t)|^2 dt.$

Then

$$\sigma_t^2(x_1) = \frac{1}{\|x_1(t)\|_{L^2}^2} \int_{\mathbb{R}} (t - \langle t \rangle_1)^2 \, |x_1(t)|^2 \, dt$$

$$= \frac{\int_{\mathbb{R}} \left[t - \frac{\int_{\mathbb{R}} t\, |x_1(t)|^2 dt}{\int_{\mathbb{R}} |x_1(t)|^2 dt} \right]^2 |x_1(t)|^2 dt}{\int_{\mathbb{R}} |x_1(t)|^2 dt}$$

$$= \frac{\int_{\mathbb{R}} \left[t - \frac{\int_{\mathbb{R}} t\, |x(t-t_0)|^2 dt}{\int_{\mathbb{R}} |x(t-t_0)|^2 dt} \right]^2 |x(t-t_0)|^2 dt}{\int_{\mathbb{R}} |x(t-t_0)|^2 dt} \qquad (*).$$

- Let $u = t - t_0$, $t = u + t_0$, $du = dt$.

Then

$$(*) = \frac{\int_{\mathbb{R}} \left[u + t_0 - \frac{\int_{\mathbb{R}} (u+t_0)|x(u)|^2 du}{\int_{\mathbb{R}} |x(u)|^2 du} \right]^2 |x(u)|^2 du}{\int_{\mathbb{R}} |x(u)|^2 du}$$

$$= \frac{1}{\|x(t)\|_{L^2}^2} \int_{\mathbb{R}} \left[u + t_0 - \frac{1}{\|x(t)\|_{L^2}^2} \left(\int_{\mathbb{R}} u|x(u)|^2 du + t_0 \int_{\mathbb{R}} |x(u)|^2 du \right) \right]^2$$

$$\times |x(u)|^2 du$$

$$= \frac{1}{\|x(t)\|_{L^2}^2} \int_{\mathbb{R}} \left[u + t_0 - \frac{1}{\|x(t)\|_{L^2}^2} \int_{\mathbb{R}} u|x(u)|^2 du - t_0 \frac{\|x(t)\|_{L^2}^2}{\|x(t)\|_{L^2}^2} \right]^2 |x(u)|^2 du$$

$$= \frac{1}{\|x(t)\|_{L^2}^2} \int_{\mathbb{R}} (u + t_0 - \langle t \rangle - t_0)^2 |x(u)|^2 du$$

write "t" instead of "u".

$$= \frac{1}{\|x(t)\|_{L^2}^2} \int_{\mathbb{R}} (t - \langle t \rangle)^2 |x(t)|^2 dt$$

$$= \sigma_t^2(x).$$

QED.

$\Rightarrow$ It follows that $\sigma_\Omega^2(x)$ is also translation invariant.

Lemma B:

Let $x(t) \in L^2(\mathbb{R})$, $\Omega_0 \in \mathbb{R}$, and $x_1(t) = x(t)e^{j\Omega_0 t}$. Then $\sigma_t^2(x_1) = \sigma_t^2(x)$, i.e., the measure σ_t^2 is modulation invariant.

Pf:

$$\sigma_t^2(x_1) = \frac{1}{\|x_1(t)\|_{L^2}^2} \int_{\mathbb{R}} (t - \langle t \rangle_1)^2 |x_1(t)|^2 dt$$

$$= \frac{\int_{\mathbb{R}} \left[t - \frac{\int_{\mathbb{R}} t |x_1(t)|^2 dt}{\int_{\mathbb{R}} |x_1(t)|^2 dt} \right]^2 |x_1(t)|^2 dt}{\int_{\mathbb{R}} |x_1(t)|^2 dt}$$

$$= \frac{\int_{\mathbb{R}} \left[t - \frac{\int_{\mathbb{R}} t |x(t)e^{j\Omega_0 t}|^2 dt}{\int_{\mathbb{R}} |x(t)e^{j\Omega_0 t}|^2 dt} \right]^2 |x(t)e^{j\Omega_0 t}|^2 dt}{\int_{\mathbb{R}} |x(t)e^{j\Omega_0 t}|^2 dt}$$

$$= \int_{\mathbb{R}} \left[t - \frac{\int_{\mathbb{R}} t |x(t)|^2 dt}{\int_{\mathbb{R}} |x(t)|^2 dt} \right]^2 |x(t)|^2 dt \Big/ \int_{\mathbb{R}} |x(t)|^2 dt$$

$$= \frac{1}{\|x(t)\|_{L^2}^2} \int_{\mathbb{R}} (t - \langle t \rangle)^2 |x(t)|^2 dt = \sigma_t^2(x).$$

QED.

$\Rightarrow$ It follows that $\sigma_\Omega^2(x)$ is also modulation invariant.

Lemma C:

Let $x(t) \in L^2(\mathbb{R})$ with

$$X(\Omega) \xleftrightarrow{\mathcal{F}} x(t)$$

$$t_0 = \langle t \rangle = \frac{1}{\|x(t)\|_{L^2}^2} \int_{\mathbb{R}} t\, |x(t)|^2 dt$$

$$\Omega_0 = \langle \Omega \rangle = \frac{1}{\|X(\Omega)\|_{L^2}^2} \int_{\mathbb{R}} \Omega\, |X(\Omega)|^2 d\Omega$$

and let $x_1(t) = x(t+t_0)e^{-j\Omega_0 t}$.

Then the following hold:

(a) $\langle t \rangle_1 = \frac{1}{\|x_1(t)\|_{L^2}^2} \int_{\mathbb{R}} t\, |x_1(t)|^2 dt = 0$

(b) $\langle \Omega \rangle_1 = \frac{1}{\|X_1(\Omega)\|_{L^2}^2} \int_{\mathbb{R}} \Omega\, |X_1(\Omega)|^2 d\Omega = 0$

(c) $\sigma_t^2(x_1) = \sigma_t^2(x)$

(d) $\sigma_\Omega^2(x_1) = \sigma_\Omega^2(x)$.

Proof: (c) and (d) follow immediately from Lemma A and Lemma B.

For (a), we have

$$\langle t\rangle_1 = \frac{1}{\|x_1(t)\|^2_{L^2}} \int_{\mathbb{R}} t\, |x_1(t)|^2 dt$$

$$= \int_{\mathbb{R}} t\, |x(t+t_0) e^{-j\Omega_0 t}|^2 dt \Big/ \int_{\mathbb{R}} |x(t+t_0) e^{-j\Omega_0 t}|^2 dt$$

$$= \int_{\mathbb{R}} t\, |x(t+t_0)|^2 dt \Big/ \int_{\mathbb{R}} |x(t+t_0)|^2 dt \qquad \begin{aligned} u &= t+t_0 \\ t &= u - t_0 \\ du &= dt \end{aligned}$$

$$= \int_{\mathbb{R}} (u-t_0)\, |x(u)|^2 du \Big/ \int_{\mathbb{R}} |x(u)|^2 du$$

$$= \frac{1}{\|x(t)\|^2_{L^2}} \left[\int_{\mathbb{R}} u\, |x(u)|^2 du - t_0 \int_{\mathbb{R}} |x(u)|^2 du \right]$$

$$= \frac{1}{\|x(t)\|^2_{L^2}} \int_{\mathbb{R}} t\, |x(t)|^2 dt - t_0 \frac{\|x(t)\|^2_{L^2}}{\|x(t)\|^2_{L^2}}$$

$$= \langle t\rangle - t_0 = t_0 - t_0 = 0.$$

For (b), we have by the F.T. time shift and frequency shift properties that (3-82)

$$X_1(\Omega) = \mathcal{F}\{x_1(t)\} = \mathcal{F}\{x(t+t_0)e^{-j\Omega_0 t}\}$$
$$= X(\Omega+\Omega_0)e^{jt_0\Omega}.$$

Then

$$\langle\Omega\rangle_1 = \frac{1}{\|X_1(\Omega)\|_{L^2}^2}\int_{\mathbb{R}} \Omega\,|X_1(\Omega)|^2\,d\Omega$$

$$= \int_{\mathbb{R}} \Omega\,|X(\Omega+\Omega_0)e^{jt_0\Omega}|^2\,d\Omega \Big/ \int_{\mathbb{R}} |X(\Omega+\Omega_0)e^{jt_0\Omega}|^2\,d\Omega$$

$$= \int_{\mathbb{R}} \Omega\,|X(\Omega+\Omega_0)|^2\,d\Omega \Big/ \int_{\mathbb{R}} |X(\Omega+\Omega_0)|^2\,d\Omega \qquad \begin{aligned}\theta &= \Omega+\Omega_0\\ \Omega &= \theta-\Omega_0\\ d\theta &= d\Omega\end{aligned}$$

$$= \int_{\mathbb{R}} (\theta-\Omega_0)\,|X(\theta)|^2\,d\theta \Big/ \int_{\mathbb{R}} |X(\theta)|^2\,d\theta$$

$$= \frac{1}{\|X(\Omega)\|_{L^2}^2}\left[\int_{\mathbb{R}} \theta\,|X(\theta)|^2\,d\theta - \Omega_0\int_{\mathbb{R}} |X(\theta)|^2\,d\theta\right]$$

$$= \frac{1}{\|X(\Omega)\|_{L^2}^2}\int_{\mathbb{R}} \Omega\,|X(\Omega)|^2\,d\Omega - \Omega_0\,\frac{\|X(\Omega)\|_{L^2}^2}{\|X(\Omega)\|_{L^2}^2}$$

$$= \Omega_0 - \Omega_0 = 0.$$

QED

Lemma D:

Let $x(t) \in L^2(\mathbb{R})$. Then

$$\int_{\mathbb{R}} \left[\frac{d}{dt}x(t)\right]\left[\frac{d}{dt}x(t)\right]^* dt = \frac{1}{2\pi}\int_{\mathbb{R}} \Omega^2 X(\Omega) X^*(\Omega)\, d\Omega .$$

Pf:

By the time differentiation property of the F.T., we have $\frac{d}{dt}x(t) \overset{\mathcal{F}}{\longleftrightarrow} j\Omega X(\Omega)$.

