DTI 516 Multimedia Processing

Chapter: 9

Introduction to Sounds

Dr. Paween Khoenkaw

Digital Technology Innovation : Maejo University

Topics

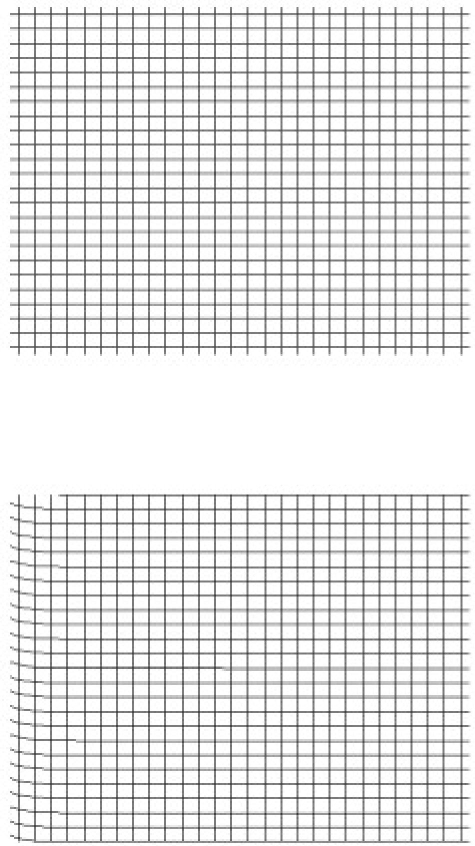
- Pure-tone Generator

- Harmonic and Envelope

- Sampling and Sample-based synthesis

What is sound

In physics, **sound** is a vibration that typically propagates as an audible wave of pressure, through a transmission medium such as a gas, liquid or solid.
*Wikipedia



The model of sound

Sound is a wave, a **sine wave** or **sinusoid wave**

$$y(t) = A\sin(2\pi f t + \varphi)$$

$$\omega = 2\pi f$$

$$y(t) = A\sin(\omega t + \varphi)$$

The model of sound

The parameter of sound

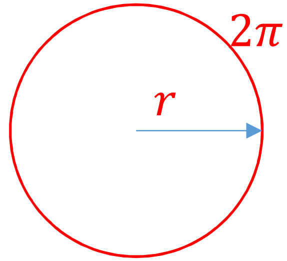
$$y(t) = A\sin(2\pi f t + \varphi)$$

What is it ?

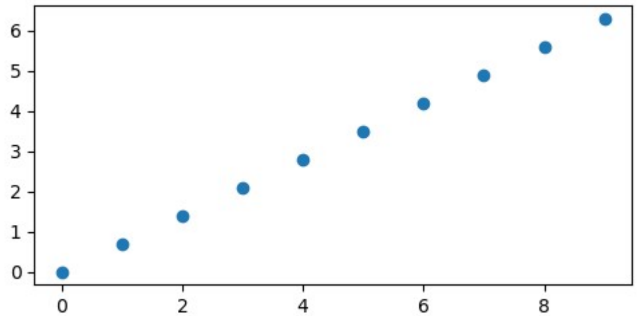
The model of sound

$$y(t) = A\sin(2\pi\, f\, t + \varphi)$$

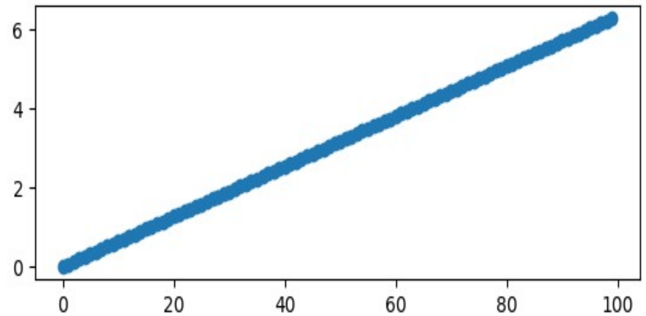
2π is the **Circumference** of the circle



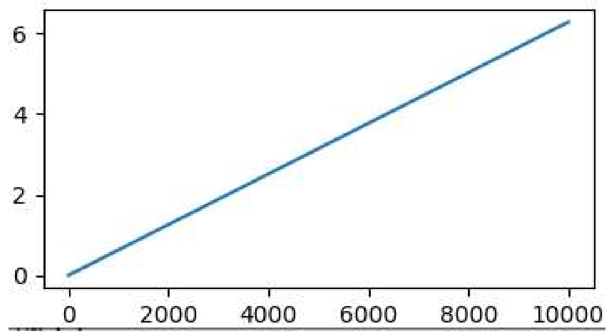
t is time or number of sample



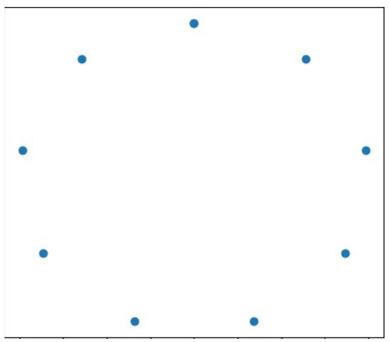
t=0,...,10

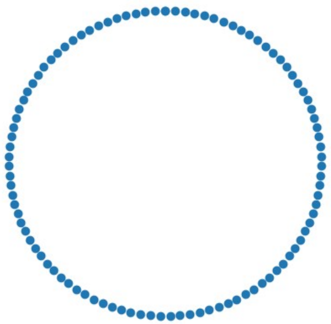


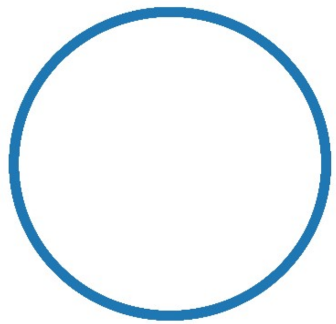
t=0,..,100



t=0,..,10000

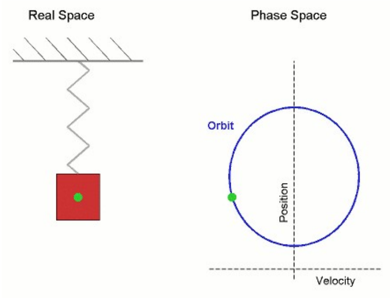




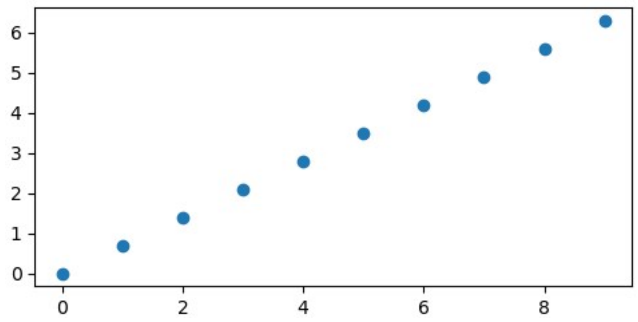


The model of sound

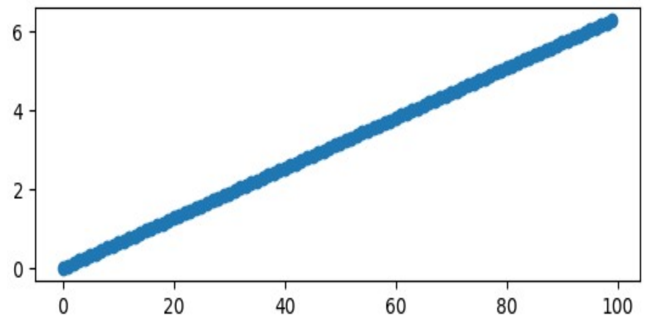
$$y(t) = A \sin(2\pi f t + \varphi)$$



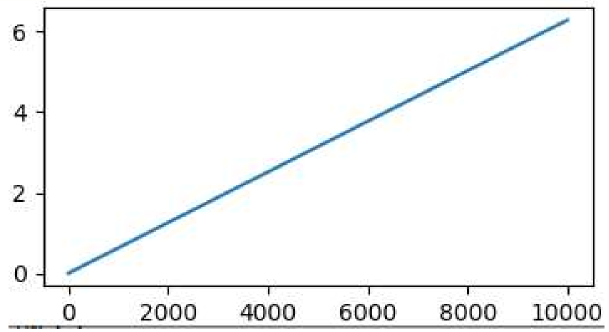
Simple harmonic motion



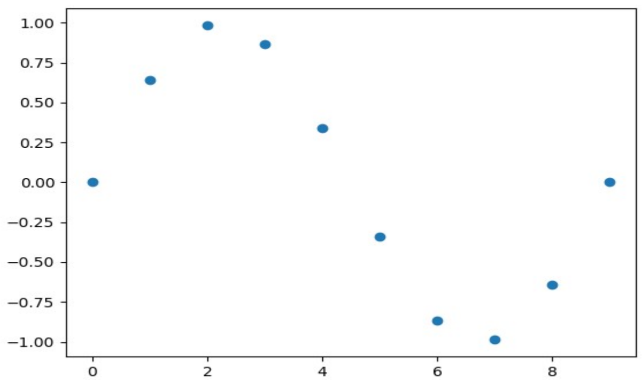
t=0,...,10

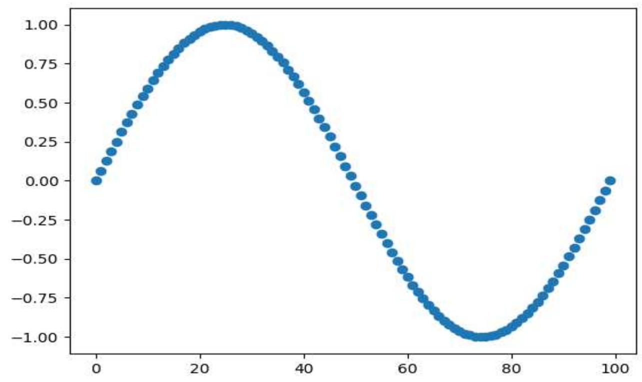


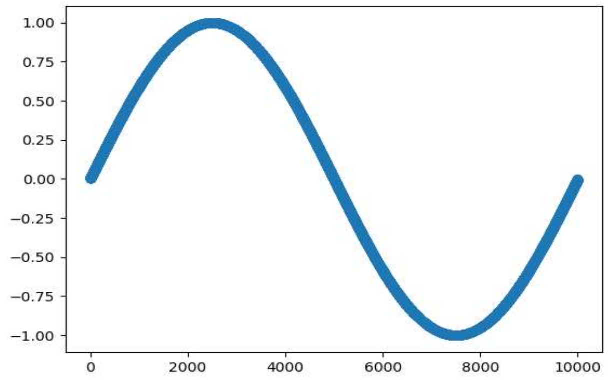
t=0,..,100



t=0,..,10000

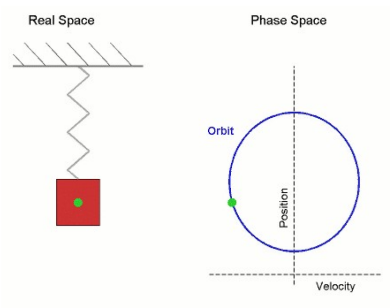




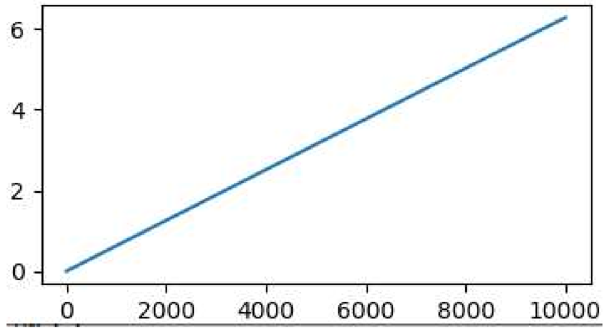


The model of sound

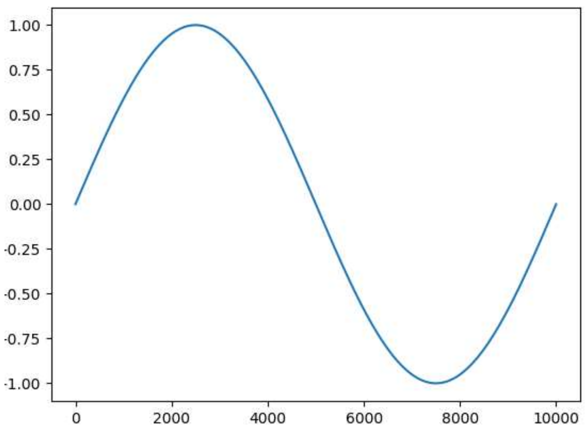
$$y(t) = A\sin(2\pi f\, t + \varphi)$$



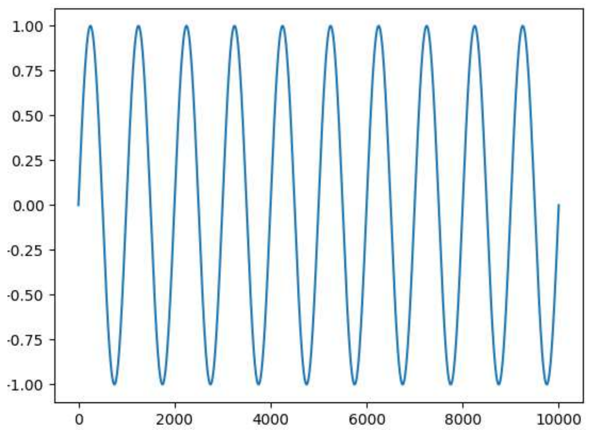
Oscillation Frequency



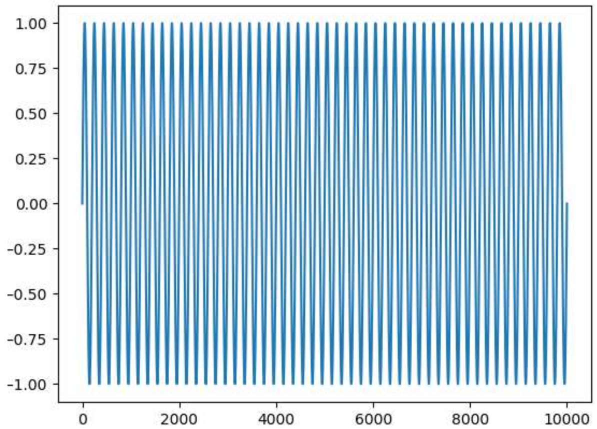
t=0,..,10000



f=1



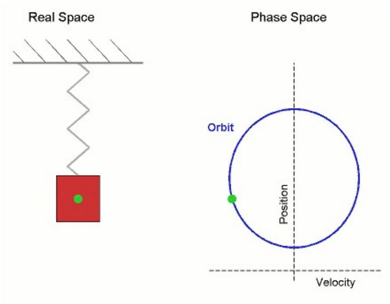
f=10



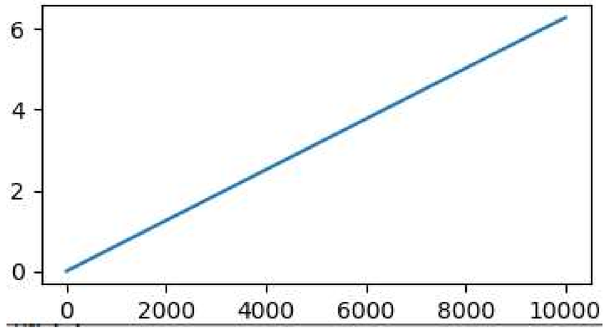
f=50

The model of sound

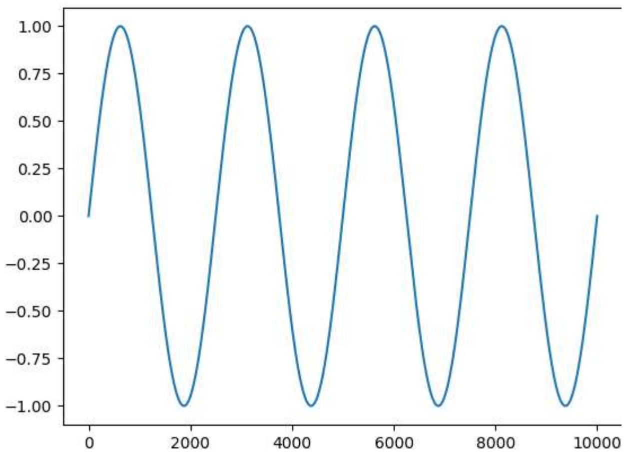
$$y(t) = A\sin(2\pi f t + \varphi)$$



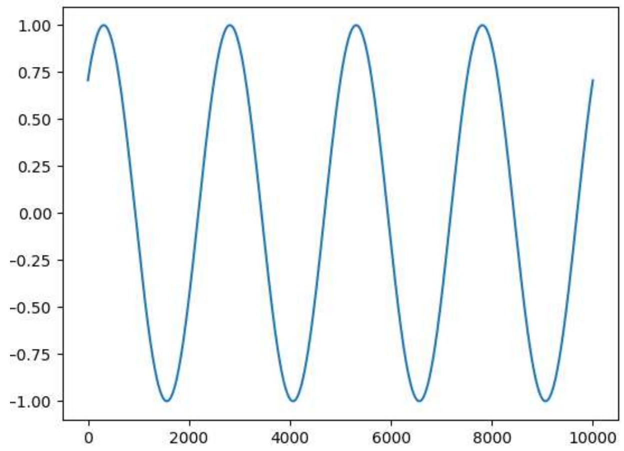
Phase shift



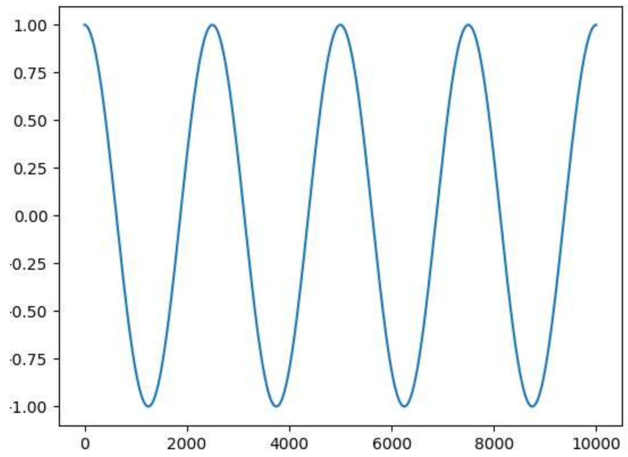
t=0,..,10000



$\varphi = 0$



$\varphi = \frac{1}{4}\pi$

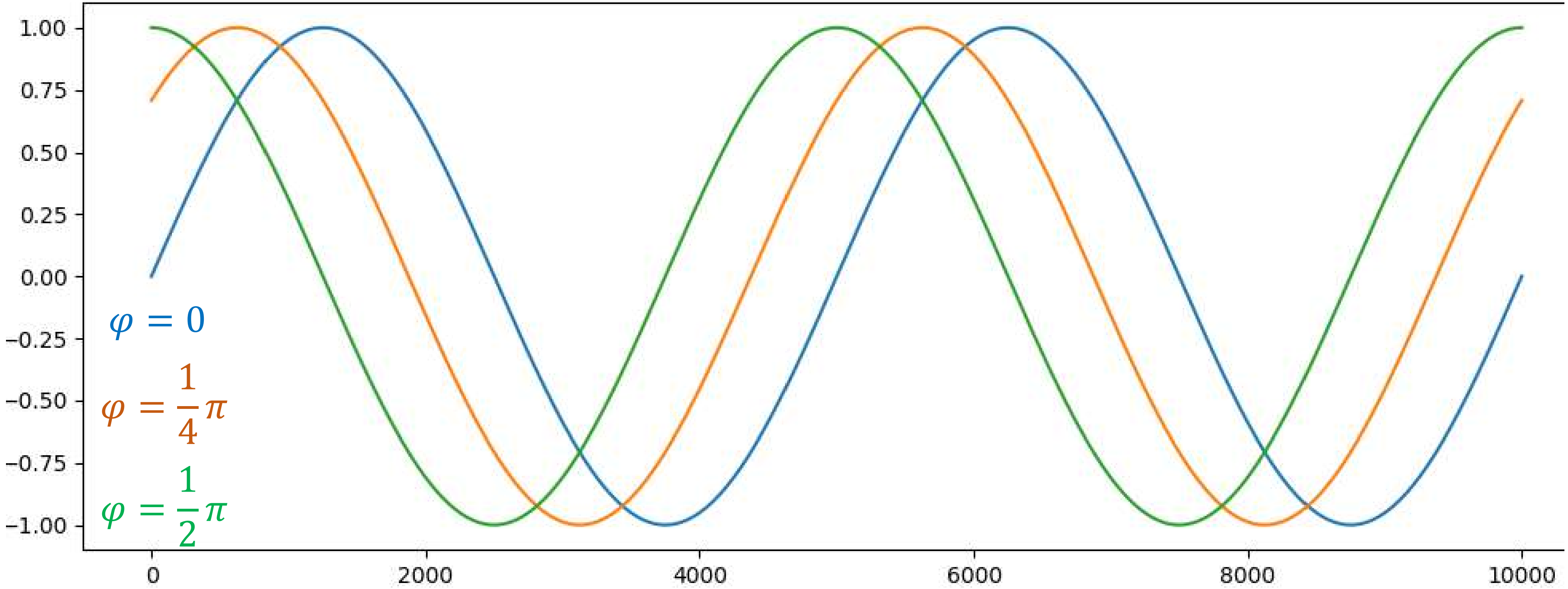


$\varphi = \frac{1}{2}\pi$

The model of sound

$$y(t) = A\sin(2\pi f t + \varphi)$$

Phase shift



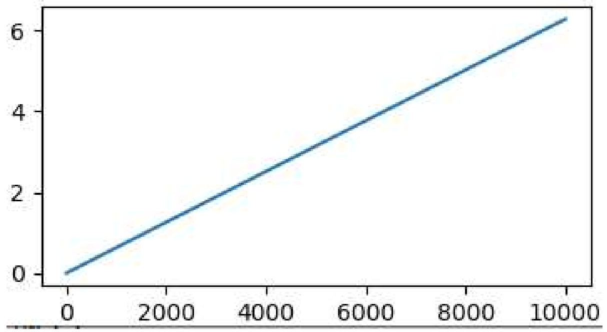
$$cos\theta = sin\left(\frac{\pi}{2} - \theta\right)$$

