LING 490 - SPECIAL TOPICS IN LINGUISTICS

Fundamentals of Digital Signal Processing

Yan Tang

Department of Linguistics, UIUC

Week 3

Last week…

$$x(t) = A \cdot \sin(2\pi \cdot f \cdot t - \phi)$$

- Properties of continuous periodic signals
  - Amplitude, frequency, period and phase
  - Relationship between regular and angular frequency: $\omega = 2\pi \cdot f$
  - Conversion between phase in degree and radians: $d = \frac{360 \cdot \phi}{2\pi}$ and $\phi = \frac{2\pi \cdot d}{360}$

Last week…

$$x(n) = A \cdot \cos(\omega_N \cdot n - \phi)$$

$$\omega_N = 2\pi \cdot f_N$$

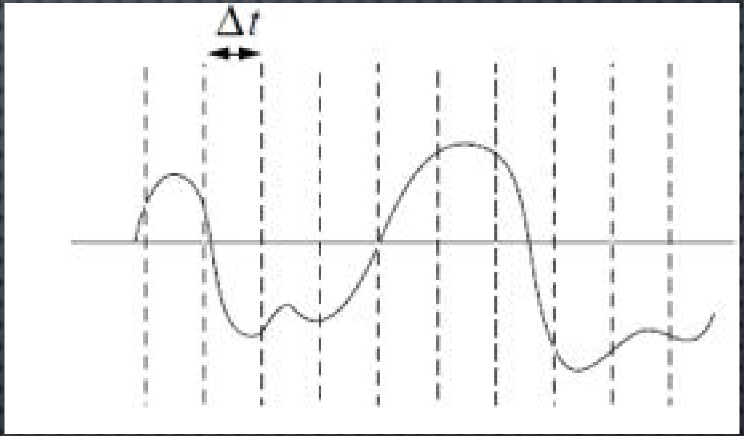
- Properties of discrete periodic signals
  - Frequency f_N : cycle/sample. Must be a rational number
  - Signals with frequency of $(\omega_N + k \cdot 2\pi)$ are identical
  - The highest effective angular frequency is $\omega = \pi$, i.e. $f = 1/2$

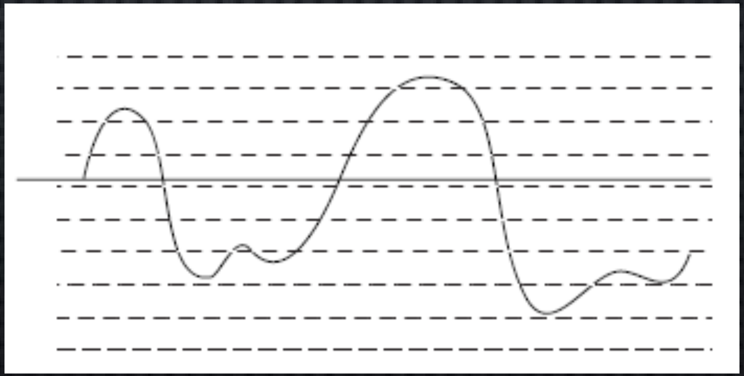
Digitalisation

- DSP requires that the signal is *sampled* **in time and** *quantised* **in amplitude**

- Both have the potential to lose information or add artefacts to the signal

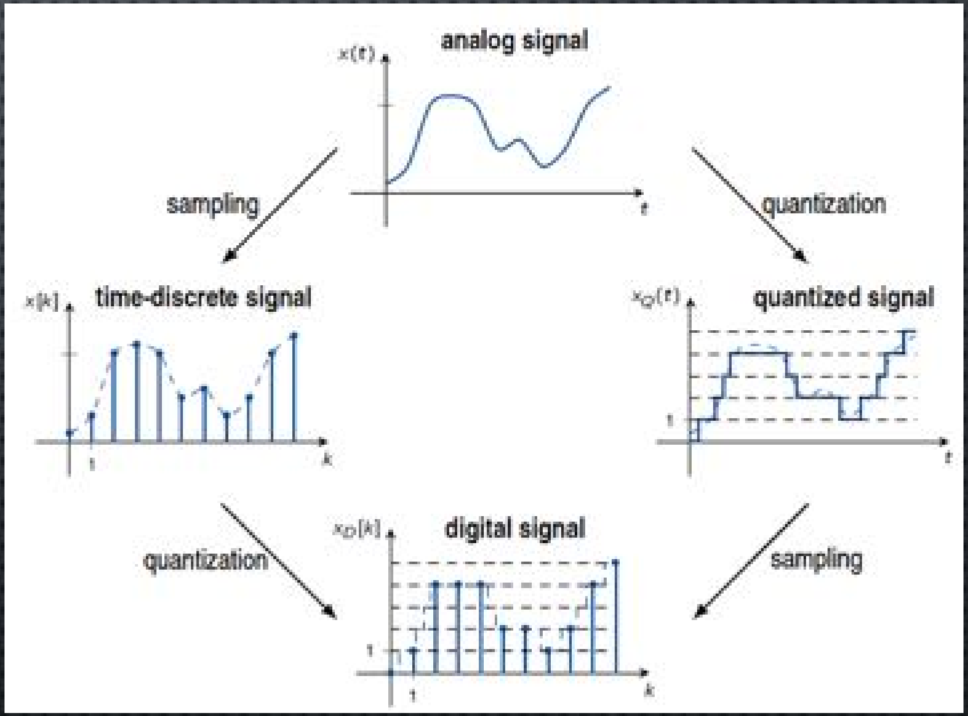
- Often inevitable, but the consequences should be aware of





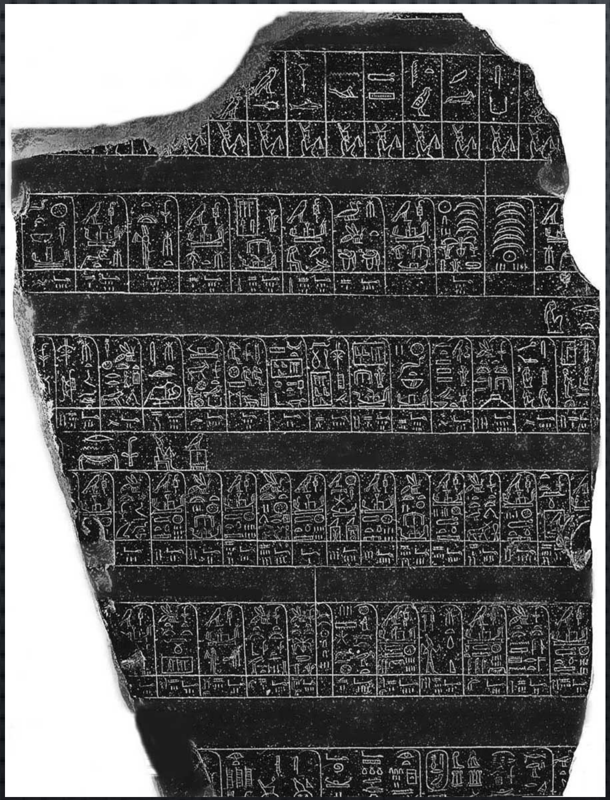
Sampling

- Sampling is the process of breaking up a signal in time
- Digital systems deal with discrete data rather than continuous data
- This process has a few consequences for our signal

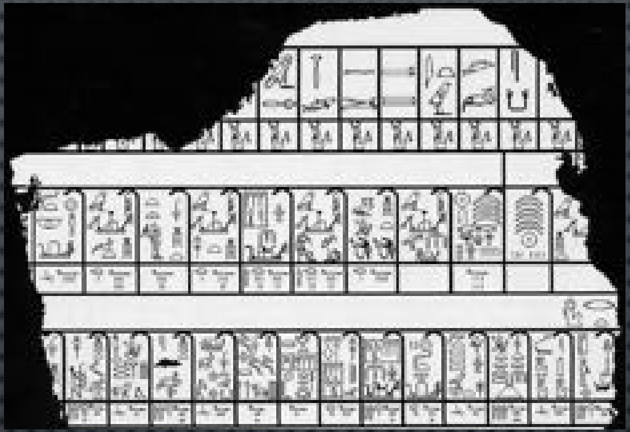


[DEMO]

The earliest application of DSP sampling



The Palermo Stone



- Ancient Egyptian back to 25th century BC
- Recorded the data of annual flood level
- Resembles the current approach of data sampling

Relationship between continuous time and sample number

- The number of samples taken in each second is called the *sampling frequency* or *sampling rate* (usually denoted as fs)

$$t = nT_s = \frac{n}{f_s}$$

$\boldsymbol{f_s}$: sampling frequency in *Hz*

$\boldsymbol{t}$: continuous time in seconds

$\boldsymbol{n}$: number of samples

$\boldsymbol{T_s}$: the sampling period

Consequence of all this?

