COM303: Digital Signal Processing

Lecture 19: Quantization

overview

- quantization
- A/D and D/A converters
- oversampling

quantization

Overview:

- Quantization
- Uniform quantization and error analysis
- Clipping, saturation, companding

Overview:

- Quantization
- Uniform quantization and error analysis
- Clipping, saturation, companding

Overview:

- Quantization
- Uniform quantization and error analysis
- Clipping, saturation, companding

Quantization

- digital devices can only deal with integers (b bits per sample)
- we need to map the range of a signal onto a finite set of values
- irreversible loss of information $\rightarrow$ quantization noise

Quantization

- digital devices can only deal with integers (b bits per sample)
- we need to map the range of a signal onto a finite set of values
- irreversible loss of information $\rightarrow$ quantization noise

Quantization

- digital devices can only deal with integers (b bits per sample)
- we need to map the range of a signal onto a finite set of values
- irreversible loss of information $\rightarrow$ quantization noise

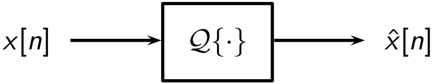
Quantization schemes

$$x[n] \longrightarrow \boxed{\mathcal{Q}\{\cdot\}} \longrightarrow \hat{x}[n]$$

Several factors at play:

- storage budget (bits per sample)
- storage scheme (fixed point, floating point)
- properties of the input
  - range
  - probability distribution

Quantization schemes



Several factors at play:

- storage budget (bits per sample)
- storage scheme (fixed point, floating point)
- properties of the input
  - range
  - probability distribution

Quantization schemes

$$x[n] \longrightarrow \boxed{\mathcal{Q}\{\cdot\}} \longrightarrow \hat{x}[n]$$

Several factors at play:

- storage budget (bits per sample)
- storage scheme (fixed point, floating point)
- properties of the input
  - range
  - probability distribution

Quantization schemes

$$x[n] \longrightarrow \boxed{\mathcal{Q}\{\cdot\}} \longrightarrow \hat{x}[n]$$

Several factors at play:

- storage budget (bits per sample)
- storage scheme (fixed point, floating point)
- properties of the input
  - range
  - probability distribution

Quantization schemes

$$x[n] \longrightarrow \boxed{\mathcal{Q}\{\cdot\}} \longrightarrow \hat{x}[n]$$

Several factors at play:

- storage budget (bits per sample)
- storage scheme (fixed point, floating point)
- properties of the input
  - range
  - probability distribution

Scalar quantization

$$x[n] \longrightarrow \boxed{\mathcal{Q}\{\cdot\}} \longrightarrow \hat{x}[n]$$

The simplest quantizer:

- each sample is encoded individually (hence *scalar*)
- each sample is quantized independently (memoryless quantization)
- each sample is encoded using R bits

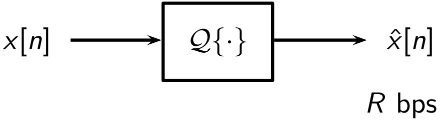
Scalar quantization

$$x[n] \longrightarrow \boxed{\mathcal{Q}\{\cdot\}} \longrightarrow \hat{x}[n]$$

The simplest quantizer:

- each sample is encoded individually (hence *scalar*)
- each sample is quantized independently (memoryless quantization)
- each sample is encoded using R bits

Scalar quantization



The simplest quantizer:

- each sample is encoded individually (hence *scalar*)
- each sample is quantized independently (memoryless quantization)
- each sample is encoded using R bits

Scalar quantization

Input signal: $A \leq x[n] \leq B$ (A, B can be ∞)

- each sample quantized over 2^R possible values $\Rightarrow 2^R$ intervals.
- each interval associated to a quantization value

Scalar quantization

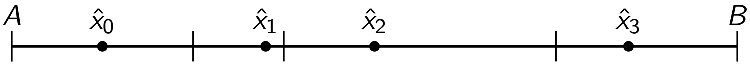
Input signal: $A \leq x[n] \leq B$ (A, B can be ∞)

- each sample quantized over 2^R possible values $\Rightarrow$ 2^R intervals.
- each interval associated to a quantization value

Scalar quantization

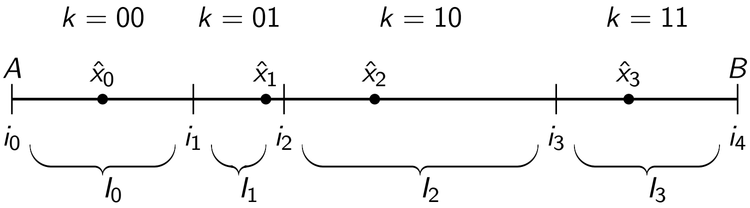
Input signal: $A \leq x[n] \leq B$ (A, B can be ∞)

- each sample quantized over 2^R possible values $\Rightarrow$ 2^R intervals.
- each interval associated to a quantization value



Scalar quantization

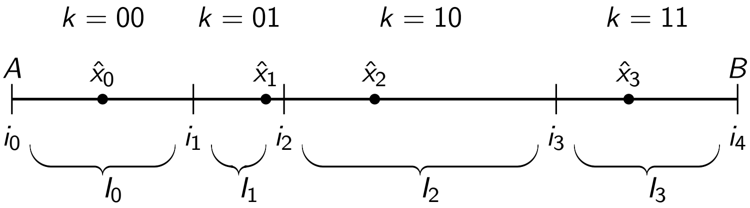
Example for $R = 2$:



- what are the optimal interval boundaries i_k?
- what are the optimal quantization values $\hat{x}_k$?

Scalar quantization

Example for $R = 2$:



- what are the optimal interval boundaries i_k?
- what are the optimal quantization values $\hat{x}_k$?

Optimal Quantization

The optimal quantizer minimizes the energy of the quantization error:

$$e[n] = \mathcal{Q}\{x[n]\} - x[n] = \hat{x}[n] - x[n]$$

- model $x[n]$ as a stochastic process
- find the optimal i_k and $\hat{x}_k$ that minimize $\sigma_e^2 = \mathrm{E}\left[e^2[n]\right]$
- optimal quantizer will depend on the input's statistics

Optimal Quantization

The optimal quantizer minimizes the energy of the quantization error:

$$e[n] = \mathcal{Q}\{x[n]\} - x[n] = \hat{x}[n] - x[n]$$

- model $x[n]$ as a stochastic process
- find the optimal i_k and $\hat{x}_k$ that minimize $\sigma_e^2 = \mathrm{E}\left[e^2[n]\right]$
- optimal quantizer will depend on the input's statistics

Optimal Quantization

The optimal quantizer minimizes the energy of the quantization error:

$$e[n] = \mathcal{Q}\{x[n]\} - x[n] = \hat{x}[n] - x[n]$$

- model $x[n]$ as a stochastic process
- find the optimal i_k and $\hat{x}_k$ that minimize $\sigma_e^2 = \mathrm{E}\left[e^2[n]\right]$
- optimal quantizer will depend on the input's statistics