Sound can be model as sine function or cosine function

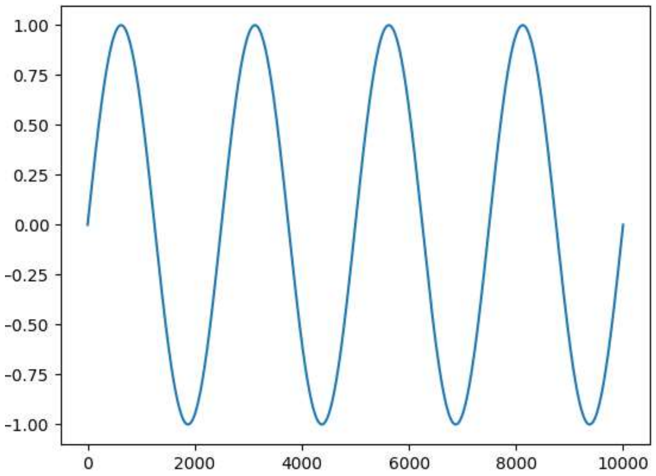
The model of sound

$$y(t) = A \sin(2\pi\, f\, t + \varphi)$$

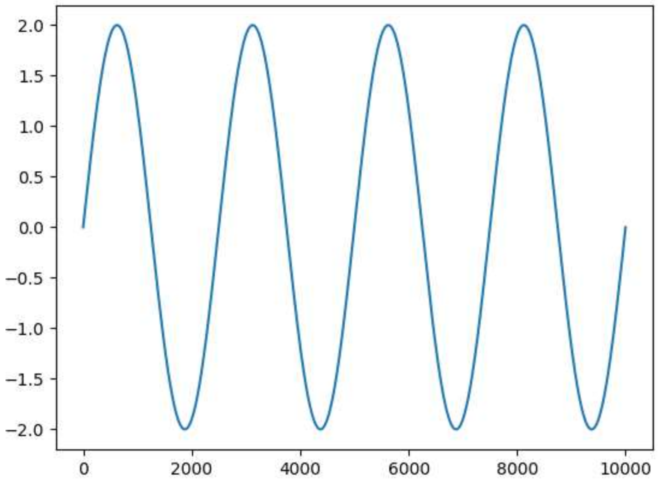
Phase shift



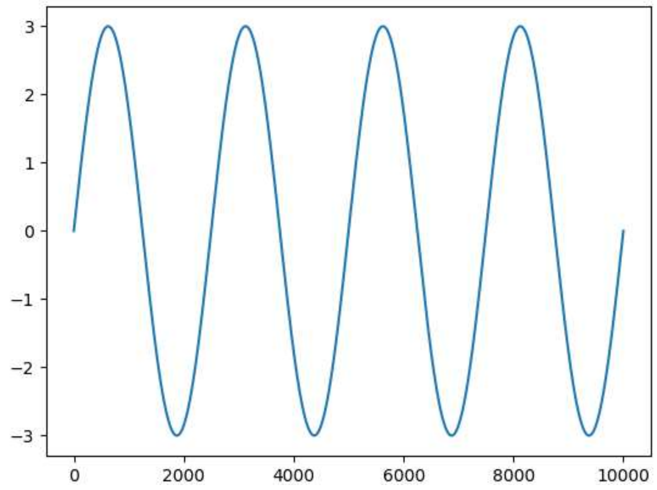
t=0,..,10000



$A = 1$



$A = 2$

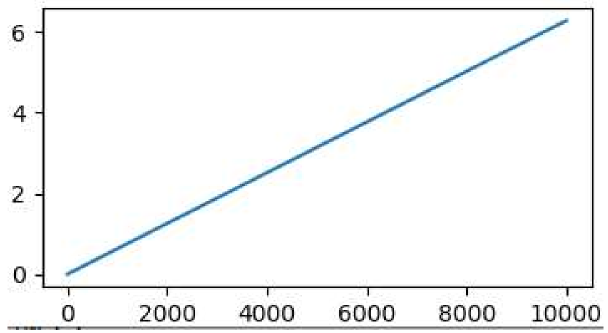


$A = 3$

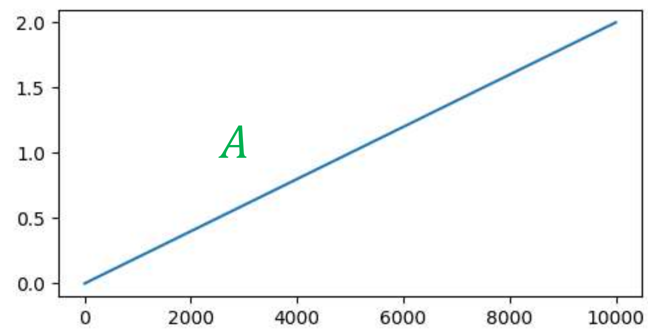
The model of sound

$$y(t) = A \sin(2\pi\, f\, t + \varphi)$$

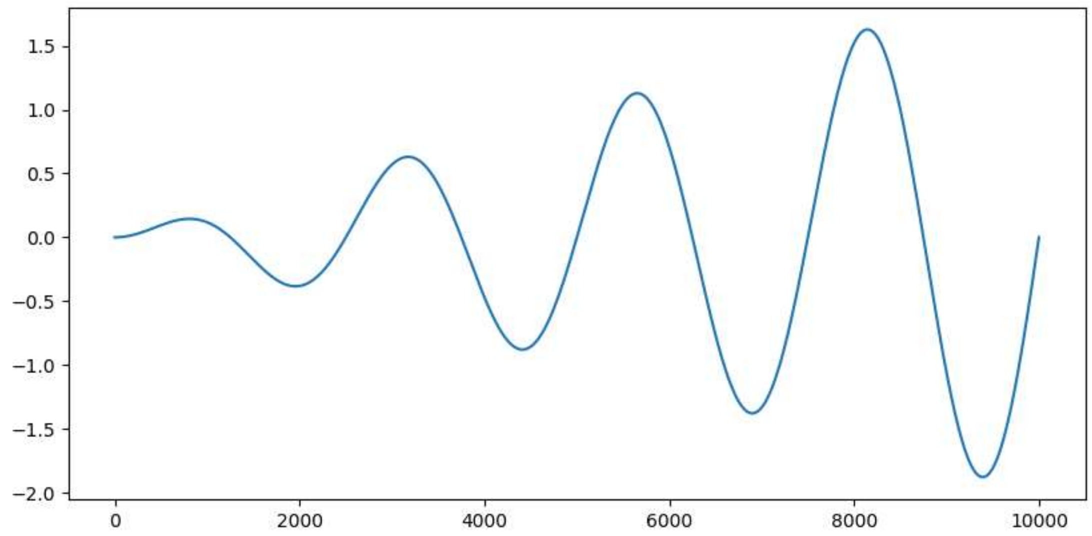
Phase shift



t=0,..,10000



$$y(t) = A(t) \sin(2\pi\, f\, t + \varphi)$$

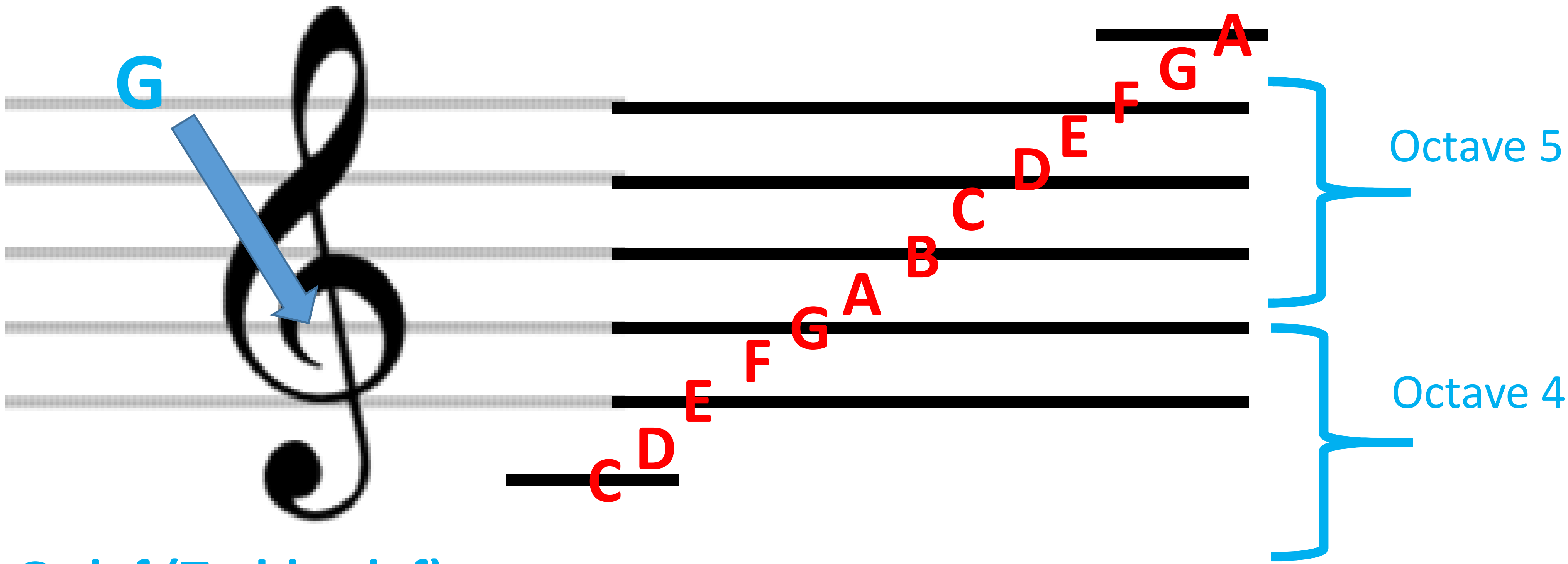


$\varphi = 0$

$f = 4$

Read the score and program the music

Clefs : Describe the pitch or frequency



G clef (Treble clef)

Read the score and program the music

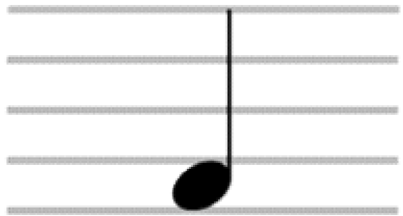
Clefs : Describe the pitch or frequency

Read the score and program the music

Note and rest : Describe length

The length in music is not define in second or millisecond , but based on tempo and ratio

Tempo is number of beat in 1 minute
measure in **beat-per-minute**

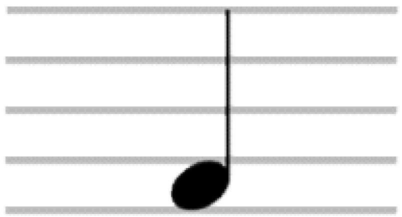


Crotchet (Quarter note)
=
1 Beat

Read the score and program the music

Note and rest : Describe length

Tempo is number of beat in 1 minute measure in **beat-per-minute**



Crotchet (Quarter note)

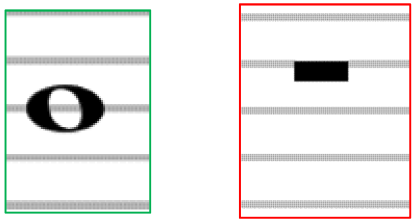
=

1 Beat

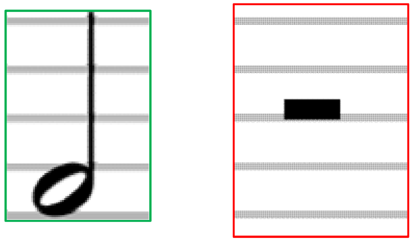
$$lenght\ of\ crotchet = \frac{60 \times 1000}{tempo}$$ in millisecond unit

Read the score and program the music

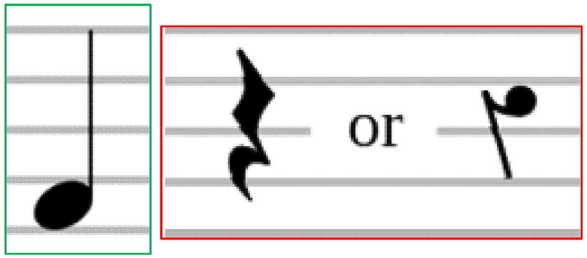
Note and rest : Describe length



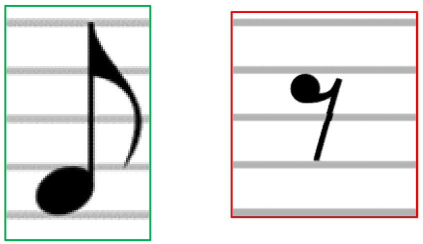
4crotchet



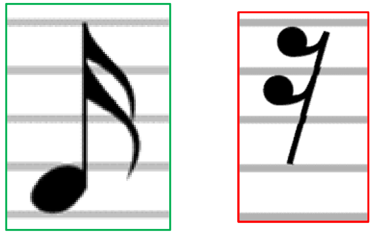
2crotchet



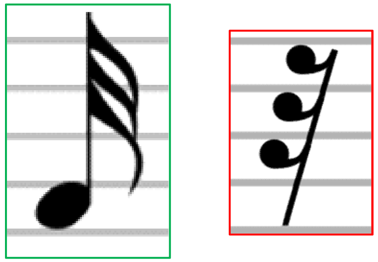
crotchet



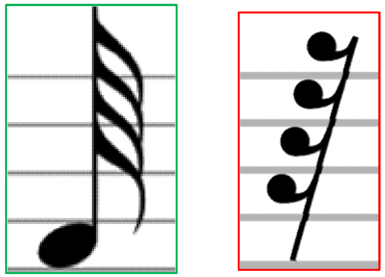
$\frac{1}{2}$ *crotchet*



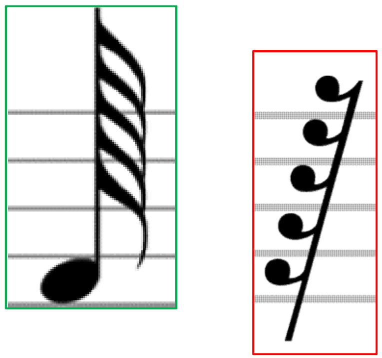
$\frac{1}{4}$ *crotchet*



$\frac{1}{8}$ *crotchet*



$\frac{1}{16}$ *crotchet*



$\frac{1}{32}$ *crotchet*

https://en.wikipedia.org/wiki/List_of_musical_symbols

Read the score and program the music

Note and rest : Describe length

https://en.wikipedia.org/wiki/List_of_musical_symbols

Read the score and program the music

What is the length of this note on the 60 bpm music score ?

$\frac{1}{2} crotchet$

$$lenght\ of\ crotchet = \frac{60 \times 1000}{tempo}$$

$$lenght\ of\ crotchet = \frac{60 \times 1000}{60} = 1000\ ms$$

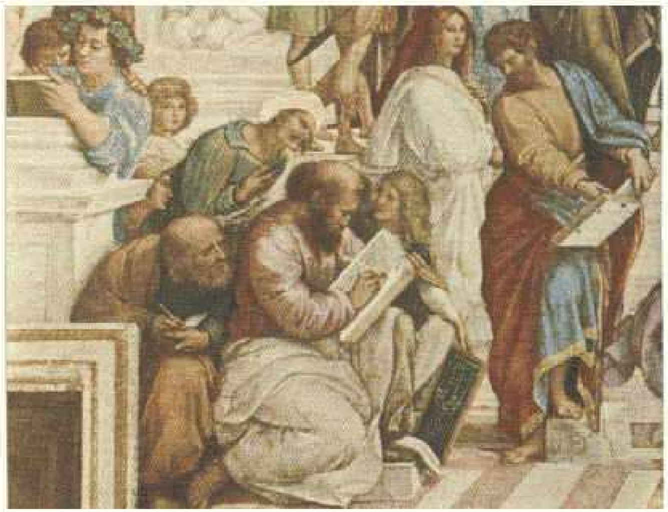
$$Lenght = \frac{1}{2} crotchet = \frac{1000}{2} = 500\ ms$$

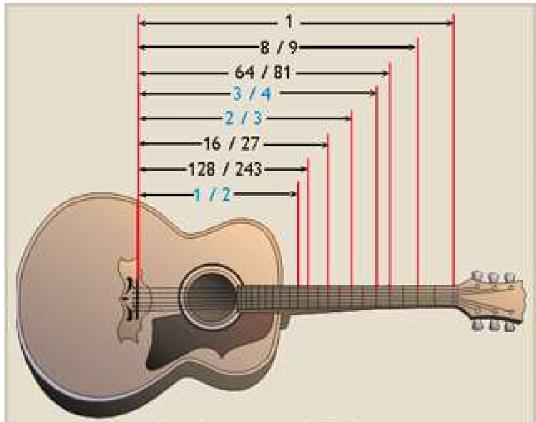
Read the score and program the music

What is the frequency of each note ?

It can be any frequency, but will it sound right ?

