Digital Audio and Speech Processing (Savısal Ses ve Konusma İsleme)

Prof. Dr. Nizamettin AYDIN

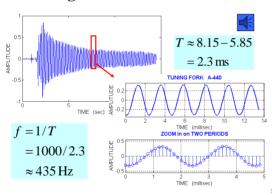
naydin@yildiz.edu.tr nizamettinavdin@gmail.com http://www3.vildiz.edu.tr/~naydin

Sinusoids

What's a signal

- A signal can be defined as
 - a pattern of variations of a physical quantity that can be manipulated, stored, or transmitted by physical process.
 - an information variable represented by physical quantity.
 - For digital systems, the variable takes on discrete values.
- In the mathematical sense it is a function of time, x(t), that carries an information.

Tuning Fork A-440 Waveform

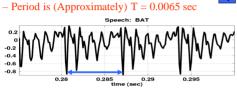


Speech Example

- More complicated signal (BAT.WAV)
- Waveform x(t) is NOT a Sinusoid
- · Theory will tell us
 - -x(t) is approximately a sum of sinusoids
 - FOURIER ANALYSIS
 - Break x(t) into its sinusoidal components
 - Called the FREQUENCY SPECTRUM

Speech Signal: BAT

- Nearly Periodic in Vowel Region



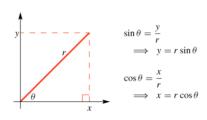
- · This speech signal is an example of one-dimensional continuous-time signal.
 - Can be represented mathematically as a function of single independent variable (t).

Two-dimensional stationary signal

- · This is a two dimensional signal (an image)
 - A spatial pattern not varying in time
 - Represented mathematically as a function of two spatial variables (x,y)
- · However, videos are timevarying images that involves three independent variables (x,y,t)



SINE and COSINE functions



SINES and COSINES

• Always use the COSINE FORM

$$A\cos(2\pi(440)t + \varphi)$$
• Sine is a special case:
$$\sin(\omega t) = \cos(\omega t - \frac{\pi}{2})$$

Sinusoidal Signal

$$A\cos(\omega t + \varphi)$$

- FREQUENCY
 - Radians/sec
 - Hertz (cycles/sec)
 - $\omega = (2\pi)$
- PERIOD (in sec)



 AMPLITUDE - Magnitude



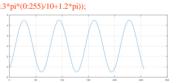
• PHASE



Example of Sinusoid

- $5\cos(0.3\pi t + 1.2\pi)$ Given the Formula
- A=5, $\omega=0.3\pi$, and $\varphi=1.2\pi$
- Make a plot
- Matlab (Octave) script:

plot(5*cos(0.3*pi*(0:255)/10+1.2*pi));



PLOTTING COSINE SIGNAL from the **FORMULA**

$$5\cos(0.3\pi t + 1.2\pi)$$

• Determine **period**:

$$T = 2\pi/\omega = 2\pi/0.3\pi = 20/3$$

• Determine a peak location by solving

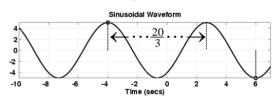
$$(\omega t + \varphi) = 0 \implies (0.3\pi t + 1.2\pi) = 0$$

- Zero crossing is T/4 before or after
- Positive & Negative peaks spaced by T/2

PLOT the SINUSOID

$$5\cos(0.3\pi t + 1.2\pi)$$

• Use T=20/3 and the peak location at t=-4



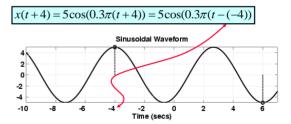
TIME-SHIFT

• In a mathematical formula we can replace *t* with *t-t*...

$$x(t-t_m) = A\cos(\omega(t-t_m))$$

- Then the t=0 point moves to $t=t_m$
- Peak value of $\cos(\omega(t-t_m))$ is now at $t=t_m$

TIME-SHIFTED SINUSOID



PHASE <--> TIME-SHIFT

· Equate the formulas:

$$A\cos(\omega(t-t_m)) = A\cos(\omega t + \varphi)$$

• and we obtain:

$$-\omega t_m = \varphi$$

• or, $t_m = -\frac{\varphi}{\varphi}$

TIME-SHIFT

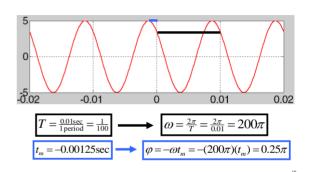
- Whenever a signal can be expressed in the form $x_1(t)=s(t-t_1)$, we say that $x_1(t)$ is time shifted version of s(t)
 - If t_1 is a + number, then the shift is to the right, and we say that the signal s(t) has been *delayed* in time.
 - If t₁ is a number, then the shift is to the left, and we say that the signal s(t) was advanced in time.

