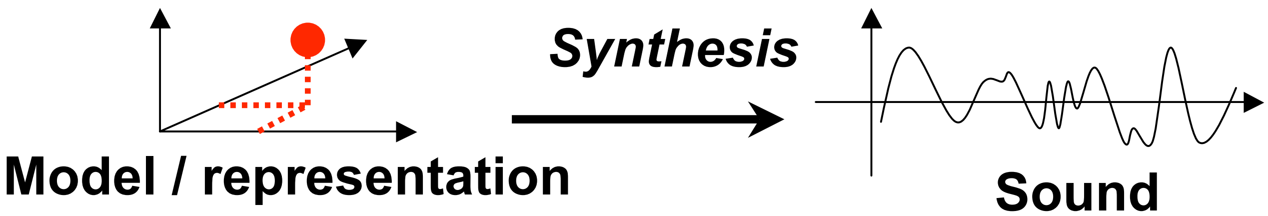
Synthesis Techniques

Juan P Bello

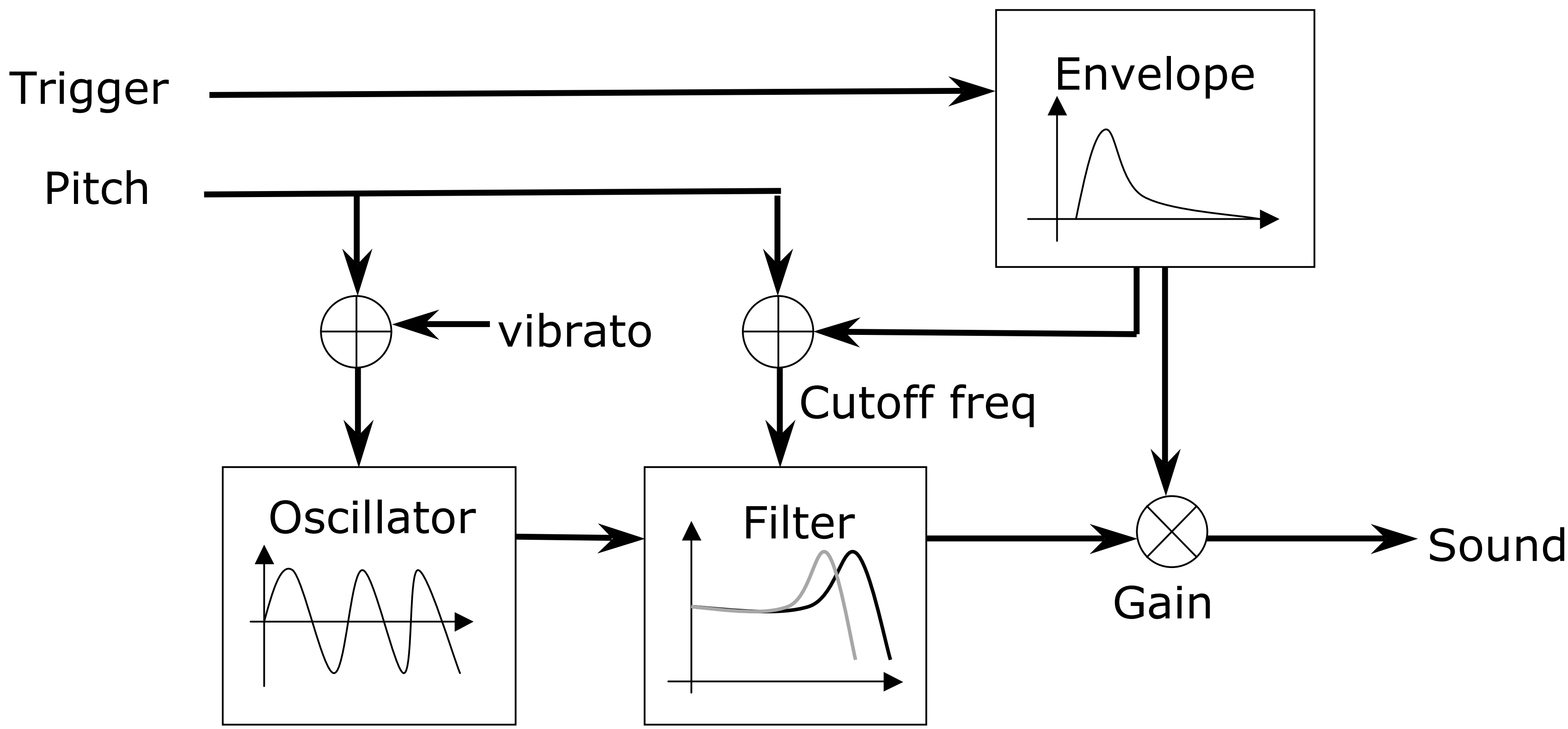
Synthesis

- It implies the artificial construction of a complex body by combining its elements.
  - ➢ Complex body: acoustic signal (sound)
  - ➢ Elements: parameters and/or “basic signals”



- Motivations:
  - ➢ Reproduce existing sounds
  - ➢ Reproduce the physical process of sound generation
  - ➢ Generate new pleasant sounds
  - ➢ Control/explore timbre

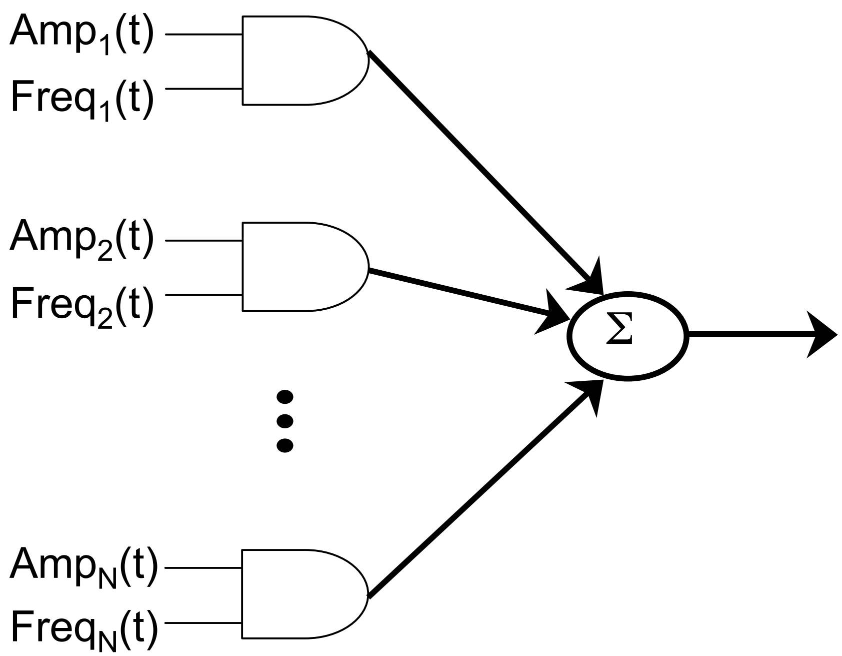
How can I generate new sounds?



