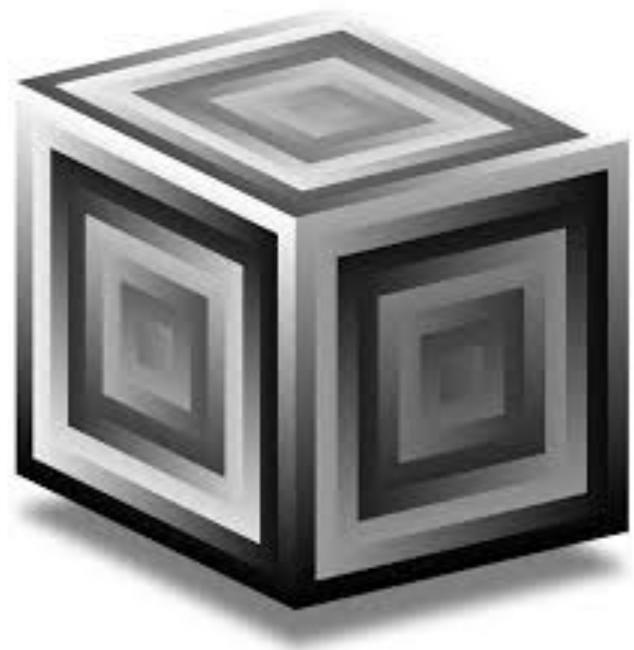
Frequency and Phase Modulation

Topics Addressed

- Frequency Modulation
- Controlling FM
- Harmonicity in FM
- Negative Frequencies
- Uses of FM Synthesis
- Phase Modulation



Amplitude Modulation

$$A \sin(2\pi f t + \phi)$$

Replace a constant amplitude

Frequency Modulation

$$A \sin(2\pi f t + \phi)$$

Replace a constant frequency

Phase Modulation

$$A \sin(2\pi f t + \phi)$$

Replace a constant phase

Frequency Modulation

- We have seen how multiplying two signals can lead to interesting effects with sidebands for both ring modulation and amplitude modulation.
- Both Frequency Modulation and Phase Modulation involve challenging math (calculus and Bessel functions of the first order), so we will take a more exploratory approach and not provide quite the level of mathematical detail as we did with AM synthesis
- Like AM synthesis, frequency modulation requires a carrier wave and a modulator wave. Here we will be modulating the frequency of the carrier wave with a modulator wave.

Frequency Modulation

- Consider a carrier wave $A_c \sin(2\pi f_c t \ + \ \phi_c)$ where A_c is the amplitude of the carrier frequency, f_c is the frequency of the carrier frequency, and ϕ_c is the phase of the carrier frequency.
- We want to modulate f_c. This means that at any given instant the value of f_c will be changing, which distinguishes it from amplitude modulation.
- Let's define a modulator signal called $m(t)$. For our purposes, $m(t)$ will be a periodic waveform scaled between -1 to +1 like a sine wave.
- Let's then define the instantaneous frequency of our signal as $f_c + \ k_f m(t)$ where k_f is a constant expressing the deviation away from the carrier frequency.
    - Why express our instantaneous frequency like this? We preserve the original carrier frequency f_c and fluctuate around that value.

Aside: Deriving an equation for FM

- The value of a sine wave is determined by its phase which is simply a number in radians. We plug in a time *t* that gives us back a number in radians that then is plugged into sine to give us an amplitude.
- *Instantaneous frequency* then is the change in phase with respect to time in the same way that instantaneous velocity is the change in distance with respect to time. With most sine waves, the instantaneous frequency is constant. For example, a sine wave with frequency 440Hz is its instantaneous frequency. These two equations then are equivalents: $A_c \sin(2\pi f_c t + \phi_c) = A_c \sin\left(2\pi \int_0^t f_i(t)dt + \phi_c\right)$ where $f_i(t)$ is the instantaneous frequency based on time. If the instantaneous frequency is constant, say 440Hz, then $A_c \sin\left(2\pi \int_0^t 440dt + \phi_c\right) =$ $A_c \sin(2\pi(440)t + \phi_c)$ which makes intuitive sense.

Aside: Deriving an equation for FM

- $A_c \sin\left(2\pi \int_0^t f_i(t)dt + \phi_c\right)$ expresses a signal based on instantaneous frequency. Here let's use the instantaneous frequency we were considering before: $f_c + k_f m(t)$
- $A_c \sin\left(2\pi \int_0^t (f_c + k_f m(t))dt + \phi_c\right)$
- $A_c \sin\left(2\pi f_c t + 2\pi k_f \int_0^t m(t)dt + \phi_c\right)$ <- final form
- Note that ϕ_c is often ignored and set to zero.

