Lecture 41, Fri 12/7/01

4 Partial Overview. ~~Review (and Preview of~~

As we saw last time and throughout this course, it is useful to model the inputs and outputs of systems as vectors, and then study representations of these vectors in bases. There are two primary reasons for this. First, the responses of a linear system to a few specific vectors completely determine its response to any vector representable as a linear combination of these few vectors.

Second, a good basis may provide a more compact, or sparse, representation for a class of signals, leading to compression, noise reduction, or some other type of enhancement. DNE

$\underline{v} = \sum_k a_k g_k \longrightarrow$ Linear system S $\longrightarrow S(\underline{v}) = \sum_k a_k S(g_k)$

- Know $S(\mathbf{g}_k)$'s $\Rightarrow$ can find the resp. of S to any $\mathbf{v}$ in the span of $\mathbf{g}_k$'s.
- Appropriate basis $\Rightarrow$
  - compression
  - noise removal
  - enhancement

We used this first property to show that, for a linear, time-invariant system, the output is the convolution of the input with the impulse response. DNE

$\underline{v} = \sum_{k=-\infty}^{\infty} v(k)\,\delta_k \longrightarrow$ LTI system S $\longrightarrow S(\underline{v}) = \sum_{k=-\infty}^{\infty} v(k)\,h_k$

where $\underbrace{h_k(n) = h(n-k)}_{\text{delayed impulse response}}$

Basis: delayed impulses

For LTI systems, we also found that representing signals as sums of complex exponentials was particularly convenient.

Definition . *A vector is an* **eigenvector** *of a linear system S if*

$$S(\mathbf{v}) = \alpha \mathbf{v}$$

The scalar α is then called the corresponding eigenvalue.

What we found was that a complex exponential function of any frequency was an eigenvector (or an eigenfunction) of any LTI system with absolutely summable impulse response.

Response to

$$\underbrace{\{e^{j\omega n},\ -\infty < n < \infty\}}_{\underline{V}_\omega = \begin{pmatrix} e^{j\omega(-1)} \\ e^{j\omega(0)} \\ e^{j\omega(1)} \\ \vdots \end{pmatrix}} = \sum_{-\infty}^{\infty} h(k) e^{j\omega(n-k)}$$

$$= e^{j\omega n} \sum_{-\infty}^{\infty} h(k) e^{-j\omega k}$$

$$= e^{j\omega n} H(e^{j\omega}), \quad -\infty < n < \infty$$

$$\boxed{S(\mathbf{v}_\omega) = \mathbf{v}_\omega H(e^{j\omega})}$$

the eigenvalue

This says that, if the input to such a system is a complex exponential, the calculation of the output is trivial: it's just the input multiplied by a complex number.

Now, which signals can be represented as linear combinations of these complex exponentials? It turns out that, for infinite-duration, non-periodic signals, a finite or even countable number of these exponentials is not enough. In other words, in order to represent, for example, ~~all~~ finite-energy signals, we have to have an uncountable ~~number of complex exponentials~~.

$$\ell^2: \quad \mathbf{v} = \frac{1}{2\pi} \int_{-\pi}^{\pi} V(e^{j\omega}) \mathbf{v}_\omega \, d\omega$$

$$v(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} V(e^{j\omega}) e^{j\omega n} \, d\omega \quad \text{- IDTFT}$$

The weights $V(e^{j\omega})$ are found from

$$V(e^{j\omega}) = \sum_{-\infty}^{\infty} v(n) e^{-j\omega n} \text{ - DTFT}$$

This situation is quite involved, and therefore we didn't go into any deep analysis here. The situation with N-periodic sequences, however, was much simpler, and we did analyze it in depth. What we saw is that we needed only N complex exponentials to represent any periodic sequence with period N:

$$\boxed{\begin{aligned} \text{All } N\text{-periodic sequences: } \mathbf{v} &= \sum_{k=0}^{N-1} V(k)\mathbf{g}_k, \\ \text{where } \mathbf{g}_k &= \frac{1}{N}\mathbf{v}_{\frac{2\pi k}{N}} \\ g_k(n) &= \frac{1}{N}e^{j\frac{2\pi k}{N}n} \end{aligned}}$$