Only frequencies below half of the sampling rate will be accurately represented after sampling!

e.g. To sample a 100 Hz sine wave, an $f_s > 200$ Hz is required

But why? What will happen if f_s doesn't not meet the above rule?

Demo

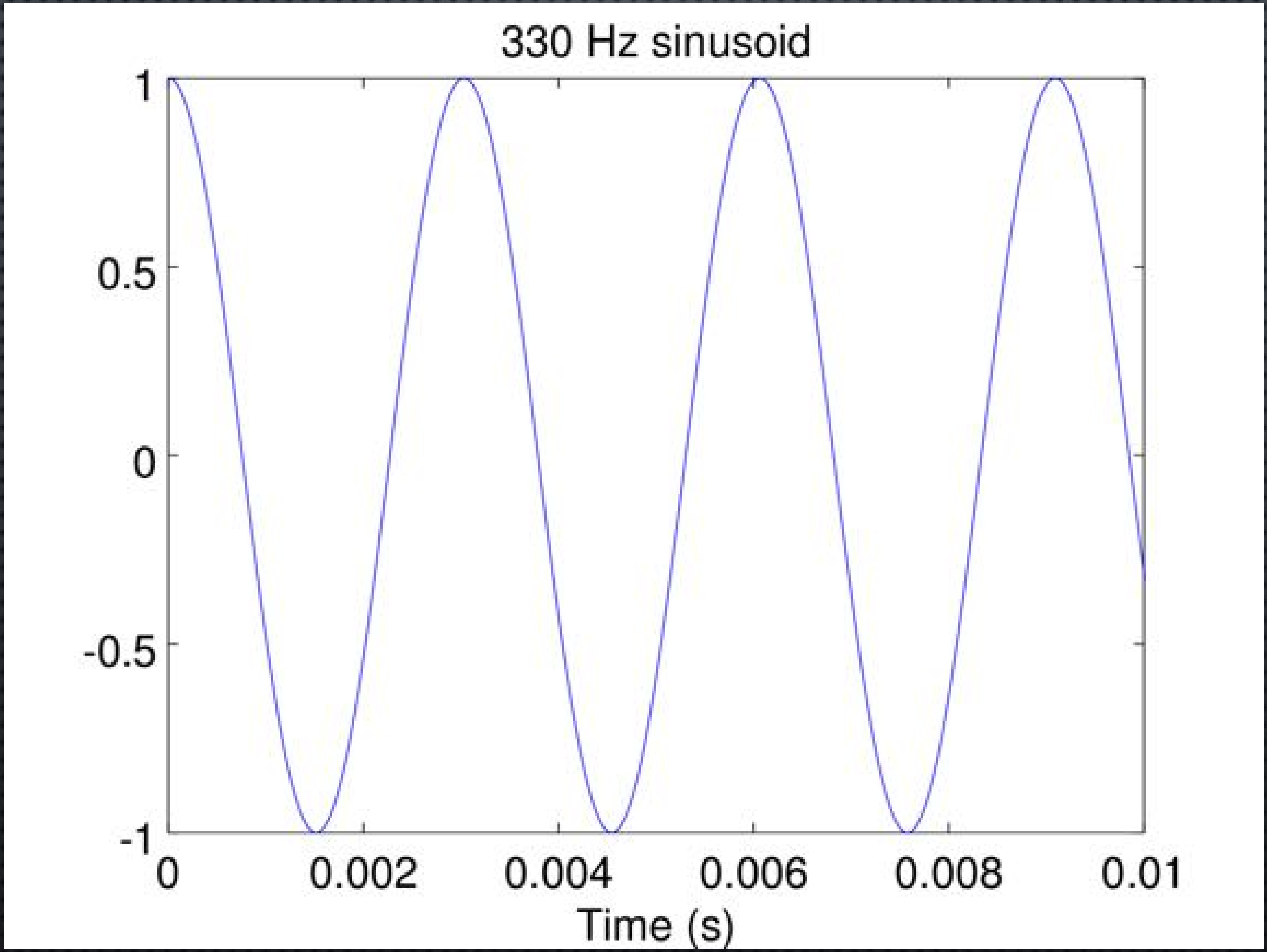
Nyquist frequency

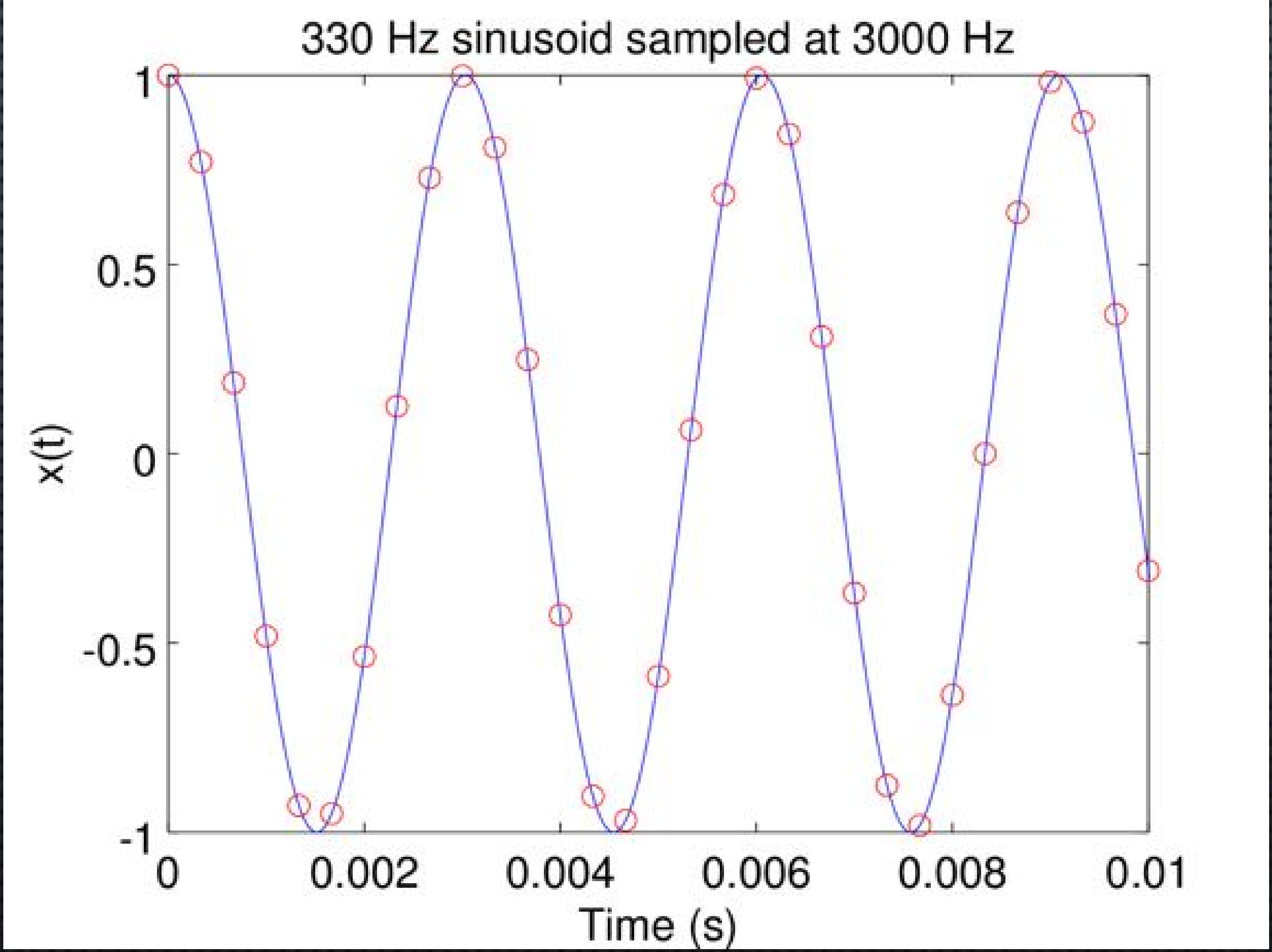
- This frequency ($f_s/2$) is called the *Nyquist frequency* or *Nyquist limit*