Applying Parseval's formula to this pair yields

$$\int_{\mathbb{R}} \left|\frac{d}{dt}x(t)\right|^2 dt = \frac{1}{2\pi}\int_{\mathbb{R}} |j\Omega X(\Omega)|^2 d\Omega = \frac{1}{2\pi}\int_{\mathbb{R}} \Omega^2 |X(\Omega)|^2 d\Omega .$$

Since $|f|^2 = ff^*$ for any complex function f, we have

$$\int_{\mathbb{R}} \left[\frac{d}{dt}x(t)\right]\left[\frac{d}{dt}x(t)\right]^* dt = \frac{1}{2\pi}\int_{\mathbb{R}} \Omega^2 X(\Omega) X^*(\Omega)\, d\Omega .$$

QED

Proof of the Heisenberg-Weyl inequality:

Let $x(t) \in L^2(\mathbb{R})$ and $X(\Omega) \overset{\mathcal{F}}{\longleftrightarrow} x(t)$.

Let $\langle t \rangle = \frac{1}{\|x(t)\|_{L^2}^2} \int_{\mathbb{R}} t\,|x(t)|^2\,dt$

and $\langle \Omega \rangle = \frac{1}{\|X(\Omega)\|_{L^2}^2} \int_{\mathbb{R}} \Omega\,|X(\Omega)|^2\,d\Omega$.

Let $x_1(t) = x(t + \langle t \rangle)\,e^{-j\langle \Omega \rangle t}$.

By Lemma C, parts (a) and (b), we have that

$$\langle t \rangle_1 = \frac{1}{\|x_1(t)\|_{L^2}^2} \int_{\mathbb{R}} t\,|x_1(t)|^2\,dt = 0$$

$$\langle \Omega \rangle_1 = \frac{1}{\|X_1(\Omega)\|_{L^2}^2} \int_{\mathbb{R}} \Omega\,|X_1(\Omega)|^2\,dt = 0.$$

By Lemma C parts (c) and (d), we have

$$\sigma_t^2(x)\,\sigma_\Omega^2(x) = \sigma_t^2(x_1)\,\sigma_\Omega^2(x_1)$$

$$= \frac{1}{\|x_1(t)\|_{L^2}^2} \int_{\mathbb{R}} (t - \langle t \rangle_1)^2\,|x_1(t)|^2\,dt$$

$$\times \frac{1}{\|X_1(\Omega)\|_{L^2}^2} \int_{\mathbb{R}} (\Omega - \langle \Omega \rangle_1)^2\,|X_1(\Omega)|^2\,d\Omega$$

$$\cdots\ \sigma_t^2(x)\,\sigma_\Omega^2(x) = \frac{1}{\|x_1(t)\|_{L^2}^2}\int_{\mathbb{R}} t^2 |x_1(t)|^2\,dt$$

$$\times \frac{1}{\|X_1(\Omega)\|_{L^2}^2}\int_{\mathbb{R}} \Omega^2 |X_1(\Omega)|^2\,d\Omega$$

$$= \frac{\int_{\mathbb{R}} t^2 |x_1(t)|^2\,dt \int_{\mathbb{R}} \Omega^2 |X_1(\Omega)|^2\,d\Omega}{\int_{\mathbb{R}} |x_1(t)|^2\,dt \int_{\mathbb{R}} |X_1(\Omega)|^2\,d\Omega}$$

$$= \frac{\int_{\mathbb{R}} t^2 x_1(t) x_1^*(t)\,dt \int_{\mathbb{R}} \Omega^2 X_1(\Omega) X_1^*(\Omega)\,d\Omega}{\int_{\mathbb{R}} x_1(t) x_1^*(t)\,dt \int_{\mathbb{R}} X_1(\Omega) X_1^*(\Omega)\,d\Omega}$$

Applying Lemma D to the numerator, we have

$$\sigma_t^2(x)\,\sigma_\Omega^2(x) = \frac{2\pi \int_{\mathbb{R}} t^2 x_1(t) x_1^*(t)\,dt \int_{\mathbb{R}} \left[\frac{d}{dt}x(t)\right]\left[\frac{d}{dt}x(t)\right]^* dt}{\int_{\mathbb{R}} x_1(t) x_1^*(t)\,dt \int_{\mathbb{R}} X_1(\Omega) X_1^*(\Omega)\,d\Omega},$$

whereupon application of Parseval's formula to the denominator results in

$$\sigma_t^2(x)\sigma_\Omega^2(x) = \frac{2\pi \int_{\mathbb{R}} [t x_1(t)][t x_1(t)]^* dt \int_{\mathbb{R}} [\frac{d}{dt} x_1(t)][\frac{d}{dt} x_1(t)]^* dt}{2\pi \left[\int_{\mathbb{R}} x_1(t)\, x_1^*(t)\, dt \right]^2}$$

Now apply the Cauchy-Schwarz inequality in Bracewell's form to the numerator to get

$$\sigma_t^2(x)\sigma_\Omega^2(x) \geqslant \frac{\left| \int_{\mathbb{R}} [t x_1(t)]^* [\frac{d}{dt} x_1(t)] + [t x_1(t)] [\frac{d}{dt} x_1(t)]^* dt \right|^2}{4 \left[\int_{\mathbb{R}} x_1(t) x_1^*(t)\, dt \right]^2}$$

$$= \frac{\left| \int_{\mathbb{R}} t \left\{ x_1^*(t) [\frac{d}{dt} x_1(t)] + x_1(t) [\frac{d}{dt} x_1^*(t)] \right\} dt \right|^2}{4 \left[\int_{\mathbb{R}} x_1(t) x_1^*(t) dt \right]^2}$$

$$= \frac{\left| \int_{\mathbb{R}} t \frac{d}{dt} [x_1^*(t) x_1(t)] dt \right|^2}{4 \left[\int_{\mathbb{R}} x_1(t) x_1^*(t) dt \right]^2} = \frac{\left| \int_{\mathbb{R}} t \frac{d}{dt} [x_1(t) x_1^*(t)] dt \right|^2}{4 \left[\int_{\mathbb{R}} x_1(t) x_1^*(t) dt \right]^2}$$

(*)

Let $u = t$ and $dv = \frac{d}{dt}[x_1(t)x_1^*(t)]\,dt$.
Then $du = dt$, $v = x_1(t)x_1^*(t)$, and the numerator of $(*)$ on page 3-86 becomes

$$\text{numerator} = \left| \int_{\mathbb{R}} v\,du \right|^2$$

$$= \left| uv\Big|_{-\infty}^{\infty} - \int_{\mathbb{R}} v\,du \right|^2$$

$$= \left| t|x_1(t)|^2\Big|_{-\infty}^{\infty} - \int_{\mathbb{R}} x_1(t)x_1^*(t)\,dt \right|^2. \quad (**)$$

It follows from the fact that $x_1(t) \in L^2(\mathbb{R})$ that $\lim_{t\to\infty} t|x_1(t)|^2 = \lim_{t\to-\infty} t|x_1(t)|^2 = 0$.

Thus, the numerator in $(**)$ is

$$\text{numerator} = \left| -\int_{\mathbb{R}} x_1(t)x_1^*(t)\,dt \right|^2$$

$$= \left| \int_{\mathbb{R}} x_1(t)x_1^*(t)\,dt \right|^2 \quad (***)$$

Plugging (***) back into (*) on page 3-86 then gives us

$$\sigma_t^2(x)\sigma_\Omega^2(x) \geq \frac{\left| \int_{\mathbb{R}} x_1(t)\, x_1^*(t)\, dt \right|^2}{4\left[\int_{\mathbb{R}} x_1(t)\, x_1^*(t)\, dt \right]^2} = \frac{1}{4}.$$

QED

DISCRETE-TIME FOURIER TRANSFORM (DTFT)

- Let $\omega_0 \in \mathbb{R}$ be a constant and let $x[n] = e^{j\omega_0 n}$.
- Let H be a discrete-time LTI system with impulse response $h[n]$:

$$x[n] \longrightarrow \boxed{\begin{matrix}\text{LTI}\\ H\end{matrix}} \longrightarrow y[n] = x[n] * h[n]$$

Then:

$$y[n] = x[n] * h[n] = \sum_{k \in \mathbb{Z}} x[n-k]\, h[k]$$

$$= \sum_{k \in \mathbb{Z}} e^{j\omega_0 (n-k)}\, h[k]$$

$$= \underbrace{e^{j\omega_0 n}}_{x[n]} \underbrace{\sum_{k \in \mathbb{Z}} h[k]\, e^{-j\omega_0 k}}_{\text{a complex number (there's no "}n\text{" in it)... call it "}\lambda\text{"}}$$

$$= \lambda\, x[n].$$

$\Rightarrow$ This shows that, for any $\omega \in \mathbb{R}$, the signal $x[n] = e^{j\omega n}$ is an eigenfunction of <u>any</u> discrete-time LTI system.

- The associated eigenvalues are given by the system frequency response $H(e^{j\omega}) = \sum_{n \in \mathbb{Z}} h[n] e^{-j\omega n}$.
- Although $H(e^{j\omega})$ is a function of ω, it is usually written as "$H(e^{j\omega})$"... the reason is that you will be doing partial fraction expansions in terms of the "character" $e^{j\omega}$.

Note: all <u>possible</u> graphs of the signal

$e^{j\omega_0 n}$ can be generated using values of ω_0 in the range $-\pi \leq \omega_0 < \pi$.

$\Rightarrow$ real choices of ω_0 outside this range just give you the same graphs over again.

$\Rightarrow$ You can also generate all the graphs by choosing $\omega_0 \in [0, 2\pi)$ or in <u>any</u> other half-open interval of length 2π.

$\Rightarrow$ This is because $e^{j\omega n} = e^{j(\omega + 2\pi k)n}$ $\forall k \in \mathbb{Z}$.

- We will see later that, for $\omega_1, \omega_2 \in [-\pi, \pi)$, the inner product (in the $\ell^2(\mathbb{Z})$ sense) of $e^{j\omega_1 n}$ and $e^{j\omega_2 n}$,

i.e., the projection of the vector $e^{j\omega_1 n}$ onto the vector $e^{j\omega_2 n}$,

is given by

$$\langle e^{j\omega_1 n}, e^{j\omega_2 n} \rangle = \sum_{n \in \mathbb{Z}} e^{j\omega_1 n} e^{-j\omega_2 n} = 2\pi \delta(\omega_2 - \omega_1).$$

- In other words,

$$\langle e^{j\omega_1 n}, e^{j\omega_2 n} \rangle = \begin{cases} 2\pi\delta(0), & \omega_1 = \omega_2 \\ 0, & \omega_1 \neq \omega_2 \end{cases}.$$

- Thus, like the continuous-time spectral basis, the set of discrete-time signals

$$\{e^{j\omega n}\}_{\omega \in [-\pi, \pi)}$$

is a basis for a very large space of discrete-time signals $x[n]$.