SINUSOID from a PLOT

3 steps

- Measure the period, T
 - Between peaks or zero crossings
 - Compute frequency: $\omega = 2\pi/T$
- Measure time of a peak: t_m
 - Compute phase: $\phi = -\omega t_{\rm m}$
- Measure height of positive peak: A

 (A, ω, ϕ) from a PLOT



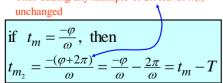
Copyright 2000 N. AYDIN. All rights reserved.

PHASE is AMBIGUOUS

- The cosine signal is periodic
 - Period is 2π

$$A\cos(\omega t + \varphi + 2\pi) = A\cos(\omega t + \varphi)$$

- Thus adding any multiple of 2π leaves x(t)

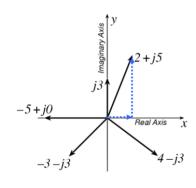


COMPLEX NUMBERS

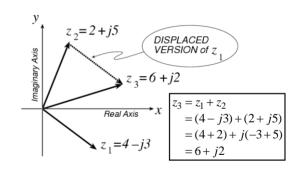
- To solve: $z^2 = -1$
 - -z=j
 - Math and Physics use z = i
- Complex number: z = x + iy



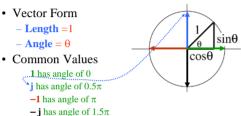
PLOT COMPLEX NUMBERS



COMPLEX ADDITION = VECTOR Addition

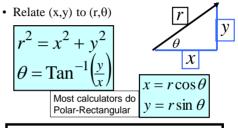


*** POLAR FORM ***



– **j** has angle of 1.5 π also, angle of –**j could** be –0.5 π = 1.5 π –2 π because the PHASE is **AMBIGUOUS**

POLAR <--> RECTANGULAR



Need a notation for POLAR FORM

Euler's FORMULA

• Complex Exponential

- Real part is cosine
- Imaginary part is sine
- Magnitude is one



$$e^{j\theta} = \cos(\theta) + j\sin(\theta)$$

$$re^{j\theta} = r\cos(\theta) + jr\sin(\theta)$$

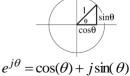
COMPLEX EXPONENTIAL

$$e^{j\omega t} = \cos(\omega t) + j\sin(\omega t)$$

• Interpret this as a Rotating Vector

 $\theta = \omega t$

Angle changes vs. time ex: $\omega = 20\pi \text{ rad/s}$ Rotates 0.2π in 0.01 secs



cos = REAL PART

Real Part of Euler's

$$\cos(\omega t) = \Re e\{e^{j\omega t}\}$$

General Sinusoid
$$x(t) = A\cos(\omega t + \varphi)$$

So,
$$A\cos(\omega t + \varphi) = \Re e\{Ae^{j(\omega t + \varphi)}\}\$$
$$= \Re e\{Ae^{j\varphi}e^{j\omega t}\}\$$

REAL PART EXAMPLE

$$A\cos(\omega t + \varphi) = \Re e \left\{ A e^{j\varphi} e^{j\omega t} \right\}$$

Evaluate:

$$x(t) = \Re e \left\{ -3je^{j\omega t} \right\}$$

Answer:

$$x(t) = \Re e \left\{ (-3j)e^{j\omega t} \right\}$$
$$= \Re e \left\{ 3e^{-j0.5\pi}e^{j\omega t} \right\} = 3\cos(\omega t - 0.5\pi)$$

COMPLEX AMPLITUDE

General Sinusoid

$$x(t) = A\cos(\omega t + \varphi) = \Re e \left\{ A e^{i\varphi} e^{j\omega t} \right\}$$

$$Complex AMPLITUDE = X$$

$$z(t) = X e^{j\omega t} \qquad X = A e^{i\varphi}$$

Then, any Sinusoid = REAL PART of Xejut

$$x(t) = \Re e \left\{ X e^{j\omega t} \right\} = \Re e \left\{ A e^{j\varphi} e^{j\omega t} \right\}$$

Basic properties of the sine and cosine functions

Property	Equation	
Equivalence	$\sin \theta = \cos(\theta - \pi/2) \text{ or } \cos(\theta) = \sin(\theta + \pi/2)$	
Periodicity	$cos(\theta + 2\pi k) = cos \theta$, when k is an integer	
Evenness of cosine	$cos(-\theta) = cos \theta$	
Oddness of sine	$\sin(-\theta) = -\sin\theta$	
Zeros of sine	$sin(\pi k) = 0$, when k is an integer	
Ones of cosine	$cos(2\pi k) = 1$ when k is an integer.	
Minus ones of cosine	$\cos[2\pi(k+\frac{1}{2})] = -1$, when k is an integer.	