The frequency table of the musical note is call "Temperament"


THE PYTHAGORIAN GROUP From the School of Athens



Pythagoras' Guitar

Note	C	D	E	F	G	A	B	C
Ratio	$1/1$	$9/8$	$81/64$	$4/3$	$3/2$	$27/16$	$243/128$	$2/1$

Step	—	$9/8$	$9/8$	$256/243$	$9/8$	$9/8$	$9/8$	$256/243$	—

Read the score and program the music

ตารางที่ 4.5 อัตราส่วนความถี่และระยะพิตช์ของเสียงขลุ่ยเพียงออ

ระดับเสียง	ความถี่ (Hz)	อัตราส่วนความถี่	ระยะพิตช์ (cents)
โด	465.39		
		1.0984 ≅ 1.10	162.4355 ≅ 162
เร	511.17		
		1.0896 ≅ 1.10	148.4936 ≅ 149
มี	556.95		
		1.0822 ≅ 1.10	136.7281 ≅ 137
ฟา	602.72		
		1.1013 ≅ 1.10	167.0085 ≅ 167
ซอล	663.76		
		1.0805 ≅ 1.10	133.9600 ≅ 134
ลา	717.16		
		1.1383 ≅ 1.14	224.2714 ≅ 224
ที	816.35		
		1.0935 ≅ 1.10	154.6688 ≅ 155
โด′	892.64		

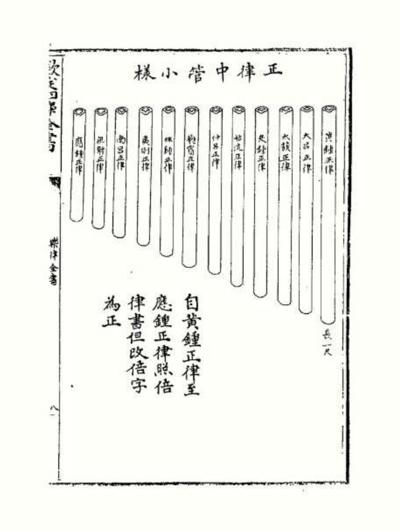
http://siambanlang.blogspot.com/2012/06/blog-post.html

Read the score and program the music

Equal temperament

J. S. Bach wrote *The Well-Tempered Clavier* to demonstrate the musical possibilities of well temperament

Zhu Zaiyu 朱載堉



Zhu Zaiyu has been credited as the first person to solve the equal temperament problem mathematically

$$Ratio = \sqrt[12]{2} \approx 1.059463094334816469336$$

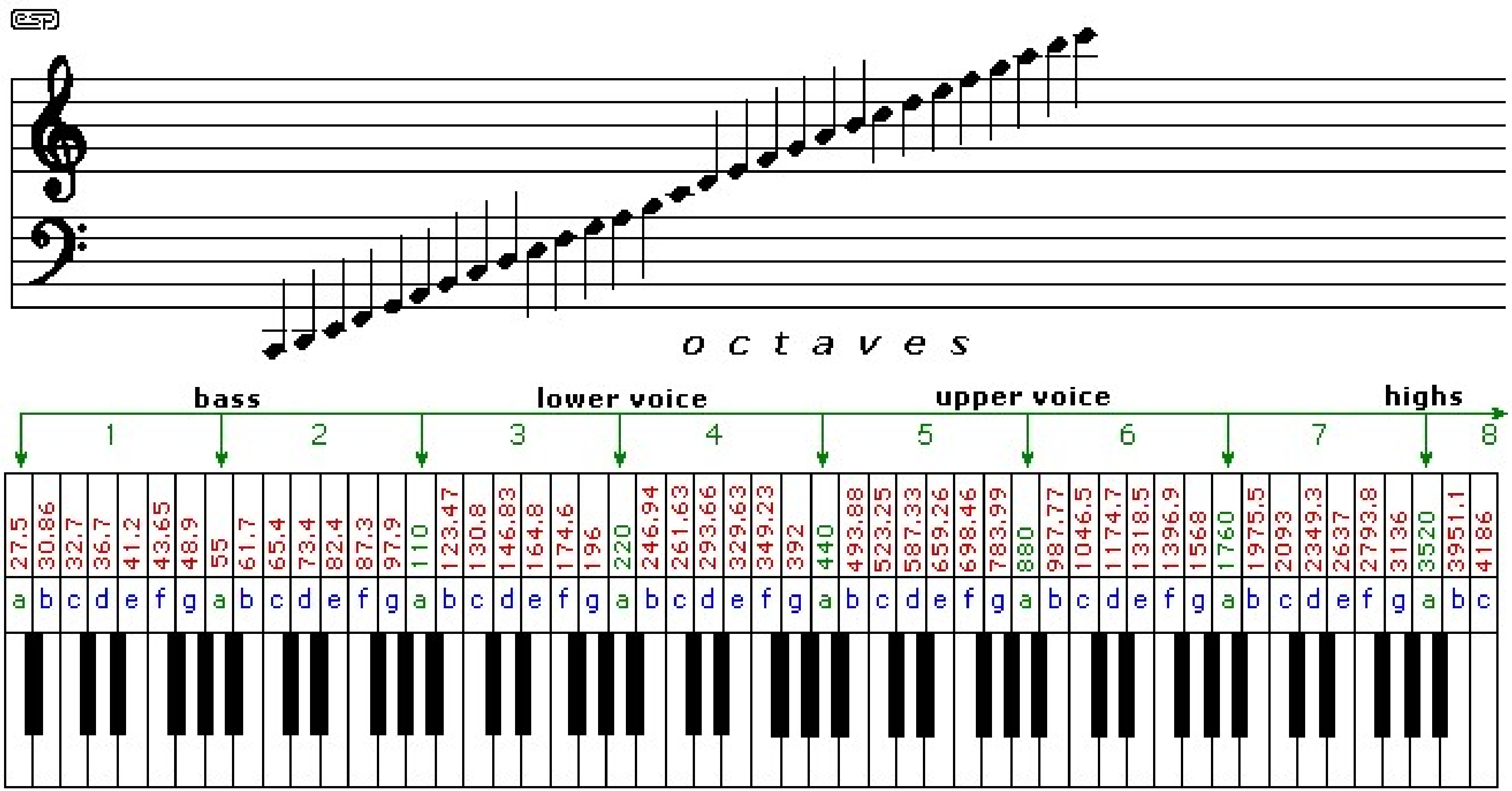
Temperament

ไทย

Equal Temperament

Read the score and program the music

Equal temperament

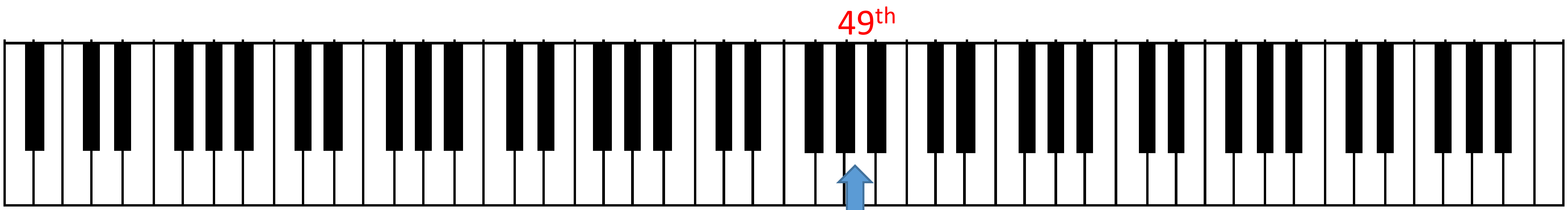


Read the score and program the music

Equal temperament

$$Ratio = \sqrt[12]{2} \approx 1.0594630943481646 9336$$

$$f_i = f_0\left(\sqrt[12]{2}\right)^{i-49}$$



ISO 16:1975
Acoustics — Standard tuning frequency
(Standard musical pitch)

440Hz

Middle A

Read the score and program the music

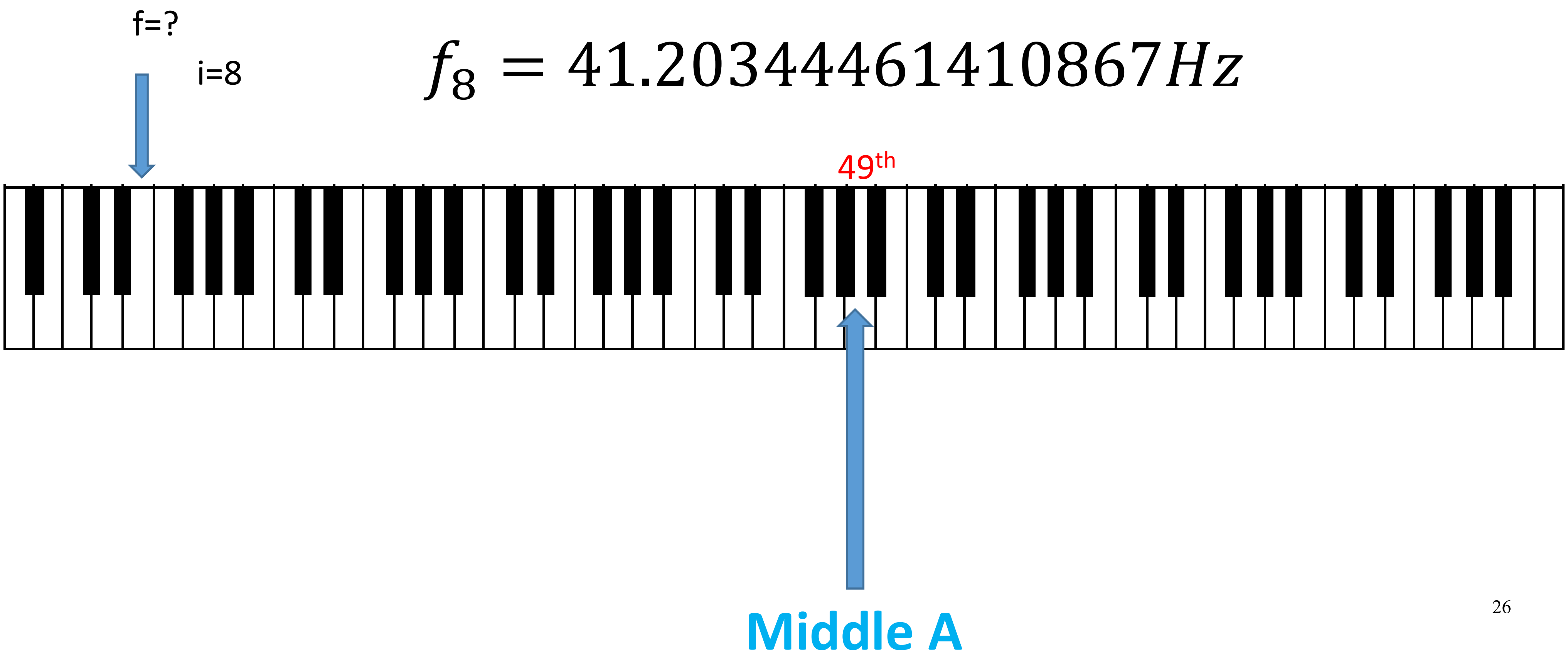
Equal temperament

$$f_i = f_0\left(\sqrt[12]{2}\right)^{i-49}$$

$$f_8 = 440\left(\sqrt[12]{2}\right)^{8-49}$$

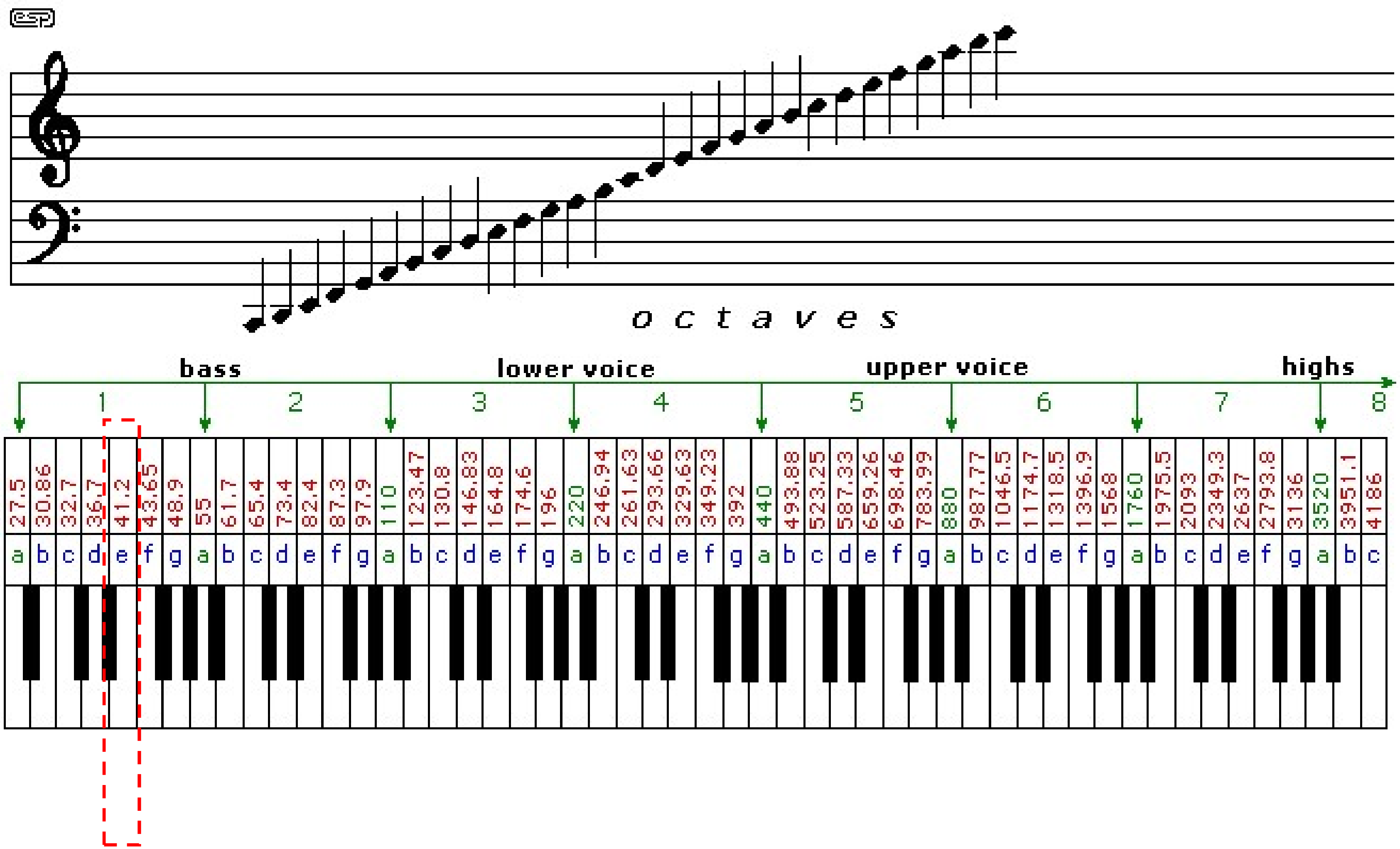
$$f_8 = 440\left(\sqrt[12]{2}\right)^{-41}$$

$$f_8 = 41.20344461410867Hz$$



Read the score and program the music

Equal temperament



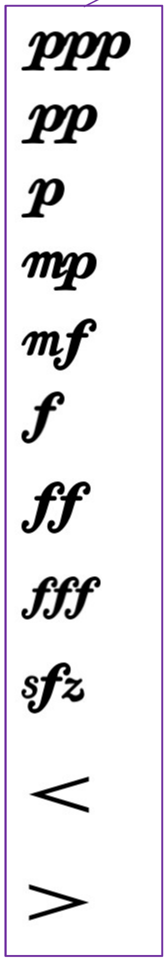
Read the score and program the music

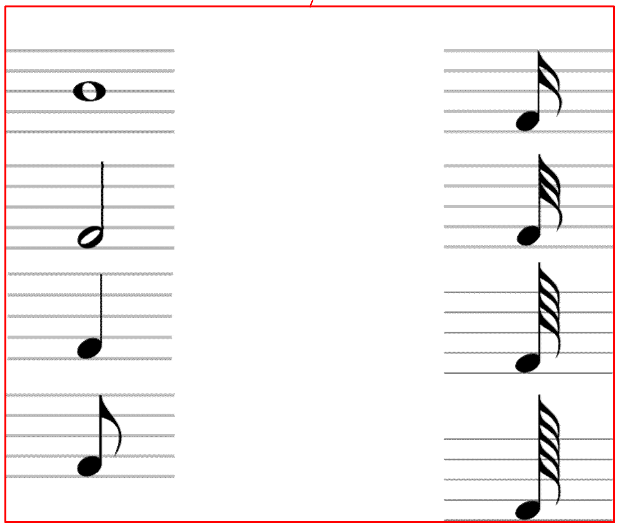
Dynamic of the note

ppp	**Pianississimo**[D 1] Extremely soft. Very infrequently does one see softer dynamics than this, which are specified with additional ***p***s.
pp	**Pianissimo** Very soft. Usually the softest indication in a piece of music, though softer dynamics are often specified with additional ***p***s.
p	**Piano** Soft; louder than *pianissimo*.
mp	**Mezzo piano** Moderately soft; louder than *piano*.
mf	**Mezzo forte** Moderately loud; softer than *forte*. If no dynamic appears, *mezzo-forte* is assumed to be the prevailing dynamic level.
f	**Forte** Loud. Used as often as *piano* to indicate contrast.
ff	**Fortissimo** Very loud. Usually the loudest indication in a piece, though louder dynamics are often specified with additional ***f***s (such as fortississimo – seen below).
fff	**Fortississimo**[D 1] Extremely loud. Very infrequently does one see louder dynamics than this, which are specified with additional ***f***s.
sfz	**Sforzando** Literally "forced", denotes an abrupt, fierce accent on a single sound or chord. When written out in full, it applies to the sequence of sounds or chords under or over which it is placed.
$<$	**Crescendo** A gradual increase in volume. Can be extended under many notes to indicate that the volume steadily increases during the passage.
$>$	**Diminuendo** Also **decrescendo** A gradual decrease in volume. Can be extended in the same manner as crescendo.

Read the score and program the music

$$y(t) = A\sin(t2\pi f + \varphi)$$



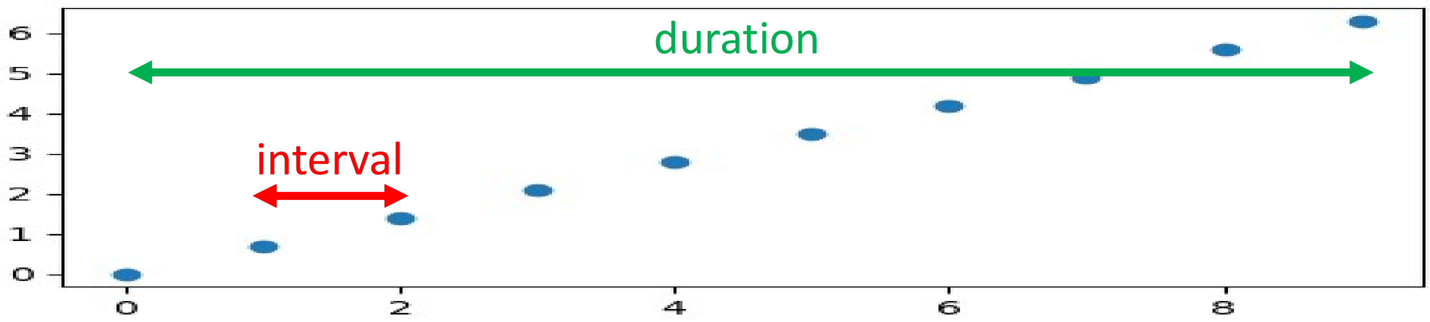


The model of sound

$$y(t) = A\sin(2\pi\, f\, t + \varphi)$$

How to make a sound using this equation

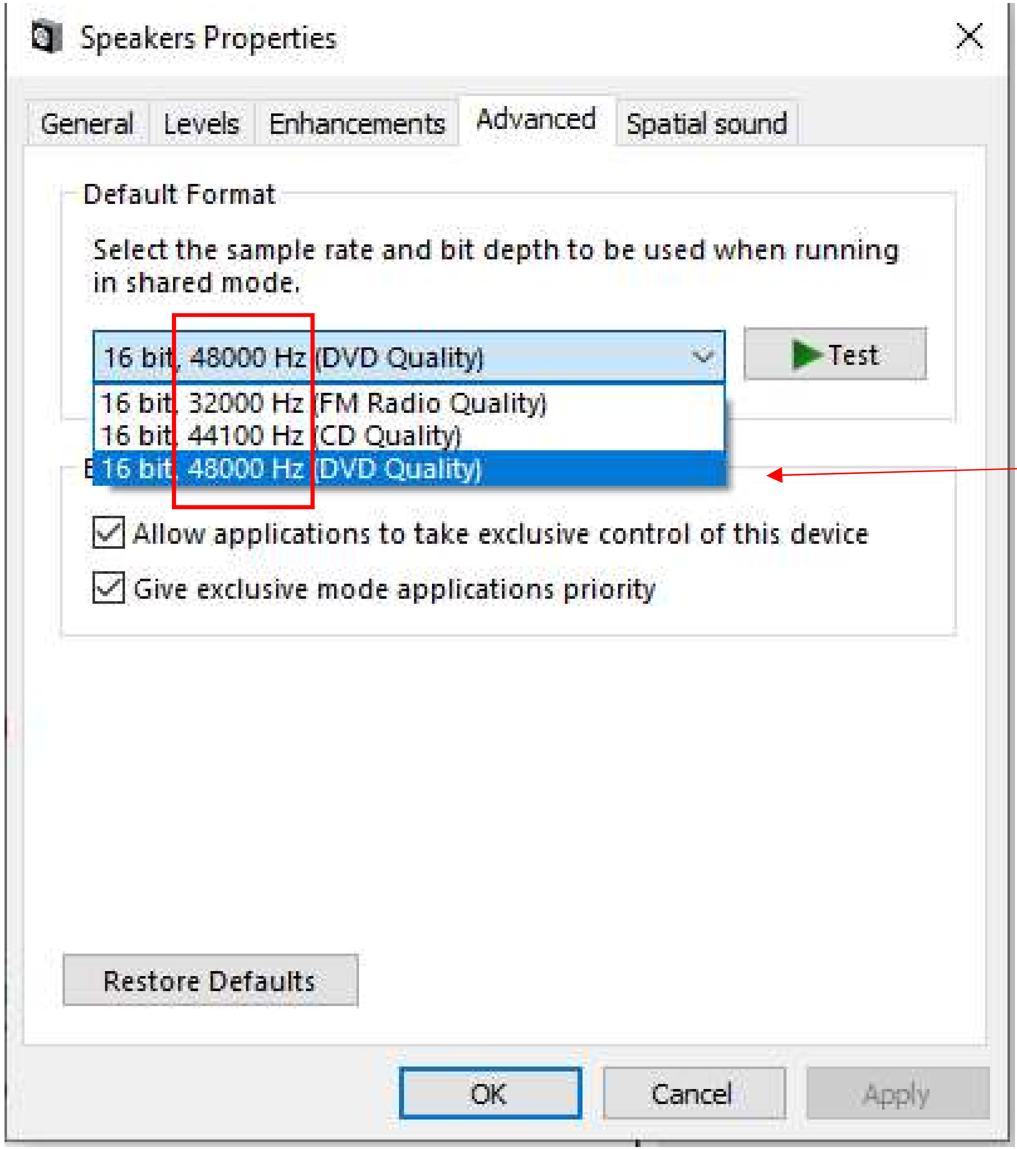
1) Duration of the sound
2) Soundcard playback interval



t=0,...,10

The model of sound

Soundcard playback interval



It plays 48000 samples in one seconds

The 1 seconds of sound requires 48000 samples

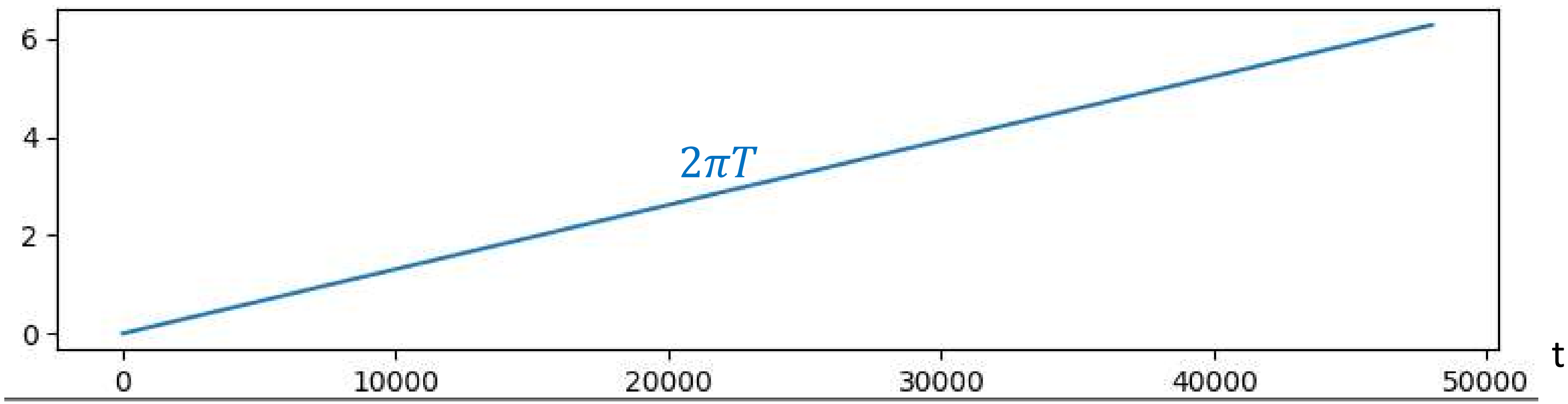
The model of sound

Let's make a 1 seconds of sound!