Quantization MSE

$$\sigma_e^2 = \int_{-\infty}^{\infty} (x - \mathcal{Q}\{x\})^2 \, f_x(x) \, dx$$

$$= \sum_{k=0}^{2^R-1} \int_{i_k}^{i_{k+1}} (x - \hat{x}_k)^2 \, f_x(x) \, dx$$

find global minimum wrt i_k, $\hat{x}_k$

Quantization MSE

$$\sigma_e^2 = \int_{-\infty}^{\infty} (x - \mathcal{Q}\{x\})^2 \, f_x(x)\, dx$$

$$= \sum_{k=0}^{2^R-1} \int_{i_k}^{i_{k+1}} (x - \hat{x}_k)^2 \, f_x(x)\, dx$$

find global minimum wrt i_k, $\hat{x}_k$

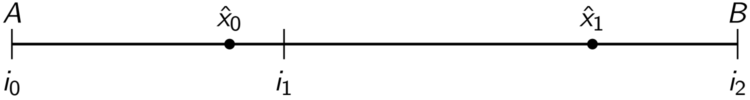
Quantization MSE

$$\sigma_e^2 = \int_{-\infty}^{\infty} (x - \mathcal{Q}\{x\})^2 \, f_x(x) \, dx$$

$$= \sum_{k=0}^{2^R-1} \int_{i_k}^{i_{k+1}} (x - \hat{x}_k)^2 \, f_x(x) \, dx$$

find global minimum wrt i_k, $\hat{x}_k$

Simple example: optimal one-bit quantizer



3 free parameters: $i_1, \hat{x}_0, \hat{x}_1$

Simple example: optimal one-bit quantizer

$$\sigma_e^2 = \int_A^{i_1} (x - \hat{x}_0)^2 \, f_x(x) \, dx + \int_{i_1}^{B} (x - \hat{x}_1)^2 \, f_x(x) \, dx$$

find $i_1, \hat{x}_0, \hat{x}_1$ such that

$$\frac{\partial \sigma_e^2}{\partial i_1} = \frac{\partial \sigma_e^2}{\hat{x}_0} = \frac{\partial \sigma_e^2}{\hat{x}_1} = 0$$

Simple example: optimal one-bit quantizer

$$\sigma_e^2 = \int_A^{i_1} (x - \hat{x}_0)^2 \, f_x(x) \, dx + \int_{i_1}^{B} (x - \hat{x}_1)^2 \, f_x(x) \, dx$$

find $i_1, \hat{x}_0, \hat{x}_1$ such that

$$\frac{\partial \sigma_e^2}{\partial i_1} = \frac{\partial \sigma_e^2}{\hat{x}_0} = \frac{\partial \sigma_e^2}{\hat{x}_1} = 0$$

little calculus reminder

$$\frac{\partial}{\partial t}\int_{\alpha}^{t} f(\tau)\,d\tau = \frac{\partial}{\partial t}\left[F(t) - F(\alpha)\right] = f(t)$$

Optimal one-bit quantizer: threshold

$$\frac{\partial \sigma_e^2}{\partial i_1} = \frac{\partial}{\partial i_1}\left[\int_A^{i_1} (x - \hat{x}_0)^2\, f_x(x)\, dx + \int_{i_1}^{B} (x - \hat{x}_1)^2\, f_x(x)\, dx\right]$$

$$= (i_1 - \hat{x}_0)^2\, f_x(i_1) - (i_1 - \hat{x}_1)^2\, f_x(i_1) = 0$$

$$\Rightarrow (i_1 - \hat{x}_0)^2 - (i_1 - \hat{x}_1)^2 = 0$$

$$\Rightarrow i_1 = \frac{\hat{x}_0 + \hat{x}_1}{2}$$

Optimal one-bit quantizer: threshold

$$\frac{\partial \sigma_e^2}{\partial i_1} = \frac{\partial}{\partial i_1}\left[\int_A^{i_1} (x - \hat{x}_0)^2\, f_x(x)\, dx + \int_{i_1}^{B} (x - \hat{x}_1)^2\, f_x(x)\, dx\right]$$

$$= (i_1 - \hat{x}_0)^2\, f_x(i_1) - (i_1 - \hat{x}_1)^2\, f_x(i_1) = 0$$

$$\Rightarrow (i_1 - \hat{x}_0)^2 - (i_1 - \hat{x}_1)^2 = 0$$

$$\Rightarrow i_1 = \frac{\hat{x}_0 + \hat{x}_1}{2}$$

Optimal one-bit quantizer: threshold

$$\frac{\partial \sigma_e^2}{\partial i_1} = \frac{\partial}{\partial i_1}\left[\int_A^{i_1} (x - \hat{x}_0)^2 \, f_x(x)\, dx + \int_{i_1}^{B} (x - \hat{x}_1)^2 \, f_x(x)\, dx\right]$$

$$= (i_1 - \hat{x}_0)^2 \, f_x(i_1) - (i_1 - \hat{x}_1)^2 \, f_x(i_1) = 0$$

$$\Rightarrow (i_1 - \hat{x}_0)^2 - (i_1 - \hat{x}_1)^2 = 0$$

$$\Rightarrow i_1 = \frac{\hat{x}_0 + \hat{x}_1}{2}$$

Optimal one-bit quantizer: threshold

$$\frac{\partial \sigma_e^2}{\partial i_1} = \frac{\partial}{\partial i_1}\left[\int_A^{i_1} (x - \hat{x}_0)^2 f_x(x)\, dx + \int_{i_1}^{B} (x - \hat{x}_1)^2 f_x(x)\, dx\right]$$

$$= (i_1 - \hat{x}_0)^2 f_x(i_1) - (i_1 - \hat{x}_1)^2 f_x(i_1) = 0$$

$$\Rightarrow (i_1 - \hat{x}_0)^2 - (i_1 - \hat{x}_1)^2 = 0$$

$$\Rightarrow i_1 = \frac{\hat{x}_0 + \hat{x}_1}{2}$$

Optimal one-bit quantizer: values

$$\frac{\partial \sigma_e^2}{\partial \hat{x}_0} = \frac{\partial}{\partial x_0} \int_A^{i_1} (x - \hat{x}_0)^2 \, f_x(x) \, dx$$

$$= \int_A^{i_1} 2(\hat{x}_0 - x) \, f_x(x) \, dx = 0$$

$$\Rightarrow \hat{x}_0 = \frac{\int_A^{i_1} x \, f_x(x) \, dx}{\int_A^{i_1} f_x(x) \, dx} \qquad \textit{(center of mass)}$$

$$\Rightarrow \hat{x}_1 = \frac{\int_{i_1}^{B} x \, f_x(x) \, dx}{\int_{i_1}^{B} f_x(x) \, dx}$$

Optimal one-bit quantizer: values

$$\frac{\partial \sigma_e^2}{\partial \hat{x}_0} = \frac{\partial}{\partial x_0} \int_A^{i_1} (x - \hat{x}_0)^2 \, f_x(x) \, dx$$

$$= \int_A^{i_1} 2(\hat{x}_0 - x) \, f_x(x) \, dx = 0$$

$$\Rightarrow \hat{x}_0 = \frac{\int_A^{i_1} x \, f_x(x) \, dx}{\int_A^{i_1} f_x(x) \, dx} \qquad \textit{(center of mass)}$$

$$\Rightarrow \hat{x}_1 = \frac{\int_{i_1}^{B} x \, f_x(x) \, dx}{\int_{i_1}^{B} f_x(x) \, dx}$$

Optimal one-bit quantizer: values

$$\frac{\partial \sigma_e^2}{\partial \hat{x}_0} = \frac{\partial}{\partial x_0} \int_A^{i_1} (x - \hat{x}_0)^2 \, f_x(x) \, dx$$

$$= \int_A^{i_1} 2(\hat{x}_0 - x) \, f_x(x) \, dx = 0$$

$$\Rightarrow \hat{x}_0 = \frac{\int_A^{i_1} x \, f_x(x) \, dx}{\int_A^{i_1} f_x(x) \, dx} \qquad \textit{(center of mass)}$$

$$\Rightarrow \hat{x}_1 = \frac{\int_{i_1}^{B} x \, f_x(x) \, dx}{\int_{i_1}^{B} f_x(x) \, dx}$$