– IDFT (or I DF Series)

The weights $V(k)$ are found from the DFT:

$$\boxed{V(k) = \sum_{n=0}^{N-1} v(n)e^{-j\frac{2\pi k}{N}n}}$$

DFT or DF Series

This leads to an alternative way to calculate the output of an LTI system

truncate, then FFT

Inf.-duration: $\underline{v} = \frac{1}{2\pi}\int_{-\pi}^{\pi} V(e^{j\omega})\underline{v}_\omega \, d\omega \longrightarrow$ LTI system S $\longrightarrow$

by linearity of S

$$S(\underline{v}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} V(e^{j\omega})S(\underline{v}_\omega)\,d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} V(e^{j\omega})H(e^{j\omega})\underline{v}_\omega\,d\omega$$

FFT

N-periodic: $\underline{v} = \sum_{k=0}^{N-1} V(k)\, g_k \longrightarrow$ LTI system S $\longrightarrow$

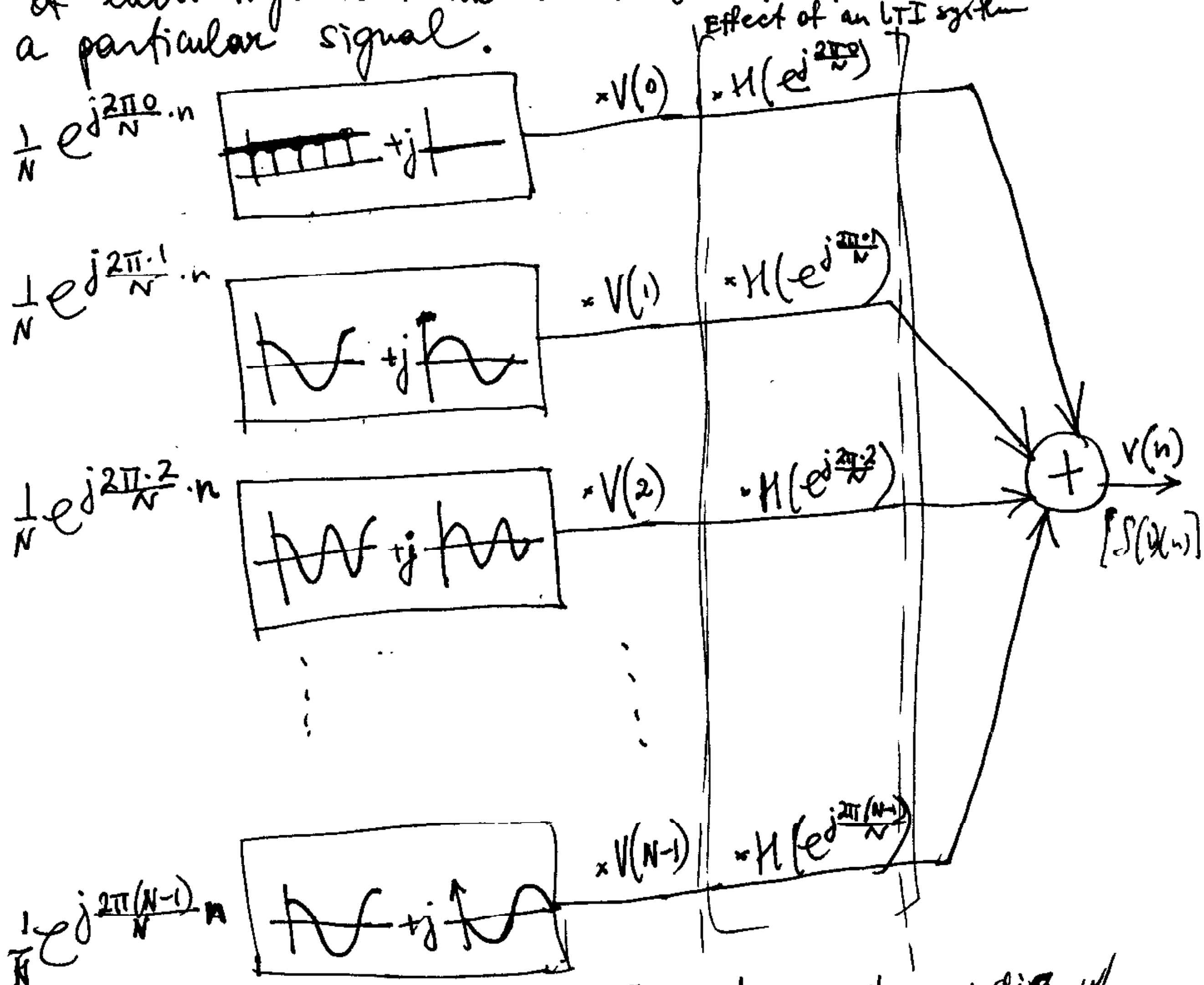
$$S(\underline{v}) = \sum_{k=0}^{N-1} V(k)\,S(g_k) = \sum_{k=0}^{N-1} V(k)\,\underbrace{H\left(e^{j\frac{2\pi k}{N}}\right)}_{\text{DFT of } h(n)} g_k$$