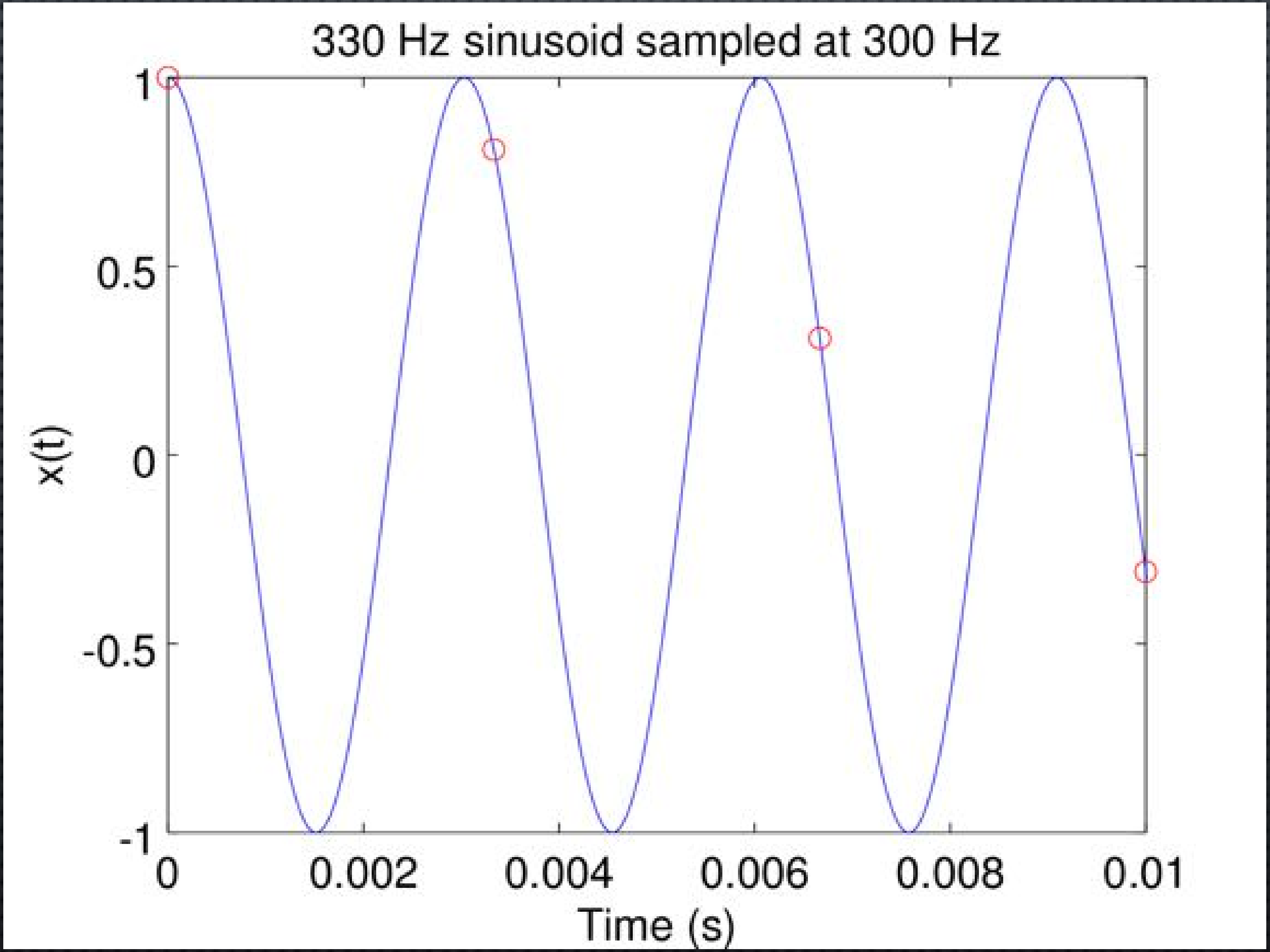
← After this guy: Harry Nyquist

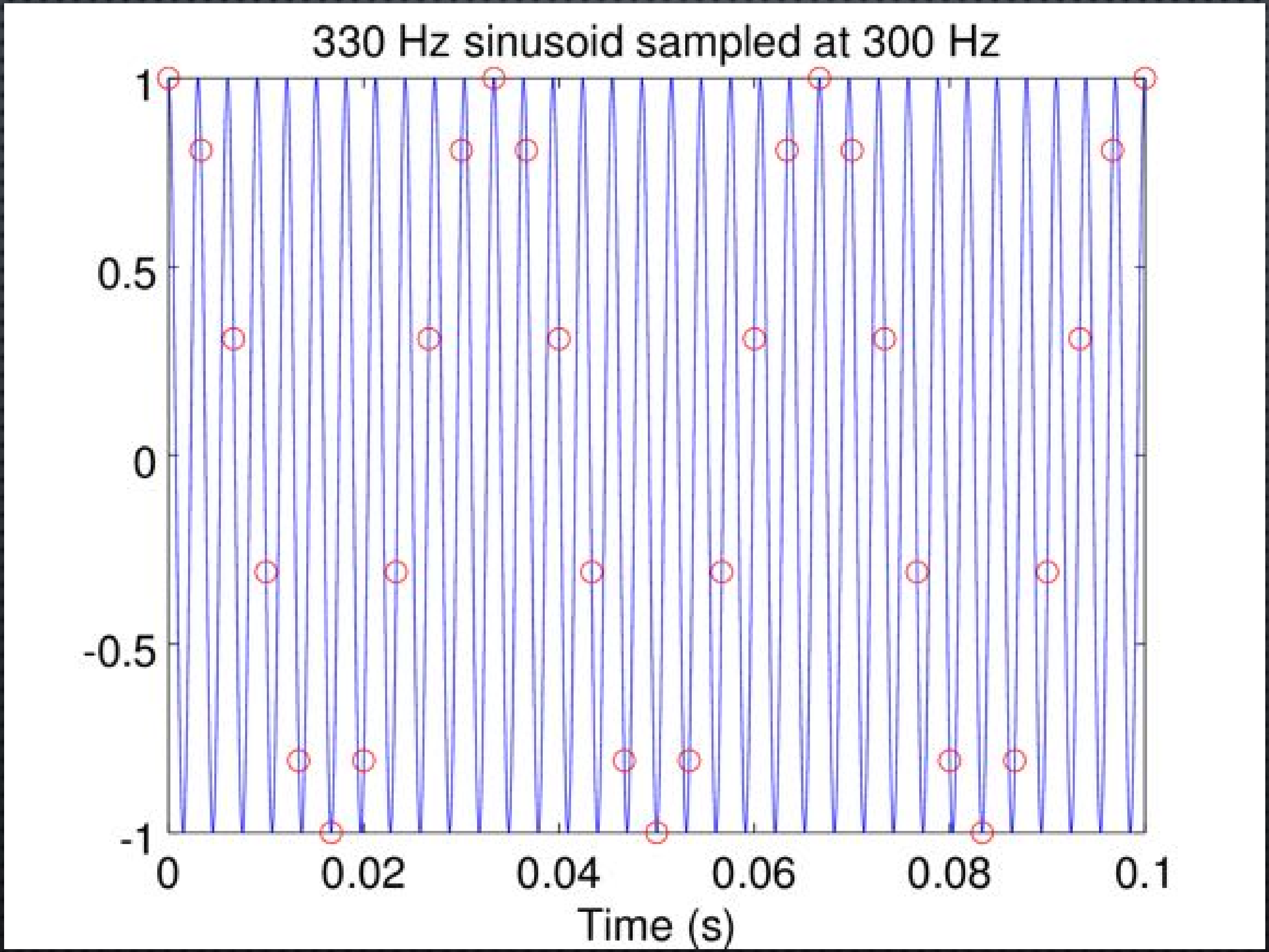
Sampling theorem

A continuous (analogue) signal can be recovered from a sampled signal if the signal is sampled at a rate $f_s > 2f_{max}$


330 Hz sinusoid
1
0.5
0
-0.5
-1
0
0.002
0.004
0.006
0.008
0.01
Time (s)


330 Hz sinusoid sampled at 3000 Hz
x(t)
1
0.5
0
-0.5
-1
0
0.002
0.004
0.006
0.008
0.01
Time (s)


330 Hz sinusoid sampled at 300 Hz
x(t)
1
0.5
0
-0.5
-1
0
0.002
0.004
0.006
0.008
0.01
Time (s)


330 Hz sinusoid sampled at 300 Hz
1
0.5
0
-0.5
-1
0
0.02
0.04
0.06
0.08
0.1
Time (s)

Demo

Without limiting the frequency of input signal, the sampled signal – which should not exist – is a mirrored image of the signal at the Nyquist Frequency

i.e. Given a sampling rate of f_s Hz, a sampled sine wave at a frequency of f (where $\frac{f_s}{2} < f \leq f_s$) is indistinguishable from a sine wave at a frequency of $(f - f_s)$ Hz

What will happen when $f > f_s$?

Sinewave aliasing

The Sampled Sinewave Theorem:

Given a sampling rate of f_s Hz, and an integer k, a sine wave at a frequency of f is indistinguishable from a sine wave at a frequency of $(f + k \cdot f_s)$ after being sampled by f_s