  - This space includes all the discrete-time signals $x[n] \in \ell^1(\mathbb{Z})$, $x[n] \in \ell^2(\mathbb{Z})$, and many others as well.

- As shown above, this basis is orthogonal, but <u>not</u> orthonormal, since each basis vector has length $\sqrt{2\pi}$ instead of 1 (in a distributional sense).

- So, let us write a signal $x[n] : \mathbb{Z} \to \mathbb{C}$ as a linear composition of this discrete-time spectral basis. The steps are the same as always!

Step 1: project $x[n]$ onto each basis signal (find the dot product):

$$\langle x[n], e^{j\omega n} \rangle = \sum_{n \in \mathbb{Z}} x[n] e^{-j\omega n} \triangleq X(e^{j\omega}).$$

Step 2: Add up these dot products times the respective basis signals... don't forget to divide by 2π to make up for the fact that the basis is orthogonal, but not orthonormal:

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \underbrace{\langle x[n], e^{j\omega n} \rangle}_{\text{The dot product, a number.}} \underbrace{e^{j\omega n}}_{\text{The basis signal (vector).}} d\omega$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega.$$

FACT: any DTFT $X(e^{j\omega})$ is 2π-periodic in ω.

Pf: Let $\omega \in \mathbb{R}$. For any $k \in \mathbb{Z}$, let us evaluate the function $X(e^{j\omega})$ at $\omega + 2\pi k$:

$$X(e^{j(\omega+2\pi k)}) = \sum_{n\in\mathbb{Z}} x[n] e^{-j(\omega+2\pi k)n}$$

$$= \sum_{n\in\mathbb{Z}} x[n] e^{-j\omega n} \underbrace{e^{-j2\pi kn}}_{\text{this is one}}$$

$$= \sum_{n\in\mathbb{Z}} x[n] e^{-j\omega n}$$

$$= X(e^{j\omega}). \quad \text{QED.}$$

- Thus, the inversion integral may be computed using the basis vectors in any 2π frequency interval... the "correct" dot products will automatically be lined up with the "correct" basis vectors due to the 2π periodicity of the DTFT.

- In other words

$$x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n}\, d\omega$$

$$= \frac{1}{2\pi}\int_{0}^{2\pi} X(e^{j\omega}) e^{j\omega n}\, d\omega$$

$$= \frac{1}{2\pi}\int_{\pi}^{3\pi} X(e^{j\omega}) e^{j\omega n}\, d\omega \quad \ldots$$

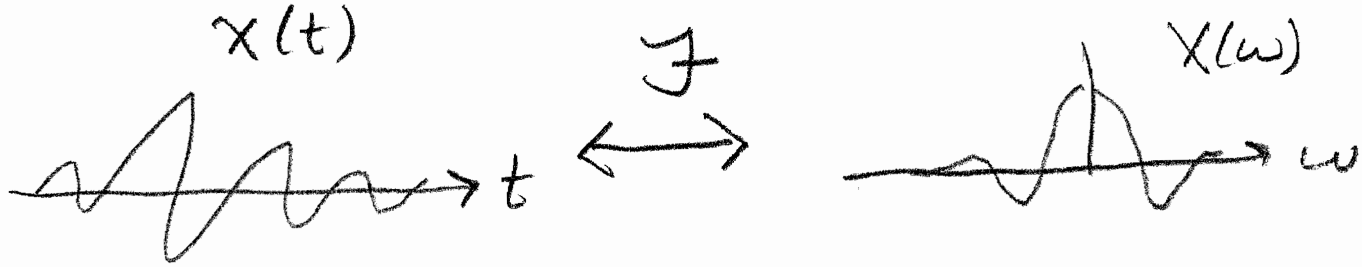
- For this reason, the inversion integral is often written

$$x[n] = \frac{1}{2\pi}\int_{2\pi} X(e^{j\omega}) e^{j\omega n}\, d\omega.$$

- Going back to page 3-89, we see that the eigenvalues corresponding to the spectral basis functions, i.e. the Frequency Response, are given by the DTFT of the impulse response for any discrete-time LTI system H:

$$H(e^{j\omega}) = \text{DTFT}\{h[n]\} = \sum_{n\in\mathbb{Z}} h[n] e^{-j\omega n}$$

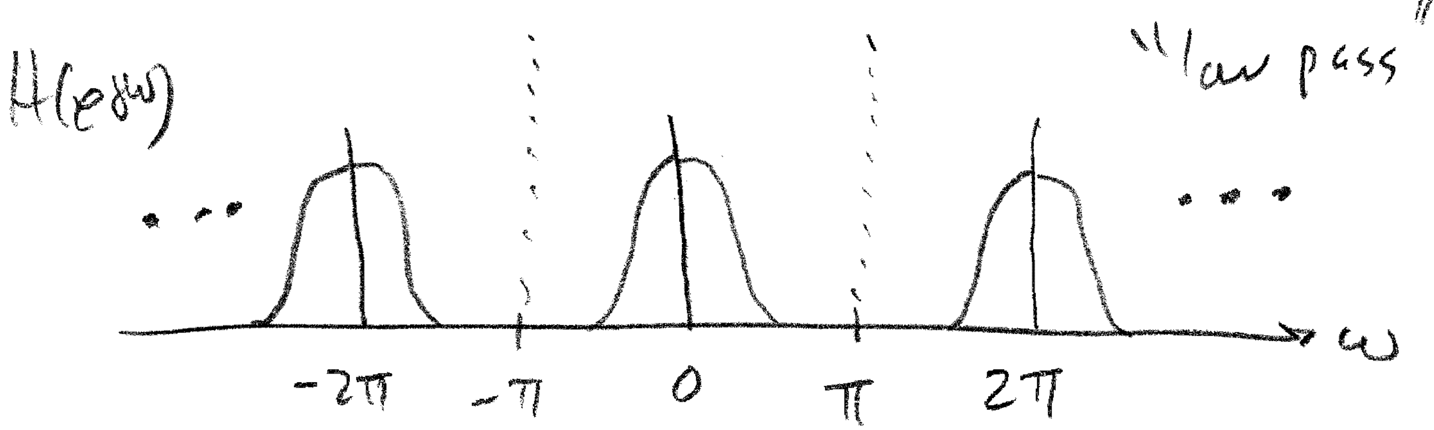
NOTE: for the continuous-time Fourier transform, $x(t)$ and $X(\Omega)$ were both "like" continuous-time signals... in the sense that they both had domain $\mathbb{R}$:



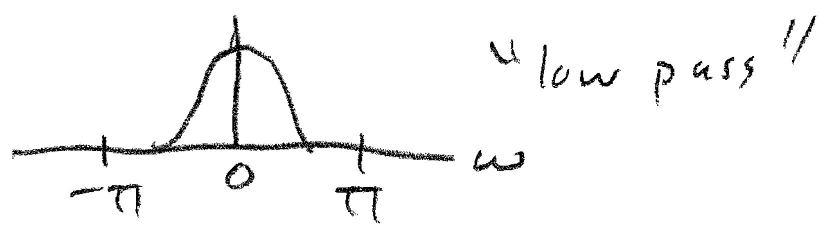
- we will see later that, for the discrete Fourier transform, or DFT, the signal $x[n]$ and it's DFT $X(k)$ are both "like" discrete-time signals... in the sense that the both have domain $[0, N-1] \subset \mathbb{Z}$.
- The DTFT is weird in this way:
  - $x[n]$ is "like" a discrete-time signal in the sense that it has domain $\mathbb{Z}$.
  - $X(e^{j\omega})$ has domain $\mathbb{R}$... it's "like" a continuous-time signal in this sense.

- The 2π-periodicity of the DTFT may take some getting used to at first.

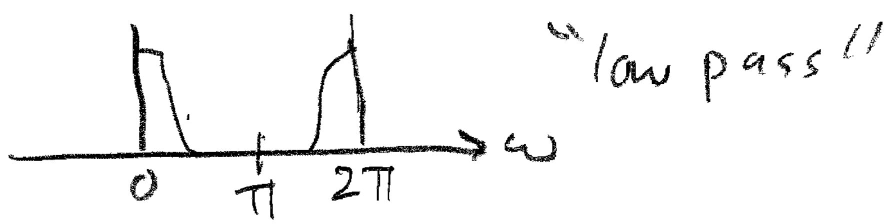
- For example, consider a low-pass filter (LTI) H. The frequency response $H(e^{j\omega})$ is 2π-periodic:



- Normally, only one period is graphed.
  - Some people prefer to graph it from $-\pi$ to π:



  - Others prefer to graph it from 0 to 2π:

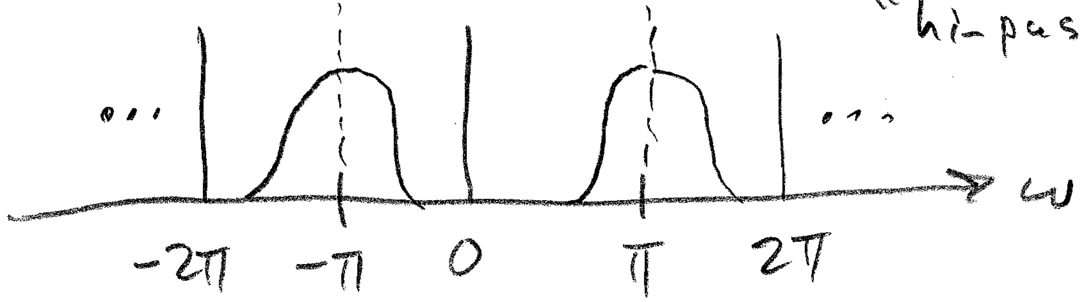


- So, the main lobe of $H(e^{j\omega})$ for a low-pass filter is concentrated about $\omega = \ldots -2\pi, 0, 2\pi \ldots$

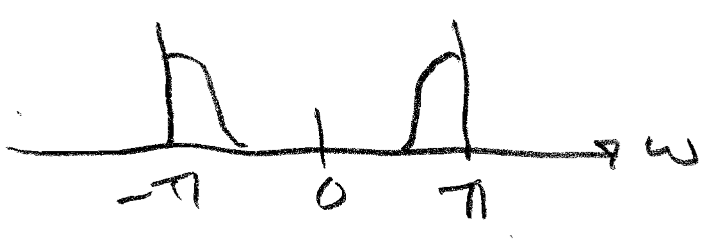
  i.e., even multiples of π.