$$y(t) = A\sin(2\pi\, f\, t + \varphi)$$

1) Spread 2π in 1 seconds

$$Y = A\sin(2\pi\, T\, f)$$




```
n=48000
T=linspace(0,2*pi,n)
plot(T)
```

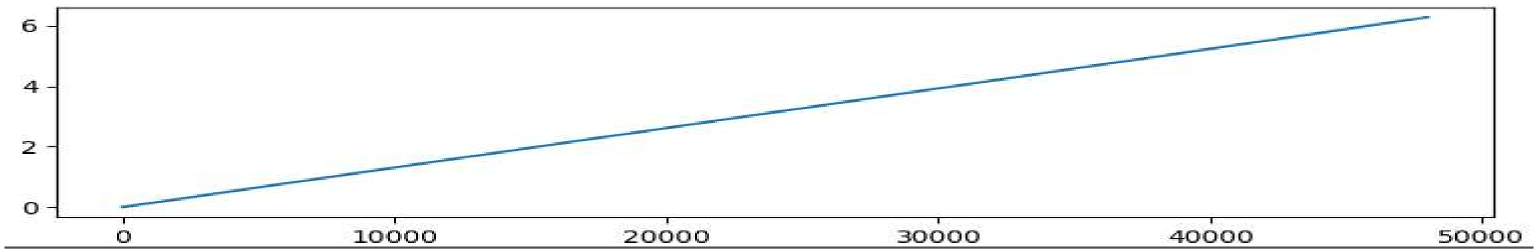
Canon In D

Johann Pachelbel

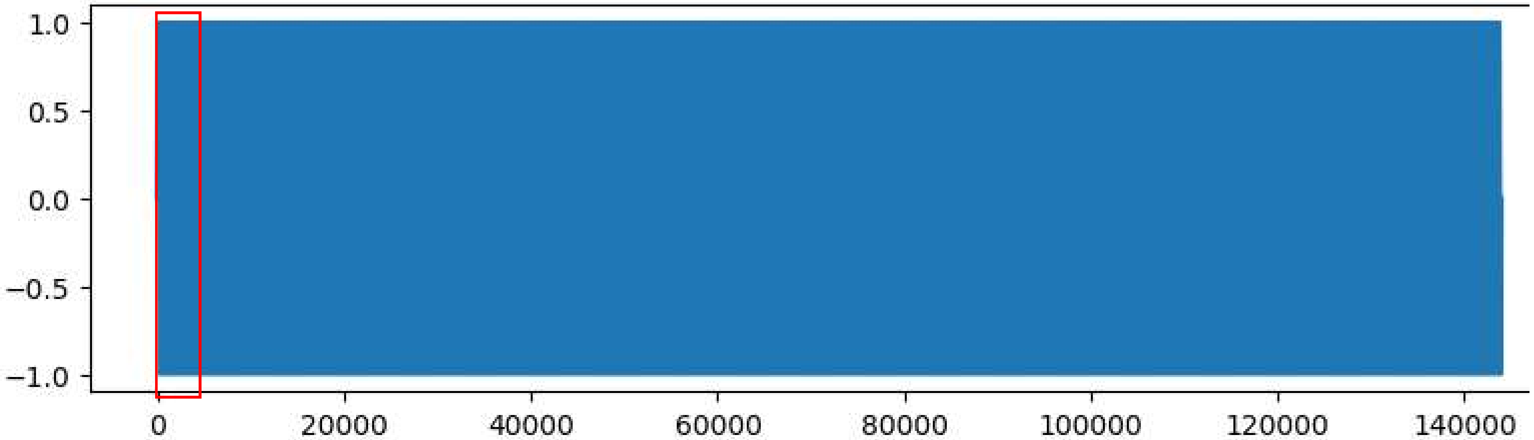
The model of sound

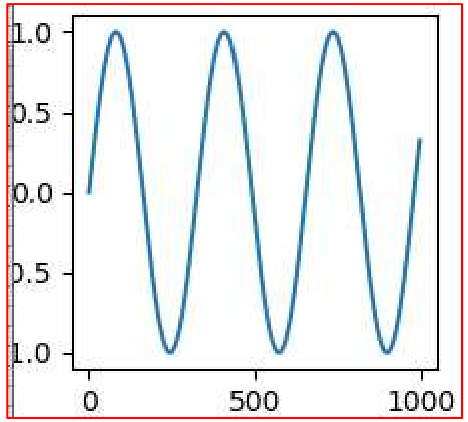
$$Y = A \sin(2\pi\, T\, f)$$

3) Create sound samples



T





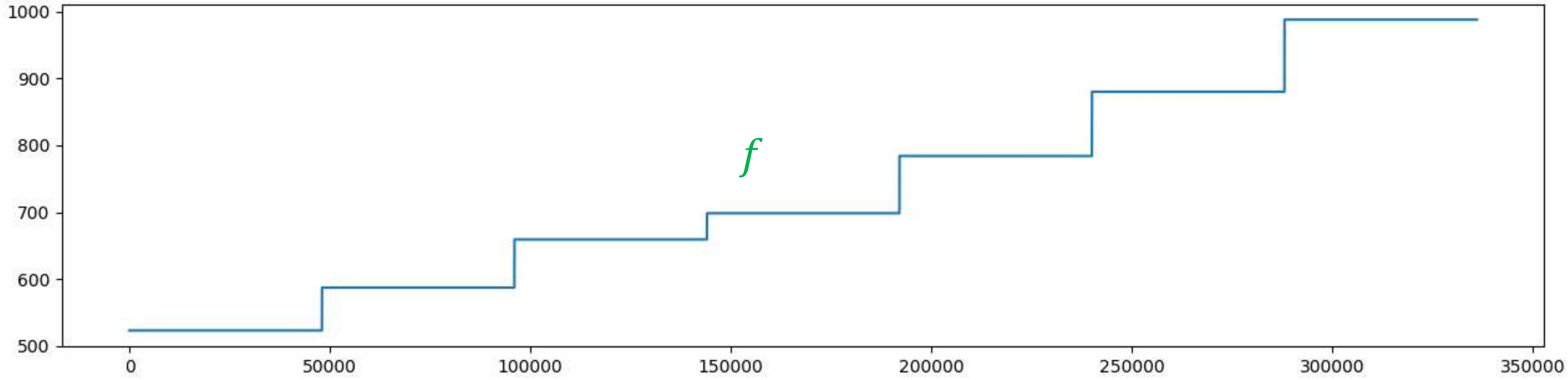
$Y = \sin(440T)$

```
from pylab import *
import sounddevice as sd
import soundfile as sf
n=48000
T=linspace(0,2*pi,n)
Y=sin(440*T)
plot(Y)
sd.play(Y,48000)
```

The model of sound

$$Y = A\sin(2\pi\, T\, f)$$

3) Create sound samples

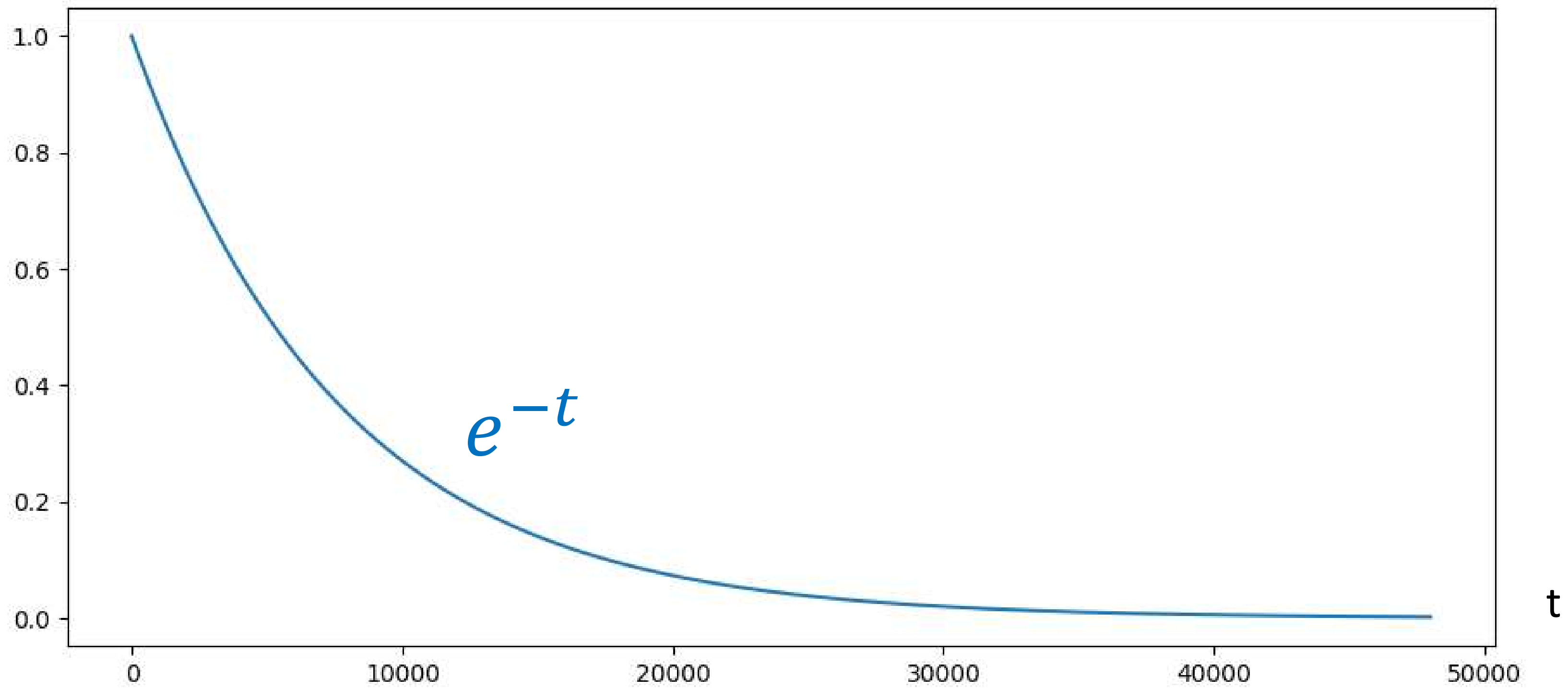


Why this sound unlike musical to us ?

Let ‘s add damping factor ?

The model of sound

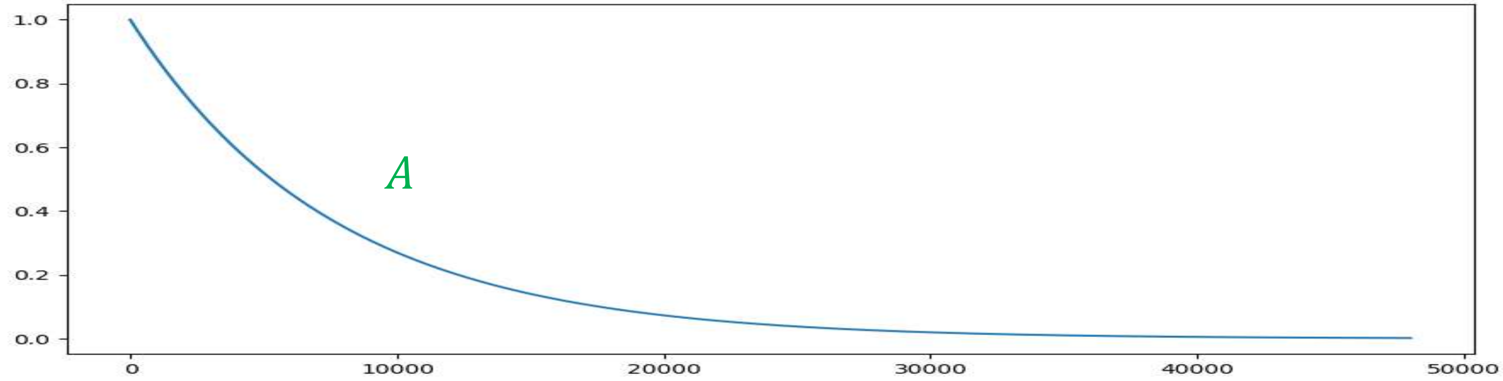
A **damped sine wave** is a sinusoidal function whose amplitude approaches zero as time increases



Damping function

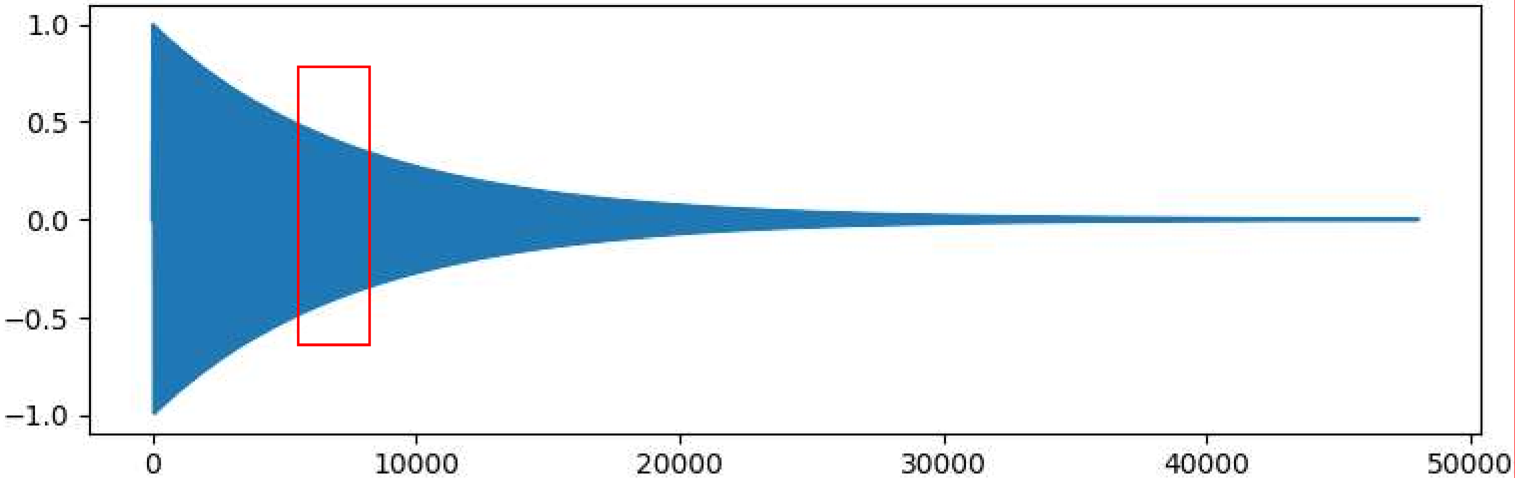
The model of sound

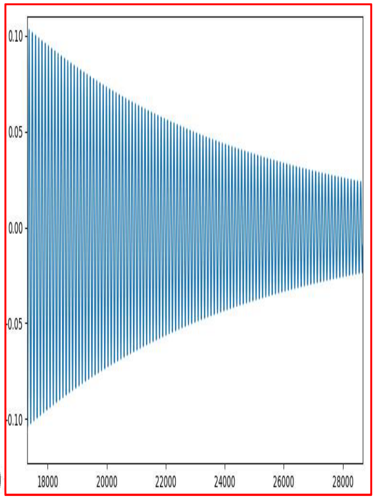
A **damped sine wave** is a sinusoidal function whose amplitude approaches zero as time increases