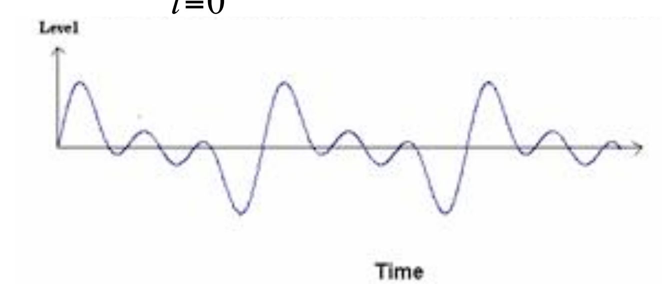
- Networks of basic elements → synthesis techniques
- Two main types: linear and non-linear

Additive Synthesis

- It is based on the idea that complex waveforms can be created by the addition of simpler ones.
- It is a linear technique, i.e. do not create frequency components that were not explicitly contained in the original waveforms
- Commonly, these simpler signals are sinusoids (sines or cosines) with time-varying parameters, according to Fourier's theory:

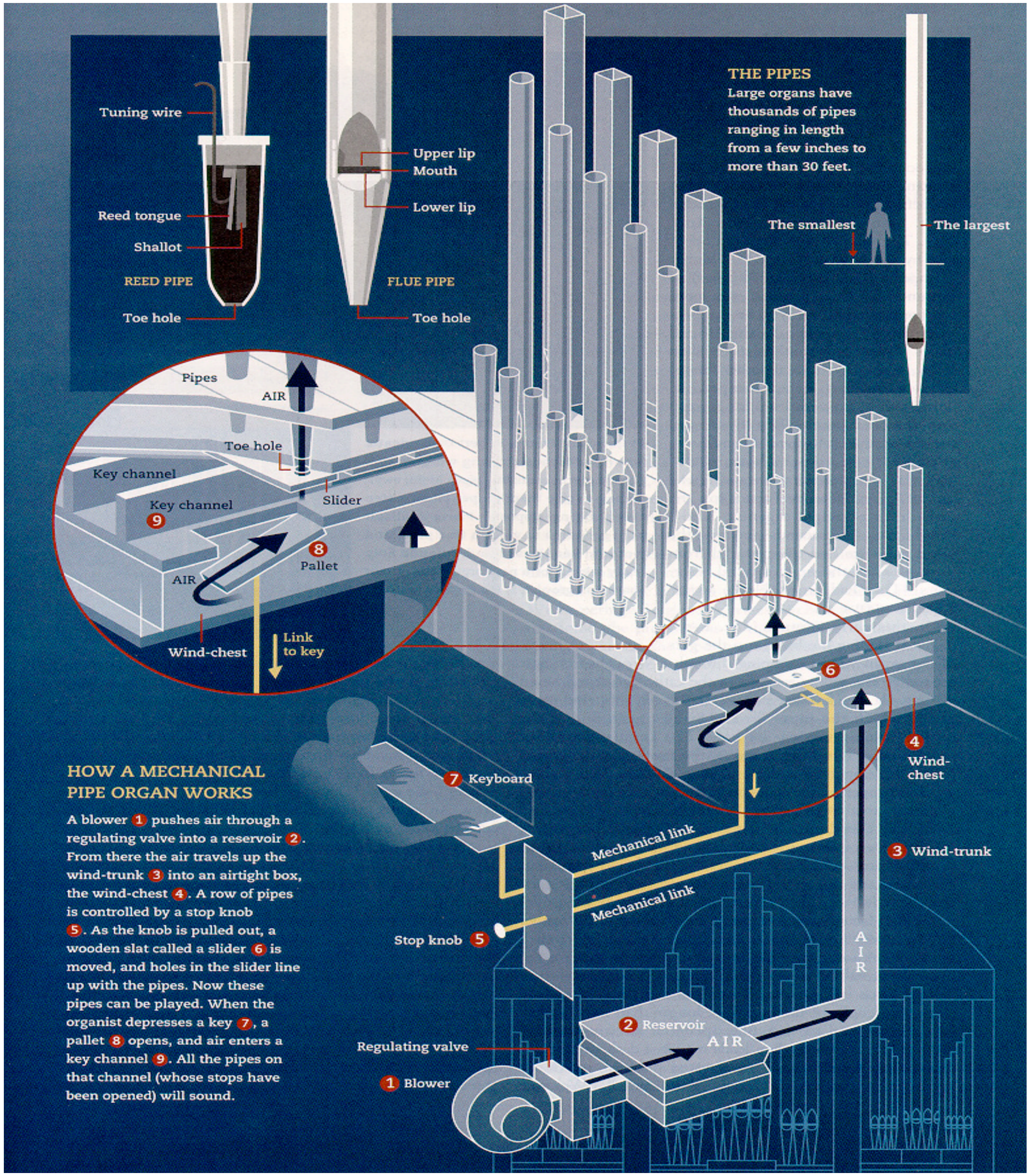


$$s(t) = \sum_{i=0} A_i \sin\left(2\pi f_i t + \varphi_i\right)$$



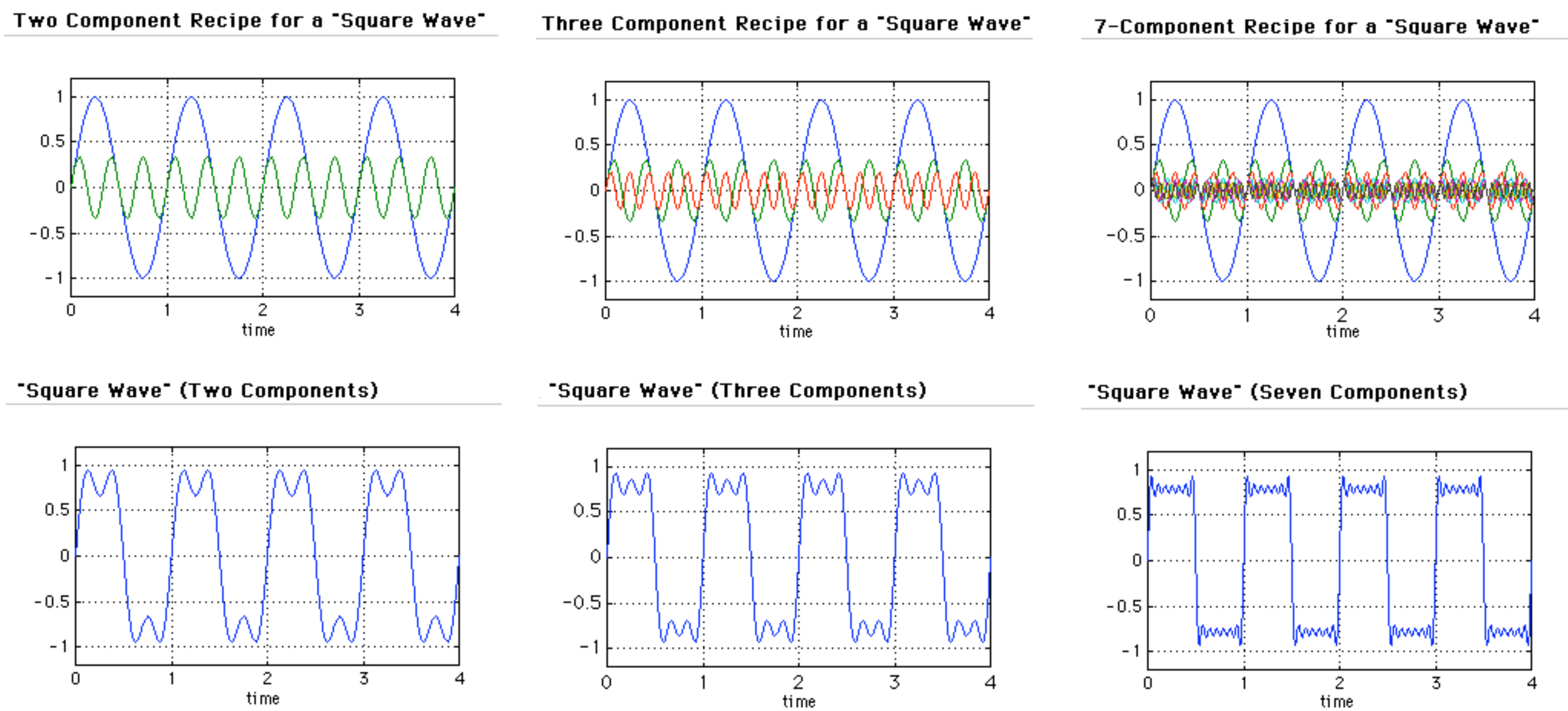
Additive Synthesis:

A Pipe Organ



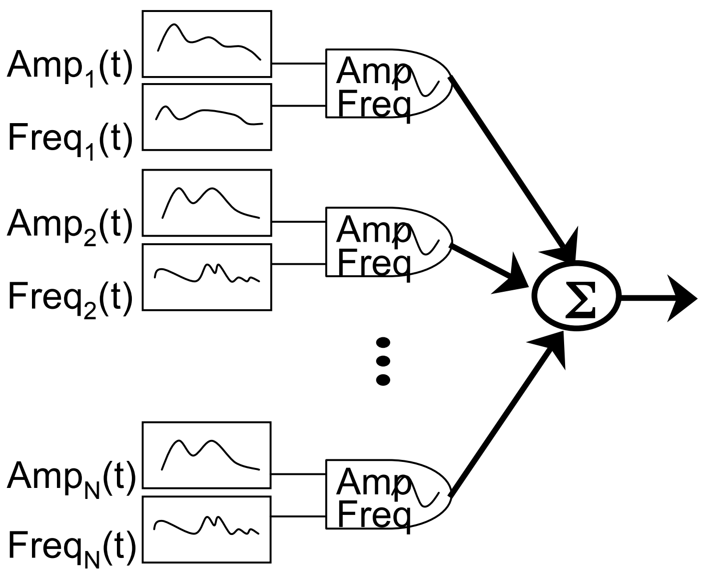
Additive Synthesis

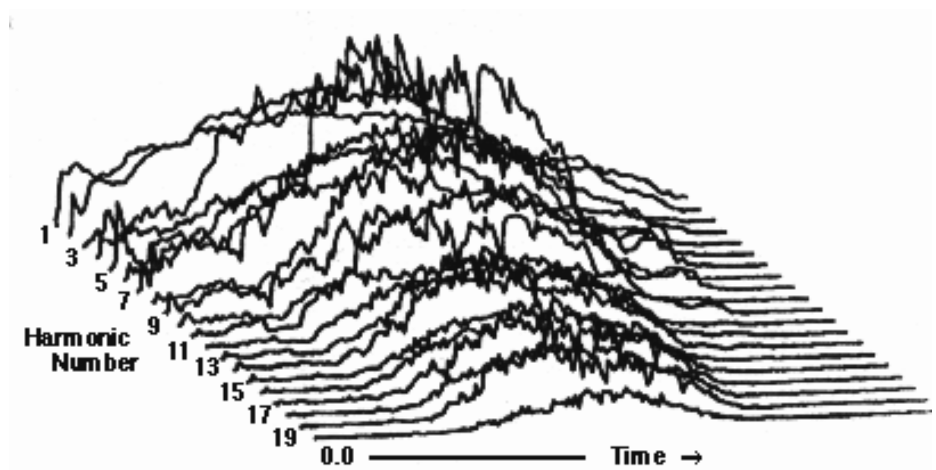
- Square wave: only odd harmonics. Amplitude of the n^{th} harmonic = **1/n**



Time-varying sounds

- According to Fourier, all sounds can be described and reproduced with additive synthesis.
- Even impulse-like components can be represented by using a short-lived sinusoid with “infinite” amplitude.





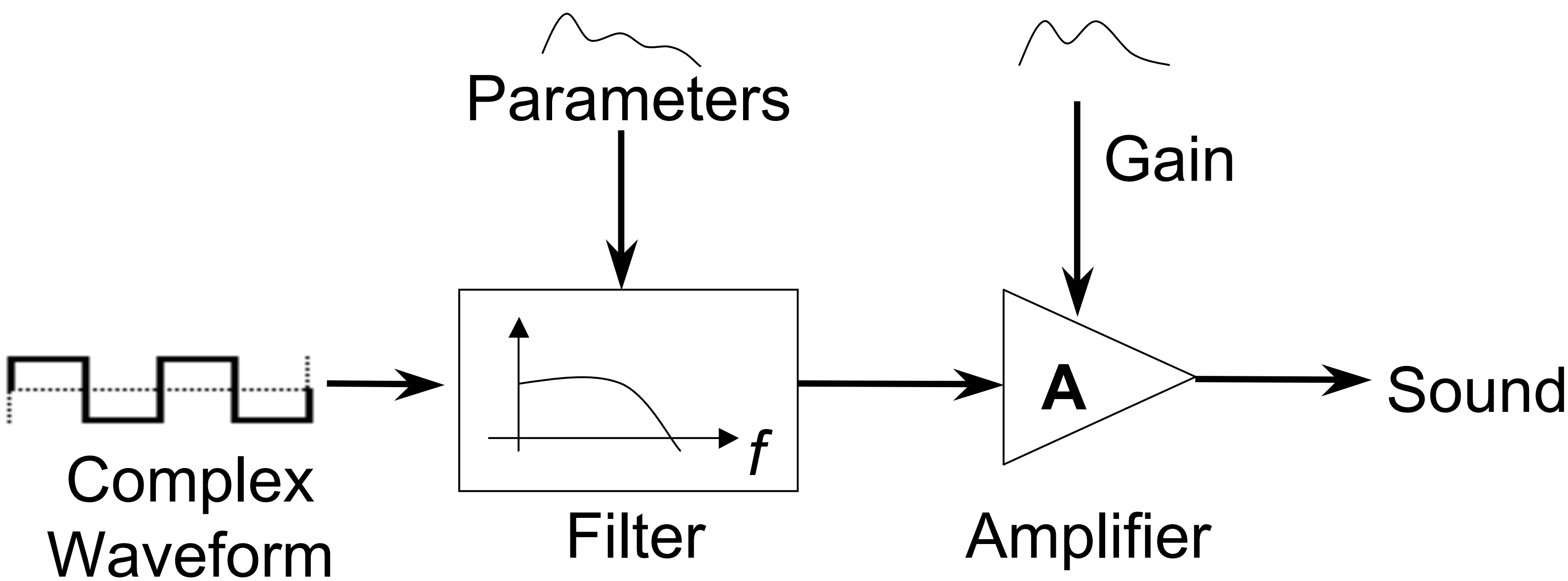
- Additive synthesis is very general (perhaps the most versatile).
- Control data hungry: large number of parameters are required to reproduce realistic sounds