For uniformly-distributed input

$$f_x(x) = \frac{1}{B - A}$$

$$\hat{x}_0 = \frac{\int_A^{i_1} x \, dx}{\int_A^{i_1} dx} = \frac{A + i_1}{2}$$

$$\hat{x}_1 = \frac{\int_{i_1}^{B} x \, dx}{\int_{i_1}^{B} dx} = \frac{i_1 + B}{2}$$

$$i_1 = \frac{\hat{x}_0 + \hat{x}_1}{2} = \frac{A + B}{2}$$

For uniformly-distributed input

$$f_x(x) = \frac{1}{B - A}$$

$$\hat{x}_0 = \frac{\int_A^{i_1} x \, dx}{\int_A^{i_1} dx} = \frac{A + i_1}{2}$$

$$\hat{x}_1 = \frac{\int_{i_1}^{B} x \, dx}{\int_{i_1}^{B} dx} = \frac{i_1 + B}{2}$$

$$i_1 = \frac{\hat{x}_0 + \hat{x}_1}{2} = \frac{A + B}{2}$$

For uniformly-distributed input

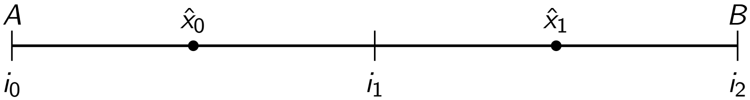
$$f_x(x) = \frac{1}{B - A}$$

$$\hat{x}_0 = \frac{\int_A^{i_1} x\,dx}{\int_A^{i_1} dx} = \frac{A + i_1}{2}$$

$$\hat{x}_1 = \frac{\int_{i_1}^{B} x\,dx}{\int_{i_1}^{B} dx} = \frac{i_1 + B}{2}$$

$$i_1 = \frac{\hat{x}_0 + \hat{x}_1}{2} = \frac{A + B}{2}$$

Optimal one-bit quantizer



Uniform quantization of uniform input

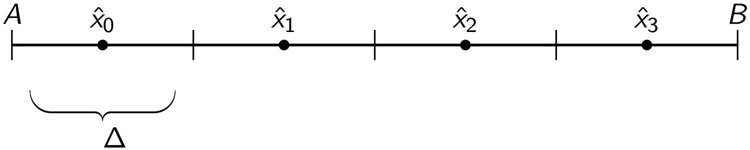
- simple but very general case
- optimal subdivision: 2^R *equal* intervals of width $\Delta = (B - A)2^{-R}$
- optimal quantization values: interval's midpoint

Uniform quantization of uniform input

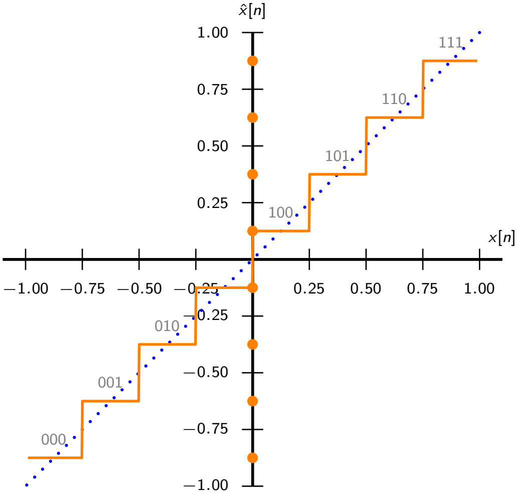
- simple but very general case
- optimal subdivision: 2^R *equal* intervals of width $\Delta = (B - A)2^{-R}$
- optimal quantization values: interval's midpoint

Uniform quantization of uniform input

- simple but very general case
- optimal subdivision: 2^R *equal* intervals of width $\Delta = (B - A)2^{-R}$
- optimal quantization values: interval's midpoint



Uniform 3-Bit quantization function



Uniform quantization of uniform input: error analysis

$$\sigma_e^2 = \int_A^B f_x(x)(\mathcal{Q}\{x\} - x)^2 \, dx$$

$$= \sum_{k=0}^{2^R-1} \int_{I_k} f_x(x)(\hat{x}_k - x)^2 \, dx$$

$$f_x(s) = \frac{1}{B-A}$$

$$\Delta = \frac{B-A}{2^R}$$

$$I_k = [A + k\Delta, A + (k+1)\Delta]$$

$$\hat{x}_k = A + k\Delta + \Delta/2$$

Uniform quantization of uniform input: error analysis

$$\sigma_e^2 = \int_A^B f_x(x)(\mathcal{Q}\{x\} - x)^2\,dx$$

$$= \sum_{k=0}^{2^R-1} \int_{I_k} f_x(x)(\hat{x}_k - x)^2\,dx$$

$$f_x(s) = \frac{1}{B-A}$$

$$\Delta = \frac{B-A}{2^R}$$

$$I_k = [A + k\Delta, A + (k+1)\Delta]$$

$$\hat{x}_k = A + k\Delta + \Delta/2$$

Uniform quantization of uniform input: error analysis

$$\sigma_e^2 = \int_A^B f_x(x)(\mathcal{Q}\{x\} - x)^2\, dx$$
$$= \sum_{k=0}^{2^R-1} \int_{I_k} f_x(x)(\hat{x}_k - x)^2\, dx$$

$$f_x(s) = \frac{1}{B-A}$$
$$\Delta = \frac{B-A}{2^R}$$
$$I_k = [A + k\Delta, A + (k+1)\Delta]$$
$$\hat{x}_k = A + k\Delta + \Delta/2$$

Uniform quantization of uniform input: error analysis

$$\sigma_e^2 = \sum_{k=0}^{2^R-1} \int_{A+k\Delta}^{A+k\Delta+\Delta} \frac{(A + k\Delta + \Delta/2 - x)^2}{B - A}\, dx$$

$$= 2^R \int_0^{\Delta} \frac{(\Delta/2 - x)^2}{B - A}\, dx$$

$$= \frac{\Delta^2}{12}$$

Uniform quantization of uniform input: error analysis

$$\sigma_e^2 = \sum_{k=0}^{2^R-1} \int_{A+k\Delta}^{A+k\Delta+\Delta} \frac{(A + k\Delta + \Delta/2 - x)^2}{B - A}\, dx$$

$$= 2^R \int_0^{\Delta} \frac{(\Delta/2 - x)^2}{B - A}\, dx$$

$$= \frac{\Delta^2}{12}$$

Uniform quantization of uniform input: error analysis

$$
\begin{aligned}
\sigma_e^2 &= \sum_{k=0}^{2^R-1} \int_{A+k\Delta}^{A+k\Delta+\Delta} \frac{(A+k\Delta+\Delta/2-x)^2}{B-A}\,dx \\
&= 2^R \int_0^{\Delta} \frac{(\Delta/2-x)^2}{B-A}\,dx \\
&= \frac{\Delta^2}{12}
\end{aligned}
$$

Error analysis of the quantization error

fundamental assumptions:

- signal and quantization error are uncorrelated (ok-ish)
- quantization error process is white (stretch)

quantization noise acts as additive white noise

Error analysis

- error energy

$$\sigma_e^2 = \Delta^2/12, \qquad \Delta = (B - A)/2^R$$

- signal energy

$$\sigma_x^2 = (B - A)^2/12$$

- signal to noise ratio

$$\mathrm{SNR} = 2^{2R}$$

- in dB

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} 2^{2R} \approx 6R \text{ dB}$$