$f_s = 20$ Hz, what are the aliases for a sine wave of 4 Hz (only considering positive values of k)?

e.g. given a $f_s = 20$ Hz, a sine wave with a frequency of 4 Hz is indistinguishable from sine waves at 24 Hz, 44 Hz, 64 Hz….

DEMO

The wagon wheel effect

Phenomenon: on film or television wheels sometimes seem to spin backwards, although the vehicle is going forward.

- Under-sampling
- Aliasing

DEMO

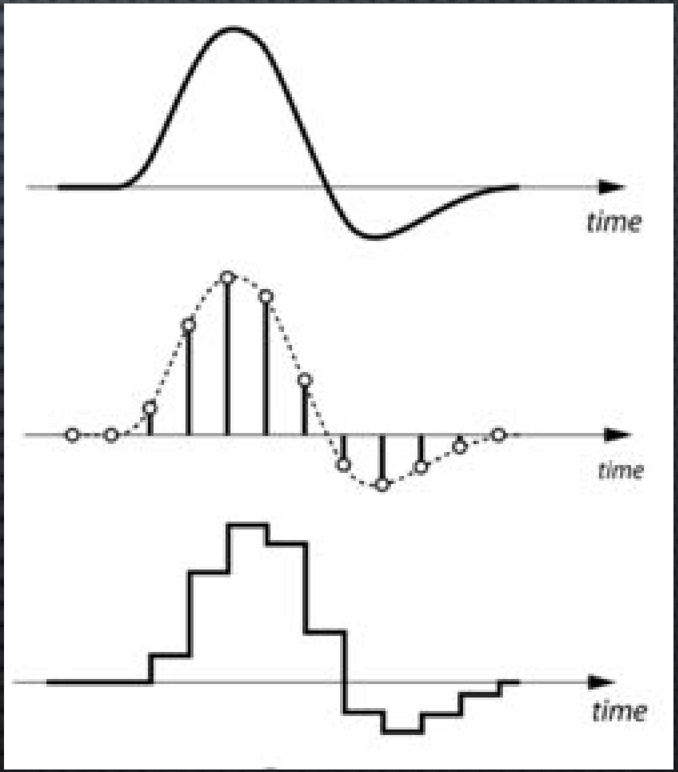
Question

What's the minimum sampling frequency we need for audio?

- The range of human hearing: 20 - 20000 Hz
- We need to sample at double that rate: at least 40 kHz
- CD audio is sampled at 44.1kHz. Professional audio often sampled at 48 kHz.

Quantisation

- As well as sampling, we also need to decimate in level.
- The process of converting a discrete-time continuous amplitude signal into a digital signal by expressing each sample as a finite number of digits is called *quantisation*.



Why quantise samples?

- Consider the signal:

$$x(n) = \begin{cases} 0.9^n, & n \geq 0 \\ 0, & n < 0 \end{cases}$$

- How many significant digits do we need to represent this signal?

Why quantise samples?