Damping function

$$Y = A\sin(2\pi T f) \quad ; f = 440$$

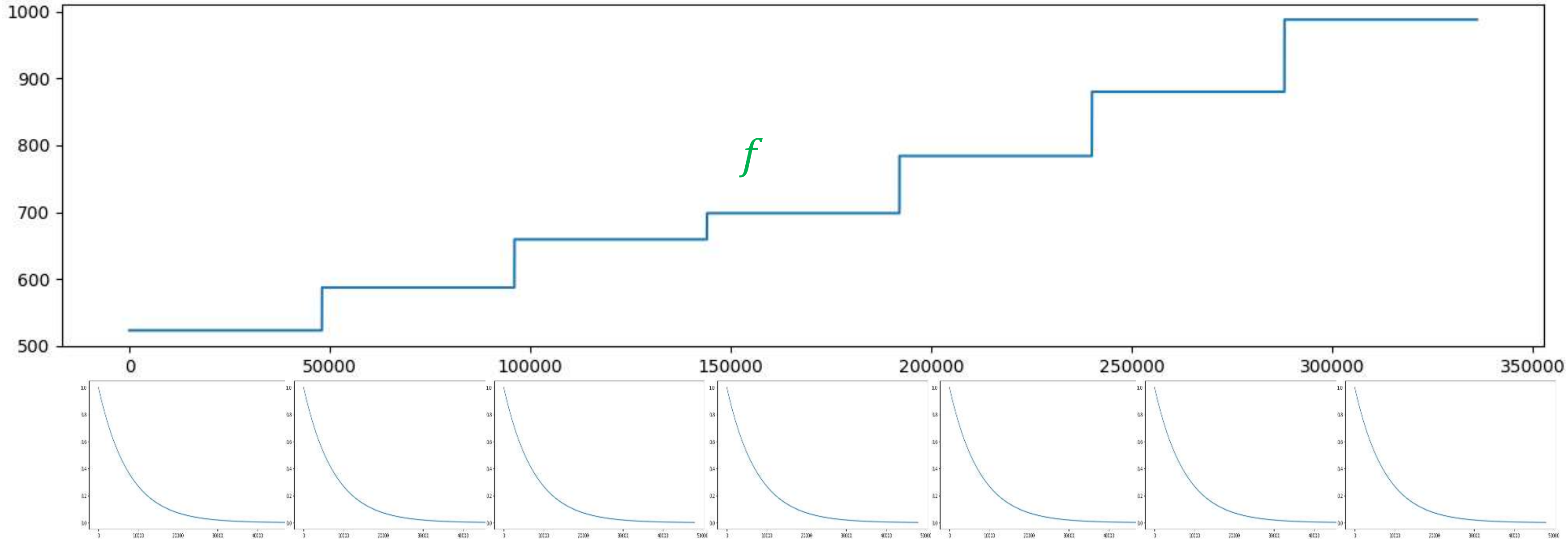


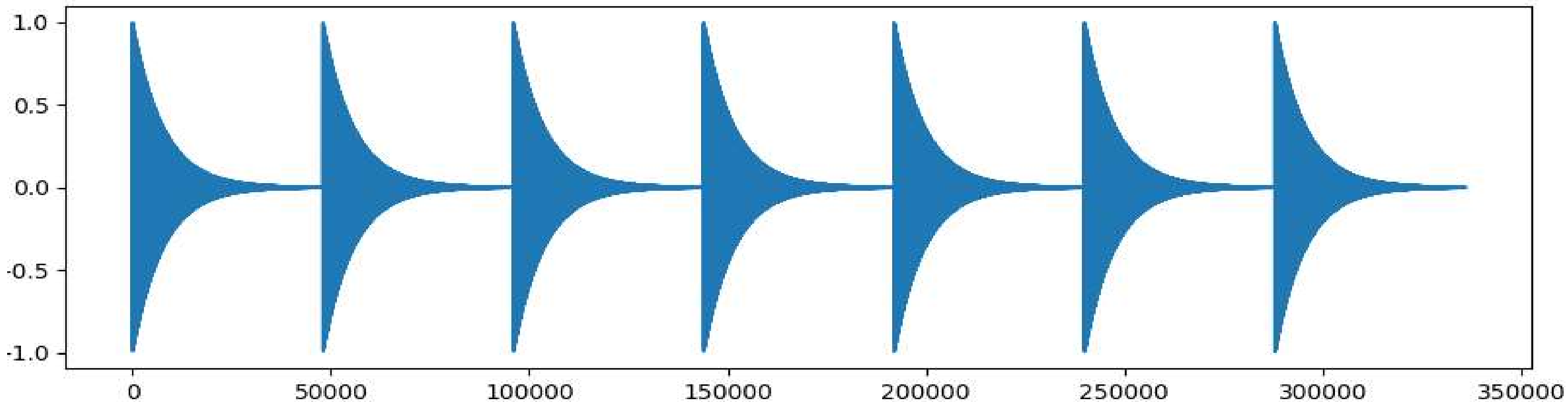


The model of sound

$$Y = A\sin(2\pi\, T\, f)$$

3) Create sound samples



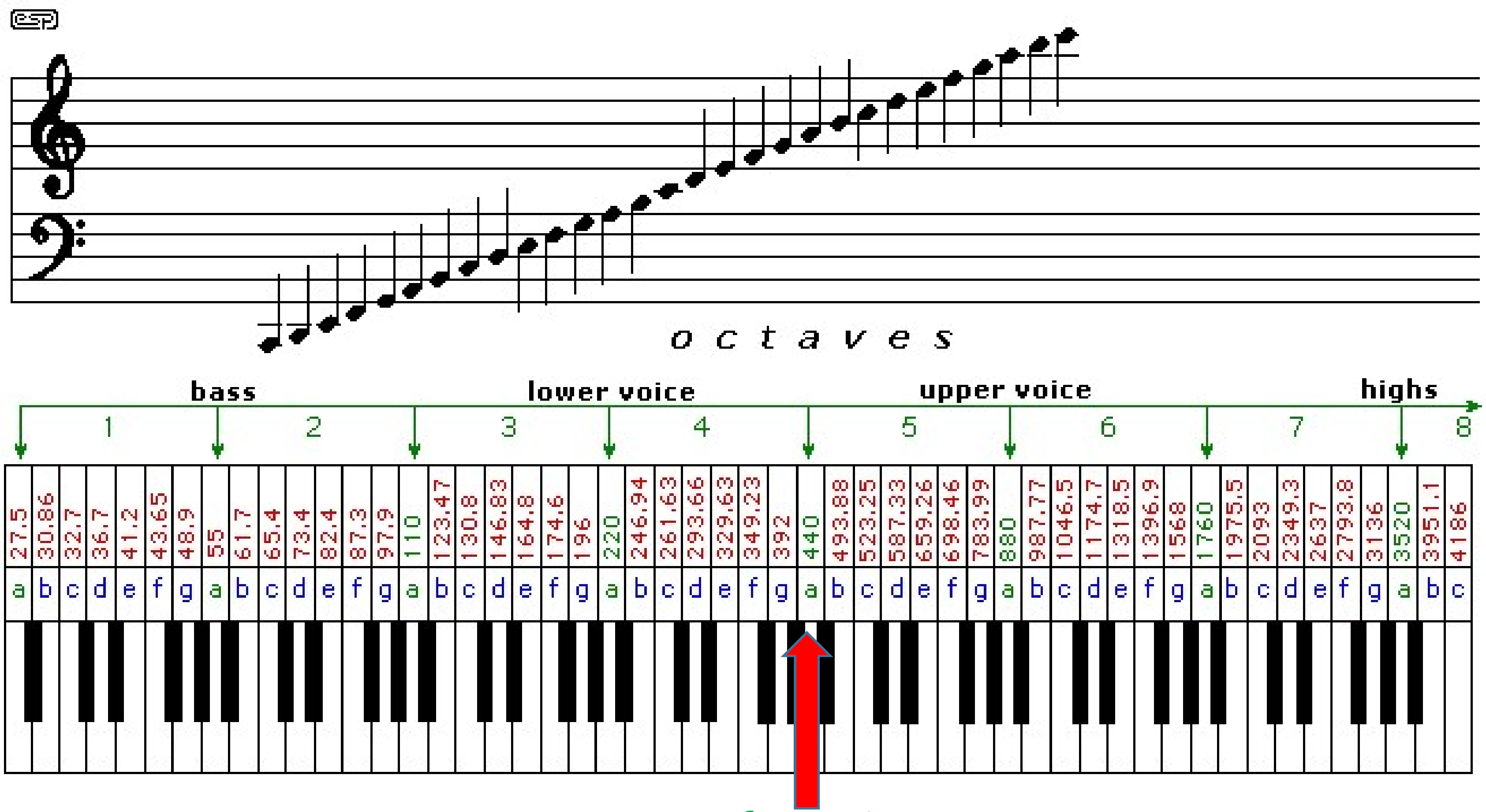


The model of sound

Let’s make a 1 seconds of sound!

$$Y = A\sin(2\pi\, T\, f)$$

2) Pick a frequency



$f = 440$

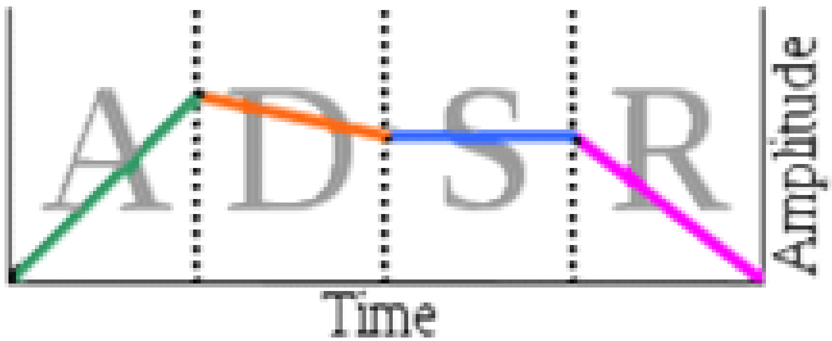
The model of sound

Envelope

In sound and music, an **envelope** describes how a sound changes over time. It normally relates to the amplitude (volume)

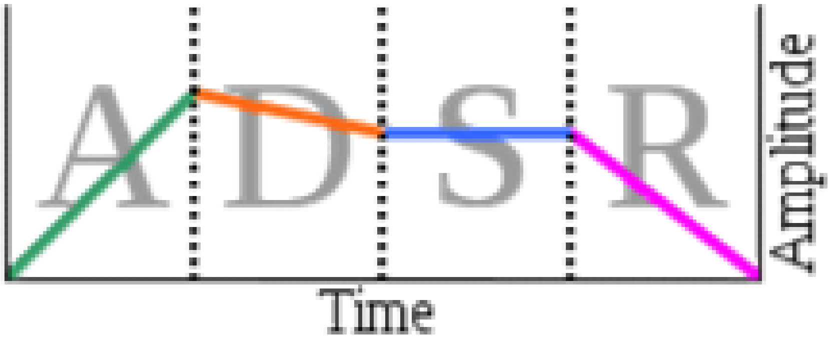
The most common kind of envelope generator has four stages: **attack**, **decay**, **sustain**, and **release** (**ADSR**).[1]

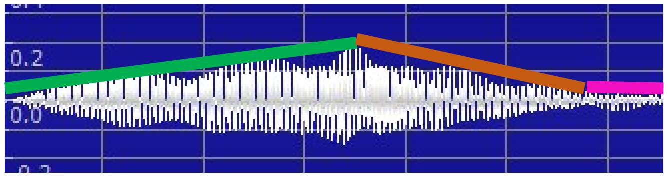
- **Attack** is the time taken for initial run-up of level from nil to peak, beginning when the key is pressed.
- **Decay** is the time taken for the subsequent run down from the attack level to the designated sustain level.
- **Sustain** is the level during the main sequence of the sound's duration, until the key is released.
- **Release** is the time taken for the level to decay from the sustain level to zero after the key is released.



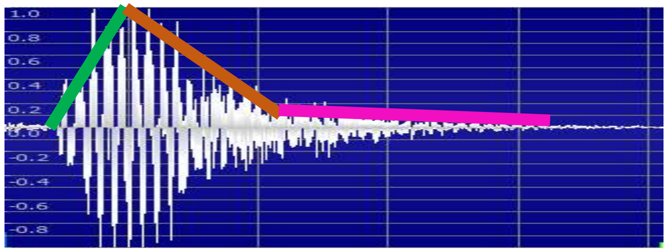
The model of sound

Envelope

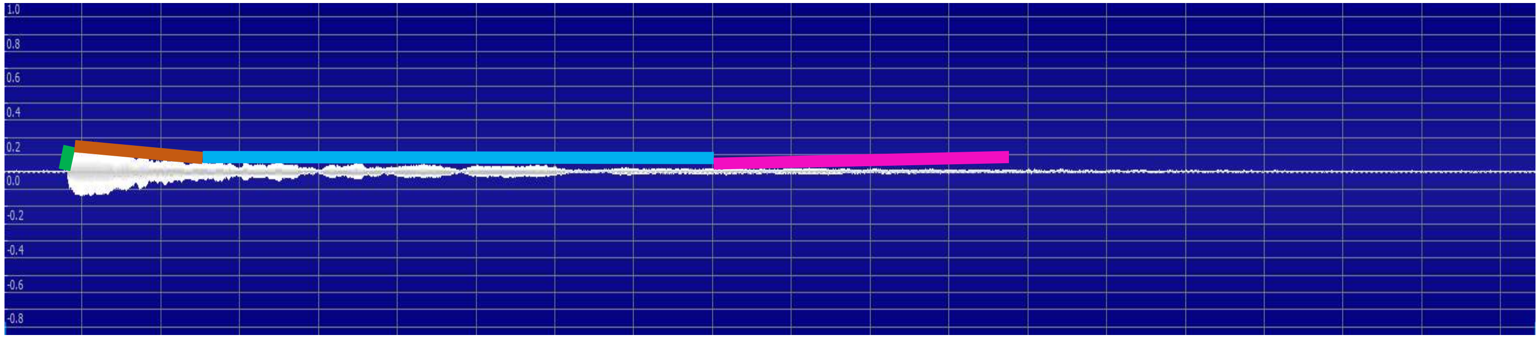




Violin

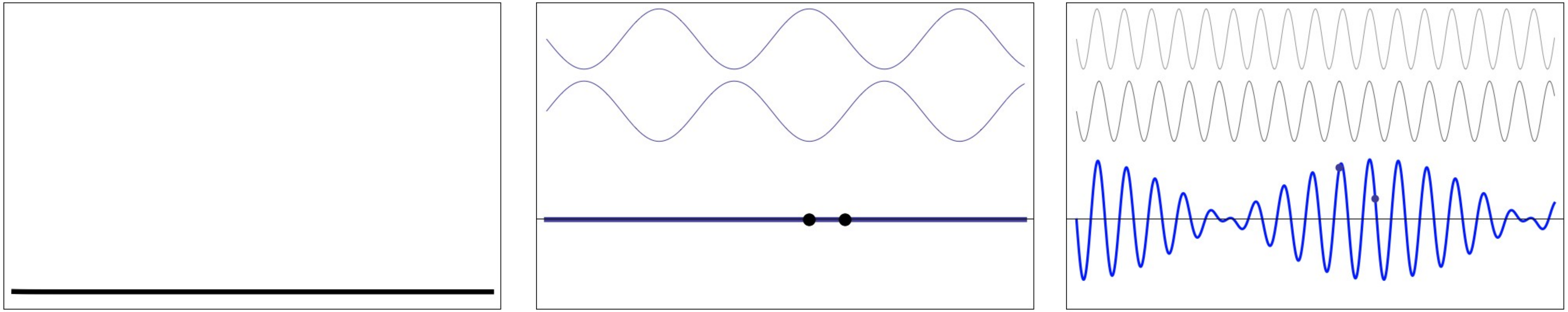


Drum



Bell

The superposition of waves

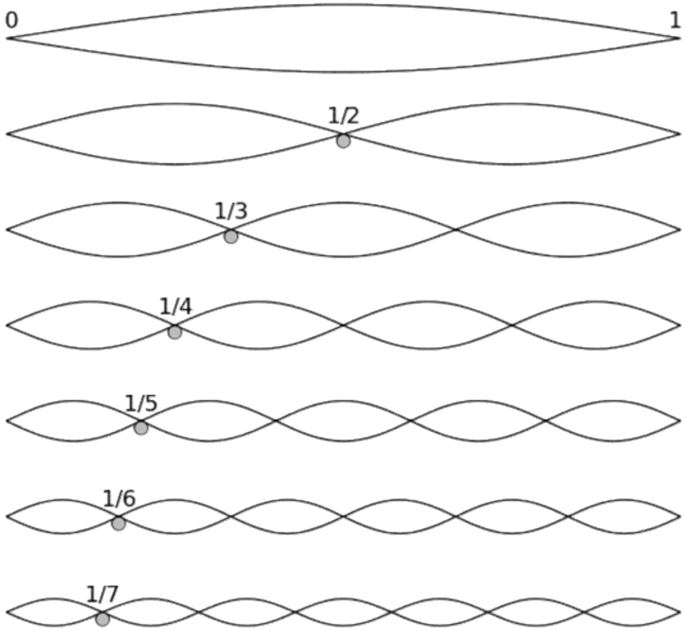


Waves can be added or subtracted

Result in a new waveform

Where the second wave come from ?

The guitar strings



Harmonic series

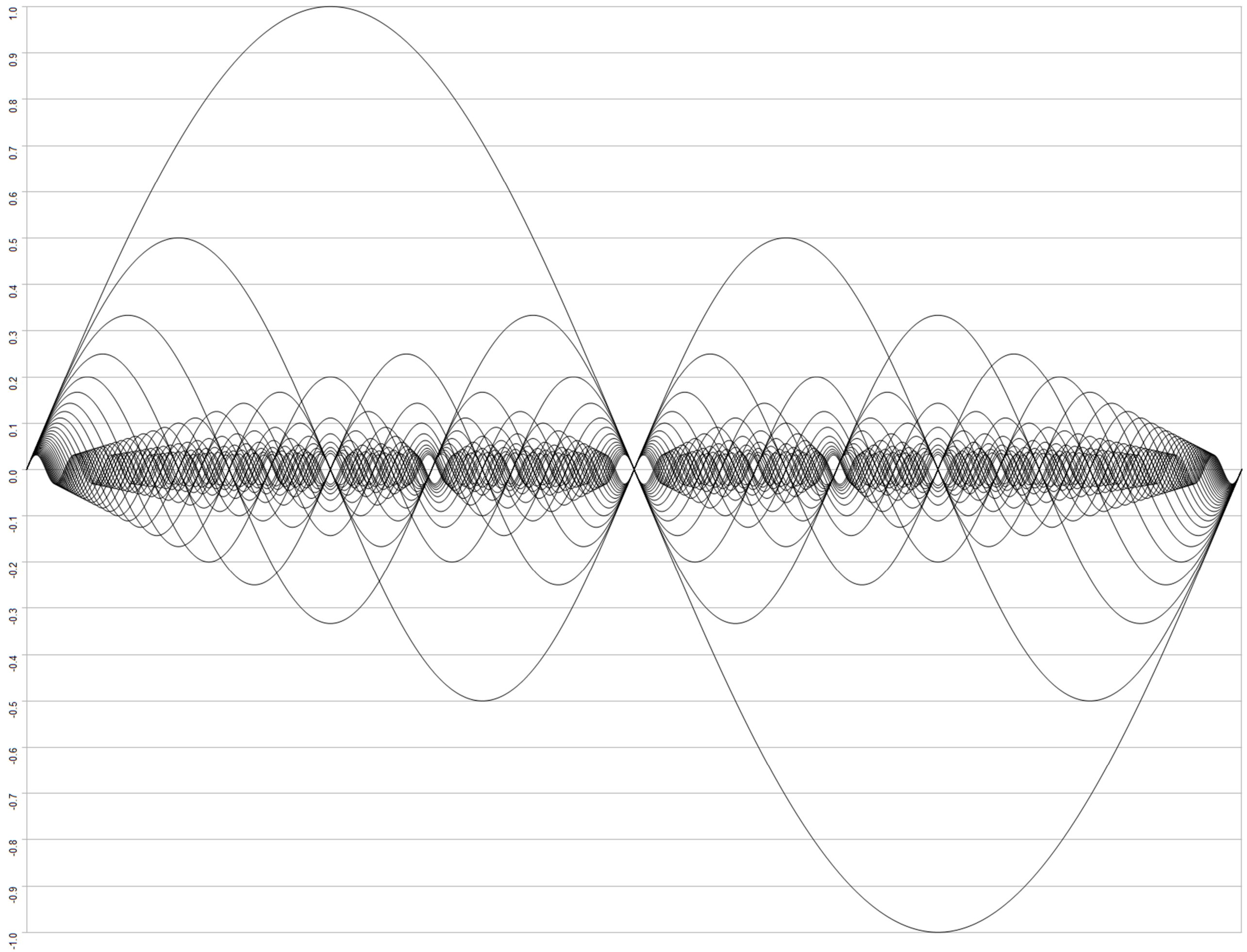
Overtone harmonic

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots$$

$$Y = 1\,(\sin(F_0 T) + \frac{1}{2}\sin(2F_0 T) + \frac{1}{3}\sin(3F_0 T) + \cdots + \frac{1}{n}\sin(nF_0 T)$$

The guitar strings

$$\sum_{n=1}^{\infty}\frac{1}{n}=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\cdots$$

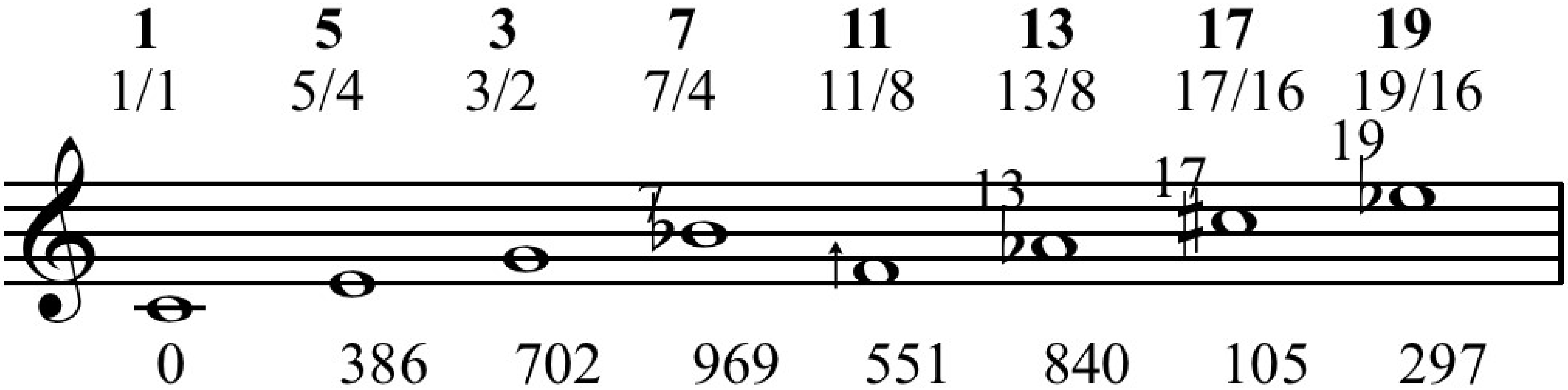


String harmonics to 32

https://en.wikipedia.org/wiki/Harmonic_series_(music)

Harmonic

Non-overtone harmonic

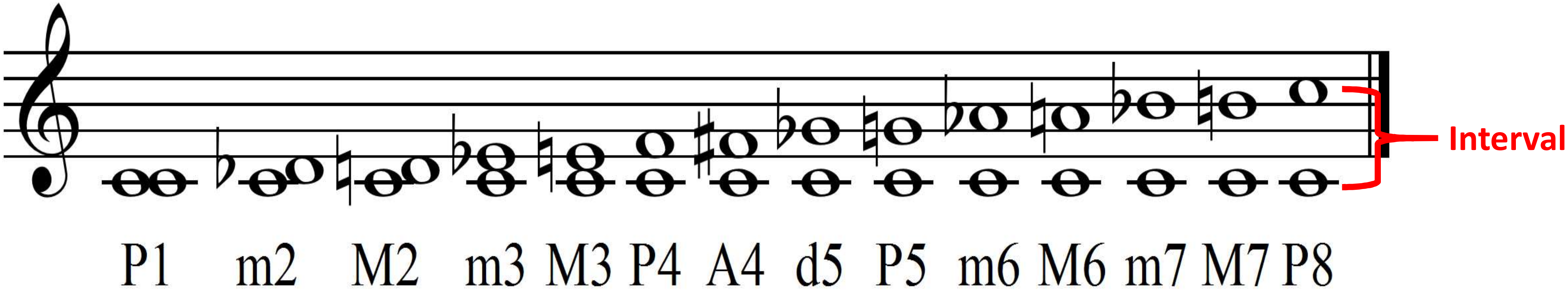


Prime harmonic

https://en.wikipedia.org/wiki/Harmonic_series_(music)

Harmonic

Interval



Number of semitones	Short	Name	5-prime limit tuning (pitch ratio)	Equal temperament
0	P1	Perfect unison	1:1	0
1	m2	Minor second	16:15,27:25	100
2	M2	Major second	9:8,10:9	200
3	m3	Minor third	6:5,32:27	300
4	M3	Major third	5:4	400
5	P4	Perfect fourth	4:3,27:20	500
6	A4	Augmented fourth ,Diminished fifth	45:32,25:18	600
7	P5	Perfect fifth	3:2,40:27	700
8	m6	Minor sixth	8:5	800
9	M6	Major sixth	5:3,27:16	900
10	m7	Minor seventh	16:9,9:5	1000
11	M7	Major seventh	15:8,50:27	1100
12	P8	Perfect octave	2:1	1200

https://en.wikipedia.org/wiki/Interval_(music)

Harmonic

Chord

A **chord**, in music, is any harmonic set of pitches consisting of multiple notes (also called "pitches") that are heard as if sounding simultaneously.[1][2] For many practical and theoretical purposes, arpeggios and broken chords (in which the notes of the chord are sounded one after the other, rather than simultaneously), or sequences of chord tones, may also be considered as chords.