Error analysis

- error energy

$$\sigma_e^2 = \Delta^2/12, \qquad \Delta = (B - A)/2^R$$

- signal energy

$$\sigma_x^2 = (B - A)^2/12$$

- signal to noise ratio

$$\mathrm{SNR} = 2^{2R}$$

- in dB

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} 2^{2R} \approx 6R \text{ dB}$$

Error analysis

- error energy

$$\sigma_e^2 = \Delta^2/12, \qquad \Delta = (B - A)/2^R$$

- signal energy

$$\sigma_x^2 = (B - A)^2/12$$

- signal to noise ratio

$$\mathrm{SNR} = 2^{2R}$$

- in dB

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} 2^{2R} \approx 6R \text{ dB}$$

Error analysis

- error energy

$$\sigma_e^2 = \Delta^2/12, \qquad \Delta = (B - A)/2^R$$

- signal energy

$$\sigma_x^2 = (B - A)^2/12$$

- signal to noise ratio

$$\mathrm{SNR} = 2^{2R}$$

- in dB

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} 2^{2R} \approx 6R \text{ dB}$$

The "6dB/bit" rule of thumb

- a compact disk has 16 bits/sample:

$$\text{max SNR} = 96\text{dB}$$

- a DVD has 24 bits/sample:

$$\text{max SNR} = 144\text{dB}$$

The "6dB/bit" rule of thumb

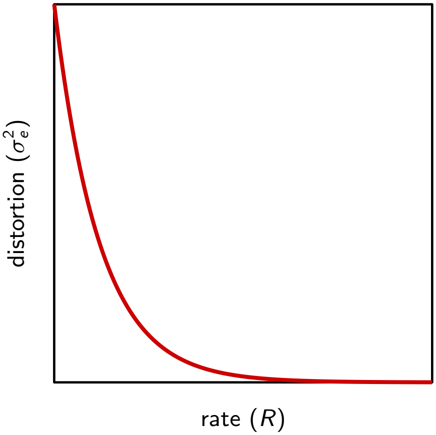
- a compact disk has 16 bits/sample:

$$\text{max SNR} = 96\text{dB}$$

- a DVD has 24 bits/sample:

$$\text{max SNR} = 144\text{dB}$$

Rate/Distortion Curve



Other quantization errors

If input is not bounded to $[A, B]$:

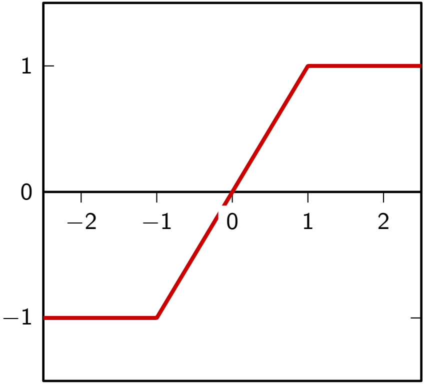
- clip samples to $[A, B]$: linear distortion (can be put to good use in guitar effects!)
- smoothly saturate input: this simulates the saturation curves of analog electronics

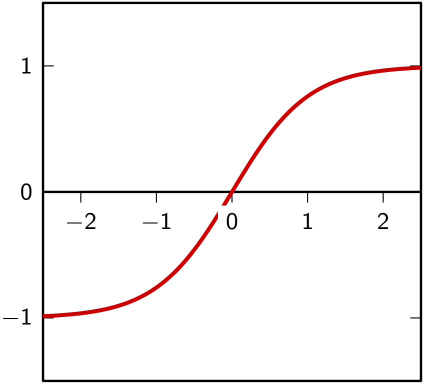
Other quantization errors

If input is not bounded to $[A, B]$:

- clip samples to $[A, B]$: linear distortion (can be put to good use in guitar effects!)
- smoothly saturate input: this simulates the saturation curves of analog electronics

Clipping vs saturation





Other quantization errors

If input is not uniform:

- use uniform quantizer and accept increased error.
  For instance, if input is Gaussian:

$$\sigma_e^2 = \frac{\sqrt{3}\pi}{2}\,\sigma^2\,\Delta^2$$

- use "companders"
- design optimal quantizer for input distribution, if known (Lloyd-Max algorithm)

Other quantization errors

If input is not uniform:

- use uniform quantizer and accept increased error.
  For instance, if input is Gaussian:

$$\sigma_e^2 = \frac{\sqrt{3}\pi}{2}\,\sigma^2\,\Delta^2$$

- use "companders"
- design optimal quantizer for input distribution, if known (Lloyd-Max algorithm)

Other quantization errors

If input is not uniform:

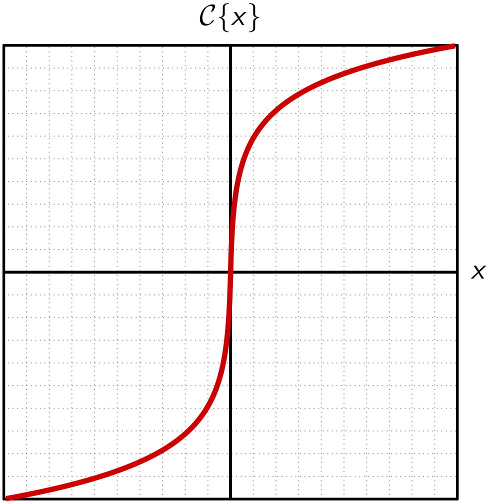
- use uniform quantizer and accept increased error.
  For instance, if input is Gaussian:

$$\sigma_e^2 = \frac{\sqrt{3}\pi}{2}\,\sigma^2\,\Delta^2$$

- use "companders"
- design optimal quantizer for input distribution, if known (Lloyd-Max algorithm)

μ-law compander

$$\mathcal{C}\{x[n]\} = \operatorname{sgn}(x[n])\frac{\ln(1+\mu|x[n]|)}{\ln(1+\mu)}$$



Lloyd-Max Quantizer design

$$\sigma_e^2 = \sum_{k=0}^{2^R-1} \int_{i_k}^{i_{k+1}} (x - \hat{x}_k)^2 \, f_x(x) \, dx$$

$$\text{A)} \quad \frac{\partial \sigma_e^2}{\partial \hat{x}_k} = 0 \Rightarrow \hat{x}_k = \frac{\displaystyle\int_{i_{k-1}}^{i_k} x \, f_x(x) \, dx}{\displaystyle\int_{i_{k-1}}^{i_k} f_x(x) \, dx}$$

$$\text{B)} \quad \frac{\partial \sigma_e^2}{\partial i_k} = 0 \Rightarrow i_k = \frac{\hat{x}_{k-1} + \hat{x}_k}{2}$$

Lloyd-Max Quantizer design

- start with a guess for the i_k
- solve A and B iteratively until convergence

A/D and D/A converters

Overview:

- Analog-to-digital (A/D) conversion
- Digital-to-analog (D/A) conversion

Overview:

- Analog-to-digital (A/D) conversion
- Digital-to-analog (D/A) conversion

From analog to digital

- sampling discretizes time
- quantization discretized amplitude
- how is it done in practice?