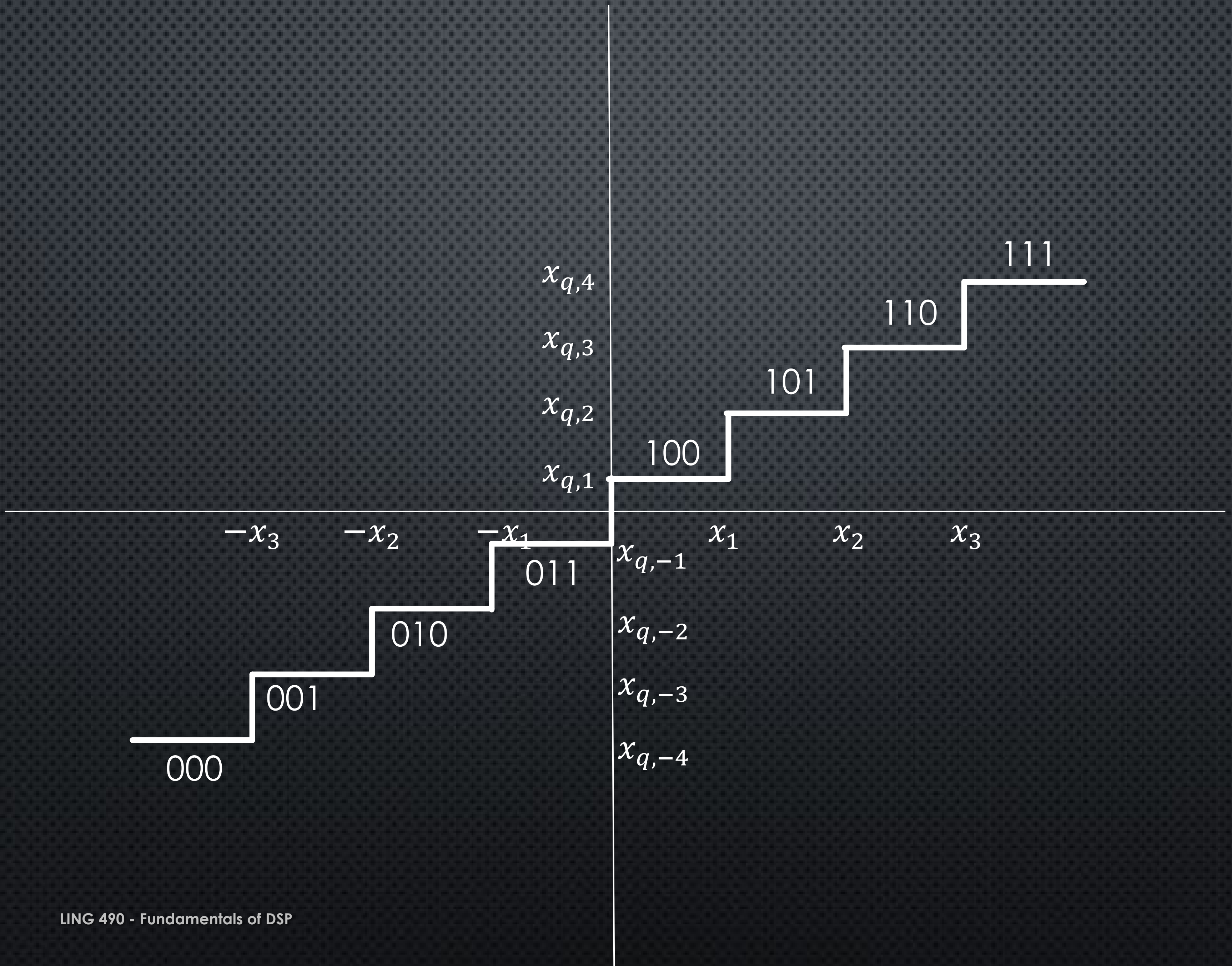
$$x(n) = \begin{cases} 0.9^n, & n \geq 0 \\ 0, & n < 0 \end{cases}$$

n	x(n)
0	1
1	0.9
2	0.81
3	0.729
4	0.6561
5	0.59049
6	0.531441
7	0.4782969
8	0.43046721
9	0.387420489

- *n* significant digits are need for each $x(n)$

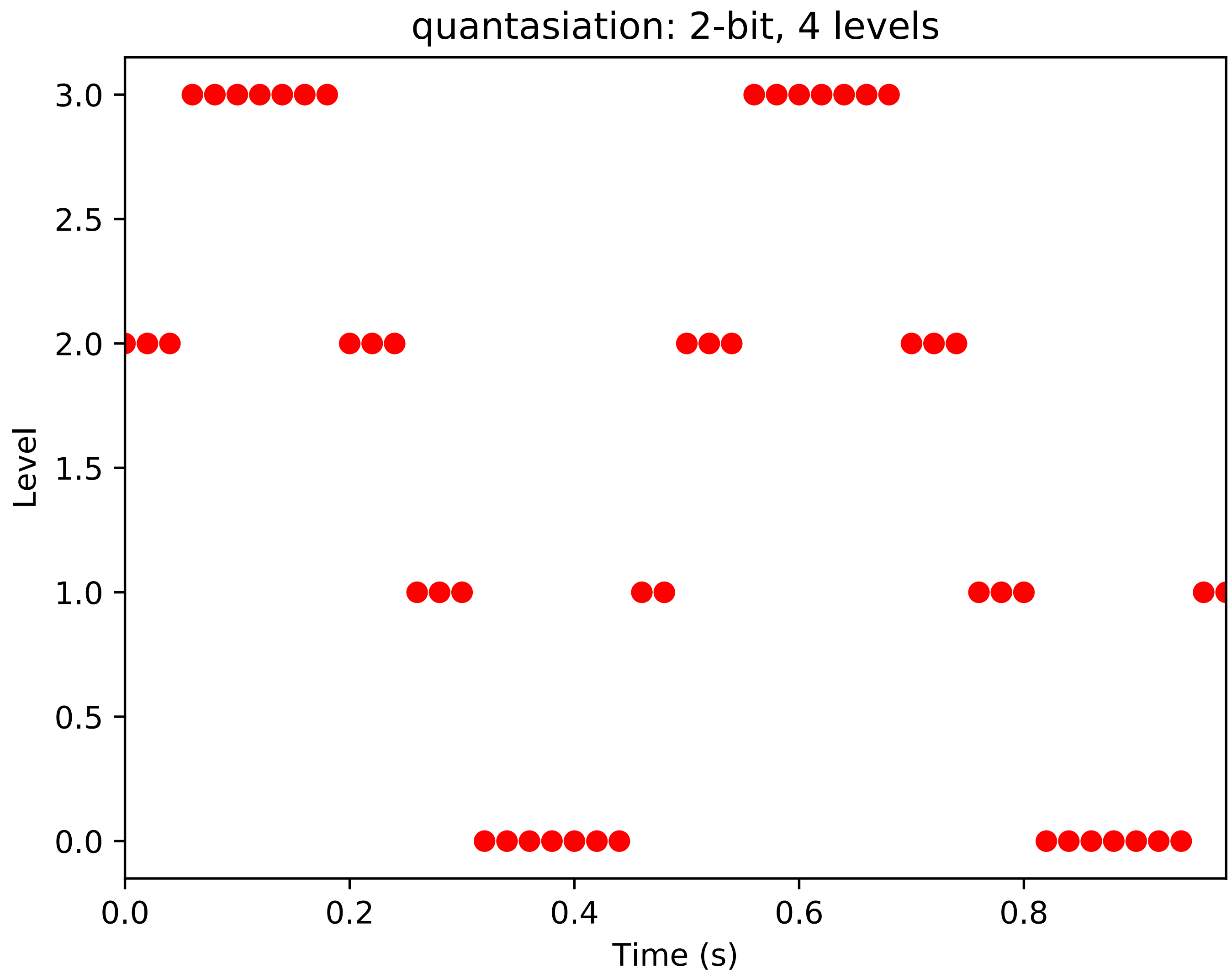
Why quantise samples?

n	x(n)	$x_q(n)$ (Rounding)	$e_q(n)$ (Rounding)
0	1	1	0
1	0.9	0.9	0
2	0.81	0.8	-0.01
3	0.729	0.7	-0.029
4	0.6561	0.7	0.0439
5	0.59049	0.6	0.00951
6	0.531441	0.5	-0.031441
7	0.4782969	0.5	0.0217031
8	0.43046721	0.4	-0.03046721
9	0.387420489	0.4	0.012579511


111
110
101
100
011
010
001
000
$x_{q,4}$
$x_{q,3}$
$x_{q,2}$
$x_{q,1}$
$-x_3$
$-x_2$
$-x_1$
x_1
x_2
x_3
$x_{q,-1}$
$x_{q,-2}$
$x_{q,-3}$
$x_{q,-4}$

Quantisation levels

- The values allowed in the digital signal
- The distance between two quantisation levels is called the *quantisation step*
- In a digital system the number of bits (quantisation resolution) determines the number of levels – ($2^{n\text{-bit}}$) levels
  - e.g. 8 bits -> 256 levels