Some basic trigonometric identities

Number	Equation
1	$\sin^2\theta + \cos^2\theta = 1$
2	$\cos 2\theta = \cos^2 \theta - \sin^2 \theta$
3	$\sin 2\theta = 2\sin\theta\cos\theta$
4	$\sin(\alpha \pm \beta) = \sin\alpha \cos\beta \pm \cos\alpha \sin\beta$
5	$cos(\alpha + \beta) = cos \alpha cos \beta \mp sin \alpha sin \beta$

Sampling and plotting sinusoids

• Plot the following function

$$20\cos(2\pi(40)t - 0.4\pi)$$

• Must evaluate x(t) at a discrete set of times, $t_n = nT_s$, where n is an integer

$$x(nT_s) = 20\cos(80\pi nT_s - 0.4\pi)$$

• T_s is called sample spacing or sampling period

SPECTRUM Representation

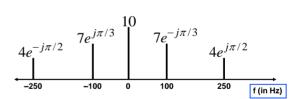
- Sinusoids with DIFFERENT Frequencies
 - SYNTHESIZE by Adding Sinusoids

$$x(t) = \sum_{k=1}^{N} A_k \cos(2\pi f_k t + \varphi_k)$$

- SPECTRUM Representation
 - Graphical Form shows **DIFFERENT** Freqs

FREQUENCY DIAGRAM

• Plot Complex Amplitude vs. Freq



Another FREQ. Diagram



MOTIVATION

- Synthesize Complicated Signals
 - Musical Notes

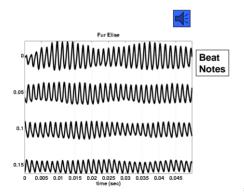


- · Piano uses 3 strings for many notes
- Chords: play several notes simultaneously
- Human Speech



- · Vowels have dominant frequencies
- · Application: computer generated speech
- Can all signals be generated this way?
 - · Sum of sinusoids?

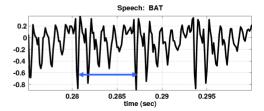
Fur Elise WAVEFORM



Speech Signal: BAT



- Nearly **Periodic** in Vowel Region
 - Period is (Approximately) T = 0.0065 sec



Euler's Formula Reversed

• Solve for cosine (or sine) $e^{j\omega t} = \cos(\omega t) + j\sin(\omega t)$ $e^{-j\omega t} = \cos(-\omega t) + j\sin(-\omega t)$ $e^{-j\omega t} = \cos(\omega t) - j\sin(\omega t)$ $e^{j\omega t} + e^{-j\omega t} = 2\cos(\omega t)$

$$\cos(\omega t) = \frac{1}{2} (e^{j\omega t} + e^{-j\omega t})$$

INVERSE Euler's Formula

• Solve for cosine (or sine)

$$\cos(\omega t) = \frac{1}{2} (e^{j\omega t} + e^{-j\omega t})$$

$$\sin(\omega t) = \frac{1}{2j} (e^{j\omega t} - e^{-j\omega t})$$

SPECTRUM Interpretation

• Cosine = sum of 2 complex exponentials:

$$A\cos(7t) = \frac{A}{2}e^{j7t} + \frac{A}{2}e^{-j7t}$$
One has a positive frequency

One has a positive frequency The other has negative freq. Amplitude of each is half as big

NEGATIVE FREQUENCY

- Is negative frequency real?
- Doppler Radar provides an example
 - Police radar measures speed by using the Doppler shift principle
 - Let's assume 400Hz ←→60 mph
 - +400Hz means towards the radar
 - -400Hz means away (opposite direction)
 - Think of a train whistle

Copyright 2000 N. AYDIN. All rights reserved.