- For a highpass filter, $H(e^{j\omega})$ has a main lobe that is concentrated about odd multiples of π:

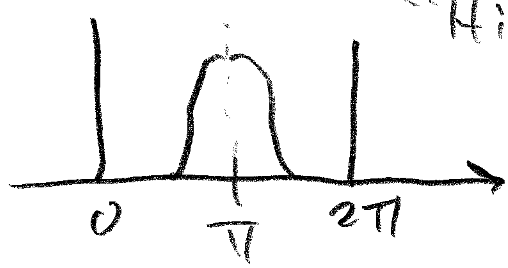
"hi-pass"



"Hi pass"

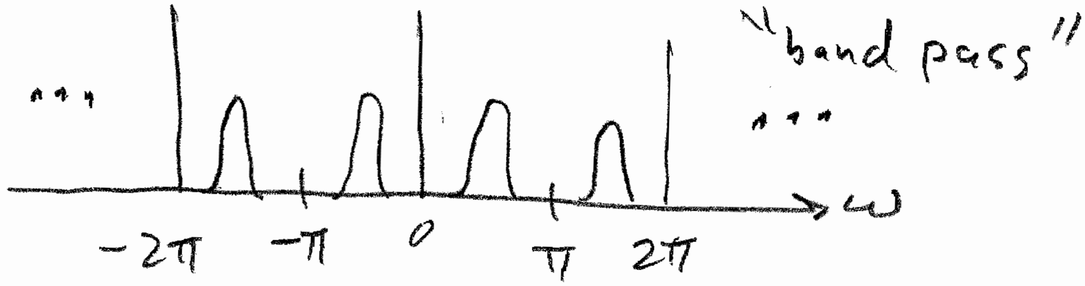


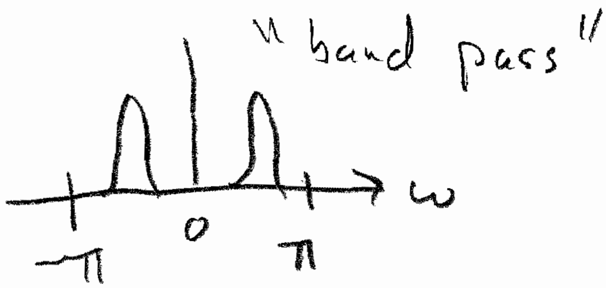
"Hipass"

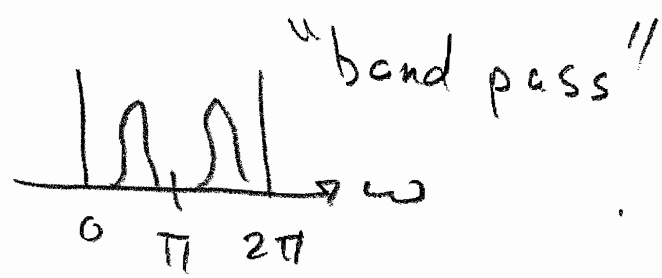


Band Pass:

3-98







Convergence:

- The Fourier pair $x[n] \overset{DTFT}{\longleftrightarrow} X(e^{j\omega})$ converges absolutely if $x[n] \in \ell^1(\mathbb{Z})$, i.e., if $x[n]$ is absolutely summable such that the ℓ^1-norm is finite:

$$\|x[n]\|_{\ell^1} = \sum_{n \in \mathbb{Z}} |x[n]| < \infty .$$

- $\forall\, x[n] \in \ell^2(\mathbb{Z})$, the DTFT pair

  $x[n] \longleftrightarrow X(e^{j\omega})$ converges either uniformly or in the mean-square sense.

- The DTFT also converges at least in a distributional sense for many $x[n]$ that are neither ℓ^1 nor ℓ^2.

- $X(e^{j\omega})$ is a complex-valued function of ω:

$$X(e^{j\omega}) = X_{re}(e^{j\omega}) + j X_{im}(e^{j\omega})$$

$$= \underbrace{A(\omega)}_{\text{spectral magnitude}} e^{j \overbrace{\theta(\omega)}^{\text{spectral phase.}}}$$

- usually, we want to have $A(\omega) \geq 0$.

- In order to improve the behavior of $\theta(\omega)$, we will sometimes allow $A(\omega)$ to be <u>signed</u> ... $\theta(\omega)$ is then called "generalized phase."

For example, suppose $x[n] = \delta[n] + \delta[n-2]$. 3-99B

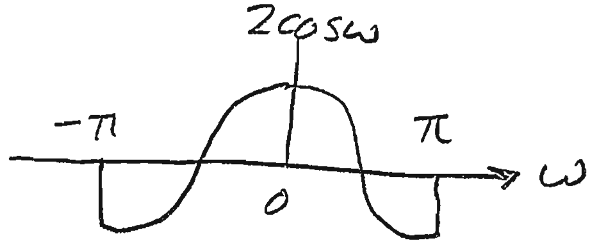
- Then $X(e^{j\omega}) = 1 + e^{-j2\omega}$

$$= [e^{j\omega} + e^{-j\omega}]e^{-j\omega}$$

$$= \left[2\cos\omega\right]e^{-j\omega}$$

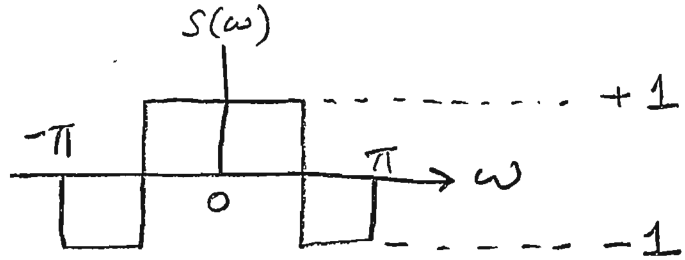
→ here, $2\cos\omega$ is a <u>real-valued</u> function,

→ But strictly speaking it can't be the spectral <u>magnitude</u> $A(\omega)$, because $2\cos\omega$ is negative for $\omega \in (\frac{\pi}{2}, \pi]$ and $\omega \in [-\pi, \pi/2)$,



⇒ whereas a <u>true</u> <u>magnitude</u> function has to be ≥ 0 for all ω.

- To find the true spectral magnitude $A(\omega)$, we would need to "fix" the sign of $2\cos\omega$ by using another function $S(\omega)$



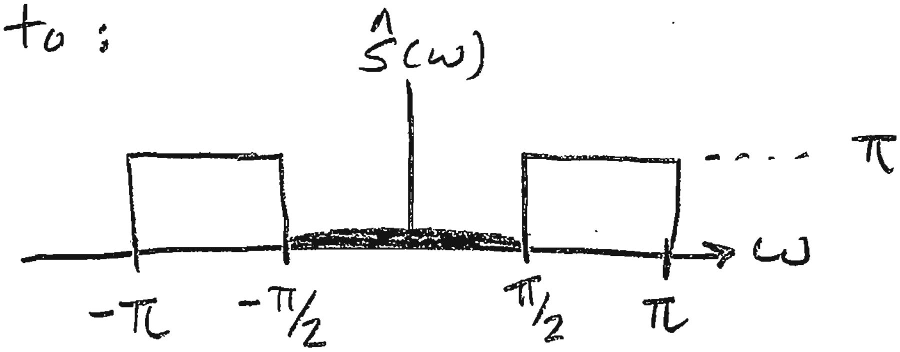
to write $X(e^{j\omega}) = 2|\cos\omega|S(\omega)e^{-j\omega}$ ⟶

Then we could let $A(\omega) = 2|\cos\omega|$, which is everywhere ≥ 0.

3-99c

- But we would also have to find a way to get $S(\omega)$ into the spectral phase function $\theta(\omega)$.

$\Rightarrow$ Because $+1 = e^{j0}$ and $-1 = e^{j\pi}$, this can be done by defining another function $\hat{S}(\omega)$ according to:



- Then we can write

$$
\begin{aligned}
X(e^{j\omega}) &= [2\cos\omega]\, e^{-j\omega} \\
&= 2|\cos\omega|\, S(\omega)\, e^{-j\omega} \\
&= 2|\cos\omega|\, e^{j\hat{S}(\omega)}\, e^{-j\omega} \\
&= 2|\cos\omega|\, e^{-j(\omega - \hat{S}(\omega))}
\end{aligned}
$$

To get

$$A(\omega) = 2|\cos\omega|$$
$$\theta(\omega) = -[\omega - \hat{S}(\omega)]$$

$\longrightarrow$

- But this is very complicated
- And it's actually undesirable for analyzing the frequency response of an LTI filter.

  ⇒ Because the spectral phase of $H(e^{j\omega})$ characterizes the delay of the filter.

  ⇒ But $\hat{s}(\omega)$ does not correspond to delay... it just accounts for sign.

- For this reason, DSP engineers will often allow $A(\omega)$ to be a signed real function.

i.e., if $h[n] = \delta[n] + \delta[n-2]$,

Then DSP engineers will often write

$$H(e^{j\omega}) = [2\cos\omega]\, e^{-j\omega}$$

$$A(\omega) = 2\cos\omega$$
$$\theta(\omega) = -\omega$$

⇒ where $A(\omega)$ is real but might be negative for some frequencies.

- When this is done, it is understood that $A(\omega)$ is not a true magnitude... but could be made ≥ 0 by adding a more complicated term $\hat{s}(\omega)$ to the phase $\theta(\omega)$ to account for sign.

Polar $\leftrightarrow$ Rectangular :

$$X_{re}(e^{j\omega}) = A(\omega)\cos\theta(\omega)$$

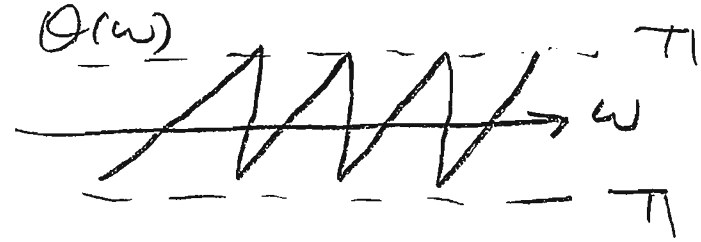
$$X_{im}(e^{j\omega}) = A(\omega)\sin\theta(\omega)$$

$$A(\omega) = \left[X_{re}^2(e^{j\omega}) + X_{im}^2(e^{j\omega})\right]^{1/2}$$

$$\theta(\omega) = \arctan\left[X_{im}(e^{j\omega}) / X_{re}(e^{j\omega})\right]$$

- for computer implementations, the "atan2" function will return a value between $-\pi$ and π.