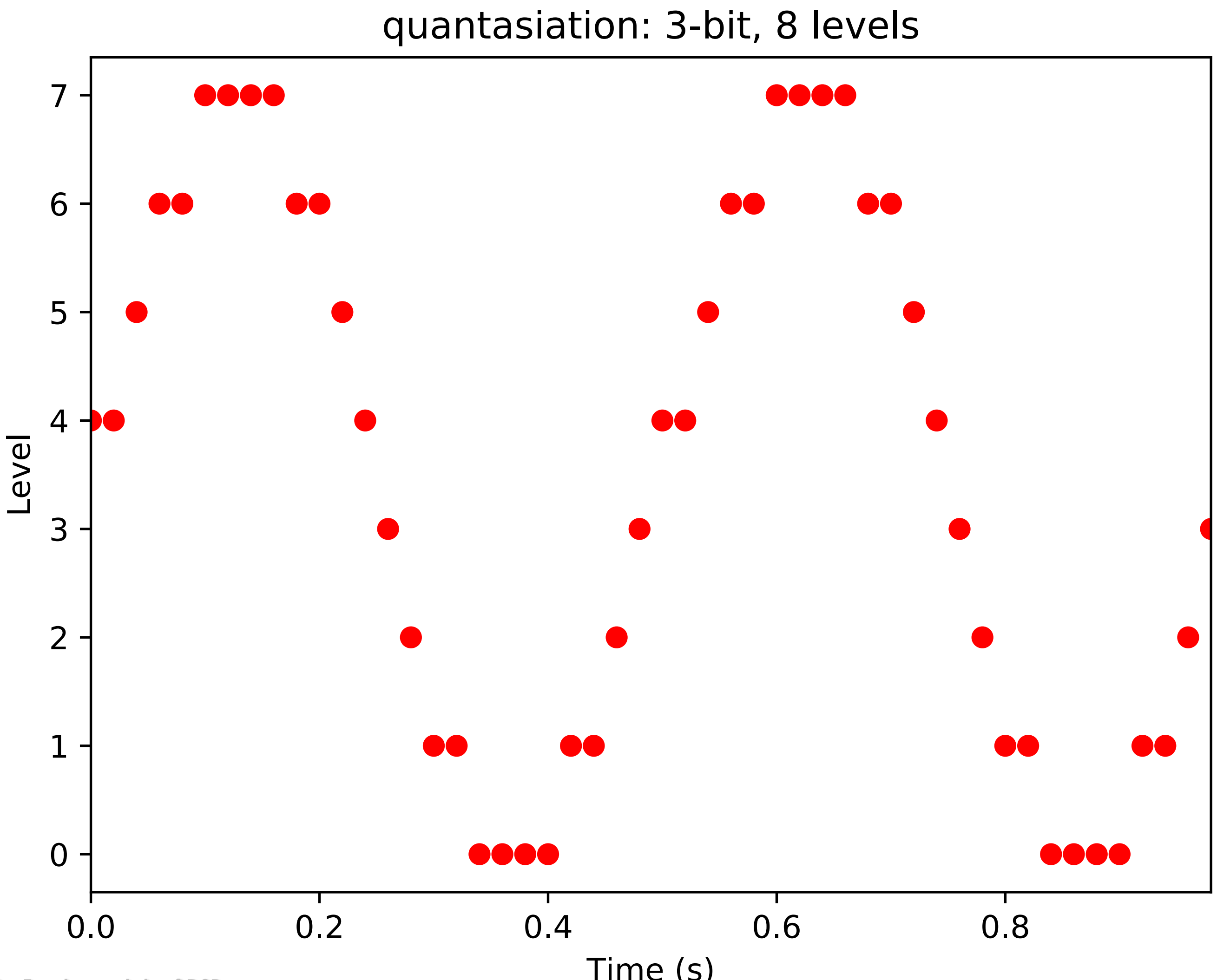

quantasiation: 2-bit, 4 levels
Level
Time (s)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0.2
0.4
0.6
0.8

1

0

1

0


quantasiation: 3-bit, 8 levels
Level
Time (s)
0
1
2
3
4
5
6
7
0.0
0.2
0.4
0.6
0.8

Quantisation error

- Quantisation error is a sequence, e_q, defined as the difference between the quantised values and the original value:

$$e_q(n) = x_q(n) - x(n)$$

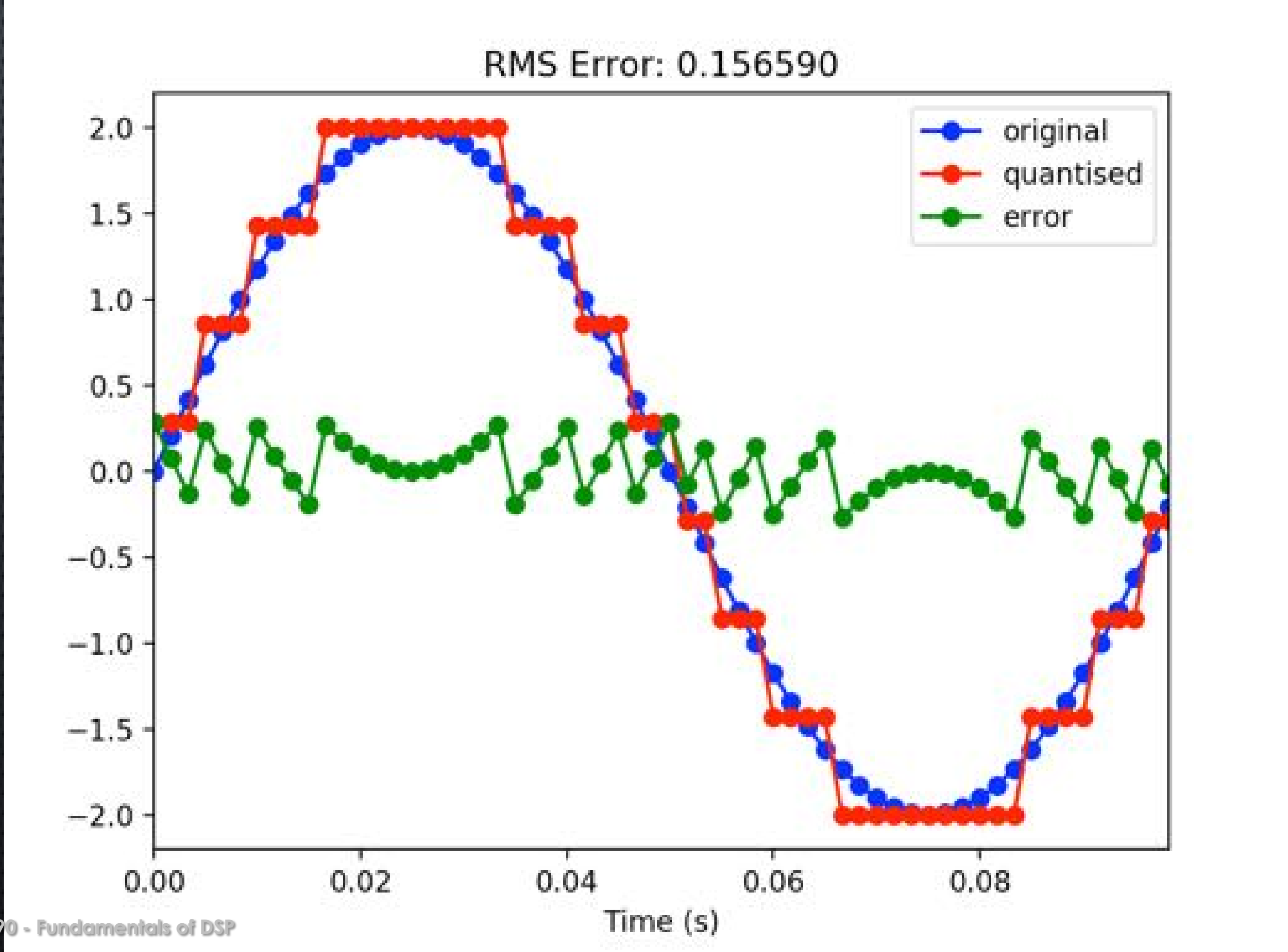
Quantisation error is unavoidable

- There will always be quantisation noise in the signal

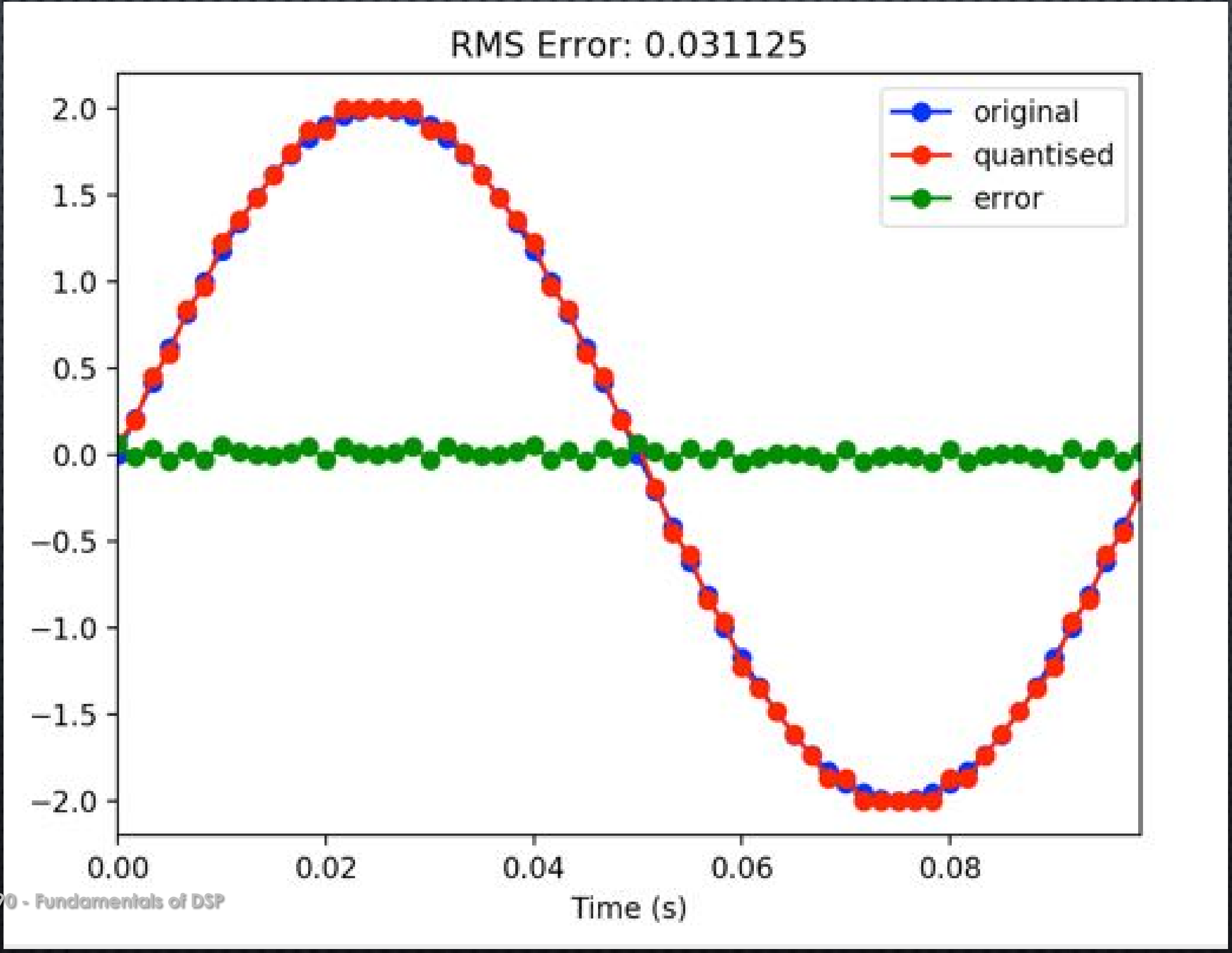
- It is always possible to talk about the signal-to-noise ratio (SNR) of a "clean" digital signal