Common types of chords

-Triads

-Seventh chords

-Extended chords

-Altered chords

-Added tone chords

-Suspended chords

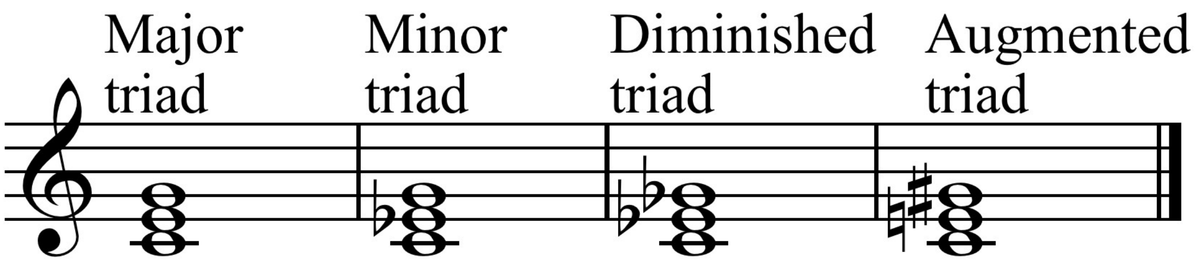
-Borrowed chords

https://en.wikipedia.org/wiki/Chord_(music)

Harmonic

Triads

In music, a **triad** is a set of three notes (or "pitch classes") that can be stacked vertically



Triad	Contains	Interval	Symbol
Major	major third perfect fifth	0-4-7	M
Minor	minor third perfect fifth	0-3-7	m
Diminished	minor third diminished fifth	0-3-6	dim
Augmented	major third and augmented fifth	0-4-8	aug

https://en.wikipedia.org/wiki/Chord_(music)

Harmonic

Triads

Triad	Contains	Interval	Symbol
Major	major third perfect fifth	0-4-7	M
Minor	minor third perfect fifth	0-3-7	m
Diminished	minor third diminished fifth	0-3-6	dim
Augmented	major third and augmented fifth	0-4-8	aug

What are the tone in Dm triad ?

Root = D   Interval= m

Dm= [root , minor third , perfect fifth]

D=146.83Hz , minor third=6:5, perfect fifth = 3:2

Dm= [root , (6/5) root, (3/2) root]

Dm= [146.83 , 176.196, 220.245]

https://en.wikipedia.org/wiki/Chord_(music)

Harmonic

Triads

Triad	Contains	Interval	Symbol
Major	major third perfect fifth	0-4-7	M
Minor	minor third perfect fifth	0-3-7	m
Diminished	minor third diminished fifth	0-3-6	dim
Augmented	major third and augmented fifth	0-4-8	aug

What are the tone in D triad ?

Root = D

Interval= M

Dm= [root , major third , perfect fifth]

D=146.83Hz , major third=5:4, perfect fifth = 3:2

Dm= [root , (5/4) root, (3/2) root]

Dm= [146.83 , 183.537, 220.245]

https://en.wikipedia.org/wiki/Chord_(music)

Harmonic

Seventh chords

A **seventh chord** is a chord consisting of a triad plus a note forming an interval of a seventh above the chord's root.

Name	Contains	Interval	Symbol
Major seventh	major third perfect fifth major seventh	0-4-7- 11	M7
Minor seventh	minor third perfect fifth minor seventh	0-3-7- 10	m7

What are the tone in Dm7 triad ?

Root = D

Interval= m7

Dm= [root , major third , perfect fifth, minor seventh]

D=146.83Hz , major third=5:4, perfect fifth = 3:2, minor seventh= 16:9

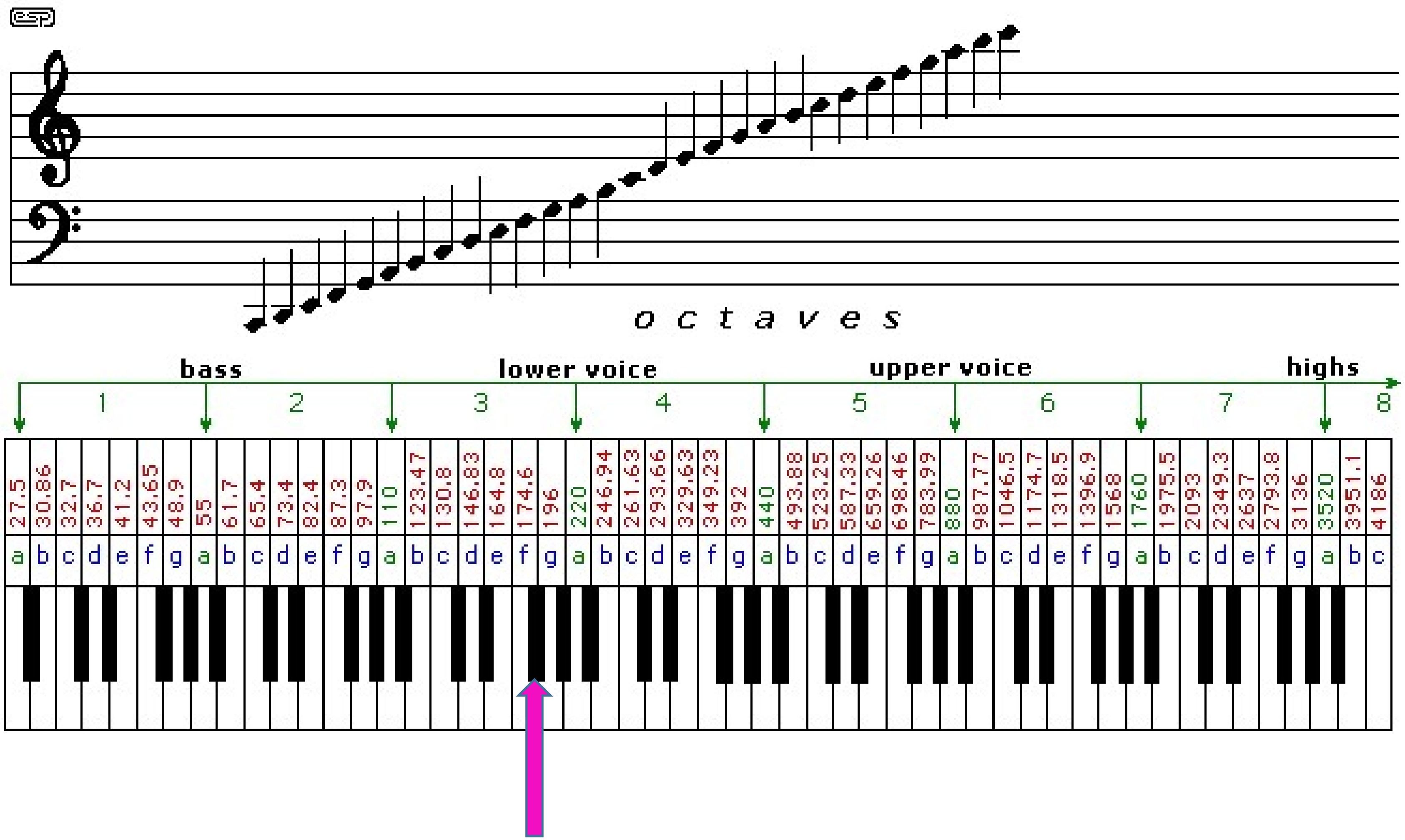
Dm= [root , (5/4) root, (3/2) root, (16/9) root]

Dm= [146.83 , 183.537, 220.245, 261.03]

https://en.wikipedia.org/wiki/Chord_(music)

Harmonic

Dm= [146.83 , 183.537, 220.245, 261.03]



https://en.wikipedia.org/wiki/Chord_(music)

Harmonic

Triads

Triad	Contains	Interval	Symbol
Major	major third perfect fifth	0-4-7	M
Minor	minor third perfect fifth	0-3-7	m
Diminished	minor third diminished fifth	0-3-6	dim
Augmented	major third and augmented fifth	0-4-8	aug

What are the tone in Dm triad ?

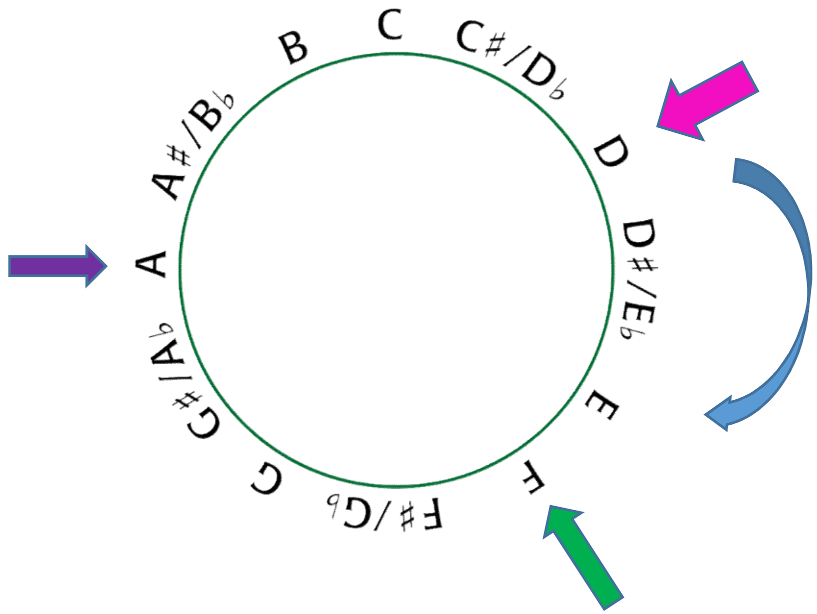
Root = D

Interval= m

Dm= [+0 , +3, +7]

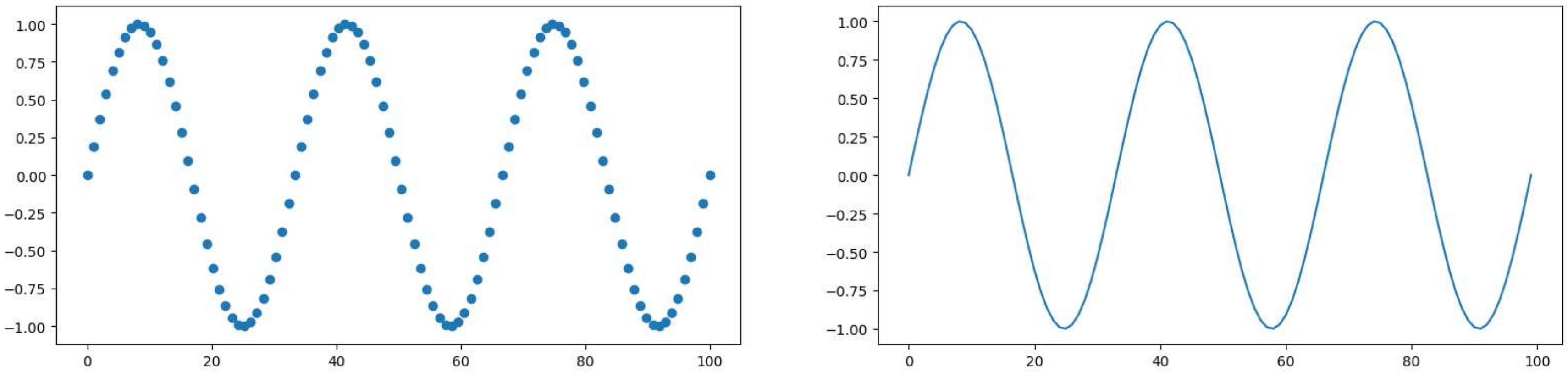
Dm= [D, F, A]

Dm= [146.83, 174.6, 220]

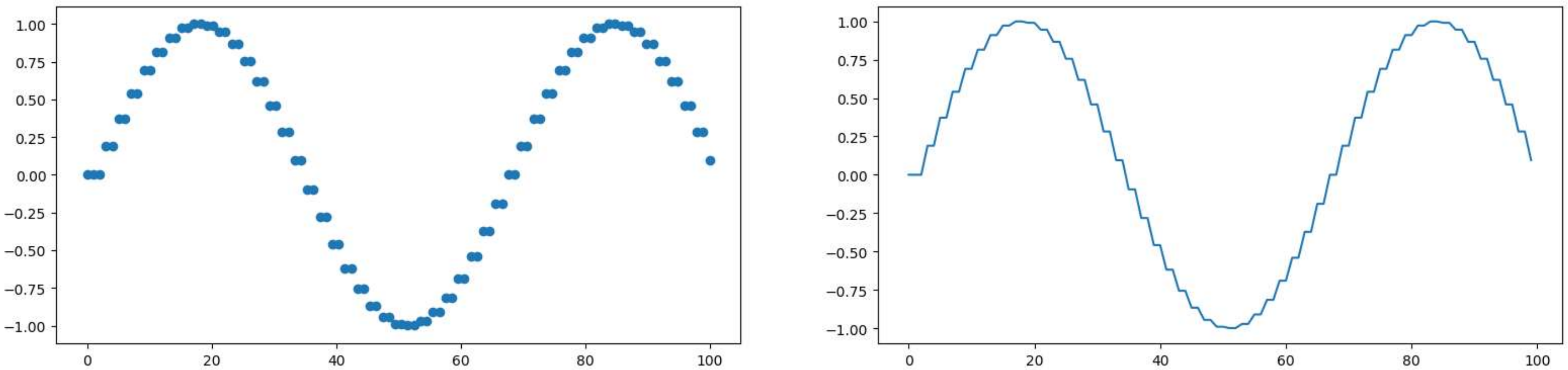


https://en.wikipedia.org/wiki/Chord_(music)

Wavetable synthesis

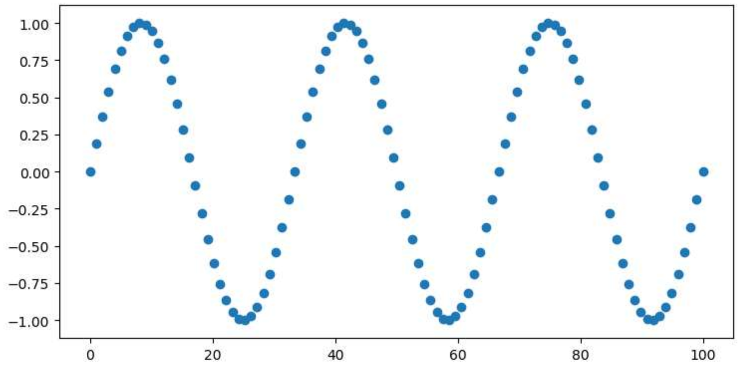


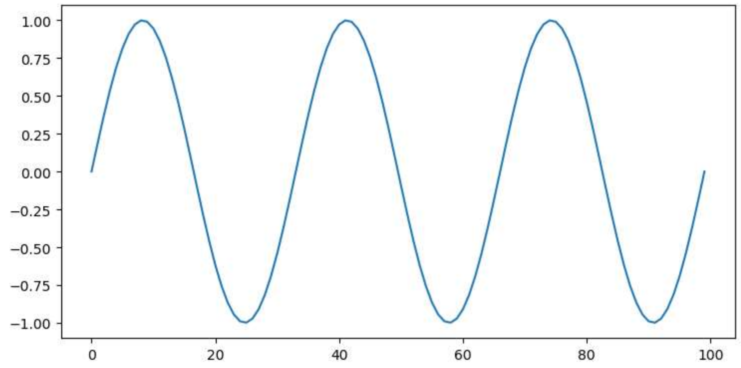
Original Waveform



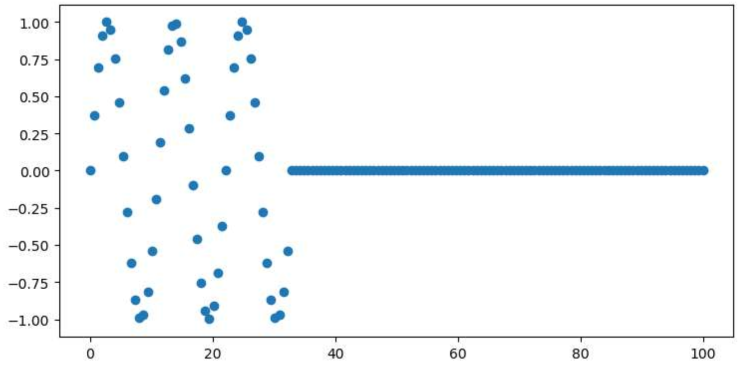
Play sample twice

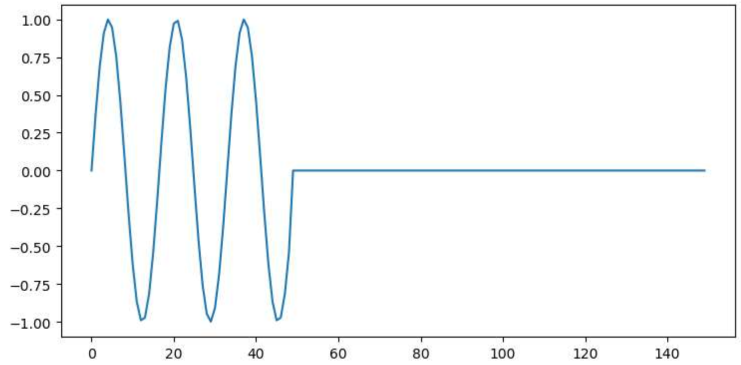
Wavetable synthesis



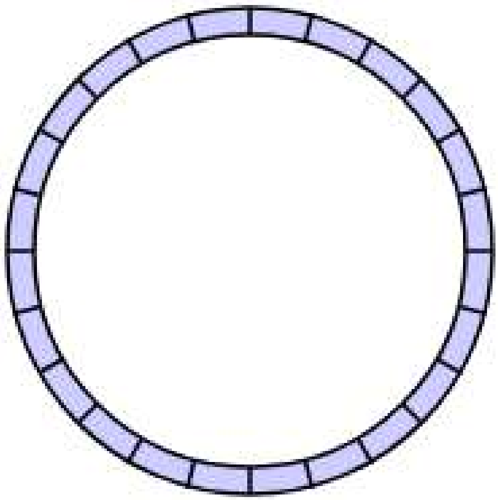


Original Waveform





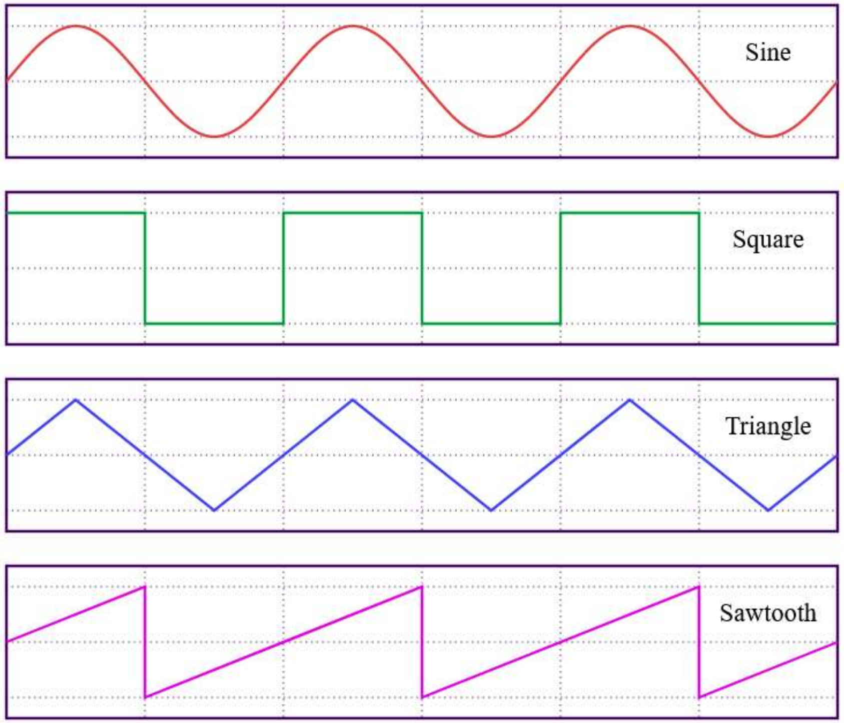
One sample Skipped



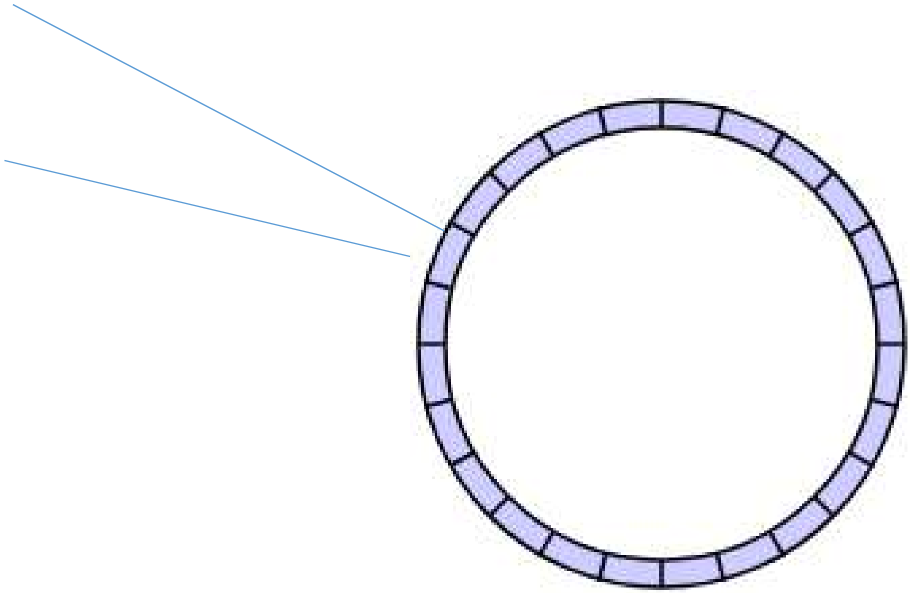
circular buffer

Wavetable synthesis

Table-lookup synthesis[9] (or **Wavetable-lookup synthesis**[10]) (Roads 1996) is a class of sound synthesis methods using the waveform tables by table-lookup, called "table-lookup oscillator" technique. The length of waveforms or samples may be varied by each sound synthesis method, from a single-cycle up to several minutes.