SPECTRUM of SINE

• Sine = sum of 2 complex exponentials:

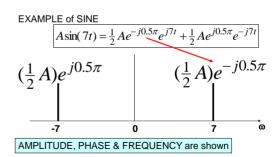
$$A\sin(7t) = \frac{A}{2j}e^{j7t} - \frac{A}{2j}e^{-j7t}$$

$$= \frac{1}{2}Ae^{-j0.5\pi}e^{j7t} + \frac{1}{2}Ae^{j0.5\pi}e^{-j7t}$$

$$\frac{-1}{j} = j = e^{j0.5\pi}$$

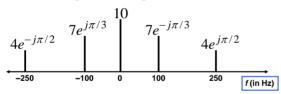
- Positive freq. has phase = -0.5π
- Negative freq. has phase = $+0.5\pi$

GRAPHICAL SPECTRUM



SPECTRUM ---> SINUSOID

• Add the spectrum components:



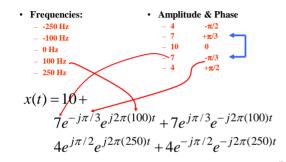
What is the formula for the signal x(t)?

Gather (A, ω, ϕ) information

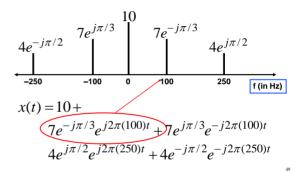
Note the conjugate phase

DC is another name for zero-freq component **DC** component always has f=0 or π (for real x(t))

Add Spectrum Components-1



Add Spectrum Components-2



Simplify Components

$$x(t) = 10 +$$

$$7e^{-j\pi/3}e^{j2\pi(100)t} + 7e^{j\pi/3}e^{-j2\pi(100)t}$$

$$4e^{j\pi/2}e^{j2\pi(250)t} + 4e^{-j\pi/2}e^{-j2\pi(250)t}$$

Use Euler's Formula to get REAL sinusoids:

$$A\cos(\omega t + \varphi) = \frac{1}{2}Ae^{-j\varphi}e^{j\omega t} + \frac{1}{2}Ae^{-j\varphi}e^{-j\omega t}$$

FINAL ANSWER

$$x(t) = 10 + 14\cos(2\pi(100)t - \pi/3) + 8\cos(2\pi(250)t + \pi/2)$$

So, we get the general form:

$$x(t) = A_0 + \sum_{k=1}^{N} A_k \cos(2\pi f_k t + \varphi_k)$$

Summary: GENERAL FORM

$$x(t) = A_0 + \sum_{k=1}^{N} A_k \cos(2\pi f_k t + \varphi_k)$$

$$x(t) = X_0 + \sum_{k=1}^{N} \Re e \left\{ X_k e^{j2\pi f_k t} \right\}$$

$$\Re e\{z\} = \frac{1}{2} z + \frac{1}{2} z^*$$

$$x(t) = X_0 + \sum_{k=1}^{N} \left\{ \frac{1}{2} X_k e^{j2\pi f_k t} + \frac{1}{2} X_k^* e^{-j2\pi f_k t} \right\}$$

Example: Synthetic Vowel

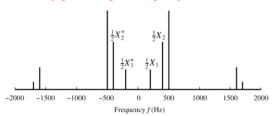
• Sum of 5 Frequency Components

f_k (Hz)	X_k	Mag	Phase (rad)
200	(771 + j12202)	12,226	1.508
400	(-8865 + j28048)	29,416	1.876
500	(48001 - j8995)	48,836	-0.185
1600	(1657 - j13520)	13,621	-1.449
1700	4723 + j0	4723	0

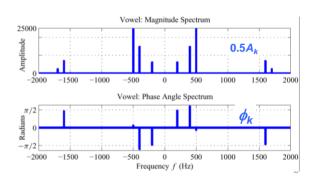
Table 3.1: Complex amplitudes for harmonic signal that approximates the vowel sound "ah".

SPECTRUM of VOWEL

- Note: Spectrum has $0.5X_{\nu}$ (except X_{DC})
- Conjugates in negative frequency



SPECTRUM of VOWEL (Polar Format)



Vowel Waveform (sum of all 5 components)

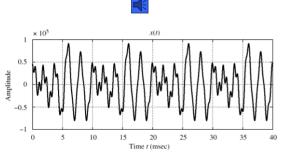


Figure 3.11 Sum of all of the terms in (3,3,4). Note that the period is 10 msec.