Frequency Modulation in SC

```
(
SynthDef(\freqMod, {
	arg out = 0, freq_c = 440, freq_m = 1, k_f = 1;
	var sig = SinOsc.ar(freq_c + (k_f * SinOsc.ar(freq_m)), 0, 1);
	Out.ar(out, sig ! 2);
}).add;
)
```

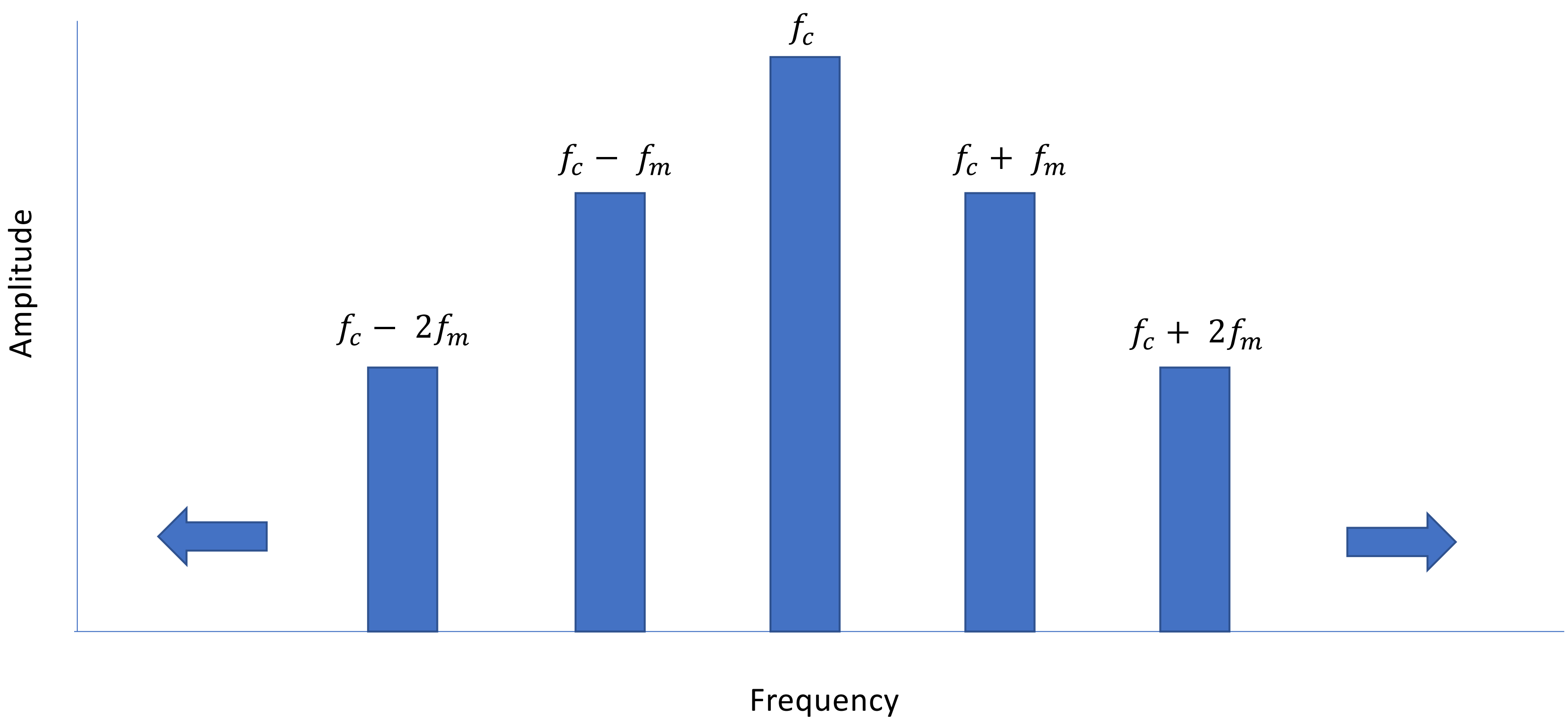
Frequency Modulation Observations

- When k_f is small and the frequency of the modulator is also small, we perceive a vibrato effect as the pitch of the carrier fluctuates subtly.
  - Note that the frequency of the modulator must be an LFO
- If we keep modulator frequency as an LFO but increase k_f to anywhere above 10, we start getting a wacky swooping sort of sound.
- When the frequency of the modulator reaches the audible range (> 20Hz), we start to perceive other notes, also called sidebands, in the same way with AM synthesis.

FM Sidebands

- Mathematically converting the equation for FM synthesis when $m(t) = \sin(2\pi f_m t)$ into a sum of sines is mathematically beyond the scope of this class (requires knowledge of Bessel functions of the first order).
- Important information about sidebands:
  - The carrier frequency is always present and generally has the largest amplitude.
  - Symmetric sidebands exist around f_c at a distance of $f_c \pm nf_m$ where n is an integer from 0 to infinity.
  - As n increases the amplitude of the sidebands decreases.