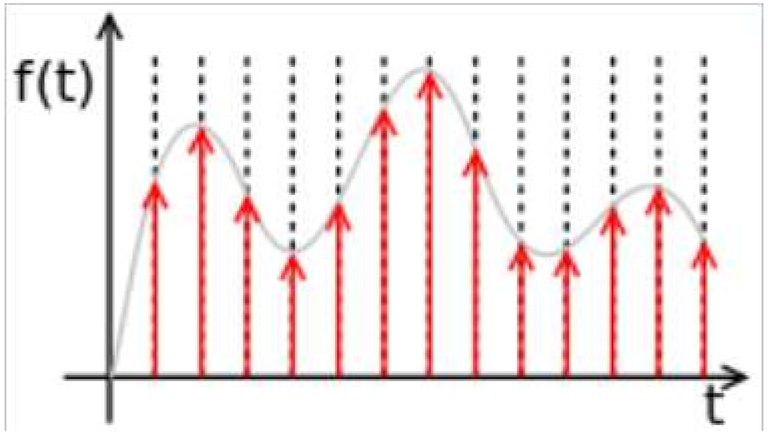
waveform



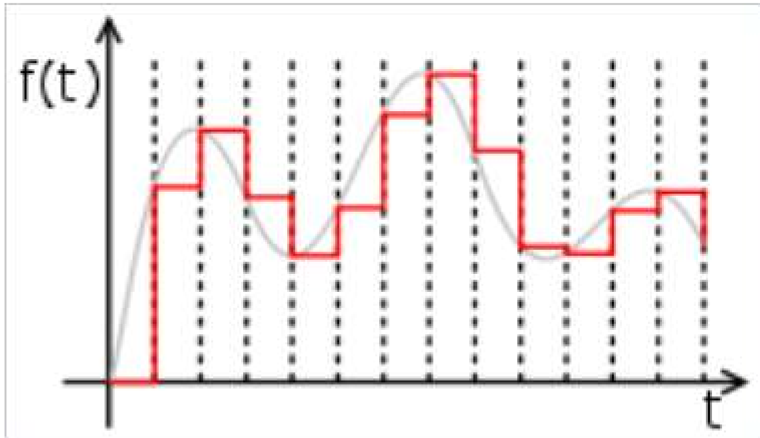
circular buffer

Digital-to-analog converter

A DAC converts an abstract finite-precision number (usually a fixed-point binary number) into a physical quantity (e.g., a voltage or a pressure). In particular, DACs are often used to convert finite-precision time series data to a continually varying physical signal.



Ideally sampled signal.



Piecewise constant output of a conventional DAC lacking a reconstruction filter.

https://en.wikipedia.org/wiki/Digital-to-analog_converter

Digital-to-analog converter

Types

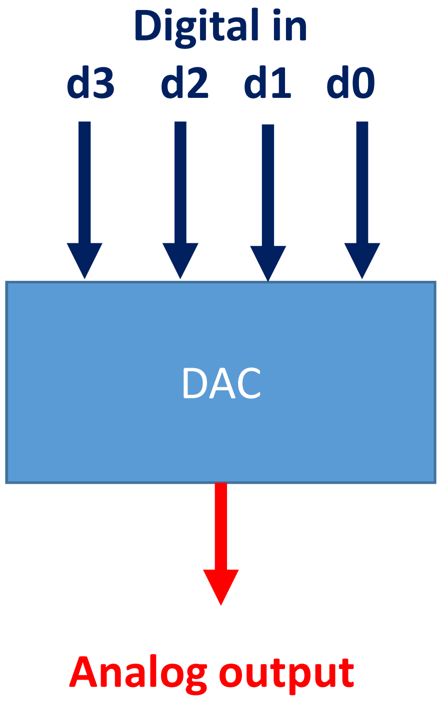
- The binary-weighted DAC
- Pulse-width modulator
- Oversampling DACs or interpolating DAC

https://en.wikipedia.org/wiki/Digital-to-analog_converter

Digital-to-analog converter

The binary-weighted DAC

DEC	HEX	d3	d2	d1	d0
0	0	0	0	0	0
1	1	0	0	0	1
2	2	0	0	1	0
3	3	0	0	1	1
4	4	0	1	0	0
5	5	0	1	0	1
6	6	0	1	1	0
7	7	0	1	1	1
8	8	1	0	0	0
9	9	1	0	0	1
10	A	1	0	1	0
11	B	1	0	1	1
12	C	1	1	0	0
13	D	1	1	0	1
14	E	1	1	1	0
15	F	1	1	1	1



Digital-to-analog converter

The binary-weighted DAC

DEC	HEX	d3	d2	d1	d0
0	0	0	0	0	0
1	1	0	0	0	1
2	2	0	0	1	0
3	3	0	0	1	1
4	4	0	1	0	0
5	5	0	1	0	1
6	6	0	1	1	0
7	7	0	1	1	1
8	8	1	0	0	0
9	9	1	0	0	1
10	A	1	0	1	0
11	B	1	0	1	1
12	C	1	1	0	0
13	D	1	1	0	1
14	E	1	1	1	0
15	F	1	1	1	1

8, 4, 2, 1

$$v = 8d_3 + 4d_2 + 2d_1 + 1d_0$$

$$v = d_3 + \frac{1}{2}d_2 + \frac{1}{4}d_1 + \frac{1}{8}d_0$$

$$v = \sum_{i=0}^{k} 2^i d_i$$

Digital-to-analog converter

The binary-weighted DAC

$$v = \sum_{i=0}^{k} 2^i d_i$$

What is the value of "1 1 0 1" ?

$$v = 8d_3 + 4d_2 + 2d_1 + 1d_0$$

$$v = 8(1) + 4(1) + 2(0) + 1(1)$$

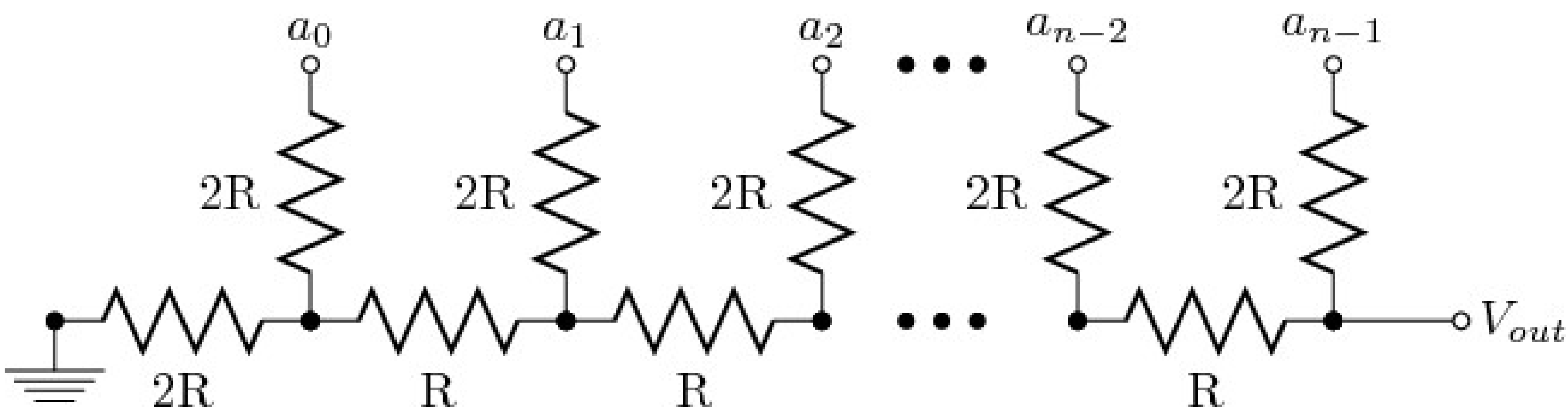
$$v = 8 + 4 + 0 + 1$$

$$v = 13$$

Digital-to-analog converter

The binary-weighted DAC

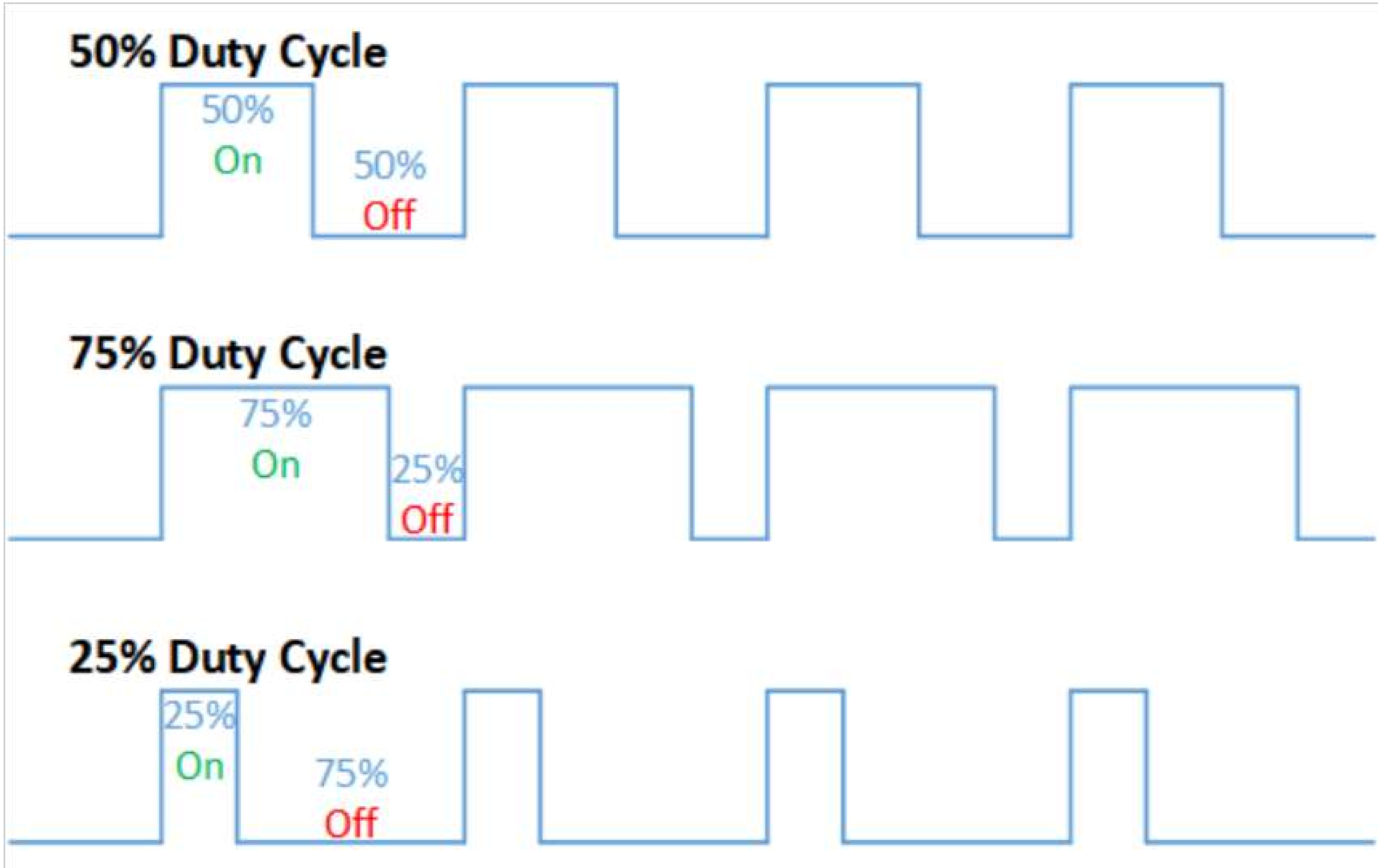
$$v = \sum_{i=0}^{k} 2^i d_i$$



$$V_{out} = V_r \frac{Value}{2^N}$$

Digital-to-analog converter

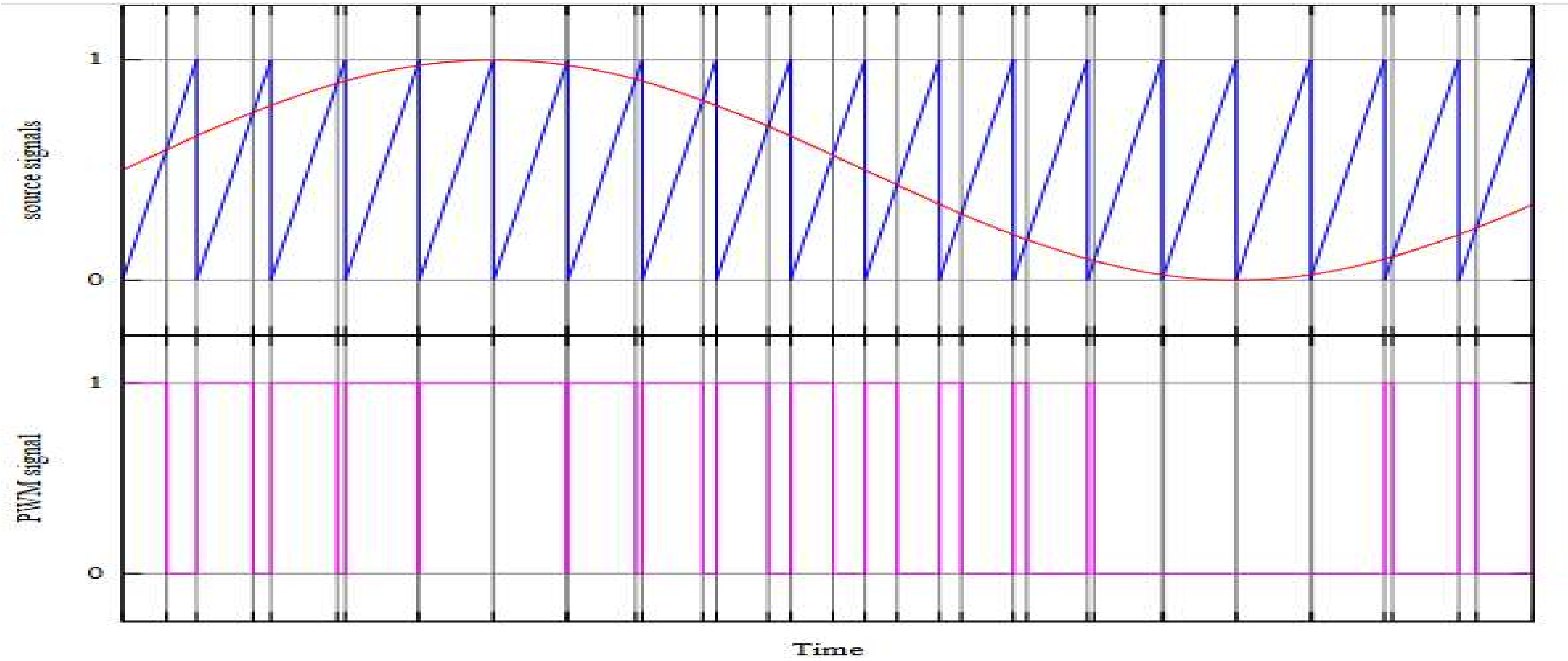
Pulse-width modulator



$$Y = \frac{1}{T}\int_0^T f(t)d(t)$$

Digital-to-analog converter

Pulse-width modulator

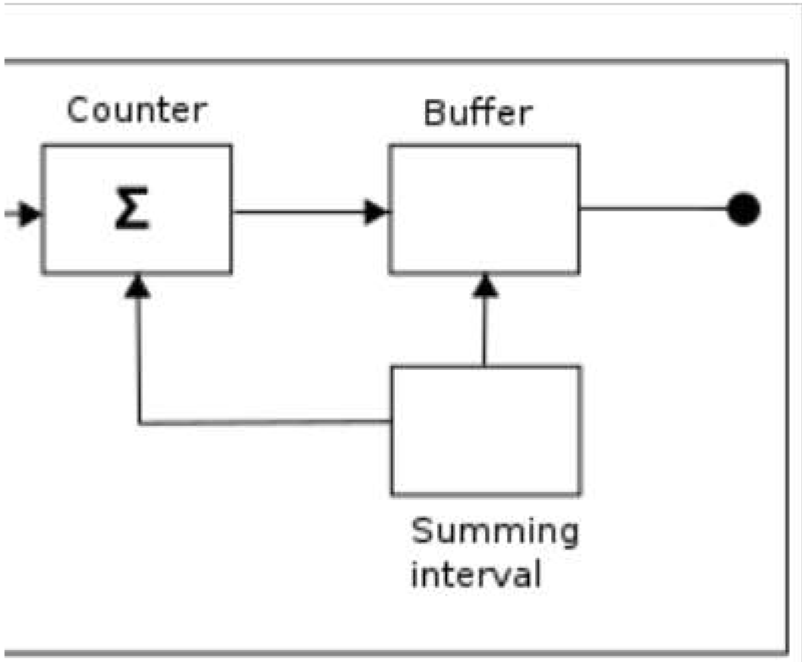


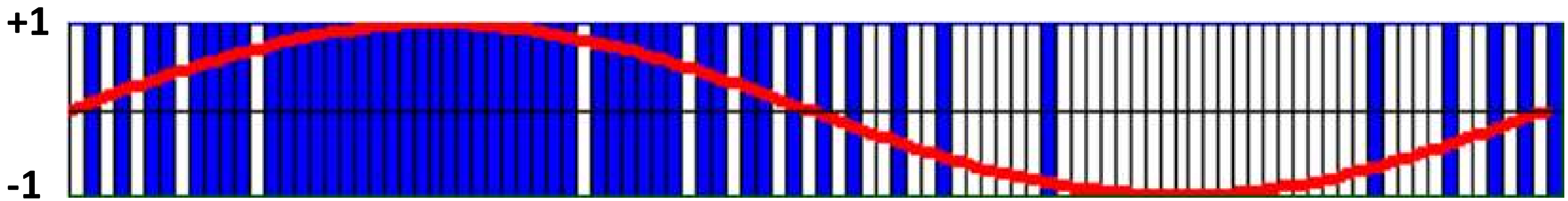
$$Y = \frac{1}{T}\int_0^T f(t)d(t)$$ ← Mean

$$Y = f * G$$ ← Mean filter (G is Gaussian)