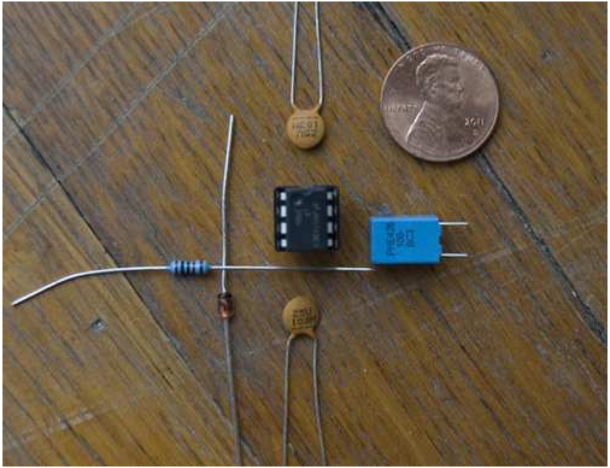
From analog to digital

- sampling discretizes time
- quantization discretized amplitude
- how is it done in practice?

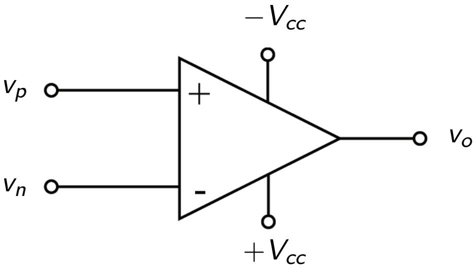
From analog to digital

- sampling discretizes time
- quantization discretized amplitude
- how is it done in practice?

From analog to digital

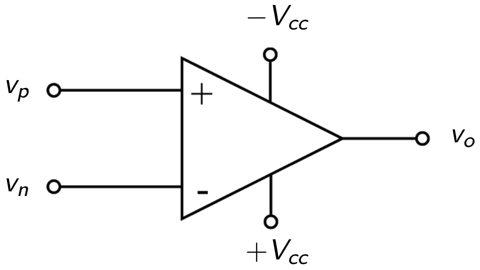


A tiny bit of electronics: the op-amp



$$v_o = G(v_p - v_n)$$

A tiny bit of electronics: the op-amp



$$v_o = G(v_p - v_n)$$

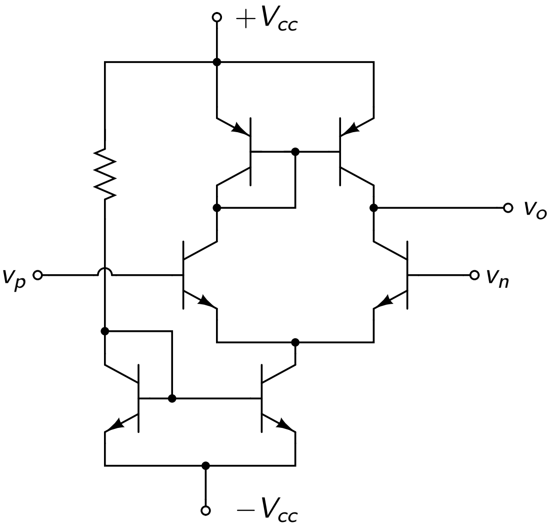
The two key properties

- infinite input gain ($G \approx \infty$)
- zero input current

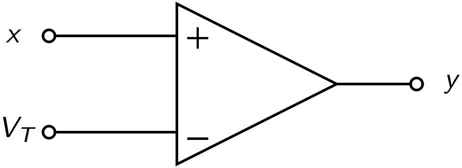
The two key properties

- infinite input gain ($G \approx \infty$)
- zero input current

Inside the box

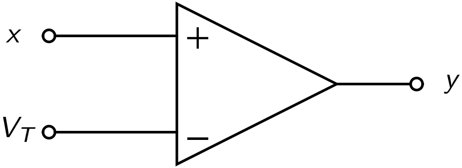


The op-amp in open loop: comparator



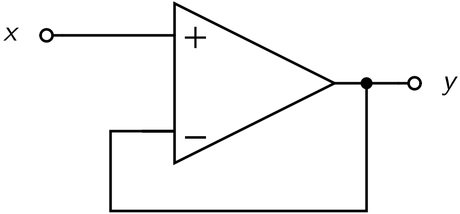
$$y = \begin{cases} +V_{cc} & \text{if } x > V_T \\ -V_{cc} & \text{if } x < V_T \end{cases}$$

The op-amp in open loop: comparator



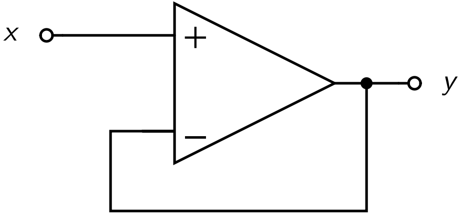
$$y = \begin{cases} +V_{cc} & \text{if } x > V_T \\ -V_{cc} & \text{if } x < V_T \end{cases}$$

The op-amp in closed loop: buffer



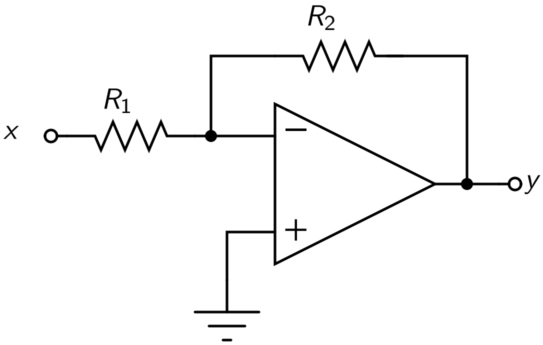
$$y = x$$

The op-amp in closed loop: buffer

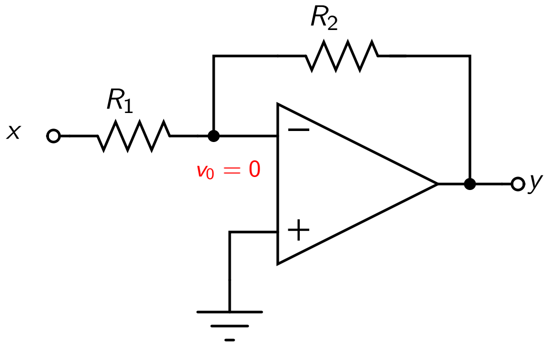


$$y = x$$

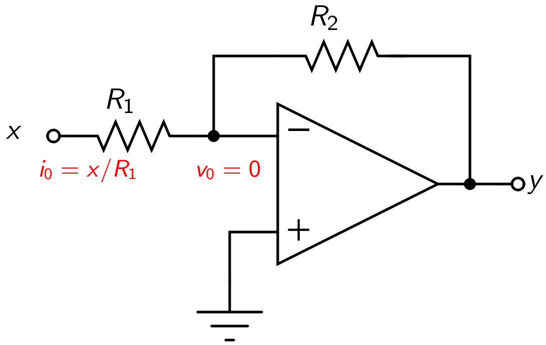
The op-amp in closed loop: inverting amplifier



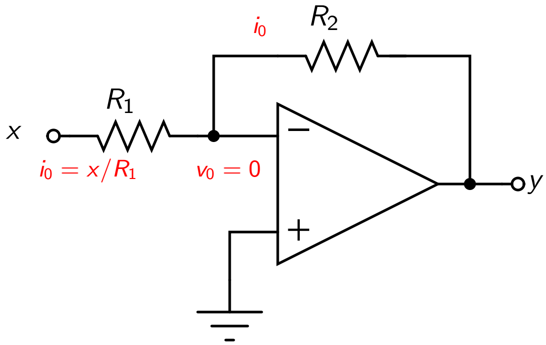
The op-amp in closed loop: inverting amplifier