Periodic Signals, Harmonics & **Time-Varying Sinusoids**

Problem Solving Skills

- Math Formula
 - Sum of Cosines
 - Amp, Freq, Phase
- · Recorded Signals
 - Speech
 - Music
 - No simple formula

- Plot & Sketches
 - -S(t) versus t

 - Spectrum
- **MATLAB**
- Numerical
- Computation
- Plotting list of numbers

- Signals with **HARMONIC** Frequencies
 - Add Sinusoids with $f_{\nu} = kf_0$

$$x(t) = A_0 + \sum_{k=1}^{N} A_k \cos(2\pi k f_0 t + \varphi_k)$$

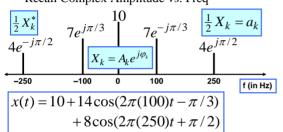
FREQUENCY can change vs. TIME

Chirps: $x(t) = \cos(\alpha t^2)$

Introduce Spectrogram Visualization (specgram.m)

SPECTRUM DIAGRAM

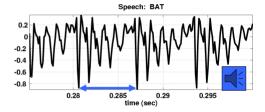
• Recall Complex Amplitude vs. Freq



SPECTRUM for PERIODIC?

• Nearly Periodic in the Vowel Region

- Period is (Approximately) T = 0.0065 sec



Copyright 2000 N. AYDIN. All rights reserved.

PERIODIC SIGNALS

• Repeat every T secs

$$x(t) = x(t+T)$$

- Example:

$$x(t) = \cos^2(3t)$$

$$T = ?$$

$$T = \frac{2\pi}{3} \quad T = \frac{\pi}{3}$$

Period of Complex Exponential

- Definition
$$x(t) = x(t+T)$$
- Example:
$$x(t) = \cos^2(3t)$$

$$T = ?$$
- Speech can be "quasi-periodic"
$$T = \frac{2\pi}{3}$$

$$T = \frac{\pi}{3}$$

$$\omega = \frac{2\pi k}{T} = \left(\frac{2\pi}{T}\right)k = \omega_0 k$$

$$k = \text{integer}$$

Harmonic Signal Spectrum

Periodic signal can only have : $f_k = kf_0$

$$x(t) = A_0 + \sum_{k=1}^{N} A_k \cos(2\pi k f_0 t + \varphi_k)$$

$$X_k = A_k e^{j\varphi_k}$$

$$x(t) = X_0 + \sum_{k=1}^{N} \left\{ \frac{1}{2} X_k e^{j2\pi k f_0 t} + \frac{1}{2} X_k^* e^{-j2\pi k f_0 t} \right\}$$

Define FUNDAMENTAL FREQ

$$x(t) = A_0 + \sum_{k=1}^{N} A_k \cos(2\pi k f_0 t + \varphi_k)$$

$$f_k = kf_0 \qquad (\omega_0 = 2\pi f_0)$$

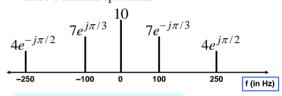
$$f_0 = \frac{1}{T_0}$$

 $f_k = k f_0$ $(\omega_0 = 2\pi f_0)$ $f_0 = \frac{1}{T_0}$ $f_0 = f_0$ (largest) T_0 = fundamenta 1 Period (shortest)

Harmonic Signal (3 Freqs)

Example

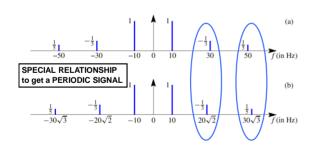
· Here's another spectrum:



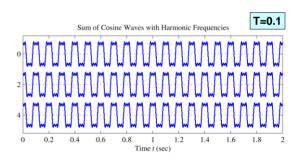
What is the fundamental frequency?

100 Hz ? 50 Hz ?

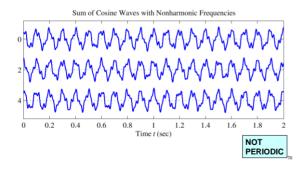
IRRATIONAL SPECTRUM



Harmonic Signal (3 Freqs)



NON-Harmonic Signal



FREQUENCY ANALYSIS

- Now, a much HARDER problem
- Given a recording of a song, have the computer write the music



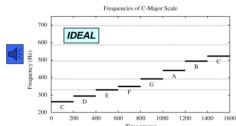
- Can a machine extract frequencies?
 - Yes, if we COMPUTE the spectrum for x(t)
 - · During short intervals

Time-Varying FREQUENCIES Diagram



SIMPLE TEST SIGNAL

- C-major SCALE: stepped frequencies
 - Frequency is constant for each note

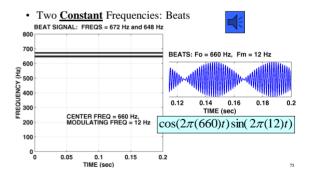


Copyright 2000 N. AYDIN. All rights reserved.