Visualizing FM Sidebands



Controlling Sidebands

- The frequency of the modulating wave determines at what frequency the sidebands occur
  - These partials can be either harmonic or inharmonic
- The value k_f (frequency deviation) from our instantaneous frequency formula $f_c + k_f m(t)$ determines the amplitude of the sidebands.
  - Like AM synthesis, FM synthesis also has a modulation index. It is defined as the ratio $\frac{k_f}{f_m}$.
  - In most analog and digital FM synthesizers sideband strength is controlled through the modulation index which is proportional to k_f but is scaled by f_m.
  - The Wikipedia article on FM has a chart explaining the amplitude of sidebands based on the modulation index. Note that as the modulation index increases, the carrier wave's frequency can have a **smaller** amplitude than some of its sidebands.

Exploring FM Synthesis

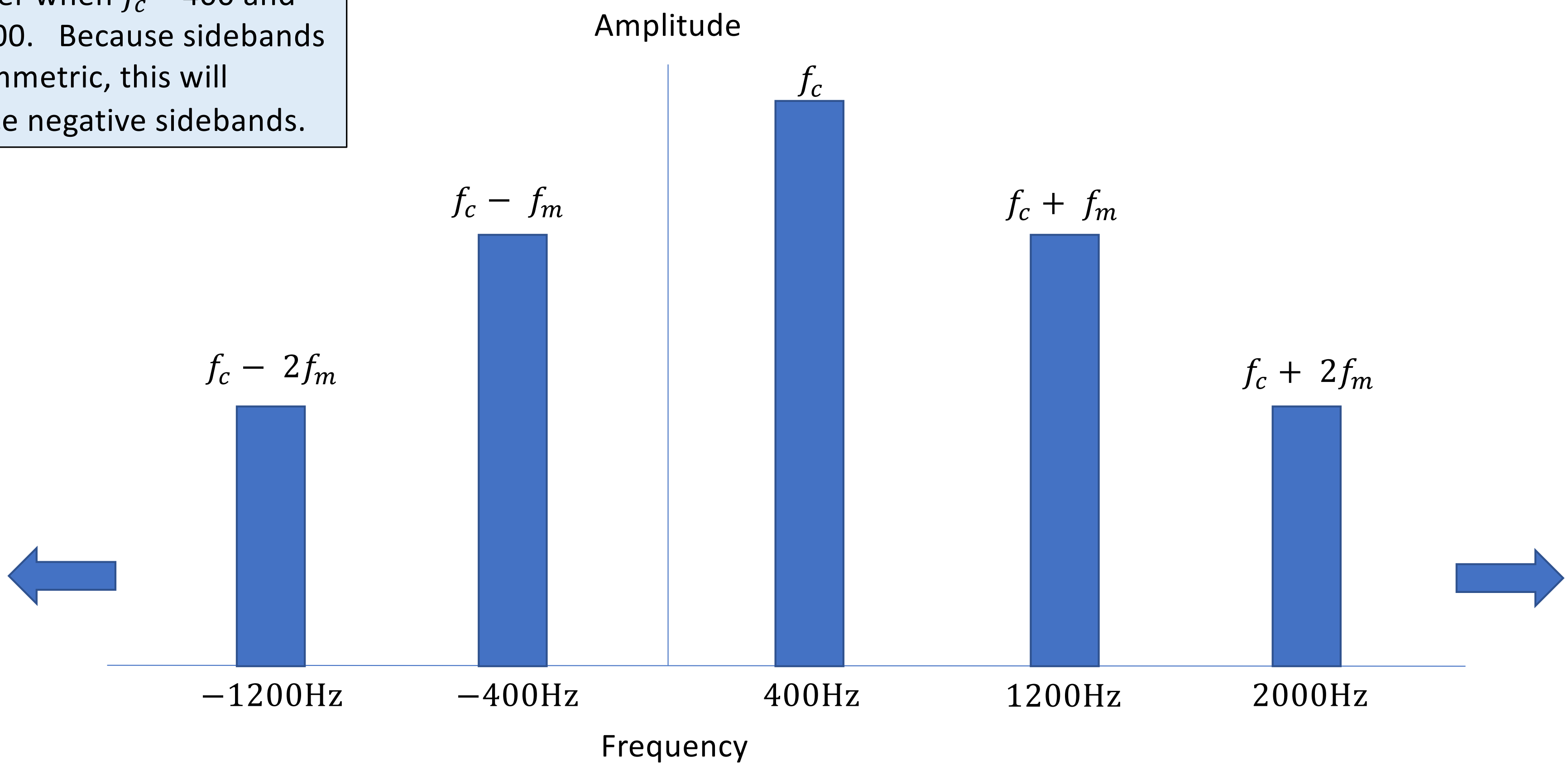
```
(
SynthDef(\freqModMouse, {
	arg out = 0, freq_c = 500;
	var car, mod, modIndex, freq_m;
	freq_m = MouseX.kr(1, 10000, 1).poll;
	modIndex = MouseY.kr(0, 10, 1);
	mod = (modIndex * freq_m) * SinOsc.ar(freq_m, 0, 1);
	car = SinOsc.ar(freq_c + mod, 0, 1);
	Out.ar(out, car ! 2);
}).add;
)
```

The poll method on a UGen posts its value every tenth of a second.