Examples

	orig	det	stoch	SMS	trans
flute					
guitar					
water					

Subtractive Synthesis

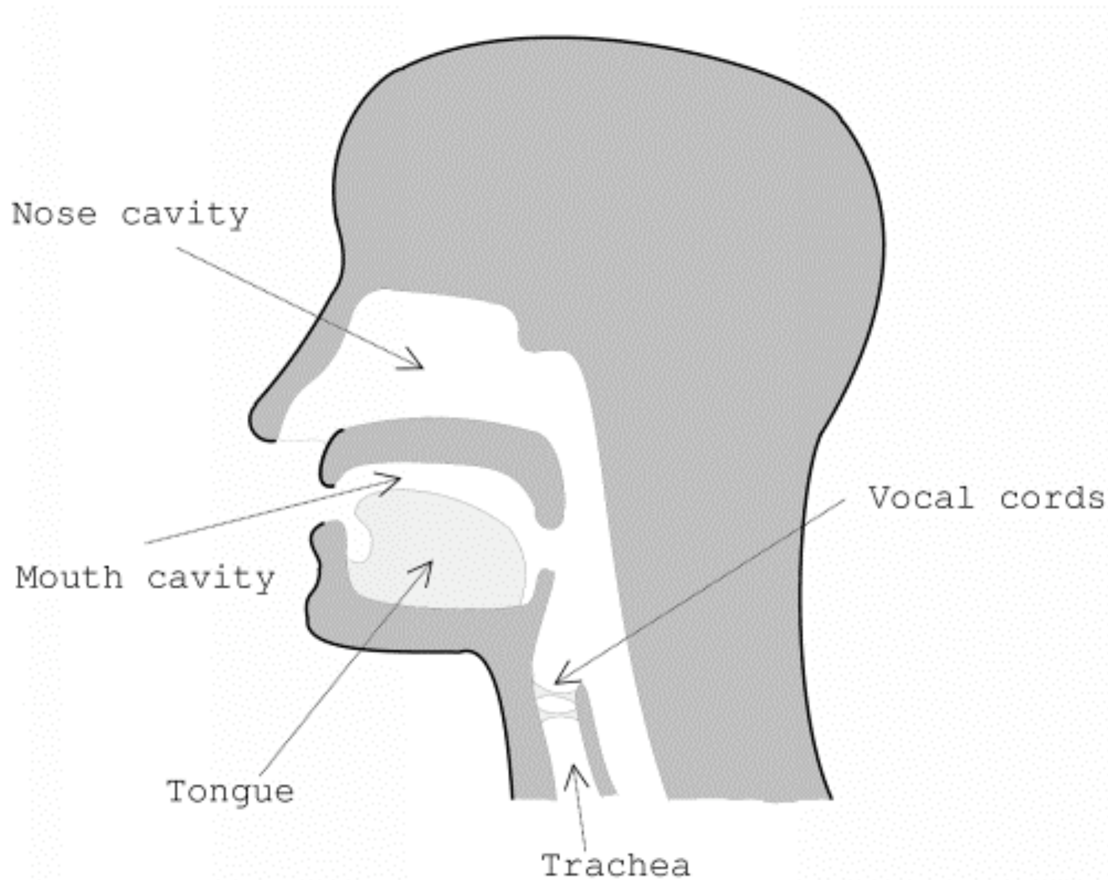
- Is another linear technique based on the idea that sounds can be generated from subtracting (filtering out) components from a very rich signal (e.g. noise, square wave).



- Its simplicity made it very popular for the design of analog synthesisers (e.g. Moog)

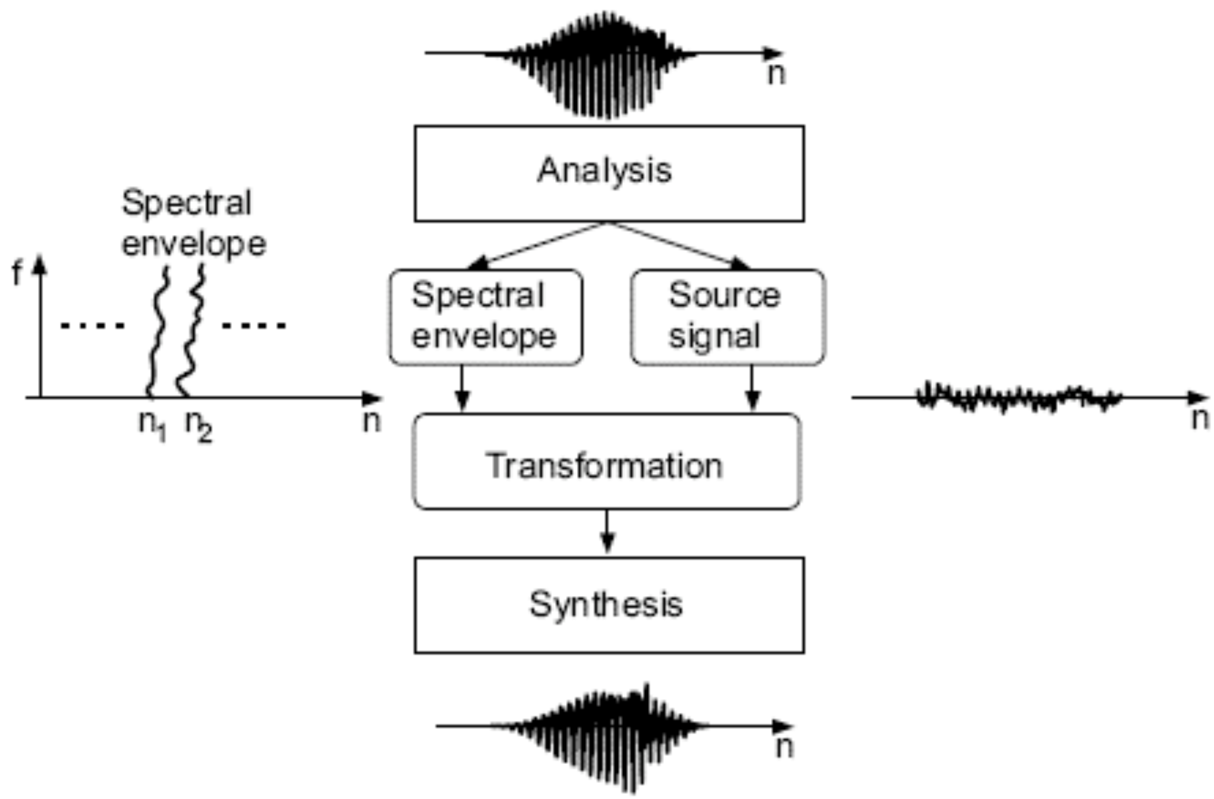
The human speech system

- The vocal chords act as an oscillator, the mouth/nose cavities, tongue and throat as filters
- We can shape a tonal sound (‘oooh’ vs ‘aaah’), we can whiten the signal (‘sssshhh’), we can produce pink noise by removing high frequencies