FFT

- Take Fourier transform of $\underline{v}$ $O(N\log N)$ with FFT
- Multiply it by the Fourier tr. of $\underline{h}$ $O(N)$
- Take inverse transform $O(N\log N)$

FFT is a clever algorithm for quickly computing the Fourier series coefficients. [Numerically, this is the same computation, because you would approximate the DTFT with DFT.]

It's important to realize that this is not just a convenient computational trick. This is a philosophy which views signals as synthesized from these basic building blocks, the complex exponentials. The Fourier transform and Fourier series is a "cookbook" which says how much of each ingredient we have to use in order to "cook" a particular signal.



(Note: these are DT, and therefore not necessarily periodic w/ period $\frac{N}{K}$.) ~~This picture accurately applies to CTFS~~).

When you put a signal through an LTI system, ~~you~~ all you do is change the quantities of these ingredients. For example, if H is a low-pass filter, it will ~~leave the low frequencies~~ preserve the low frequencies, and remove the high frequencies.

~~Also~~ We saw that LTI systems and Fourier basis can be very useful; moreover, they are fundamental. Once we understand this picture very well, we can go further.

* Non-TI systems

  E.g., our model of vocal tract. LTI model ⇒ we would always be saying the same phoneme.

* Non-linear systems

  Perona-Malik filter – can be locally low pass, high-pass, or all-pass, depending on the local ~~properties~~ contrast of the image – i.e., we are doing different things to the same frequency depending on where we are in the image.

~~* Non-Fourier bases~~

(Note that, in both cases, the point of departure for our design was an ~~linear~~ LTI system. ~~For an~~ AR model, and then we let its coefficients change in time. Here, it was a linear averaging filter, and then we ~~~~ modified it. So, in our analysis and design of non-time-inv., non-lin. systems, our intuition about LTI systems will typically play a major role.)

* ~~Non-Fourier Bases and~~ Non-Fourier transforms.

~~Here, we've laid some groundwork, which will be~~ ~~developed in much~~ more depth in EE648.

For example, what do we do to analyze the frequency content of a signal like this:

We would like to have a tool, a transform, which would tell us that these three time segments of this signal have different frequencies. The Fourier transform is not time-dependent, so it would mix all these frequencies together

Windowed (or short-time) Fourier transform: Partition the signal into small segments (perhaps, overlapping segments), and take the Fourier transform of each piece.