- FM synthesis can produce a wide range of different sounds. Here the modulation index is mapped to y-axis of the mouse which will strengthen the amplitude of the sidebands. The frequency of the modulator is mapped to the x-axis to control the location of the sidebands.
- Question: for what frequencies and/or modulation indices does FM synthesis produce *harmonic* partials?

Negative Frequencies

Consider when f_c = 400 and f_m = 800. Because sidebands are symmetric, this will produce negative sidebands.

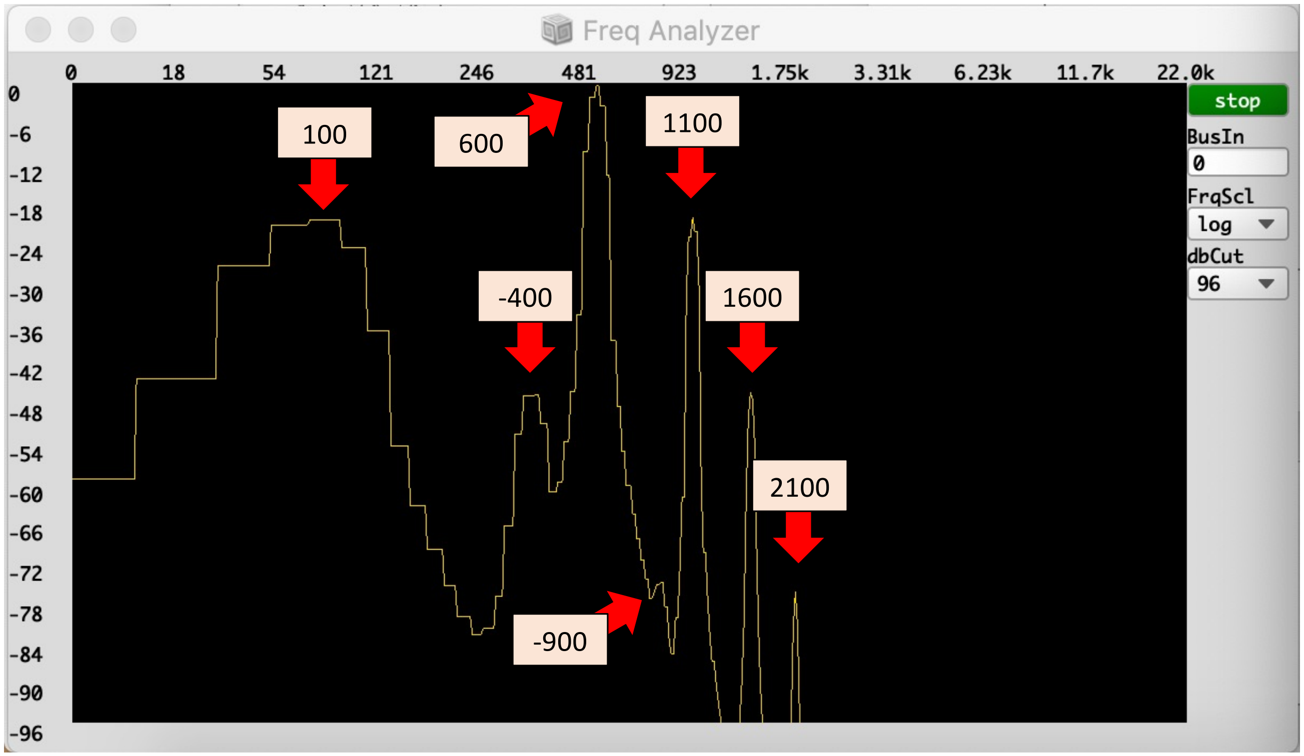


Negative Frequencies

- To consider the case of negative frequencies, let’s rely upon our trigonometric identities.
  - $\sin(-\theta) = -\sin\theta$
  - $-\sin\theta = \sin(\theta + \pi)$
  - Therefore, $\sin(-\theta) = \sin(\theta + \pi)$
- If we consider any sideband with negative frequency represented as $A_s \sin(-2\pi f_s t)$, then this is equivalent to $A_s \sin(2\pi f_s t + \pi)$
  - Negative frequencies are nothing special to the ear other than phase shifted versions of the absolute value of that frequency
  - When using a frequency scope, this means negative frequencies “wrap around” to positive frequencies.