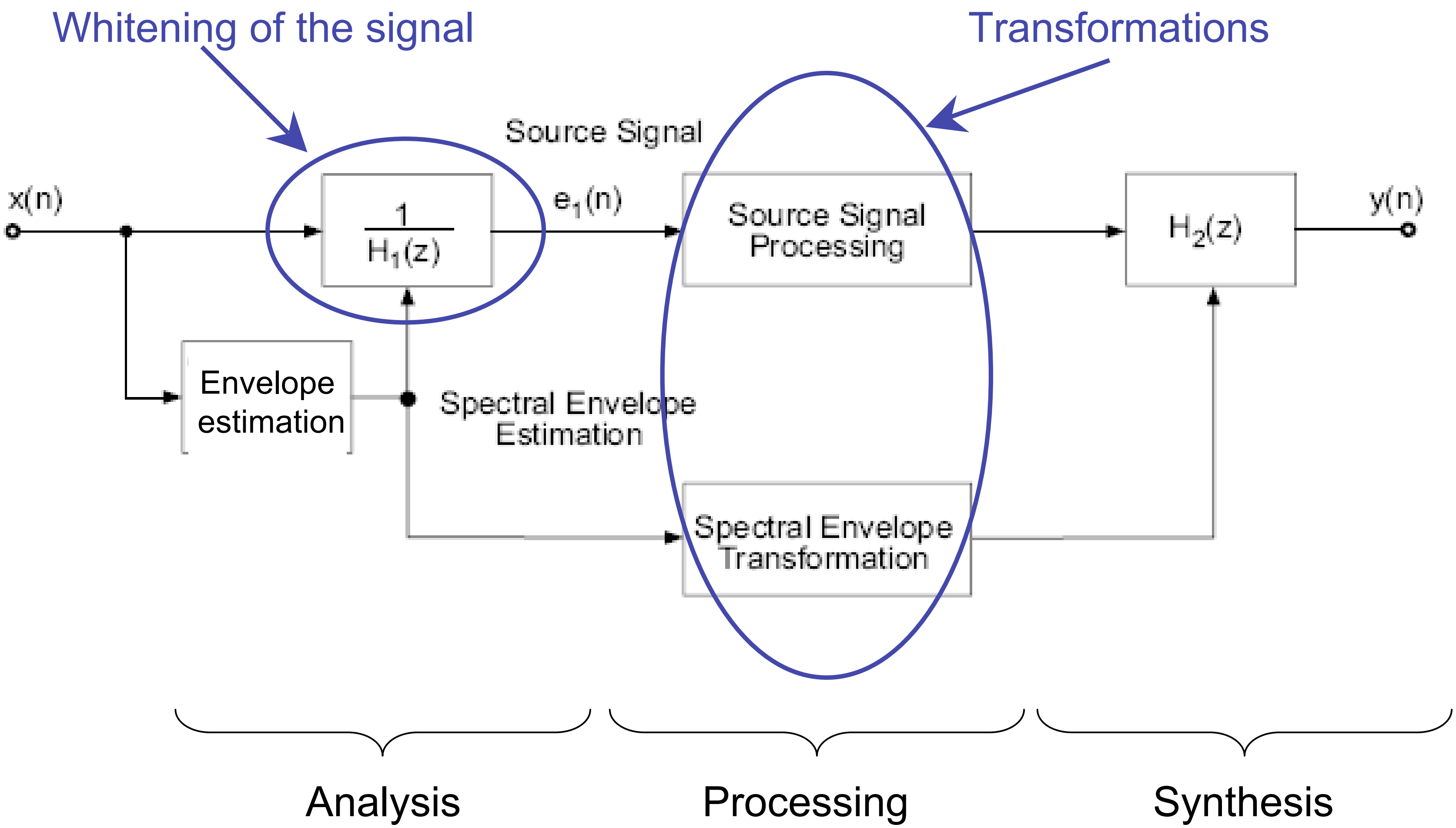


Source-Filter model

- Subtractive synthesis can be seen as a excitation-resonator or source-filter model
- The resonator or filter shapes the spectrum, i.e. defines the spectral envelope

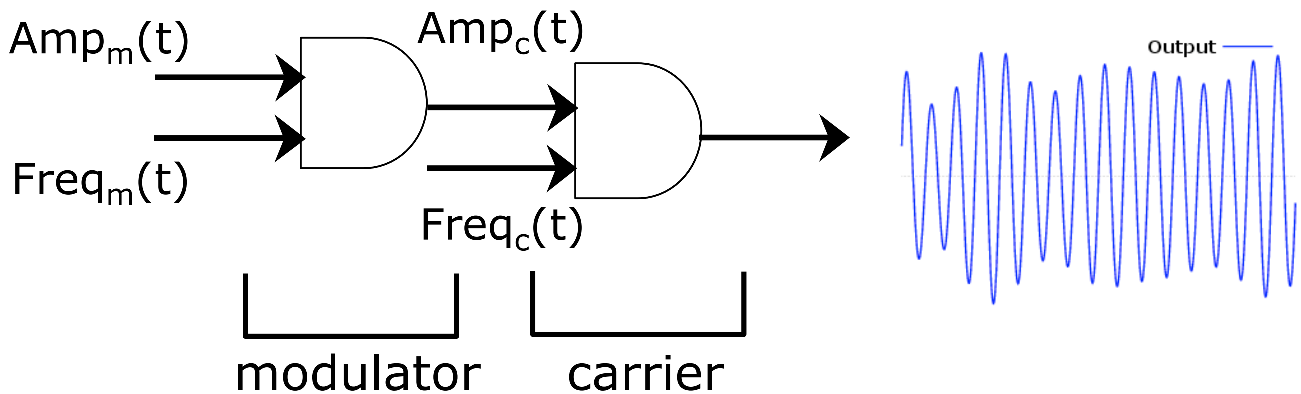


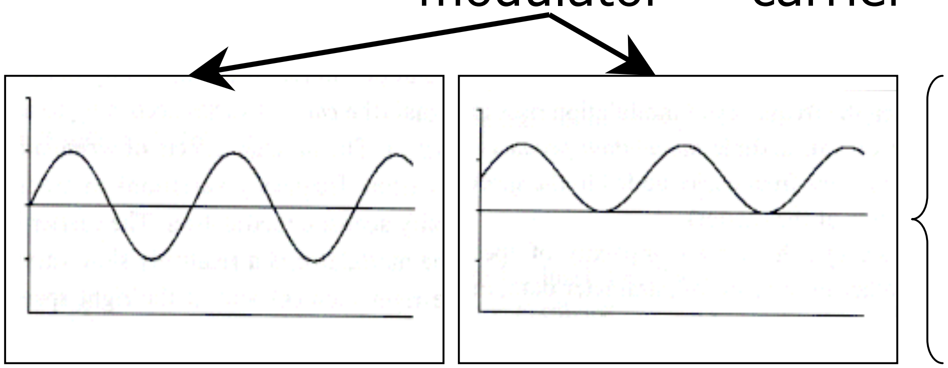
Source-Filter model



Amplitude modulation

- Non-linear technique, i.e. results on the creation of frequencies which are not produced by the oscillators.
- In AM the amplitude of the carrier wave is varied in direct proportion to that of a modulating signal.