Wave function describing a particle: a finite-energy signal $x(t)$,

$$\mathcal{E}(x) = \int_{-\infty}^{\infty} |x(t)|^2 dt < \infty$$

Pdf of the location of the particle, T: $p_T(t) = \dfrac{|x(t)|^2}{\mathcal{E}(x)}$

Pdf of the momentum of the particle, F: $p_F(f) = \dfrac{|X(f)|^2}{\mathcal{E}(x)}$

Then

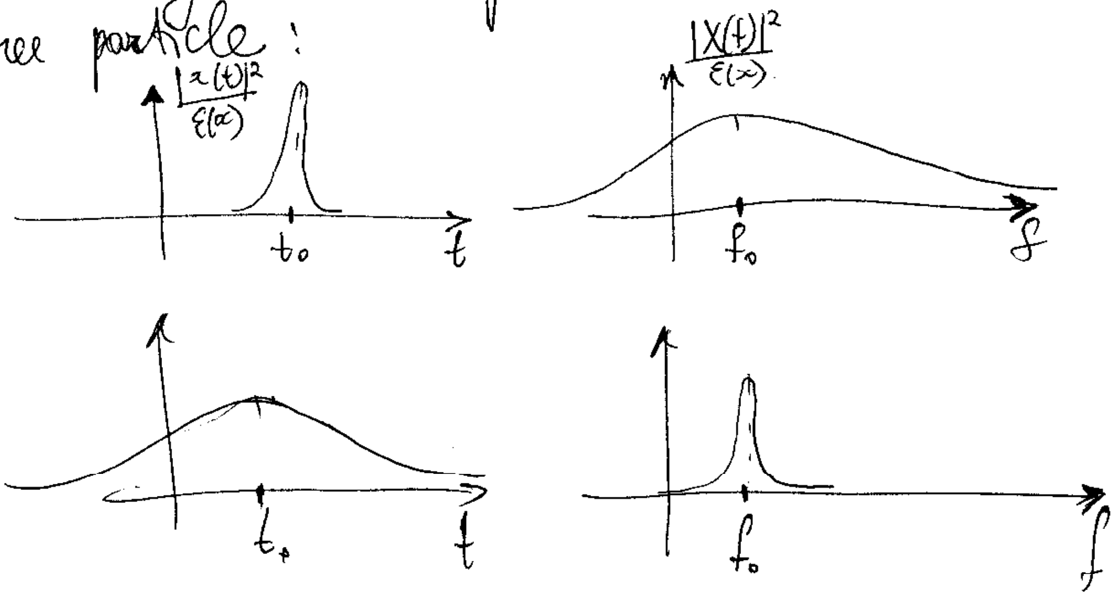
$$\sigma_T \sigma_F \geq \frac{1}{4\pi}$$

and the equality is achieved if and only if x is a modulated Gaussian:

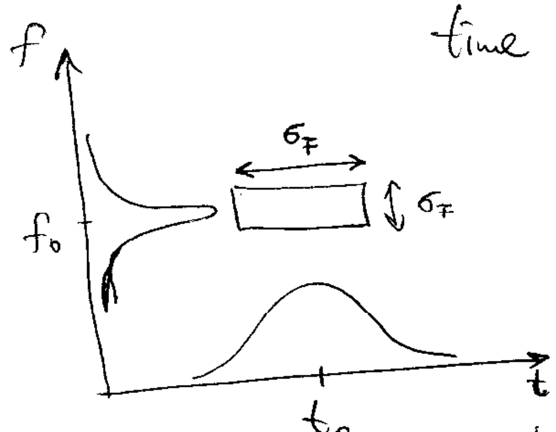
$$x(t) = a e^{j2\pi f_0 t} e^{-b(t-t_0)^2},$$

for some $a, b \in \mathbb{C}$, $t_0, f_0 \in \mathbb{R}$.