Negative Frequencies

$f_c = 600$
$f_m = 500$

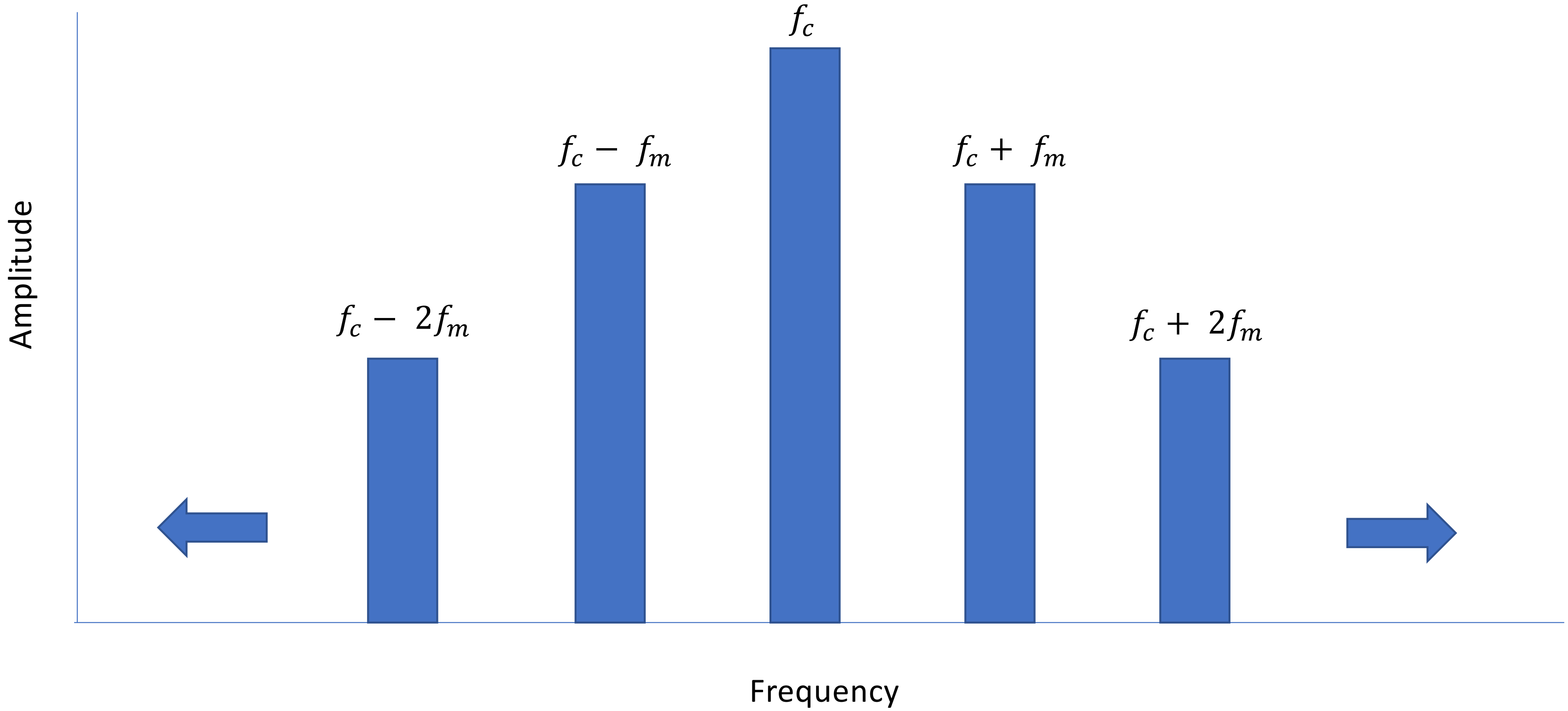


Note how -400Hz wraps to 400Hz and -900Hz wraps to 900Hz

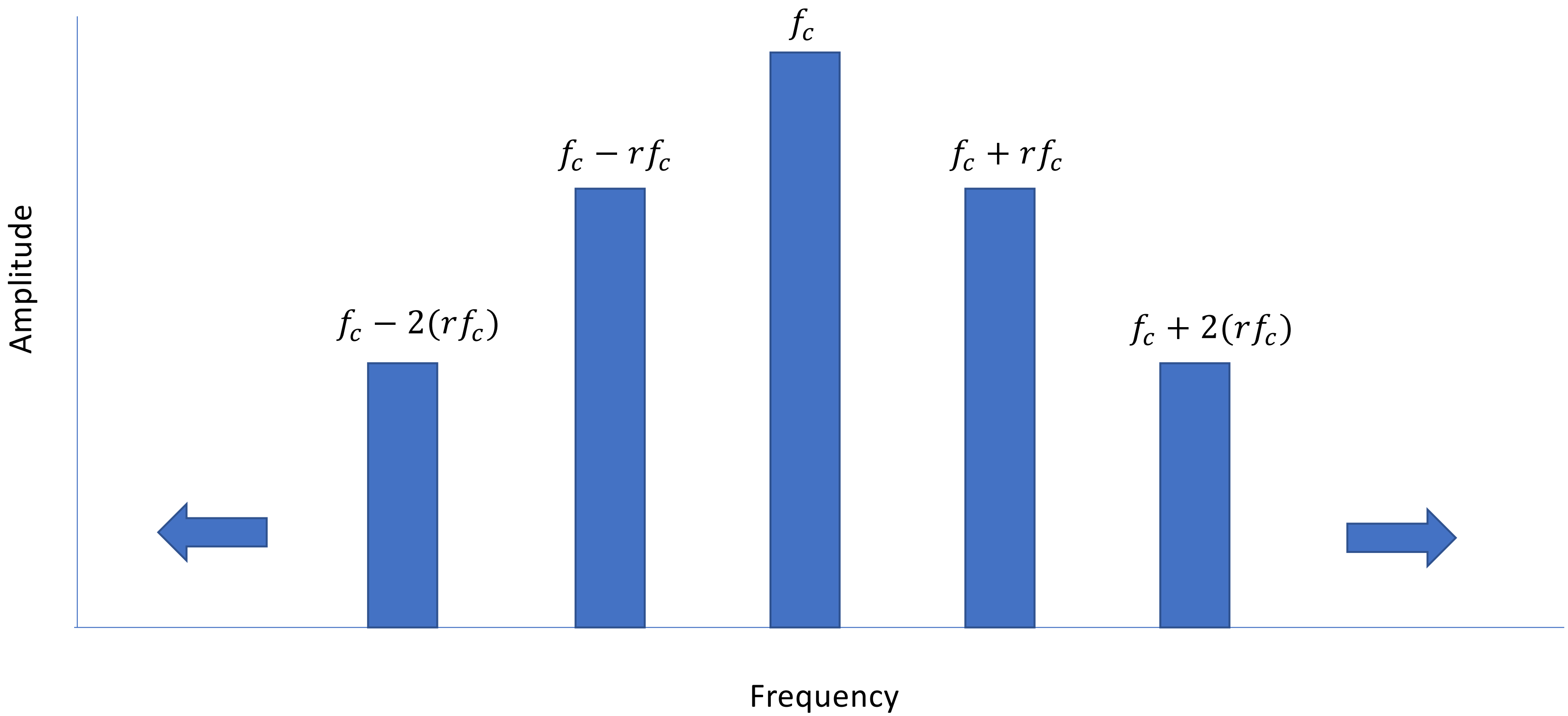
Harmonicity Ratio

- We can define the harmonicity ratio as $\frac{f_m}{f_c}$ or $\frac{f_c}{f_m}$ depending on the text. When either produces an integer, the partials of FM synthesis will be harmonic (i.e., partials from the harmonic series). For our purposes, we will prefer $\frac{f_m}{f_c}$ because f_c will always be the fundamental. Remember the harmonic series is produced by integer multiples of some fundamental. Same thing!
- When the harmonicity ratio is not an integer then FM synthesis will produce inharmonic partials.
- We can say then that the harmonicity ratio r gives us the equation for the frequency of the modulator as $f_m = r * f_c$ where r is a positive integer.