bipolar unipolar

Bipolar -> Ring modulation
Unipolar -> Amplitude modulation

Ring Modulation

- Let us define the carrier signal as:

$$c(t) = A_c \cos(\omega_c t)$$

- And the (bipolar) modulator signal as:

$$m(t) = A_m \cos(\omega_m t)$$

- The Ring modulated signal can be expressed as:

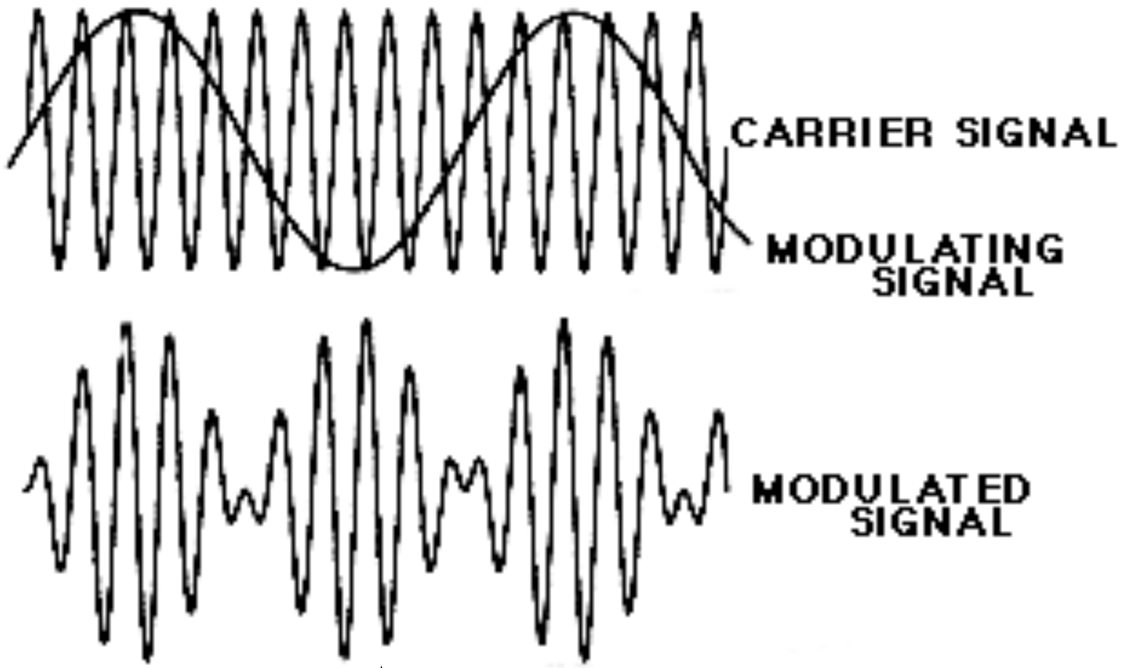
$$s(t) = A_c \cos\left(\omega_c t\right) \cdot A_m \cos\left(\omega_m t\right)$$

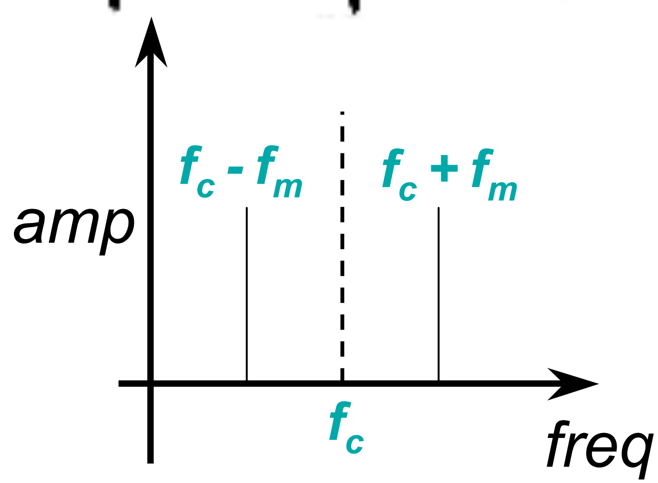
- Which can be re-written as:

$$s(t) = \frac{A_c A_m}{2}\left[\cos\left(\left[\omega_c - \omega_m\right]t\right) + \cos\left(\left[\omega_c + \omega_m\right]t\right)\right]$$

- $s(t)$ presents two sidebands at frequencies: $\omega_c - \omega_m$ and $\omega_c + \omega_m$

Ring Modulation





Amplitude Modulation

- Let us define the carrier signal as:

$$c(t) = \cos(\omega_c t)$$

- And the (unipolar) modulator signal as:

$$m(t) = A_c + A_m \cos(\omega_m t)$$

- The amplitude modulated signal can be expressed as:

$$s(t) = \left[A_c + A_m \cos(\omega_m t)\right]\cos(\omega_c t)$$

- Which can be re-written as:

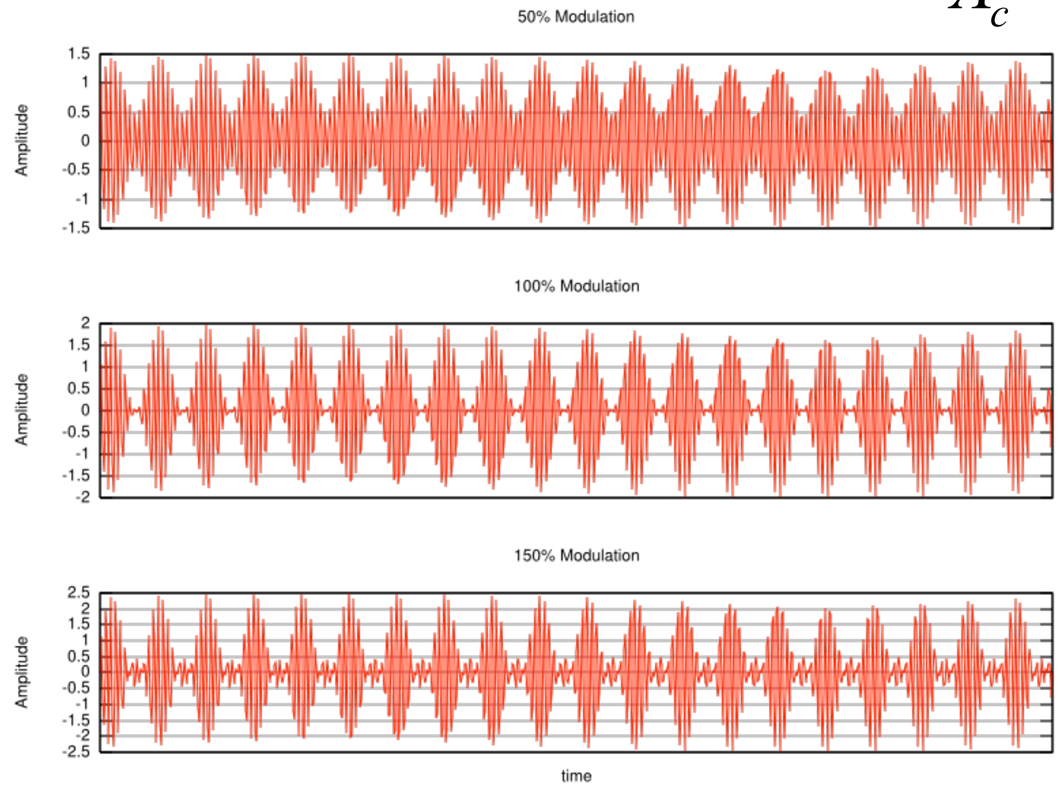
$$s(t) = A_c \cos(\omega_c t) + \frac{A_m}{2}\left[\cos([\omega_c - \omega_m]t) + \cos([\omega_c + \omega_m]t)\right]$$

- $s(t)$ presents components at frequencies: ω_c , ω_c - ω_m and ω_c + ω_m

Modulation index

- In modulation techniques a modulation index is usually defined such that it indicates how much the modulated variable varies around its original value.
- For AM this quantity is also known as modulation depth: $\beta = \frac{A_m}{A_c}$

- If β = 0.5 then the carrier's amplitude varies by 50% around its unmodulated level.
- For β = 1 it varies by 100%.
- β > 1 causes distortion and is usually avoided



C/M frequency ratio

- Lets define the carrier to modulator frequency ratio c/m (= ω_c/ ω_m) for a pitched signal m(t)

- If c/m is an integer n, then ω_c, and all present frequencies, are multiples of ω_m (which will become the fundamental)

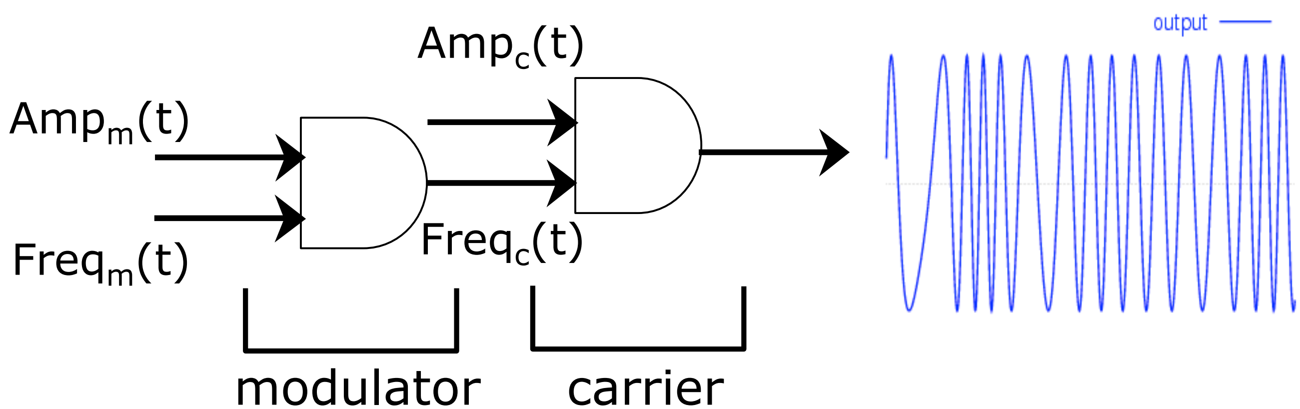
- If c/m = 1/n, then ω_c will be the fundamental

- When c/m deviates from n or 1/n (or more generally, from a ratio of integers), then the output frequencies becomes more inharmonic

- Example of C/M frequency variation

Frequency Modulation

- Frequency modulation (FM) is a form of modulation in which the frequency of a carrier wave is varied in direct proportion to the amplitude variation of a modulating signal.



- When the frequency modulation produces a variation of less than 20Hz this results on a vibrato.

Frequency Modulation

- Let us define the carrier signal as:

$$c(t) = \cos(\omega_c t)$$

- And the modulator signal as:

$$m(t) = \beta \sin(\omega_m t)$$

- The Frequency modulated signal can be expressed as:

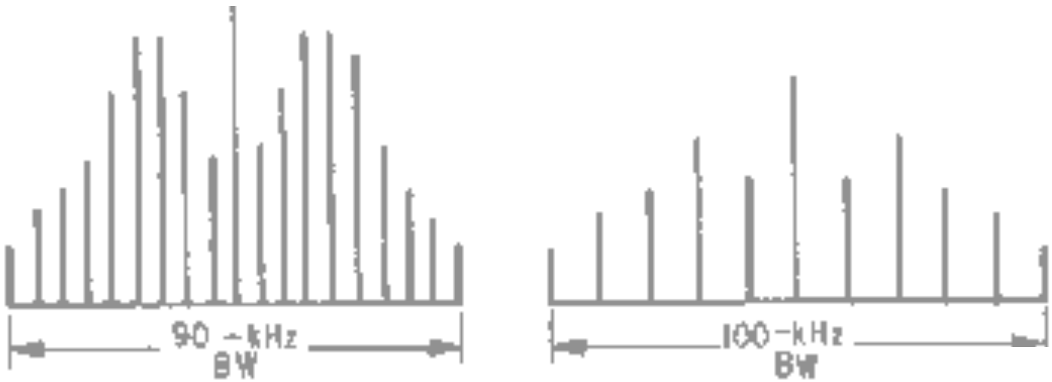
$$s(t) = \cos\left(\omega_c t + \beta \sin\left(\omega_m t\right)\right)$$