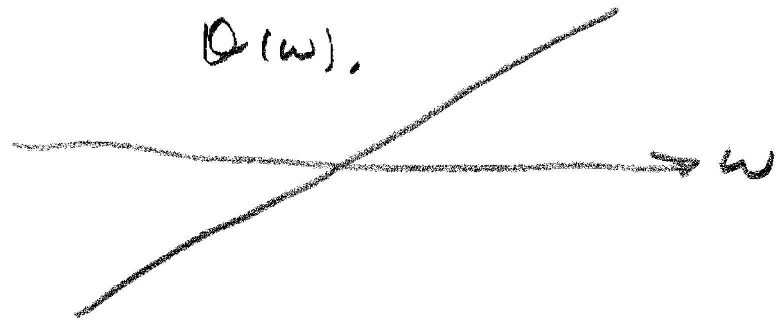
EX :



- This is called the "principal" value of the phase.

NOTE: the book writes "ARG $X(e^{j\omega})$" with capital letters to indicate the principal value of the phase... see Section 5.1.1 on page 275.

- Sometimes, we would prefer an interpretation where the phase $\theta(\omega)$ is more continuous:



- Both models are mathematicall equivalent, since the function $e^{j\theta(\omega)}$ is 2π-periodic in $\theta(\omega)$... for any $\omega \in \mathbb{R}$, you can add any integer multiple of 2π to $\theta(\omega)$ and the function $X(e^{j\omega})$ is not changed.

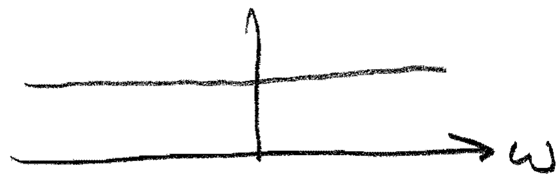
- "Phase unwrapping" means:

  take the principal phase returned by arctan and add multiples of 2π as needed to obtain an equivalent but smoother phase function $\theta(\omega)$.

<u>EX</u>: $x[n] = \delta[n]$

$$X(e^{j\omega}) = \sum_{n \in \mathbb{Z}} x[n]\, e^{j\omega n}$$

$$= \sum_{n \in \mathbb{Z}} \delta[n]\, e^{j\omega n} = \underline{\underline{1}}$$



$$DTFT^{-1}\{1\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\omega n}\, d\omega$$

$$= \frac{1}{j2\pi n} e^{j\omega n} \Big|_{\omega=-\pi}^{\pi}$$

$$= \frac{1}{n\pi} \frac{e^{j\pi n} - e^{-j\pi n}}{2j}$$

$$= \frac{\sin n\pi}{n\pi} \stackrel{?}{=} \delta[n]$$

<u>$n \neq 0$</u>

$$\frac{\sin n\pi}{n\pi} = \frac{0}{\text{not zero}} = 0$$

$\longrightarrow$

when $n=0$, we have by L'Hospital's rule:

$$\frac{\sin n\pi}{n\pi} = \lim_{A\to 0} \frac{\sin A\pi}{A\pi}$$

$$= \lim_{A\to 0} \frac{\frac{d}{dA}\sin A\pi}{\frac{d}{dA} A\pi}$$

$$= \lim_{A\to 0} \frac{\pi\cos A\pi}{\pi} = \lim_{A\to 0} \cos A\pi = 1.$$

So,

$$DTFT^{-1}\{1\} = \frac{\sin n\pi}{n\pi} = \delta[n] \checkmark$$

Symmetry of the DTFT:

- All similar to the symmetries of the continuous-time Fourier transform.

if $x[n] = x_{re}[n] + j x_{im}[n]$ real/imaginary parts

$= x_{cs}[n] + x_{ca}[n]$ conjugate symmetric/antisymmetric parts

with $x[n] \overset{DTFT}{\longleftrightarrow} X(e^{j\omega})$, then

$$x_{re}[n] \longleftrightarrow X_{cs}(e^{j\omega})$$

$$j x_{im}[n] \longleftrightarrow X_{ca}(e^{j\omega})$$

$$x_{cs}[n] \longleftrightarrow X_{re}(e^{j\omega})$$

$$x_{ca}[n] \longleftrightarrow j X_{im}(e^{j\omega})$$

$$x[-n] \longleftrightarrow X(e^{-j\omega})$$

$$x^*[-n] \longleftrightarrow X^*(e^{j\omega})$$

- If $x[n]$ is real, then

$$x[n] = x_{ev}[n] + x_{od}[n]$$

and:

$X(e^{j\omega})$ is conjugate symmetric:

real part even
imaginary part odd
magnitude even
phase odd.

$$x_{ev}[n] \longleftrightarrow X_{re}(e^{j\omega})$$
$$x_{od}[n] \longleftrightarrow jX_{im}(e^{j\omega}).$$

- if $x[n]$ is real and even, then $X(e^{j\omega})$ is real and even.

- if $x[n]$ real and odd, then $X(e^{j\omega})$ is pure imaginary and odd

DTFT Pairs: you should already know how to prove these from your undergrad Signals & Systems course. 3-106

$$a^n u[n], \ |a|<1 \longleftrightarrow \frac{1}{1-ae^{-j\omega}}$$

$$x[n] = \begin{cases} 1, & -N \le n \le N \\ 0, & \text{otherwise} \end{cases} \longleftrightarrow \frac{\sin\left[\omega\left(N+\frac{1}{2}\right)\right]}{\sin\frac{\omega}{2}}$$

$$x[n] = \frac{\sin \omega_c n}{\pi n}, \quad 0 < \omega_c < \pi$$

$$\longleftrightarrow \begin{cases} 1, & 0 \le |\omega| \le \omega_c \\ 0, & \omega_c < |\omega| \le \pi \end{cases}$$

and, outside the interval $-\pi < \omega \le \pi$, $X(e^{j\omega})$ is 2π periodic in ω.

$$\delta[n] \longleftrightarrow 1$$

$$\delta[n-n_0] \longleftrightarrow e^{-j\omega n_0}$$

$$(n+1)a^n u[n], \ |a|<1 \longleftrightarrow \frac{1}{(1-ae^{-j\omega})^2}$$

$$-a^n u(-n-1), \ |a|>1 \longleftrightarrow \frac{1}{1-ae^{-j\omega}}$$

$$\frac{(n+r-1)!}{n!(r-1)!} a^n u[n], \ |a|<1 \longleftrightarrow \frac{1}{(1-ae^{-j\omega})^r}$$

Distributional DTFT Pairs:

$$e^{j\omega_0 n} \longleftrightarrow 2\pi \sum_{k=-\infty}^{\infty} \delta(\omega - \omega_0 - 2\pi k)$$

$$\cos \omega_0 n \longleftrightarrow \pi \sum_{k=-\infty}^{\infty} \left[\delta(\omega - \omega_0 - 2\pi k) + \delta(\omega + \omega_0 - 2\pi k) \right]$$

$$\sin \omega_0 n \longleftrightarrow -j\pi \sum_{k=-\infty}^{\infty} \left[\delta(\omega - \omega_0 - 2\pi k) - \delta(\omega + \omega_0 - 2\pi k) \right]$$

$$x[n] = 1 \longleftrightarrow 2\pi \sum_{k=-\infty}^{\infty} \delta(\omega - 2\pi k)$$

$$\sum_{k=-\infty}^{\infty} \delta[n - kN] \longleftrightarrow \frac{2\pi}{N} \sum_{k=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi k}{N}\right)$$

$$u[n] \longleftrightarrow \frac{1}{1 - e^{-j\omega}} + \sum_{k=-\infty}^{\infty} \pi \delta(\omega - 2\pi k)$$

DTFT Properties:

linearity: if $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$, $y[n] \xleftrightarrow{DTFT} Y(e^{j\omega})$, and $a, b \in \mathbb{C}$ are constants, then

$$ax[n] + by[n] \xleftrightarrow{DTFT} aX(e^{j\omega}) + bY(e^{j\omega}).$$

Time Shift: if $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$ and $n_0 \in \mathbb{Z}$,

then $x[n-n_0] \xleftrightarrow{DTFT} e^{-j\omega n_0} X(e^{j\omega})$.

Time Reversal: If $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$, then

$$x[-n] \xleftrightarrow{DTFT} X(e^{-j\omega}).$$

Frequency Shifting: if $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$, then

$$e^{j\omega_0 n} x[n] \xleftrightarrow{DTFT} X(e^{j(\omega-\omega_0)}).$$

Conjugation: if $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$, then

$$x^*[n] \xleftrightarrow{DTFT} X^*(e^{-j\omega}).$$

Time Difference: if $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$, then

$$x[n] - x[n-1] \xleftrightarrow{DTFT} (1-e^{-j\omega})X(e^{j\omega}).$$

Time Accumulation:

if $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$, then

$$y[n] = \sum_{k=-\infty}^{n} x[k] \xleftrightarrow{DTFT} \frac{1}{1-e^{-j\omega}} X(e^{j\omega}).$$

DTFT Properties...

Frequency Differentiation:

$$x[n] \xleftrightarrow{DTFT} X(e^{j\omega}) \text{ then}$$

$$n x[n] \xleftrightarrow{DTFT} j\frac{d}{d\omega} X(e^{j\omega}).$$

Parseval Relation: if $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$ and $x[n] \in \ell^2(\mathbb{Z})$, then

$$\|x[n]\|_{\ell^2}^2 = \sum_{n=-\infty}^{\infty} |x[n]|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |X(e^{j\omega})|^2 d\omega$$

- The right-hand side is like an L^2-norm for the 2π-periodic spectrum $X(e^{j\omega})$... it computes a Hilbert space norm over only one period.