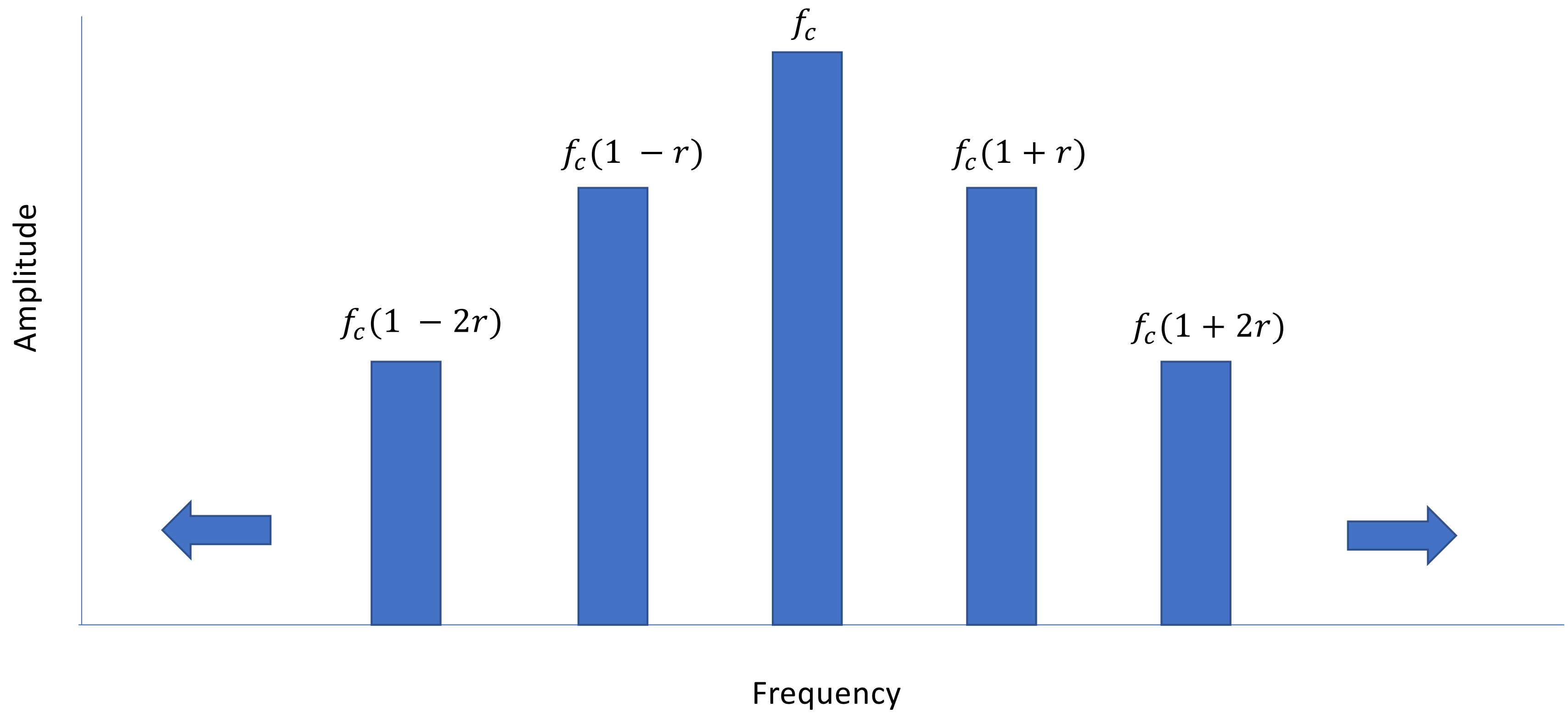
Visualizing FM Sidebands with Harmonicity Ratio



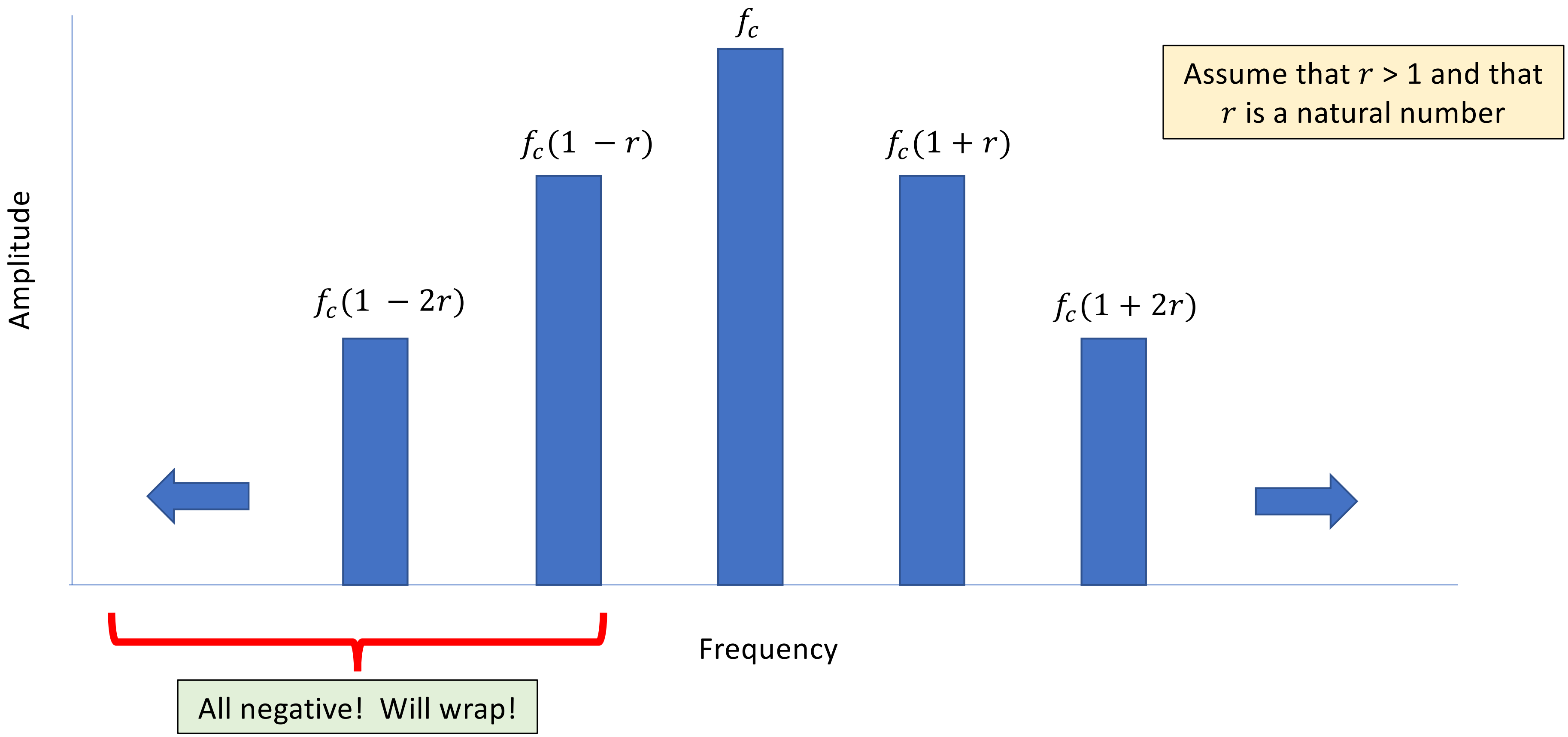
Visualizing FM Sidebands with Harmonicity Ratio