- This can be re-written as

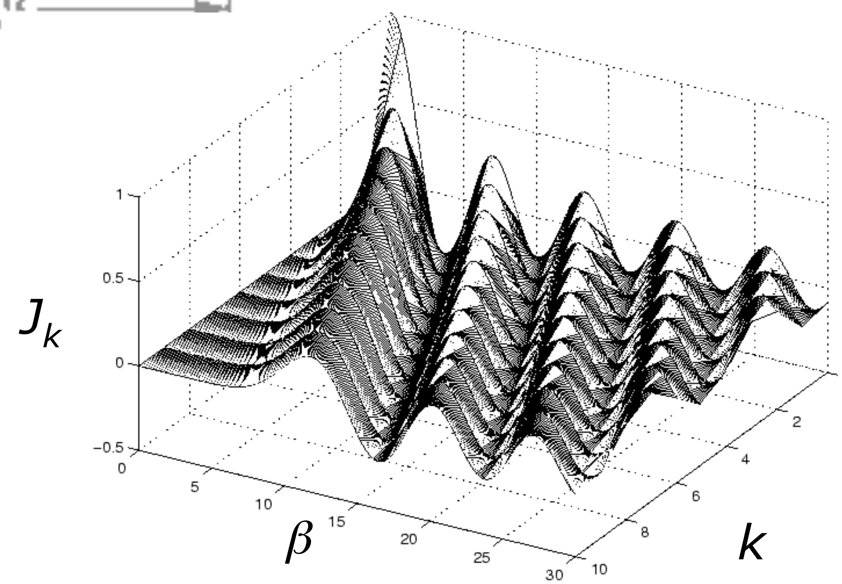
$$s(t) = \sum_{k=-\infty}^{\infty} J_k(\beta) \cos\left[\left(\omega_c + k\omega_m\right)t\right]$$

Frequency Modulation

- If $\beta \neq 0$ then the FM spectrum contains infinite sidebands at positions $\omega_c \pm k\omega_m$.



- The amplitudes of each pair of sidebands are given by the J_k coefficients which are functions of β



Modulation index

- As in AM we define a FM modulation index that controls the modulation depth.
- In FM synthesis this index is equal to β, the amplitude of the modulator and is directly proportional to Δf.
- As we have seen the value of β determines the amplitude of the sidebands of the FM spectrum
- Furthermore the amplitude decreases with the order k.
- Thus, although theoretically the number of sidebands is infinite, in practice their amplitude makes them inaudible for higher orders.
- The number of audible sidebands is a function of β, and is approximated by $2\beta+1$
- Thus the bandwidth increases with the amplitude of $m(t)$, like in some real instruments

C/M frequency ratio

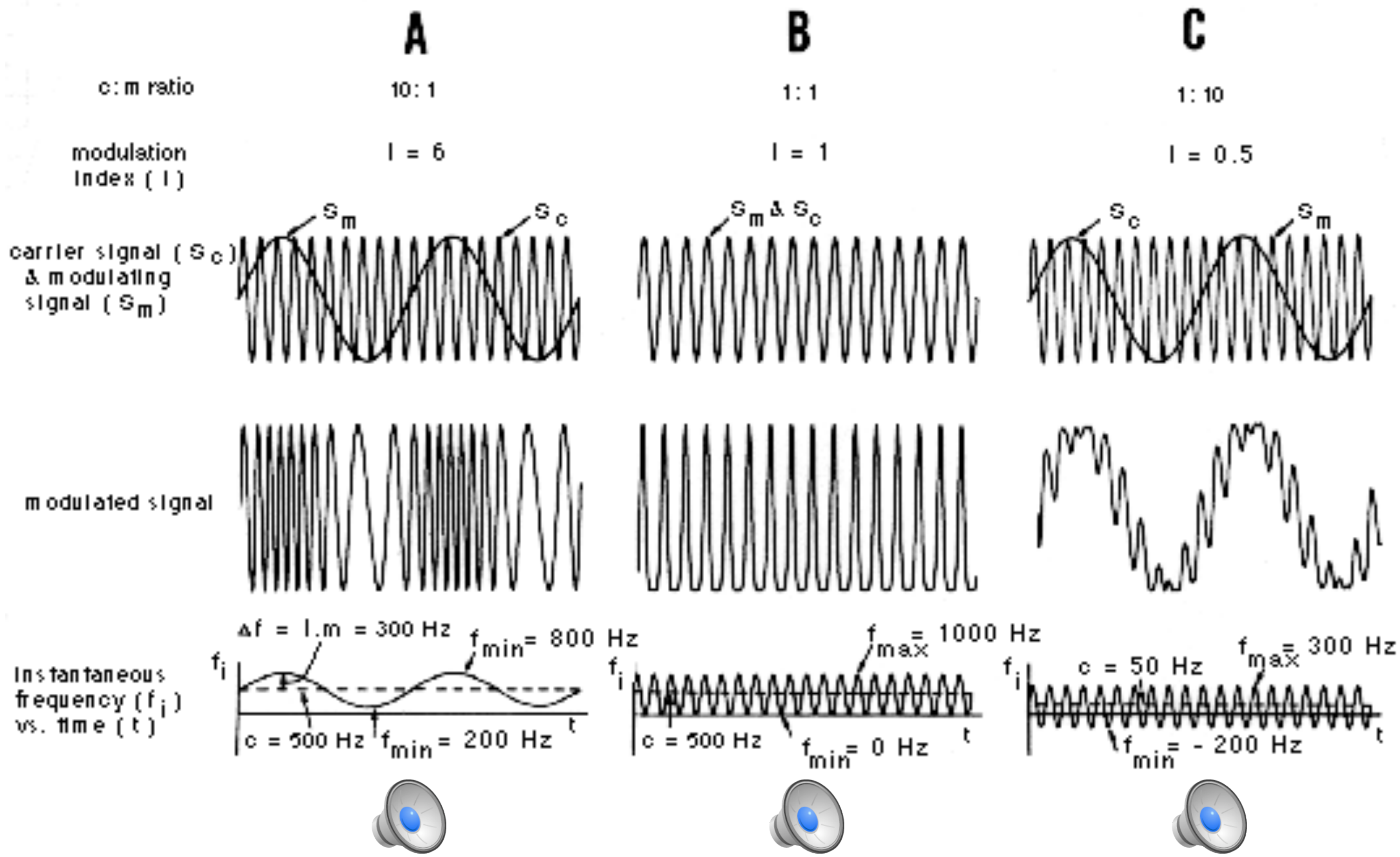
- The ratio between the carrier and modulator frequencies c/m is relevant to define the (in)harmonic characteristic of s(t).

- The sound is pitched (harmonic) if c/m is a ratio of positive integers: $\omega_c / \omega_m = N_c / N_m$
- E.g. for $f_c = 800$ Hz and $f_m = 200$ Hz, we have sidebands at 600Hz and 1kHz, 400Hz and 1.2kHz, 200Hz and 1.4kHz, etc
- Thus the fundamental frequency of the harmonic spectrum responds to: $f_0 = f_c / N_c = f_m / N_m$

- If c/m is not rational an inharmonic spectrum is produced

- If f_0 is below the auditory range, the sound will not be perceived as having definitive pitch.

FM examples



A
B
C
c:m ratio
10:1
1:1
1:10
modulation index (I)
I = 6
I = 1
I = 0.5
carrier signal (S_c) & modulating signal (S_m)
S_m
S_c
S_m & S_c
modulated signal
instantaneous frequency (f_i) vs. time (t)
Δf = I.m = 300 Hz
f_{min} = 800 Hz
c = 500 Hz
f_{min} = 200 Hz
f_{max} = 1000 Hz
f_{min} = 0 Hz
c = 50 Hz
f_{max} = 300 Hz
f_{min} = - 200 Hz
f_i
t