<u>Quantum mechanics</u>: cannot arbitrarily reduce uncertainty as to the position and the momentum of a free particle:



Signal processing: cannot get a window which has arbitrarily good resolution <u>both</u> in time and in frequency (8)



Best time-frequency resolution: Gaussian windows.

Last ~1/3 of the course: probability and random sequences.
Motivation: in real life, things are rarely deterministic.
There is always channel noise, measurement noise, quantization noise, etc.
$\Rightarrow$ very often, cannot precisely model a signal, but can model <u>average</u> behavior.

Our main thrust was considering various estimation or inference problems.

$$\mathcal{A} \longrightarrow \mathcal{B}$$

$\mathcal{A}$: object(s) of interest which we do not have access to

$\mathcal{B}$: observation(s)

we have a probabilistic model of the relationship between $\mathcal{A}$ and $\mathcal{B}$, in the form of $f_{\mathcal{B}|\mathcal{A}}$ or at least some parameters of $f_{\mathcal{B}|\mathcal{A}}$.
I.e., we have a model of the way that $\mathcal{A}$ gives rise to the observation $\mathcal{B}$.

We would like to infer a "best" estimate of $\mathcal{A}$ from $\mathcal{B}$:

$$\underbrace{\hat{\mathcal{A}}}_{\text{estimate of } \mathcal{A}} = \underbrace{\mathcal{E}}_{\text{estimator}}(\mathcal{B})$$

This is a <u>variety</u> of different problems, depending on:
- measurement model: full or partial specification of $f_{\mathcal{B}|\mathcal{A}}$.
- prior information about $\mathcal{A}$
- estimation ~~criterion~~ strategy:
  - optimize a criterion
  - choose a reasonable estimator, then show that it works

~~Ex 1. $\mathcal{B} = \{X_1, X_2, \dots, X_N\}$, indep~~

Ex 1. $\mathcal{B}$ is a r.v. Y

$\mathcal{A}$ is a r.v. X

prior info. on X: $f_X(x) =$ 

measurement model: $f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}}$

strategy: minimize $\text{Prob}(\hat{X} \neq X)$.

Ex 2. $\mathcal{B}$ is a sequence of r.v. $Y_1, Y_2, \dots, Y_N$ - iid

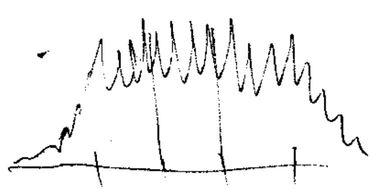
$\mathcal{A}$ is their pdf f_Y

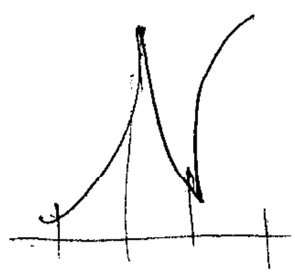
measurement model: given f_Y, the pdf of Y_n is f_Y!

strategy: histogram

BUT; depending on the bin size, can get anything

$\Rightarrow$ need prior info on f_Y (e.g., how slowly varying it is)





Ex 3. $\mathcal{B}$ is a seq. $Y_0, Y_1, \dots, Y_N$

$\mathcal{A}$ is a seq. $S_0, S_1, \dots, S_N$

prior info on S_n: ~~the seq~~ 1-st order AR sequence,

$S_n = aS_{n-1} + X_n$

measurement model: $Y_n = S_n + W_n$, uncorrelated, zero-mean, white

i.e., $E(Y_n|S_n) = S_n$

$\text{Var}(Y_n|S_n) = \text{Var}(S_n) + \text{Var}(W_n)$

criterion: find $\hat{S}_{n|n} = \sum_{k=0}^{n} a_k Y_k$ to minimize $E\left[(S_n - \hat{S}_{n|n})^2\right]$ $\rightarrow$ Kalman filter