SPECTROGRAM

- SPECTROGRAM Tool
 - MATLAB function is specgram.m
- ANALYSIS program
 - Takes x(t) as input &
 - Produces spectrum values X₁.
 - Breaks x(t) into SHORT TIME SEGMENTS
 - Then uses the FFT (Fast Fourier Transform)

SPECTROGRAM EXAMPLE

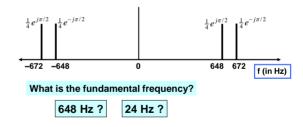


AM Radio Signal

• Same as BEAT Notes $\frac{\cos(2\pi(660)t)\sin(2\pi(12)t)}{\left[\frac{1}{2}\left(e^{j2\pi(660)t} + e^{-j2\pi(660)t}\right)\frac{1}{2j}\left(e^{j2\pi(12)t} - e^{-j2\pi(12)t}\right)\right]}$ $\frac{1}{4j}\left(e^{j2\pi(672)t} - e^{-j2\pi(672)t} - e^{j2\pi(648)t} + e^{-j2\pi(648)t}\right)$ $\frac{1}{2}\cos(2\pi(672)t - \frac{\pi}{2}) + \frac{1}{2}\cos(2\pi(648)t + \frac{\pi}{2})$

SPECTRUM of AM (Beat)

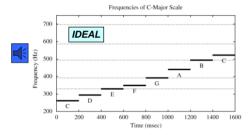
• 4 complex exponentials in AM:



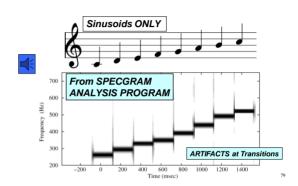
STEPPED FREQUENCIES

• C-major SCALE: successive sinusoids

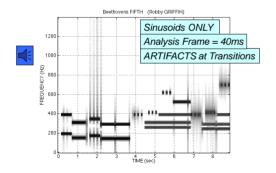
- Frequency is constant for each note



SPECTROGRAM of C-Scale



Spectrogram of LAB SONG



Time-Varying Frequency

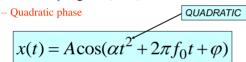
- Frequency can change vs. time
 - Continuously, not stepped
- FREQUENCY MODULATION (FM)

$$x(t) = \cos(2\pi f_c t + v(t))$$

- CHIRP SIGNALS
 - Linear Frequency Modulation (LFM)

New Signal: Linear FM

• Called Chirp Signals (LFM)



- Freq will change LINEARLY vs. time
 - Example of Frequency Modulation (FM)
 - Define "instantaneous frequency"

INSTANTANEOUS FREQ

Definition

$$x(t) = A\cos(\psi(t))$$

$$\Rightarrow \omega_i(t) = \frac{d}{dt}\psi(t)$$

Derivative of the "Angle"

• For Sinusoid:

$$x(t) = A\cos(2\pi f_0 t + \varphi)$$

$$\psi(t) = 2\pi f_0 t + \varphi$$

Makes sense

$$\Rightarrow \omega_i(t) = \frac{d}{dt}\psi(t) = 2\pi f_0$$

INSTANTANEOUS FREQ of the Chirp

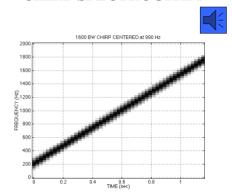
- Chirp Signals have Quadratic phase
- Freq will change LINEARLY vs. time

$$x(t) = A\cos(\alpha t^{2} + \beta t + \varphi)$$

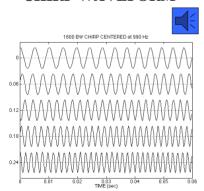
$$\Rightarrow \psi(t) = \alpha t^{2} + \beta t + \varphi$$

$$\Rightarrow \omega_{i}(t) = \frac{d}{dt}\psi(t) = 2\alpha t + \beta$$

CHIRP SPECTROGRAM



CHIRP WAVEFORM



OTHER CHIRPS

 $\psi(t)$ can be anything:

$$x(t) = A\cos(\alpha\cos(\beta t) + \varphi)$$

$$\Rightarrow \omega_i(t) = \frac{d}{dt}\psi(t) = -\alpha\beta\sin(\beta t)$$

 $\psi(t)$ could be speech or music:

- FM radio broadcast

SINE-WAVE FREQUENCY MODULATION (FM)