Digital-to-analog converter

Oversampling DACs or interpolating DAC



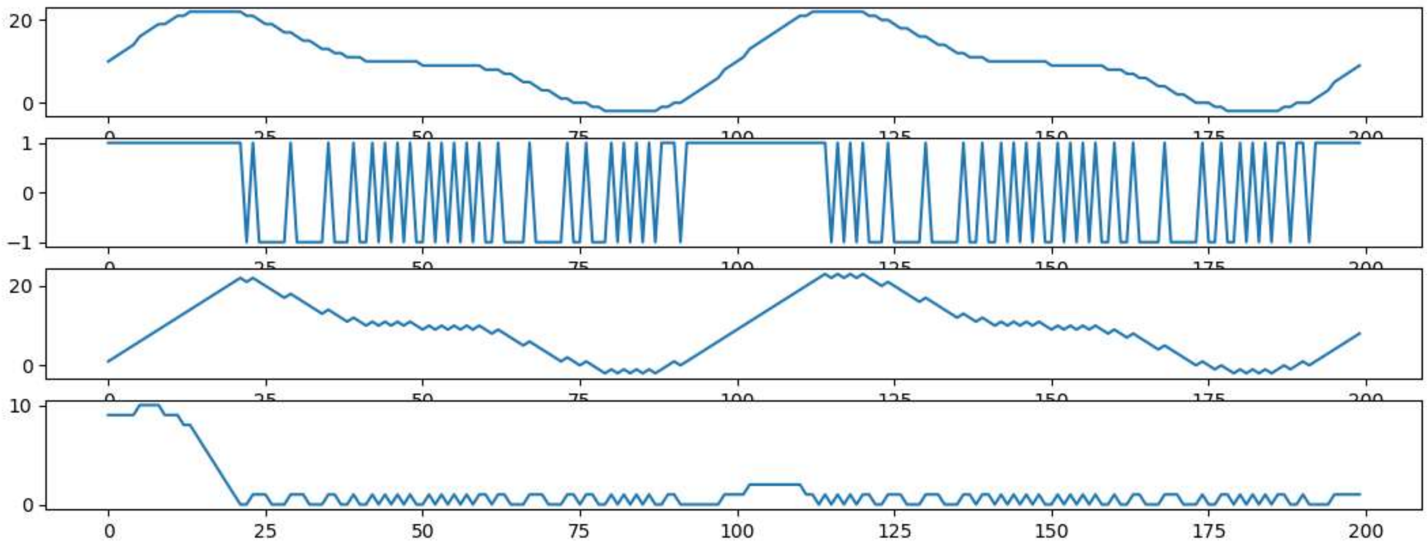


$$Y(t) = \sum_{0}^{t} D(t)$$

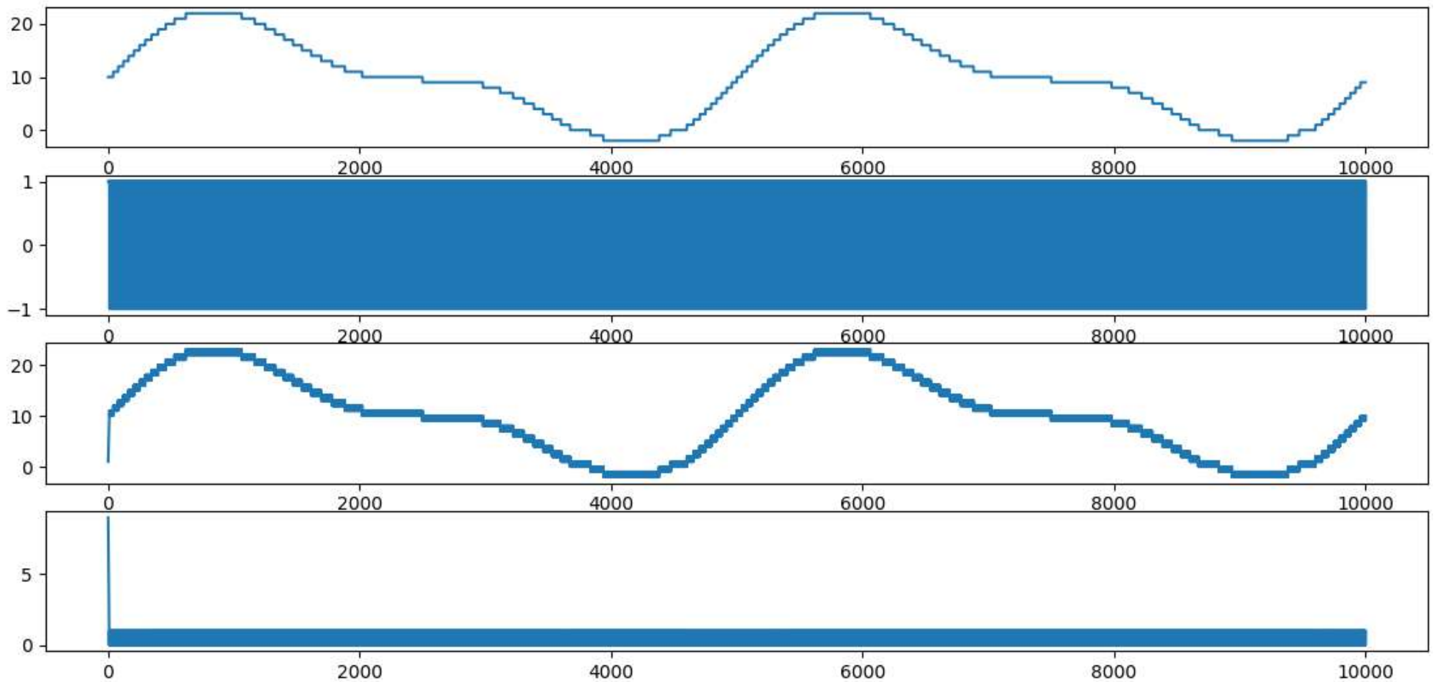
Cumulative sum

Digital-to-analog converter

Oversampling DACs or interpolating DAC



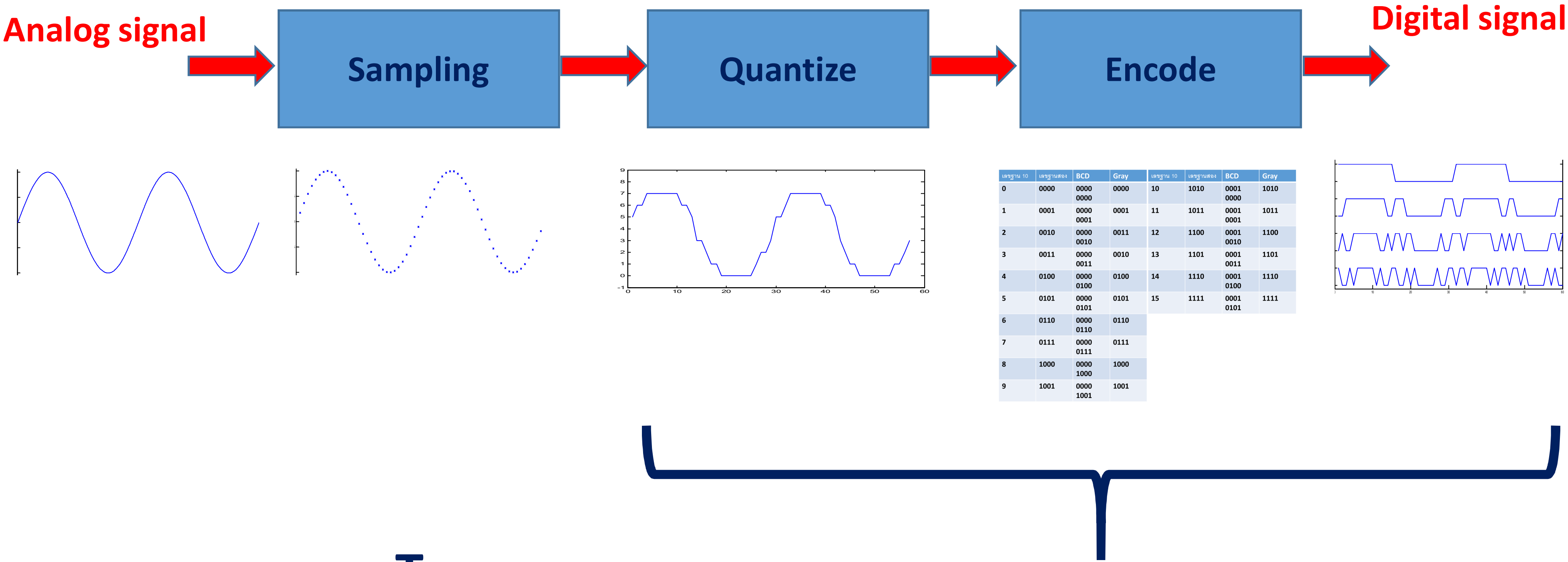
Under sampling



Over sampling

Analog-to-digital converter

An ADC converts a continuous-time and continuous-amplitude analog signal to a discrete-time and discrete-amplitude digital signal.



เลขฐาน 10	เลขฐานสอง	BCD	Gray	เลขฐาน 10	เลขฐานสอง	BCD	Gray
0	0000	0000 0000	0000	10	1010	0001 0000	1010
1	0001	0000 0001	0001	11	1011	0001 0001	1011
2	0010	0000 0010	0011	12	1100	0001 0010	1100
3	0011	0000 0011	0010	13	1101	0001 0011	1101
4	0100	0000 0100	0100	14	1110	0001 0100	1110
5	0101	0000 0101	0101	15	1111	0001 0101	1111
6	0110	0000 0110	0110				
7	0111	0000 0111	0111				
8	1000	0000 1000	1000				
9	1001	0000 1001	1001				

Types

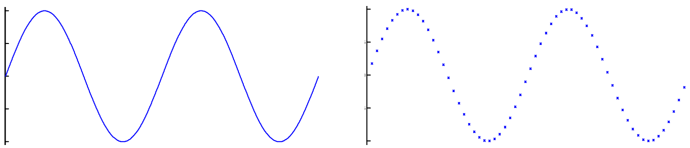
- A direct-conversion ADC or flash ADC

- Counter type ADC

- A sigma-delta ADC

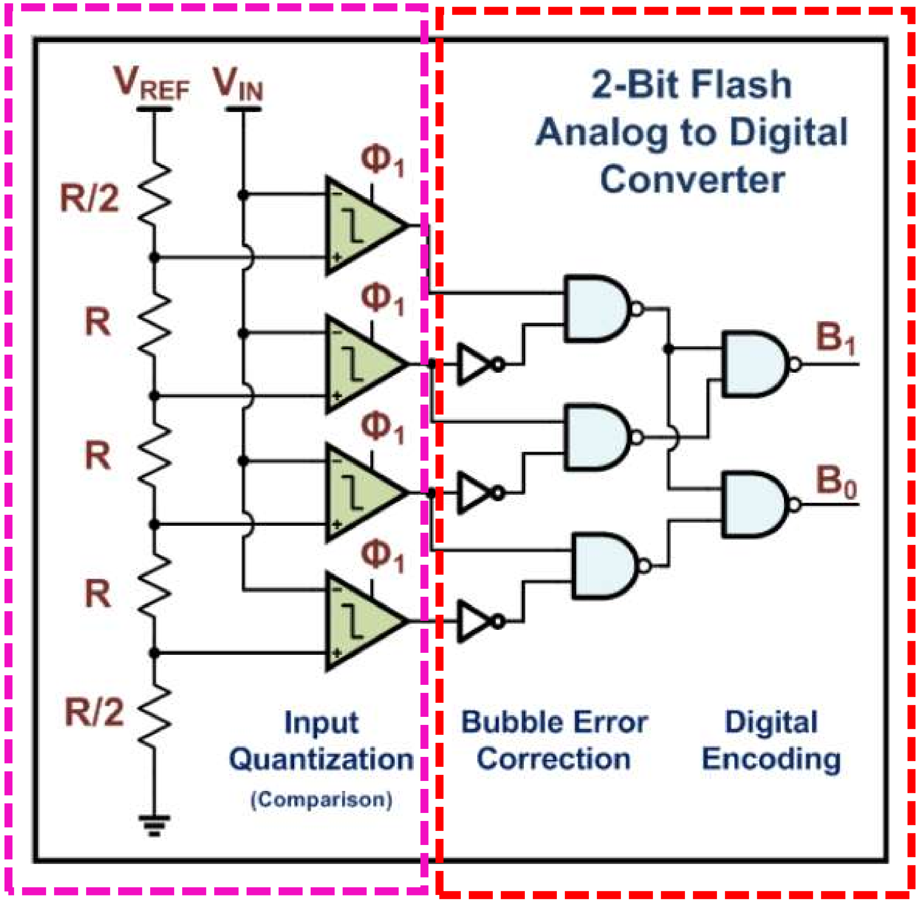
Analog-to-digital converter

A direct-conversion ADC or flash ADC



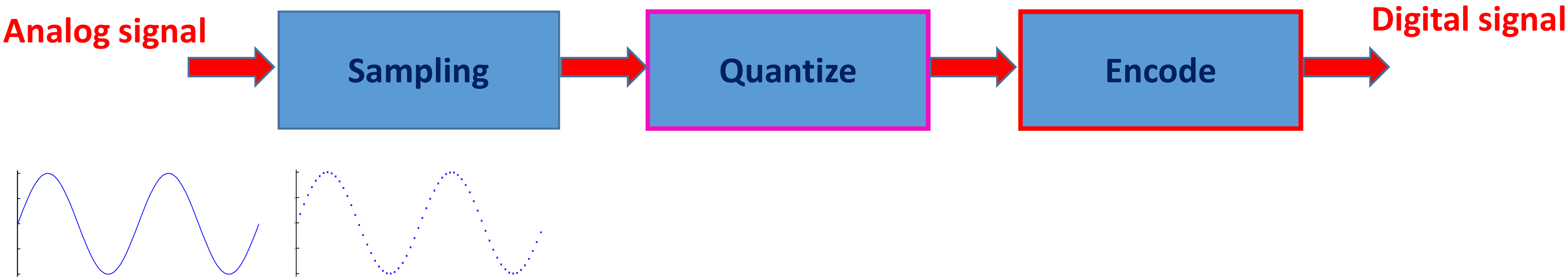
Fast

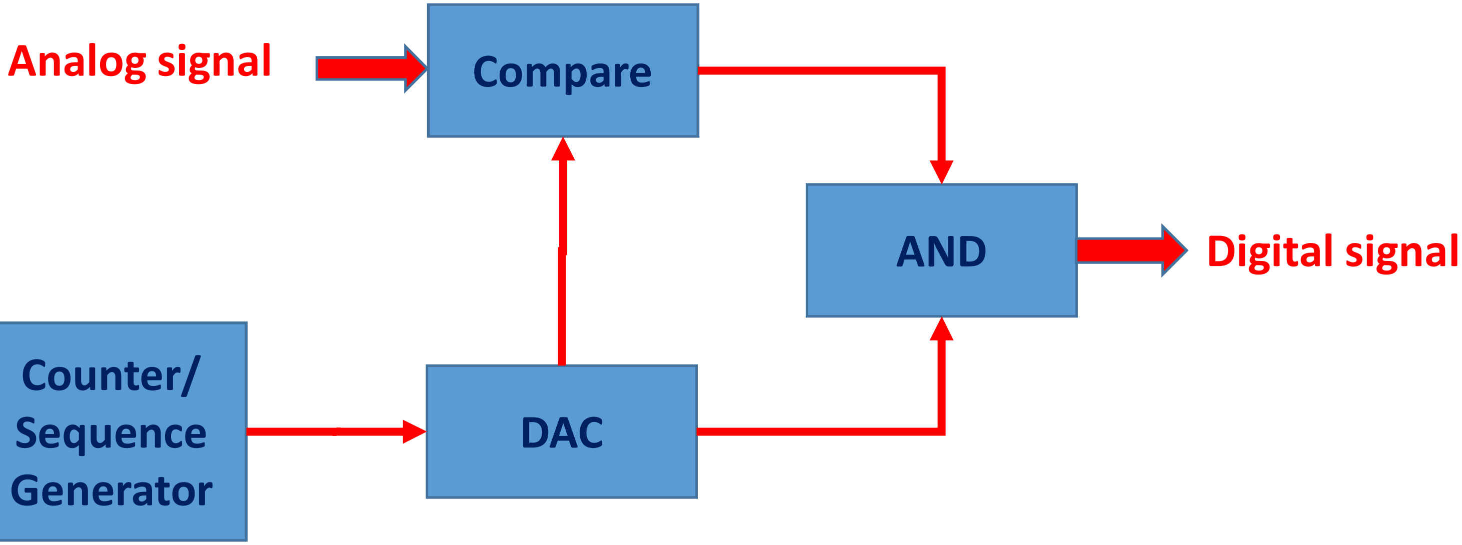
Expensive



Analog-to-digital converter

Counter type ADC

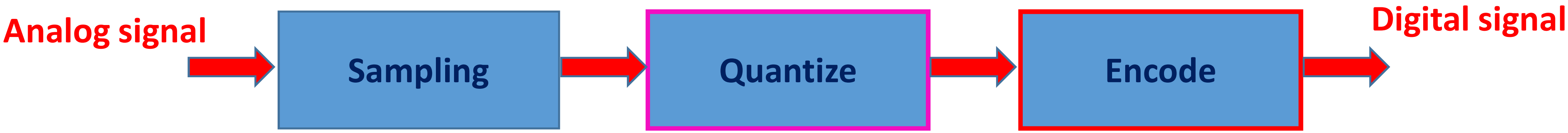


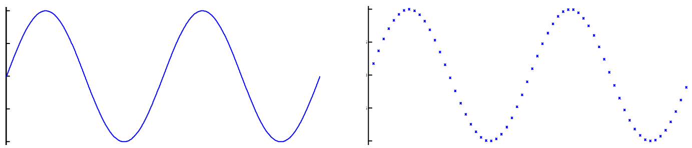


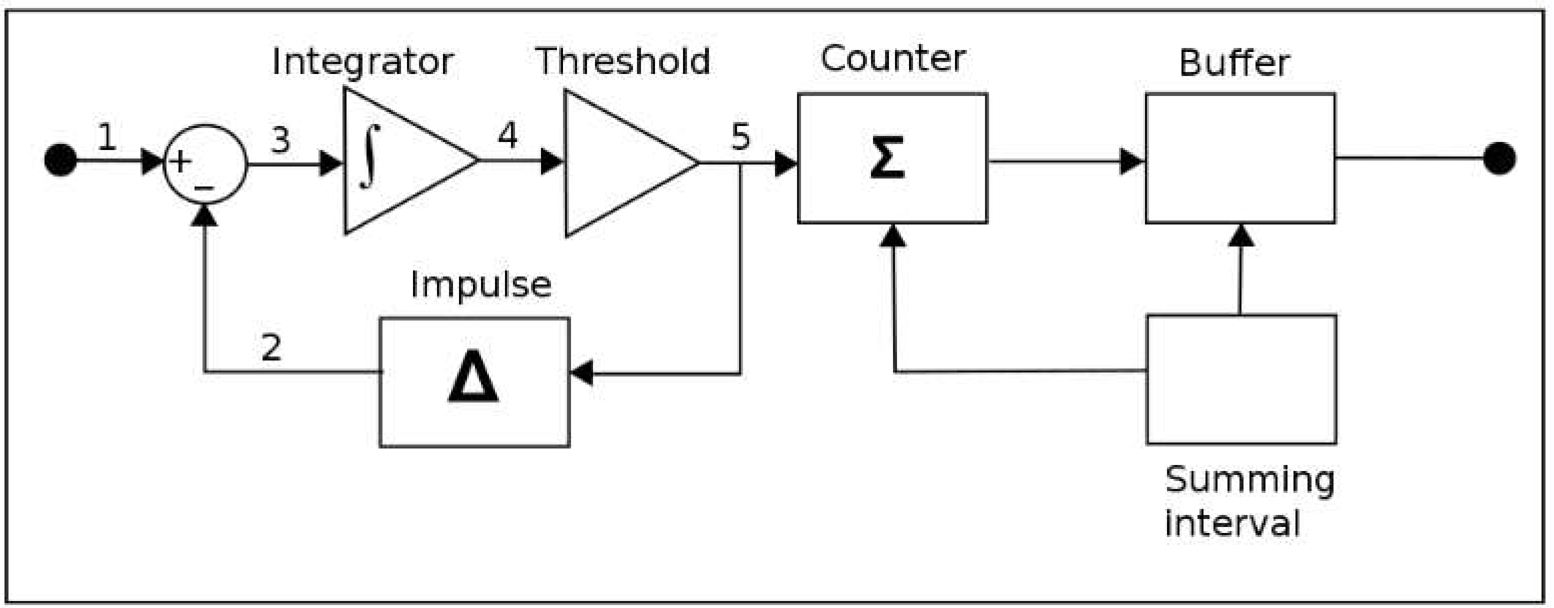
Accurate
Slow

Analog-to-digital converter

A sigma-delta ADC







Low-cost

Bandwidth-limited