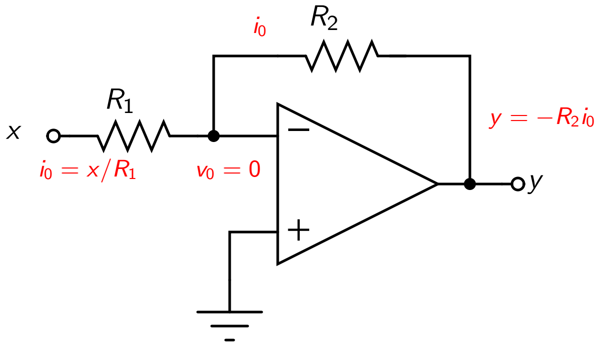
The op-amp in closed loop: inverting amplifier



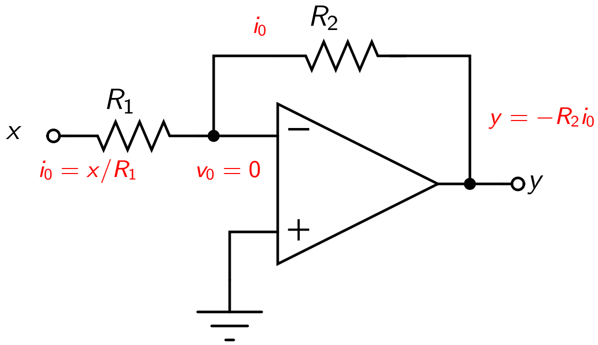
The op-amp in closed loop: inverting amplifier