Upsampling (time expansion):

if $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$ and

$$x_{(k)}[n] = \begin{cases} x[n/k], & n \text{ is an integer multiple of } k. \\ 0, & \text{otherwise,} \end{cases}$$

then

$$x_{(k)}[n] \xleftrightarrow{DTFT} X(e^{jk\omega}).$$

Frequency Convolution: if $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$ and $y[n] \xleftrightarrow{DTFT} Y(e^{j\omega})$, then

$$x[n]y[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\theta})Y(e^{j(\omega-\theta)})\,d\theta$$

Time Convolution: if $x[n] \xleftrightarrow{DTFT} X(e^{j\omega})$ and $h[n] \xleftrightarrow{DTFT} H(e^{j\omega})$, then

$$x[n] * h[n] \xleftrightarrow{DTFT} X(e^{j\omega})H(e^{j\omega}).$$

→ As in the continuous-time case, this implies that:

For an LTI discrete-time system H with impulse response $h[n]$ and frequency response $H(e^{j\omega})$,

- if we represent the input as a sum of weighted spectral basis functions:

$$x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})e^{j\omega n}\,d\omega,$$

- then each term comes through the system multiplied by a complex eigenvalue $H(e^{j\omega})$ ...

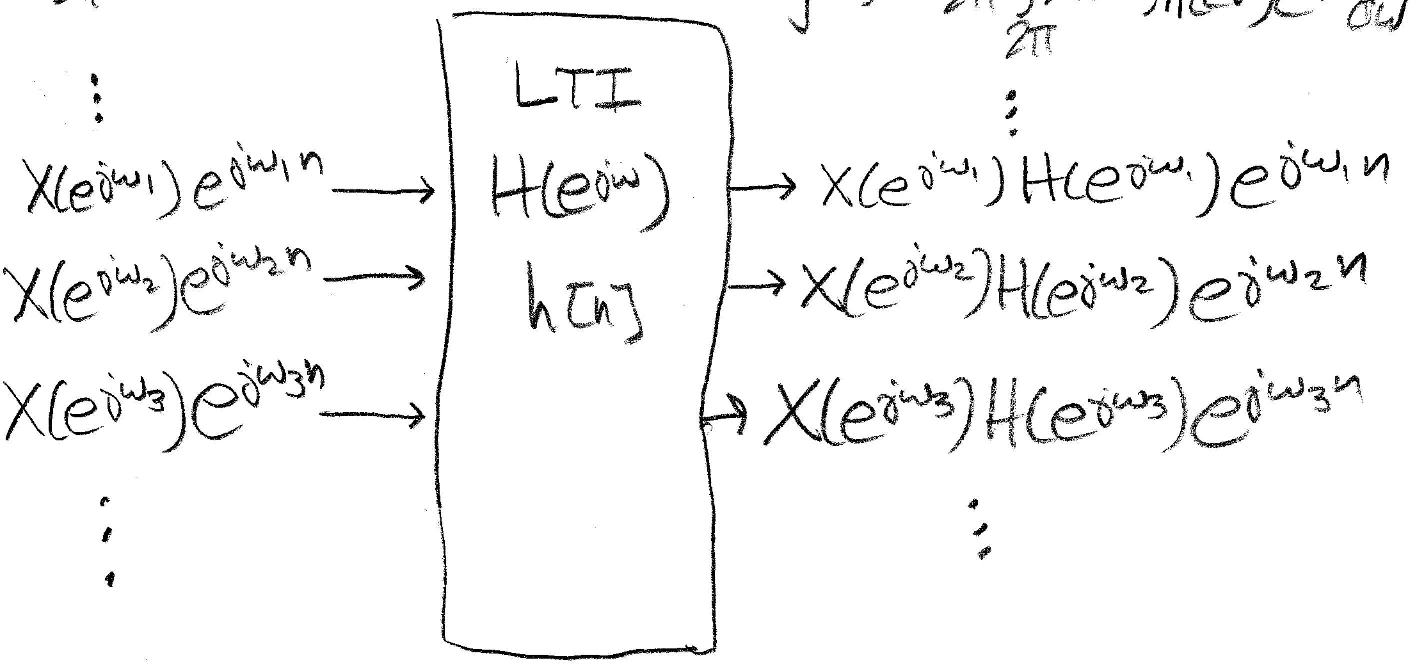
- in other words, the coordinates of $y[n]$ with respect to the spectral basis are given by $X(e^{j\omega})H(e^{j\omega})$.

- in other words, $Y(e^{j\omega}) = X(e^{j\omega})H(e^{j\omega})$

and

$$y[n] = x[n] * h[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})H(e^{j\omega})e^{j\omega n}\,d\omega$$

$$x[n] = \frac{1}{2\pi}\int_{2\pi} X(e^{j\omega})e^{j\omega n}\,d\omega \qquad\qquad y[n] = \frac{1}{2\pi}\int_{2\pi} X(e^{j\omega})H(e^{j\omega})e^{j\omega n}\,d\omega$$



- So, what happens when a complex sinusoid gets multiplied by a complex eigenvalue?

  - Consider the signal $x[n] = K e^{j(\omega_0 n + \phi)}$, where $K, \varphi \in \mathbb{R}$ and $K \geq 0$.

$$x[n] \longrightarrow \boxed{\begin{array}{c}\text{LTI}\\ H\end{array}} \longrightarrow y[n]$$

- Let $A(\omega) = |H(e^{j\omega})|$ and $\theta(\omega) = \arg H(e^{j\omega})$.
- Then

$$\begin{aligned}
y[n] &= x[n] * h[n] \\
&= H(e^{j\omega_0})\, x[n] \\
&= A(\omega_0) e^{j\theta(\omega_0)} K e^{j(\omega_0 n + \phi)} \\
&= K A(\omega_0) e^{j(\omega_0 n + \phi + \theta(\omega_0))}
\end{aligned}$$

$\Rightarrow$ The <u>magnitude</u> of the signal is scaled by the magnitude of the eigenvalue.

$\Rightarrow$ The <u>phase</u> of the signal is shifted by the angle of the eigenvalue.

- if the system impulse response $h[n]$ is real, then $H(e^{j\omega})$ is conjugate symmetric and a similar result holds for real-valued sinusoids.

$$x[n] \longrightarrow \boxed{\begin{matrix}\text{LTI}\\ H\end{matrix}} \longrightarrow y[n]$$

- Let $h[n] \in \mathbb{R}$. Then $H(e^{j\omega}) = H^*(e^{-j\omega})$,

e.g. $$H(e^{j\omega}) = A(\omega)\, e^{j\theta(\omega)}$$

$$A(\omega) = A(-\omega)$$
$$\theta(\omega) = -\theta(-\omega)$$

- Let $K, \phi, \omega_0 \in \mathbb{R}$ and let $K \geq 0$.
- if $x[n] = K\cos[\omega_0 n + \phi]$, then

$$y[n] = x[n] * h[n] = h[n] * K\cos[\omega_0 n + \phi]$$

$$= h[n] * \left\{ \frac{K}{2} e^{j[\omega_0 n + \phi]} + \frac{K}{2} e^{-j[\omega_0 n + \phi]} \right\}$$

$$= \frac{K}{2} h[n] * e^{j[\omega_0 n + \phi]} + \frac{K}{2} h[n] * e^{j[-\omega_0 n - \phi]}$$

$$= \frac{K}{2} H(e^{j\omega_0}) e^{j[\omega_0 n + \phi]} + \frac{K}{2} H(e^{-j\omega_0}) e^{j[-\omega_0 n - \phi]}$$

$$= \frac{K}{2} A(\omega_0)\, e^{j\theta(\omega_0)} e^{j[\omega_0 n + \phi]} + \frac{K}{2} A(-\omega_0)\, e^{j\theta(-\omega_0)} e^{j[-\omega_0 n - \phi]}$$

...

$$\ldots y[n] = \frac{K}{2} A(\omega_0) e^{j[\omega_0 n + \phi + \theta(\omega_0)]}$$

$$+ \underbrace{\frac{K}{2} A(\omega_0)}_{\text{spectral magnitude is even}} e^{\overbrace{-j\theta(\omega_0)}^{\text{spectral phase is odd}}} e^{j[-\omega_0 n - \phi]}$$

$$= KA(\omega_0)\left\{ \frac{1}{2} e^{j[\omega_0 n + \phi + \theta(\omega_0)]} + \frac{1}{2} e^{-j[\omega_0 n + \phi + \theta(\omega_0)]} \right\}$$

$$= K A(\omega_0) \cos[\omega_0 n + \phi + \theta(\omega_0)]$$

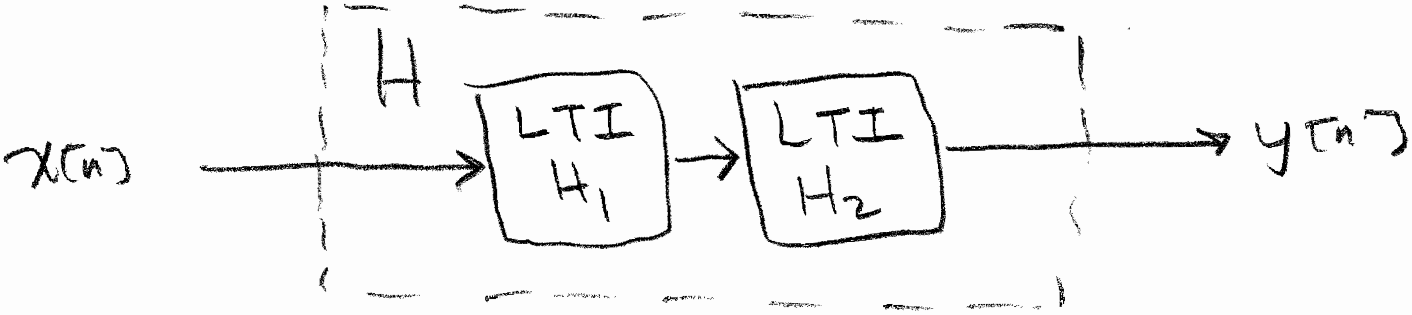
$\Rightarrow$ Amplitude is scaled by the magnitude of the eigenvalue.

$\Rightarrow$ Phase is shifted by the angle of the eigenvalue.