Visualizing FM Sidebands with Harmonicity Ratio



Visualizing FM Sidebands with Harmonicity Ratio



Harmonicity Ratio

- Sidebands generated when $r > 1$ and r is a natural number will be f_c plus the nonpositive sidebands that wrap $f_c(1 - r)$, $f_c(1 - 2r)$, … plus the positive sidebands $f_c(1 + r)$, $f_c(1 + 2r)$, …
- When $r = 1$, we get all partials.
- When $r = 2$, what partials are generated?
  - f_c
  - Wrapped negative sidebands : $f_c(1 - r)$, $f_c(1 - 2r)$, … or $f_c, 3f_c, 5f_c$, …
  - Positive sidebands: $f_c(1 + r)$, $f_c(1 + 2r)$, … or $3f_c, 5f_c, 7f_c$, …
  - All odd harmonic partials! Like a triangle or square wave.
- What about $r = 3$? Or higher ratios?

FM Synthesis Brief History

- Frequency Modulation was first used in radio in 1933. It was a better, less noisy way of transmitting signals compared to AM Radio.
- FM Synthesis in music was developed by John Chowning at Stanford University starting in 1967 and patented in 1975
  - It was a "cheap" way to produce complex sounds with simple unit generators (i.e., two sinusoids)
- FM Synthesis was used in the early development of digital synthesizers by Yamaha
- The patent for FM Synthesis expired in 1995 and can now be freely used in any digital synthesizer

Uses of FM Synthesis

- FM Synthesis is capable of producing both harmonic and inharmonic complex sounds including percussive, metallic, instrumental, and bell-like sounds.
  - Instrument modeling, the technique of imitating a physical musical instrument via sound synthesis
  - FM Synthesis was the basis of the Yamaha DX7, one of the most influential digital synthesizers in the 80s – Check out this video
- See `\fmEnv` and `\fmShimmer` in lecture code to see some interesting sounds that can be produced from FM Synthesis

Aside: Phase Modulation

- Phase modulation is closely related to frequency modulation and many of the conclusions that we can draw from frequency modulation also apply to phase modulation.
  - Phase modulation is sometimes called indirect FM or Chowning-style FM.
- Given a carrier wave, $A_c \sin(2\pi f_c t + \phi_c)$, and a modulating signal, $m(t)$, we can express phase modulation as $A_c \sin\left(2\pi f_c t + k_p m(t) + \phi_c\right)$ where k_p is the frequency deviation that determines the “strength” of the modulating wave.
  - Note that for simplicity’s sake ϕ_c is usually assumed to be 0. Thus, we can also write $A_c \sin\left(2\pi f_c t + k_p m(t)\right)$

Aside: Phase Modulation in SC

```
(
SynthDef(\phaseMod, {
	arg out = 0, freq_c = 440, freq_m = 1, k_p = 1;
	var sig = SinOsc.ar(freq_c, SinOsc.ar(freq_m, 0, k_p).mod(2pi), 1);
	Out.ar(out, sig ! 2);
}).add;
)

~phaseMod = Synth(\phaseMod);
~phaseMod.set(\freq_c, 500);
~phaseMod.set(\freq_m, 750);
~phaseMod.set(\k_p, 0.01);
~phaseMod.free;
```

- Note that mod(2pi) is necessary here because SinOsc must be wrapped between +/- 8pi
- The partials and their organization in PM synthesis is the same as in FM synthesis. We can't perceive any aural difference between the two and it's impossible to tell the difference between PM/FM synthesis
- There are differences though in terms of radio because the signal needs to be demodulated and the original modulating signal needs to be retrieved.