- Make use of enough bits for high resolution

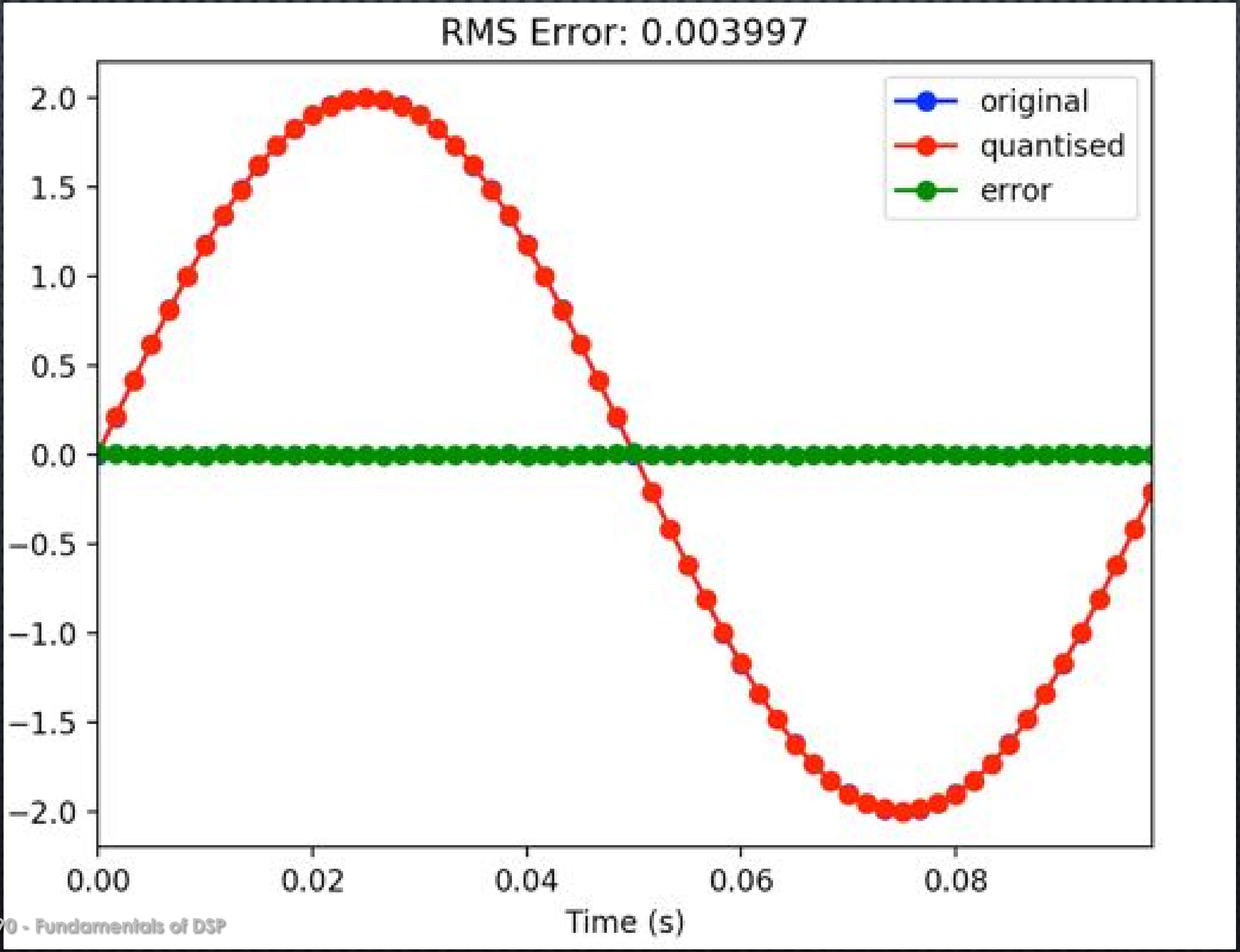
ERRORS: 3-BIT 8-LEVEL QUANTISATION



ERRORS: 5-BIT 32-LEVEL QUANTISATION



ERRORS: 8-BIT 256-LEVEL QUANTISATION

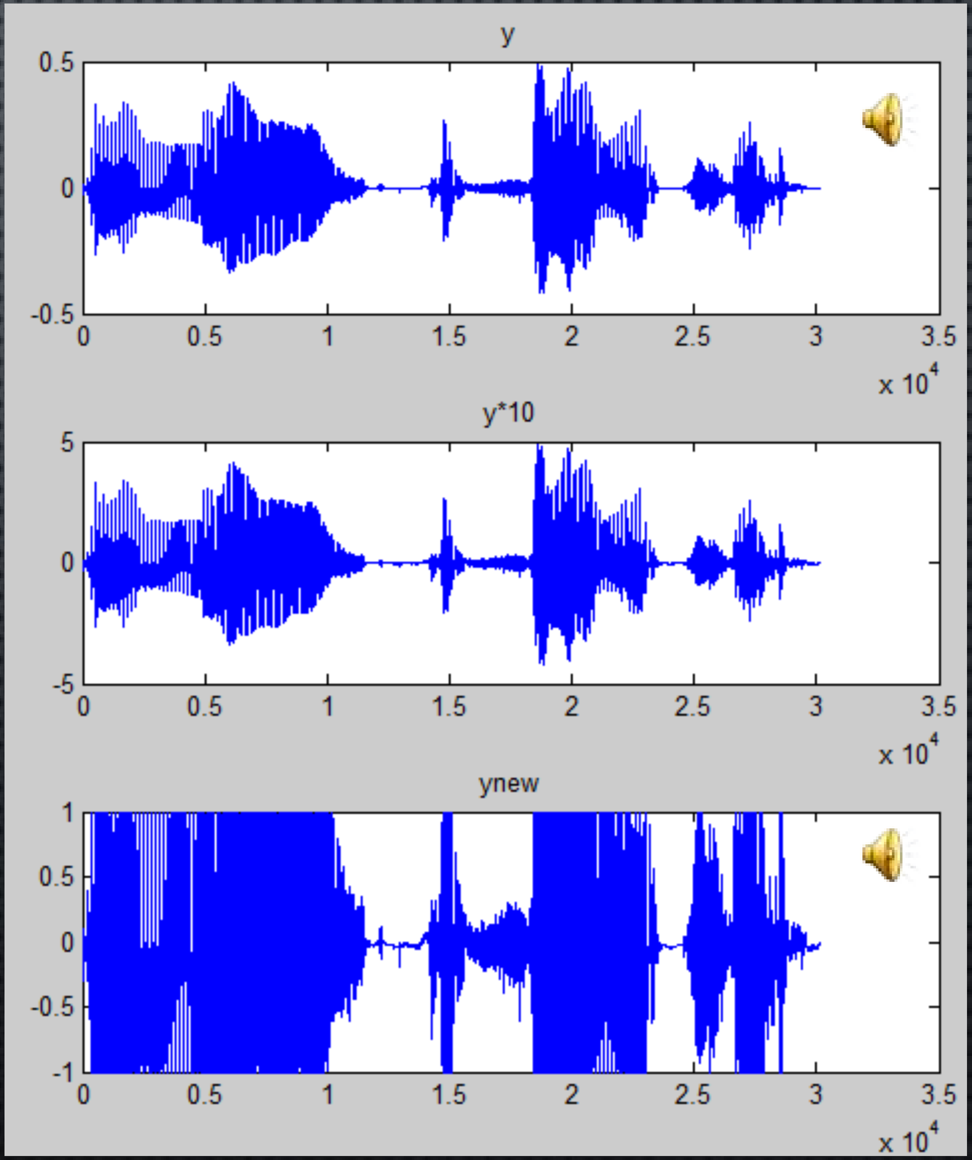


How many steps do we need?

- Make sure that there are sufficient number of steps so that the difference is inaudible to our ears
  - 16-bit resolution is mostly used: 65536 levels!!!
- In reality, even the best analogue system cannot truly represent an irrational voltage level as the noise floor will always limit the accuracy of the representation.
- What makes the errors produced by quantisation more difficult to handle, is that it is in fact distortion, e.g. clipping

Clipping in waveform

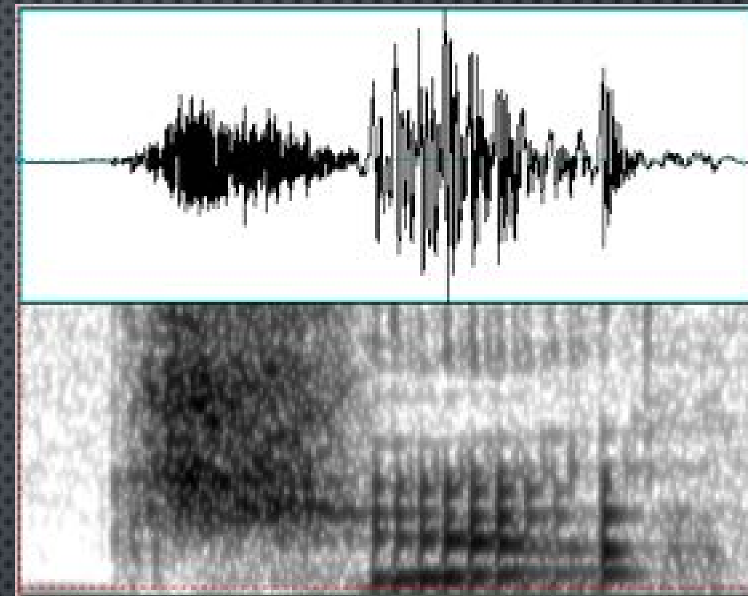
- Clipping leads to the loss of dynamic range of amplitude
- Reduction of modulation depth in speech

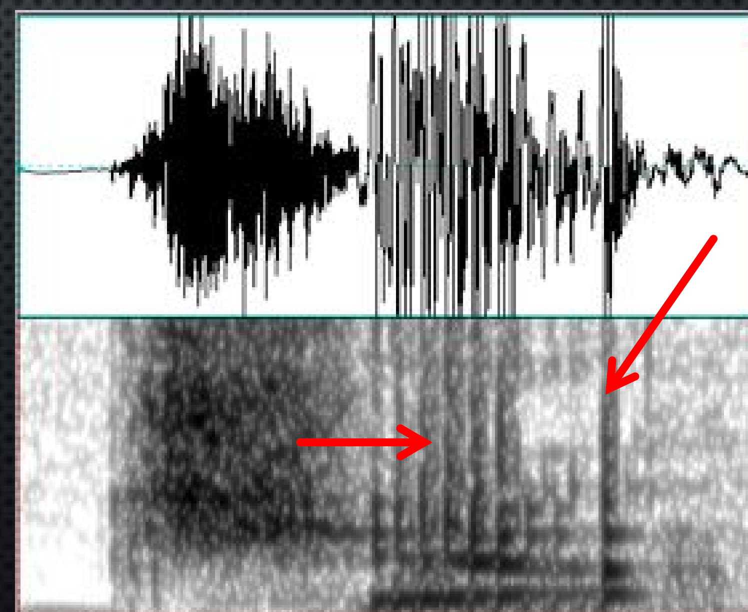


Clipping in spectrum

Clipping produces visible artefacts in the spectrogram, particularly at high frequencies

- Spurious appearance of bursts
- Extra 'vertical' noise can reduce salience of 'horizontal' structure, disguising evidence of formants





Signal acquisition: from analogue to digital

- Set gain level
  - High enough to make best use of limited number of bits
  - Low enough to avoid clipping

- Set sampling frequency high enough to capture information that human ear can perceive
  - Set it high: can always downsample – carefully! – later
  - Too high and excessive storage requirements