A nearly identical proof shows that when $x[n] = K\sin[\omega_0 n + \phi]$, the output is given by

$$y[n] = KA(\omega_0)\sin[\omega_0 n + \phi + \theta(\omega_0)].$$

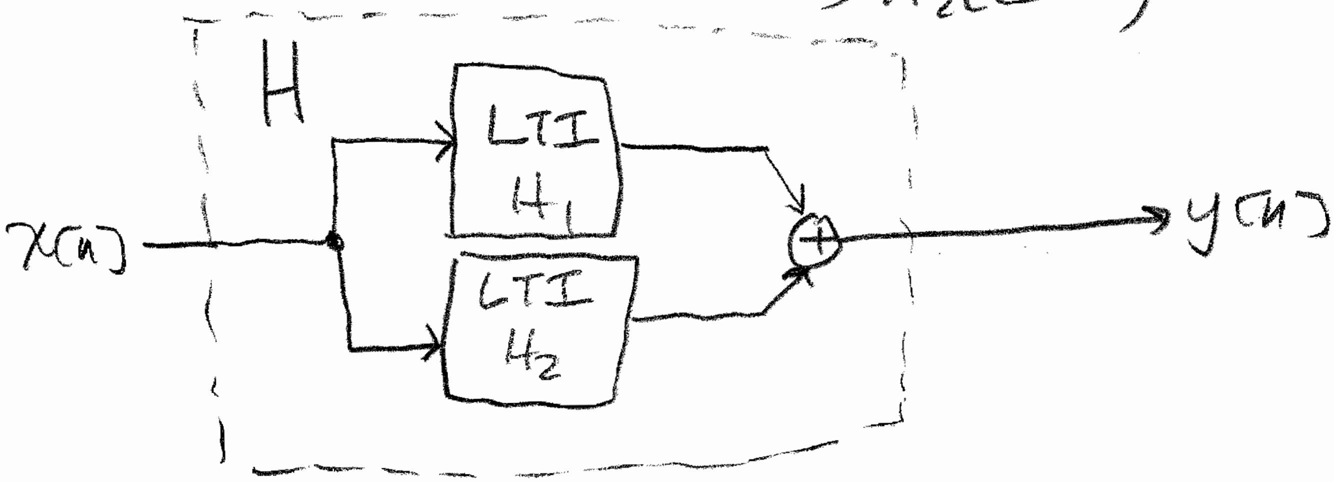
LTI System Connections



"Series" or "cascade" connection:

$$h[n] = h_1[n] * h_2[n]$$

$$H(e^{j\omega}) = H_1(e^{j\omega})\, H_2(e^{j\omega})$$

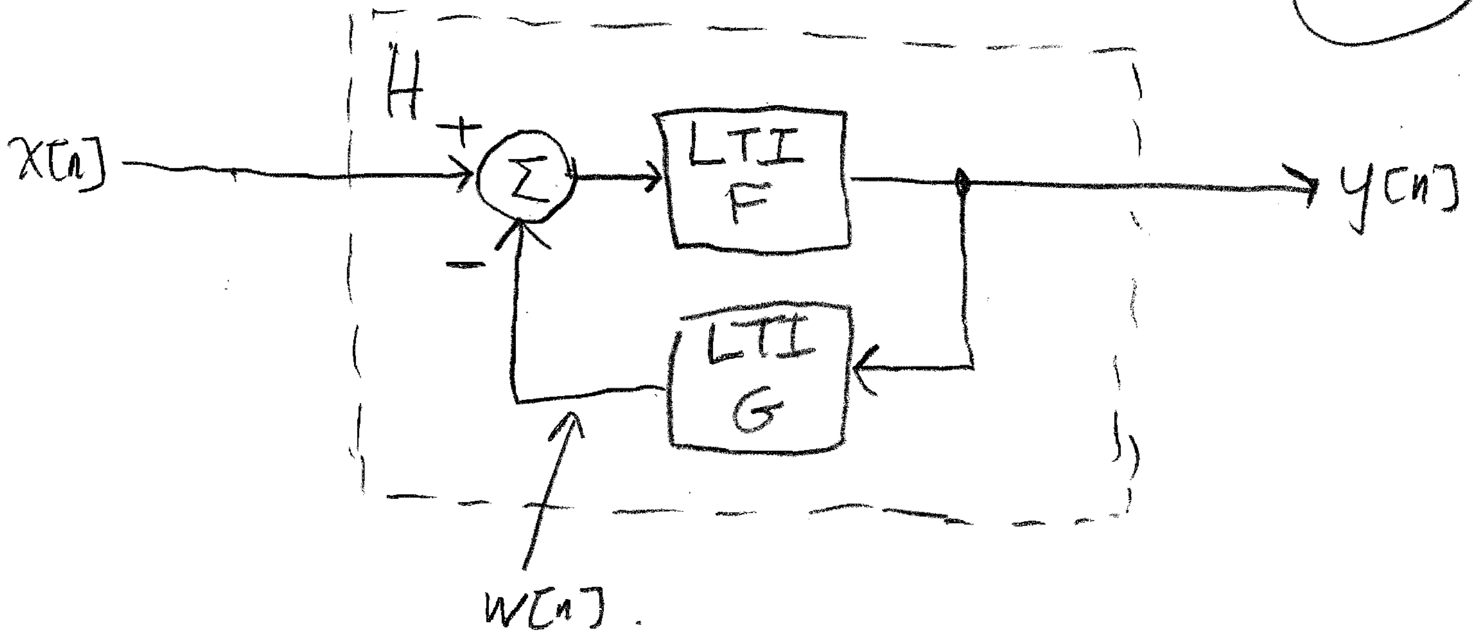


"parallel" connection:

$$h[n] = h_1[n] + h_2[n]$$

$$H(e^{j\omega}) = H_1(e^{j\omega}) + H_2(e^{j\omega})$$

Negative Feedback Connection:



- Call the output of G "w[n]".
- The input of G is y[n], so

$$W(e^{j\omega}) = Y(e^{j\omega})\, G(e^{j\omega}) \quad (*)$$

- The input of F is $x[n] - w[n] \longleftrightarrow X(e^{j\omega}) - W(e^{j\omega})$
- The output of F is y[n], so

$$Y(e^{j\omega}) = F(e^{j\omega})\left[X(e^{j\omega}) - W(e^{j\omega})\right]$$

$$= F(e^{j\omega})X(e^{j\omega}) - F(e^{j\omega})W(e^{j\omega}).$$

- plug in (*):

$$Y(e^{j\omega}) = F(e^{j\omega})X(e^{j\omega}) - F(e^{j\omega})G(e^{j\omega})Y(e^{j\omega})$$

So

$$Y(e^{j\omega})\left[1+F(e^{j\omega})G(e^{j\omega})\right] = F(e^{j\omega})X(e^{j\omega}).$$

Then

$$H(e^{j\omega}) = \frac{Y(e^{j\omega})}{X(e^{j\omega})} = \frac{F(e^{j\omega})}{1+F(e^{j\omega})G(e^{j\omega})}$$

Gain Function:

- For an LTI filter H, it is often desirable to look at a semilog plot of the spectral magnitude $A(\omega) = |H(e^{j\omega})|$. The "y-axis" is usually plotted in units of deciBells.
- The function that is plotted is called the <u>Gain Function</u> given by

$$\mathcal{G}(\omega) = 20\log_{10}|H(e^{j\omega})| \text{ dB}.$$

- Since $|H(e^{j\omega})|$ is even and 2π-periodic, you will sometimes see this plotted only for $-\pi \le \omega < \pi$ or $0 \le \omega \le \pi$.

- The negative of the gain function is called the "Attenuation function"
- In the Lab Manual, Mitra uses a script capital A for this:

$$H(e^{j\omega}) = A(\omega) e^{j\theta(\omega)}$$

$$|H(e^{j\omega})| = A(\omega)$$

$$\mathcal{A}(\omega) = 20 \log_{10} \frac{1}{A(\omega)} = -20 \log_{10} A(\omega) = -\mathcal{G}(\omega).$$

Phase Delay & Group Delay

- Let H be a discrete-time LTI system with a real impulse response $h[n]$.
- Then, as we just saw, if the input is

$$x[n] = K \cos[\omega_0 n + \phi],$$

the output is given by

$$y[n] = K A(\omega_0) \cos[\omega_0 n + \theta(\omega_0) + \phi]$$

$$= K A(\omega_0) \cos\left[\omega_0 \left(n - \frac{-\theta(\omega_0)}{\omega_0}\right) + \phi\right]$$

→ The input is scaled, and also delayed by $\frac{-\theta(\omega_0)}{\omega_0}$ samples, where $H(e^{j\omega}) = A(\omega) e^{j\theta(\omega)}$.

- The quantity $\frac{-\theta(\omega)}{\omega}$ is called the <u>phase delay</u> of the LTI system H.

- Note: if the spectral phase $\theta(\omega)$ is <u>linear</u> in ω, then the phase delay is <u>constant</u> and <u>does not depend on ω</u>.

- This implies that,
  - if we consider the input to be a sum of sinusoids,
  - All of the input sinusoids will be delayed by the <u>same</u> amount when they come through the system.
  - The time alignment between these sinusoids will be the same at the output as it was at the input.
  - The overall action of the system is to scale each input sinusoid by the magnitude of it's individual eigenvalue, and the whole signal also gets delayed by a fixed amount.

- Suppose that $\theta(\omega)$ is linear in ω, so that $\exists\, \alpha \in \mathbb{R}$ s.t. $\theta(\omega) = \alpha\omega$.

- Input: $x[n] = K\cos[\omega_0 n + \phi]$

  for some constants $K, \omega_0, \phi \in \mathbb{R}$ with $K > 0$.

$$x[n] \longrightarrow \boxed{\begin{matrix}\text{LTI}\\ H\end{matrix}} \longrightarrow y[n]$$

Then

$$
\begin{aligned}
y[n] &= KA(\omega_0)\cos[\omega_0 n + \theta(\omega_0) + \phi] \quad \text{as on p. 3-118}\\
&= KA(\omega_0)\cos[\omega_0 n + \alpha\omega_0 + \phi]\\
&= KA(\omega_0)\cos[\omega_0(n - \underbrace{-\alpha}_{\text{a time delay}}) + \phi]\\
&= A(\omega_0)\, x[n - -\alpha] \quad ///
\end{aligned}
$$

- The amplitude is scaled by $A(\omega_0)$ as on p. 3-118.
- And there is a constant time delay of $-\alpha$ samples... this delay does not depend on ω_0.

$\Rightarrow$ Since the time delay does not depend on ω, all input sinusoids are delayed by the same amount of time.

$\Rightarrow$ For an input that is a sum of sinusoids, the relative time alignment of the sinusoids is not disturbed when the signal passes through the system H... they are all delayed by the same amount.

- For this reason, it is HIGHLY DESIRABLE to design LTI filters H such that the spectral phase $\theta(\omega) = \arg H(e^{j\omega})$ is linear.

- Such systems are called "Linear phase LTI filters."

- The group delay of the LTI filter H is given by $\tau(\omega) = -\frac{d}{d\omega}\theta(\omega)$. [Book Equation (5.7) on p. 277]

  - if $\theta(\omega)$ is linear, then the group delay is constant. Input sinusoids are delayed by $-\frac{d}{d\omega}\theta(\omega)$ samples.

  - For general $\theta(\omega)$, the interpretation is that a group of sinusoids with frequency near ω_0 will all be delayed by approximately $-\frac{d}{d\omega}\theta(\omega)\Big|_{\omega=\omega_0}$ samples.

FIR filter: if an LTI filter H has an impulse response $h[n]$ that is nonzero for only a finite number of "n", then H is called a "finite impulse response filter."

- This implies that H is an "all zero" filter.

EX: Suppose H is causal and that $h[n]$ has length N, so that

$$h[n] = a_0\delta[n] + a_1\delta[n-1] + \cdots + a_{N-1}\delta[n-(N-1)]$$

for some constants $a_k \in \mathbb{C}$, $0 \le k < N$.

- Then the transfer function is

$$H(z) = \mathcal{Z}\{h[n]\} = \sum_{n\in\mathbb{Z}} h[n] z^{-n}$$

$$= a_0 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{N-1} z^{-(N-1)}$$

$$= \frac{a_0 z^{N-1} + a_1 z^{N-2} + \cdots + a_{N-2} z^1 + a_{N-1}}{z^{N-1}}$$

$\longrightarrow$

- Since the numerator is a polynomial in z of order $N-1$, it has $N-1$ roots z_k and $H(z)$ can be factored as

$$H(z) = \frac{\prod_{k=1}^{N-1} (z - z_k)}{z^{N-1}} \cdot (\text{a constant})$$

- In other words:

  - H has $N-1$ trivial poles at $z=0$.
  - The ROC of $H(z)$ is $|z| > 0$ ... i.e. all z except the point $z=0$.
  - The filter has $N-1$ zeros that determine the shape of the spectral magnitude $A(\omega) = |H(e^{j\omega})|$.

- In terms of the language of "ARMA" systems, an FIR filter is an "all MA" filter: it does not have any nontrivial AR terms.

→

- To see this, consider again a causal FIR filter H with

$$h[n] = a_0 \delta[n] + a_1 \delta[n-1] + \ldots + a_{N-1} \delta[n-(N-1)],$$

- The output is given by

$$y[n] = x[n] * h[n]$$

$$= a_0 x[n] + a_1 x[n-1] + \ldots + a_{N-1} x[n-(N-1)]$$

→ N nontrivial MA terms
  0 nontrivial AR terms.

- An LTI filter that is not FIR is called IIR... "infinite impulse response."

<u>EX</u>: $h[n] = \left(\frac{1}{2}\right)^n u[n]$.

- Since an infinite collection of MA terms can be traded for a finite collection of AR terms,
  - an IIR filter generally has both nontrivial AR terms and nontrivial MA terms.

- This implies that the transfer function $H(z)$ will generally have both a nontrivial numerator and a nontrivial denominator.

$\Rightarrow$ There will generally be nontrivial zeros and nontrivial poles that determine the shape of $A(\omega) = |H(e^{j\omega})|$.

<u>EX</u> : $h[n] = \left(\frac{1}{2}\right)^n u[n]$

$$y[n] = x[n] * h[n]$$

$$= \sum_{k \in \mathbb{Z}} h[k]\, x[n-k]$$

$$= \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k x[n-k]$$

$$= x[n] + \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k x[n-k]$$

$$= x[n] + \frac{1}{2} \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k x[(n-1)-k]$$

$$y[n] = x[n] + \frac{1}{2} y[n-1]$$ Two AR terms.

$$y[n] - \frac{1}{2} y[n-1] = x[n] \quad (*)$$

- To find the transfer function, take z-transform on both sides of (*):

$$Y(z) - \frac{1}{2} z^{-1} Y(z) = X(z)$$

$$Y(z)\left[1 - \frac{1}{2} z^{-1}\right] = X(z)$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 - \frac{1}{2} z^{-1}}$$

$\Rightarrow$ This filter has a nontrivial pole at $z = \frac{1}{2}$.

- In general, finding the spectral phase $\theta(\omega)$ of an LTI filter $H(e^{j\omega})$ can be a laborious calculation, because

$$\theta(\omega) = \arg H(e^{j\omega}) = \arctan\left\{\frac{\mathrm{Im}[H(e^{j\omega})]}{\mathrm{Re}[H(e^{j\omega})]}\right\}.$$

- In many cases it may actually be simpler to find the group delay directly using the formula

$$\frac{d}{dx}\arctan u = \frac{1}{1+u^2}\frac{du}{dx}$$

(see p. 1 of the course formula sheet... good for $-\pi/2 < \arctan u < \frac{\pi}{2}$

- However, if H is an FIR filter,
  - and if the impulse response $h[n]$ has even or odd symmetry <u>about the midpoint</u>,

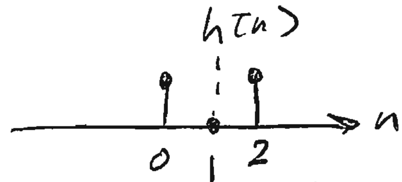
    $\Rightarrow$ Then there is a trick that makes it easy to find $\theta(\omega)$.

  - It is a fact that we will prove later that: if H is FIR and $h[n]$ has symmetry about the midpoint, then H has <u>linear phase</u>.

- I call this trick the "linear phase FIR half power trick!"

The Trick: if H is an FIR filter and $h[n]$ has even or odd symmetry about the midpoint, then you can easily write $H(e^{j\omega})$ in the form $A(\omega)e^{j\theta(\omega)}$ by factoring out half the highest power of $e^{-j\omega}$ that appears.

EX: $h[n] = \delta[n] + \delta[n-2]$



even symmetric about the midpoint $n=1$.

$$H(e^{j\omega}) = 1 + e^{-j2\omega}$$

highest power $= 2$
half the highest power $= 1$
$\rightarrow$ Factor out $e^{-j\omega}$

$$H(e^{j\omega}) = 1 + e^{-j2\omega}$$
$$= \left[e^{j\omega} + e^{-j\omega}\right]e^{-j\omega}$$
$$= \left[2\cos\omega\right]e^{-j\omega}$$

$$A(\omega) = 2\cos\omega \qquad \theta(\omega) = -\omega$$

NOTE: this is generalized linear phase, because $A(\omega) < 0$ for $\omega \in [-\pi, \pi/2)$ and for $\omega \in (\pi/2, \pi]$ ...
see pp. 3-99 through 3-99D.

EX: $h[n] = \delta[n] + 2\delta[n-1] + \delta[n-2]$

$$H(e^{j\omega}) = 1 + 2e^{-j\omega} + e^{-j2\omega}$$

highest power $= 2$
half the highest power $= 1$
factor out $e^{-j\omega}$

$$= [e^{j\omega} + 2 + e^{-j\omega}]e^{-j\omega}$$

$$= [2 + 2\cos\omega]e^{-j\omega}$$

$$A(\omega) = 2 + 2\cos\omega \qquad \theta(\omega) = -\omega$$

EX: $h[n] = \frac{1}{4}\delta[n] + \frac{1}{4}\delta[n-1] + \frac{1}{4}\delta[n-2] + \frac{1}{4}\delta[n-3]$

$$H(e^{j\omega}) = \frac{1}{4}[1 + e^{-j\omega} + e^{-j2\omega} + e^{-j3\omega}]$$

highest power $= 3$
half the highest power $= 3/2$
factor out $e^{-j\frac{3}{2}\omega}$

$$= \frac{1}{4}[e^{j\frac{3}{2}\omega} + e^{j\frac{1}{2}\omega} + e^{-j\frac{1}{2}\omega} + e^{-j\frac{3}{2}\omega}]e^{-j\frac{3}{2}\omega}$$

$$= \frac{1}{4}[2\cos(\tfrac{3}{2}\omega) + 2\cos(\tfrac{1}{2}\omega)]e^{-j\frac{3}{2}\omega}$$

$$= \frac{1}{2}[\cos(\tfrac{1}{2}\omega) + \cos(\tfrac{3}{2}\omega)]e^{-j\frac{3}{2}\omega}$$

$$A(\omega) = \frac{1}{2}[\cos(\tfrac{1}{2}\omega) + \cos(\tfrac{3}{2}\omega)] \qquad \theta(\omega) = -\frac{3}{2}\omega$$

- If $h[n]$ has odd symmetry about the midpoint, then you will get sines instead of cosines. (3-129)

<u>EX</u>: $h[n] = \delta[n] - \delta[n-1]$

$$H(e^{j\omega}) = 1 - e^{-j\omega}$$

highest power = 1
half the highest power = ½
factor out $e^{-j\frac{1}{2}\omega}$

$$= \left[e^{j\frac{1}{2}\omega} - e^{-j\frac{1}{2}\omega}\right]e^{-j\frac{1}{2}\omega}$$

$$= 2j\left[\frac{e^{j\frac{1}{2}\omega} - e^{-j\frac{1}{2}\omega}}{2j}\right]e^{-j\frac{1}{2}\omega}$$

$$= 2j\sin\left(\tfrac{1}{2}\omega\right)e^{-j\frac{1}{2}\omega}$$

Note: $j = e^{j\pi/2}$

$$= 2\sin\left(\tfrac{1}{2}\omega\right)e^{-j\frac{1}{2}\omega}e^{j\pi/2}$$

$$= 2\sin\left(\tfrac{1}{2}\omega\right)e^{-j\left(\frac{1}{2}\omega - \pi/2\right)}$$

$$A(\omega) = 2\sin\left(\tfrac{1}{2}\omega\right)$$

$$\theta(\omega) = -\left(\tfrac{1}{2}\omega - \pi/2\right) = -\tfrac{1}{2}\omega + \tfrac{\pi}{2}$$