The op-amp in closed loop: inverting amplifier

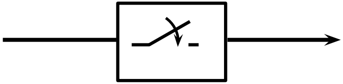


The op-amp in closed loop: inverting amplifier

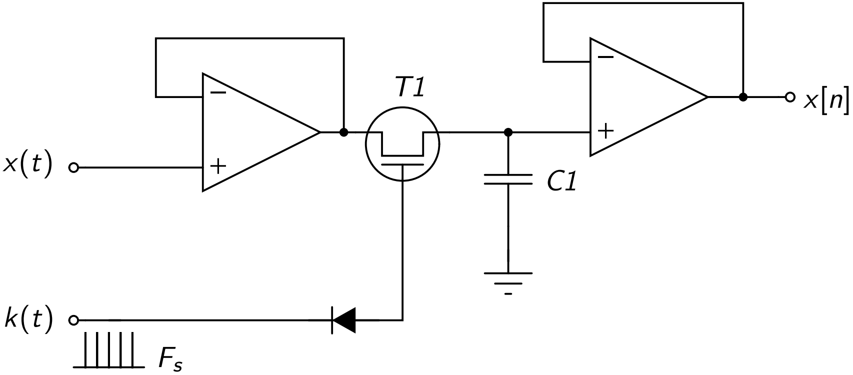


$$y = -(R_2/R_1)x$$

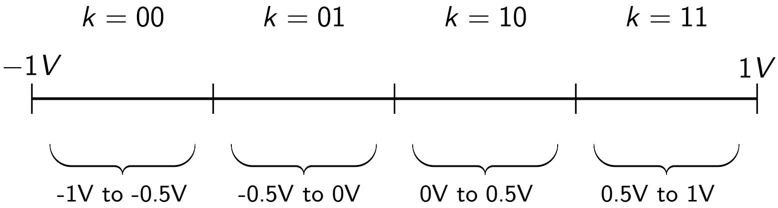
A/D Converter: Sampling

$x(t)$  $x[n]$

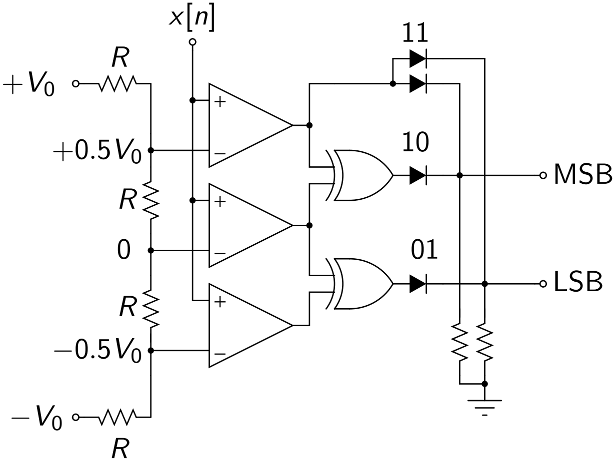
A/D Converter: Sample & Hold



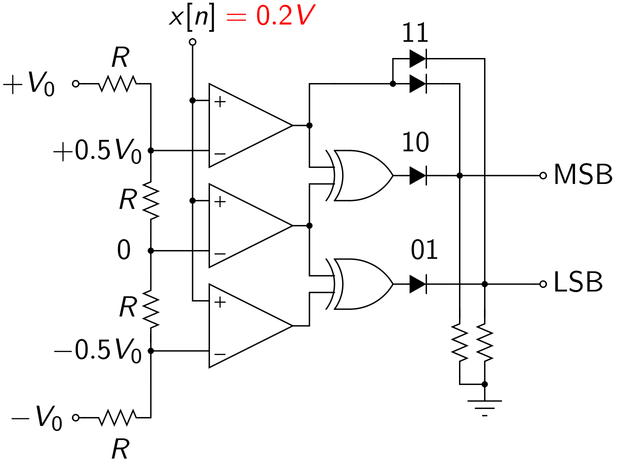
A/D Converter: Quantization



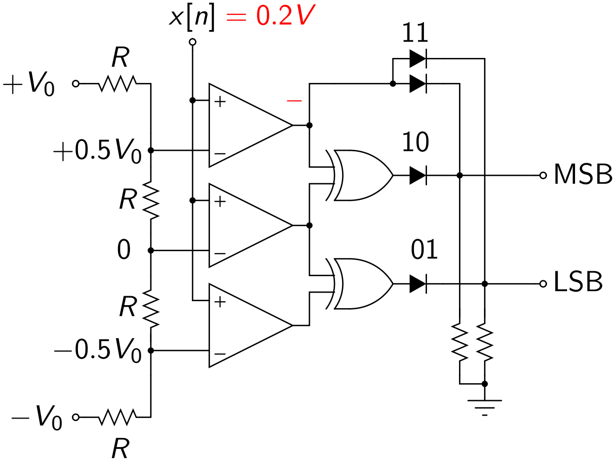
A/D Converter: 2-Bit Quantizer



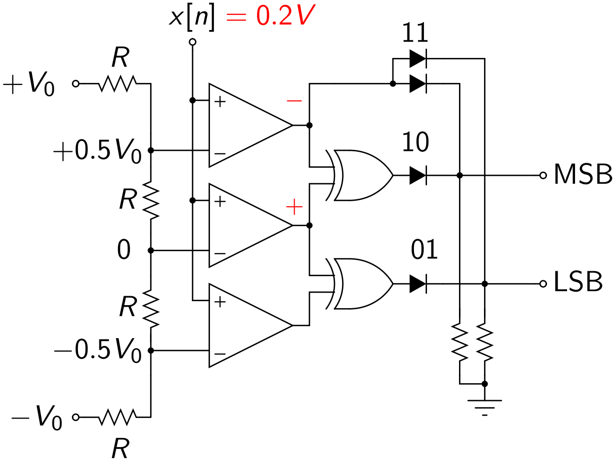
A/D Converter: 2-Bit Quantizer



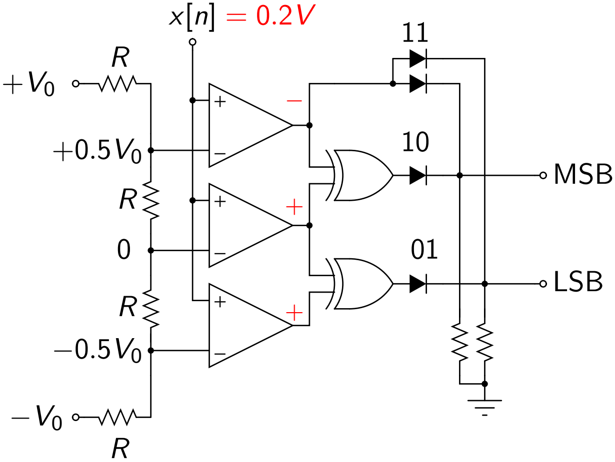
A/D Converter: 2-Bit Quantizer



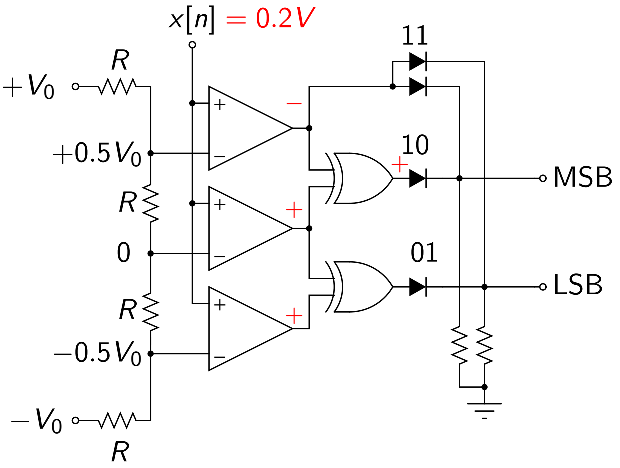
A/D Converter: 2-Bit Quantizer



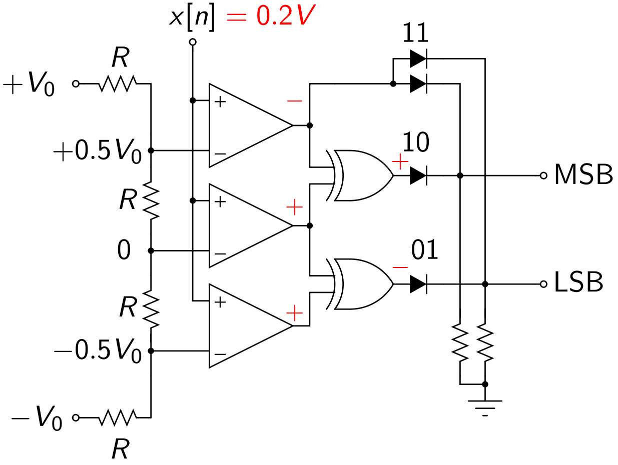
A/D Converter: 2-Bit Quantizer



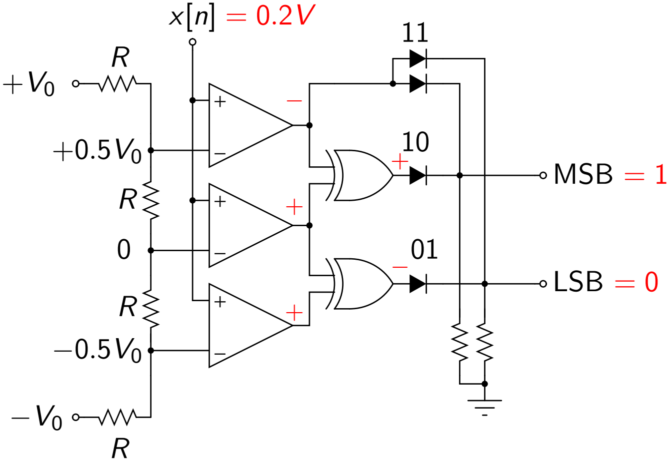
A/D Converter: 2-Bit Quantizer



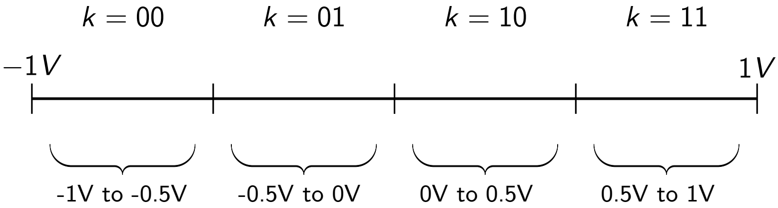
A/D Converter: 2-Bit Quantizer



A/D Converter: 2-Bit Quantizer



A/D Converter: Quantization

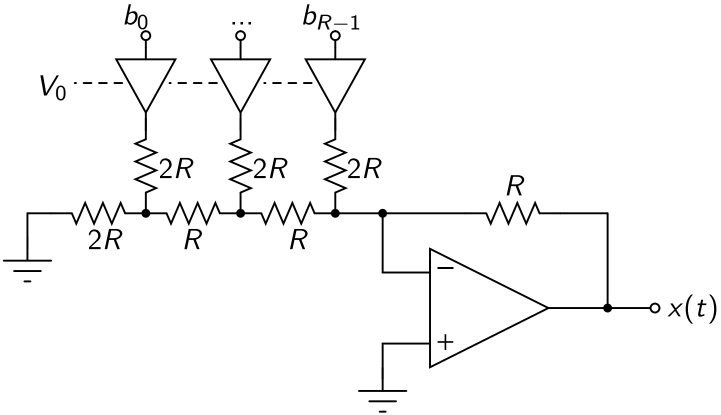


D/A Converter

$$x_B[n] = b_{R-1} b_{R-2} \ldots b_1 b_0$$

$$\hat{x}[n] = \sum_{k=0}^{R-1} \frac{V_0}{2^k} b_k$$

D/A Converter



oversampling

Oversampling

- oversampled A/D
  - reduce quantization error
- oversampled D/A
  - use cheaper hardware for interpolation

Oversampling

- oversampled A/D
  - reduce quantization error
- oversampled D/A
  - use cheaper hardware for interpolation

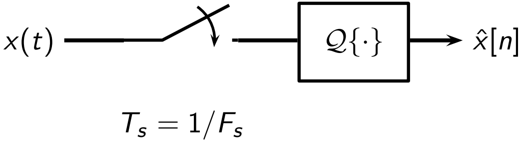
Oversampling

- oversampled A/D
  - reduce quantization error
- oversampled D/A
  - use cheaper hardware for interpolation

Oversampling

- oversampled A/D
  - reduce quantization error
- oversampled D/A
  - use cheaper hardware for interpolation

Oversampled A/D



$$\hat{x}[n] = x[n] + e[n]$$

Oversampled A/D

Key assumptions:

$$e[n] \text{ i.i.d. process, independent of } x[n]$$

$$P_e(e^{j\omega}) = \frac{\Delta^2}{12} \qquad \text{over } [-\pi, \pi] \text{ (no aliasing)}$$

Key observation:

$$X(e^{j\omega}) = F_s X\left(\frac{\omega}{2\pi} F_s\right)$$

Oversampled A/D

Key assumptions:

$e[n]$ i.i.d. process, independent of $x[n]$

$$P_e(e^{j\omega}) = \frac{\Delta^2}{12} \qquad \text{over } [-\pi, \pi] \text{ (no aliasing)}$$

Key observation:

$$X(e^{j\omega}) = F_s X\left(\frac{\omega}{2\pi} F_s\right)$$

Oversampled A/D

Key assumptions:

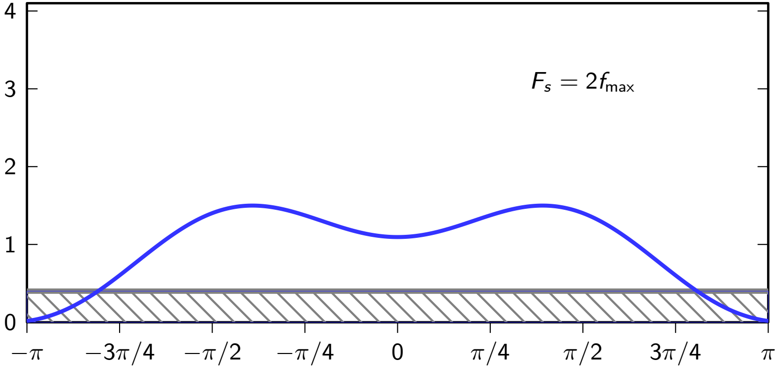
$$e[n] \text{ i.i.d. process, independent of } x[n]$$

$$P_e(e^{j\omega}) = \frac{\Delta^2}{12} \qquad \text{over } [-\pi, \pi] \text{ (no aliasing)}$$

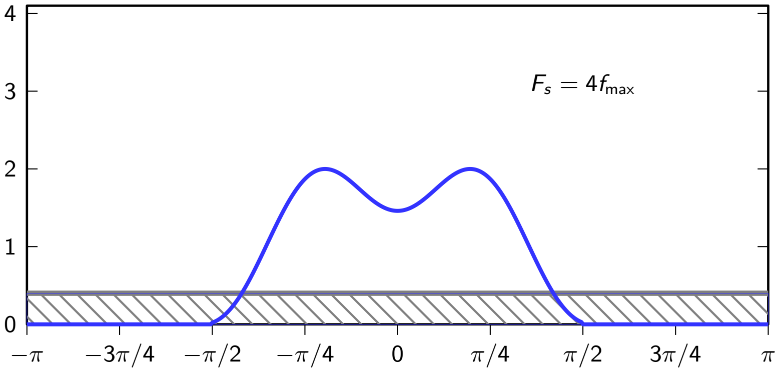
Key observation:

$$X(e^{j\omega}) = F_s\, X\left(\frac{\omega}{2\pi}F_s\right)$$

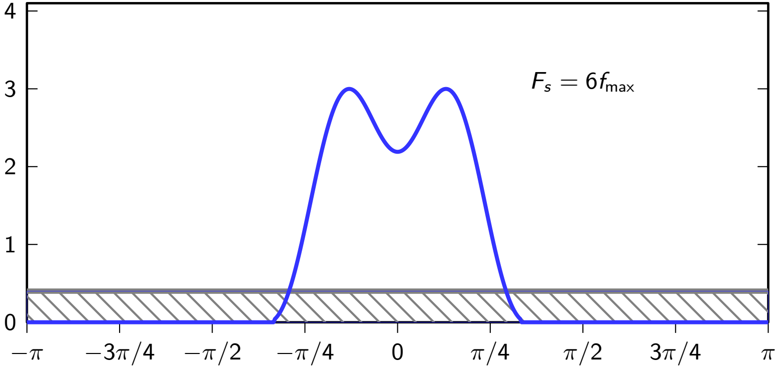
Oversampled A/D



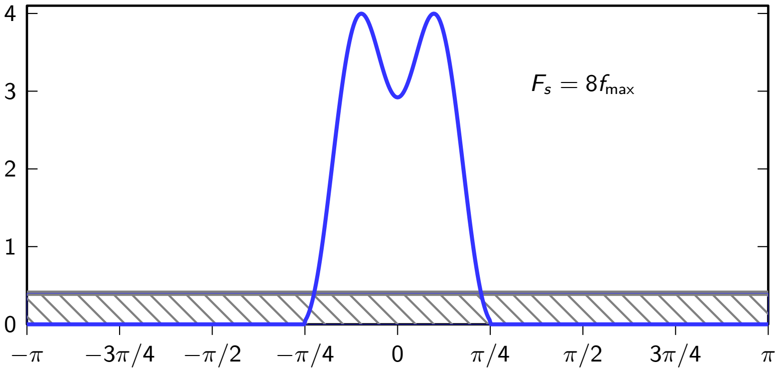
Oversampled A/D



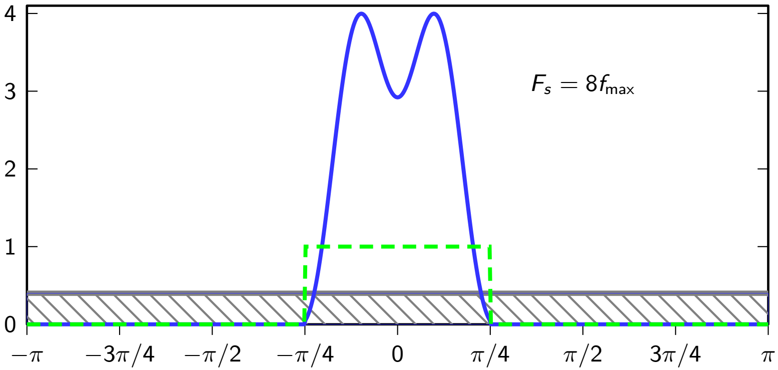
Oversampled A/D



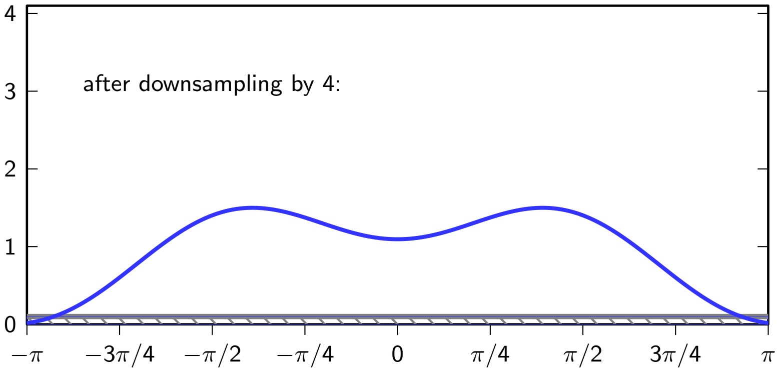
Oversampled A/D



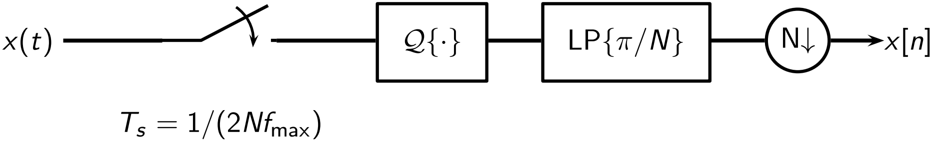
Oversampled A/D



Oversampled A/D



Oversampled A/D



- $\mathrm{SNR}_O \approx N\,\mathrm{SNR}$
- 3dB per octave (doubling of F_s)
- but key assumption (independence) breaks down fast...

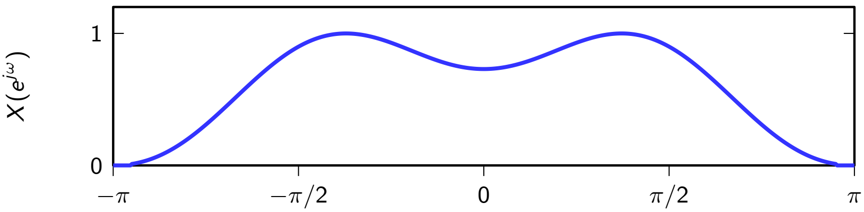
Oversampled D/A

Oversampled D/A

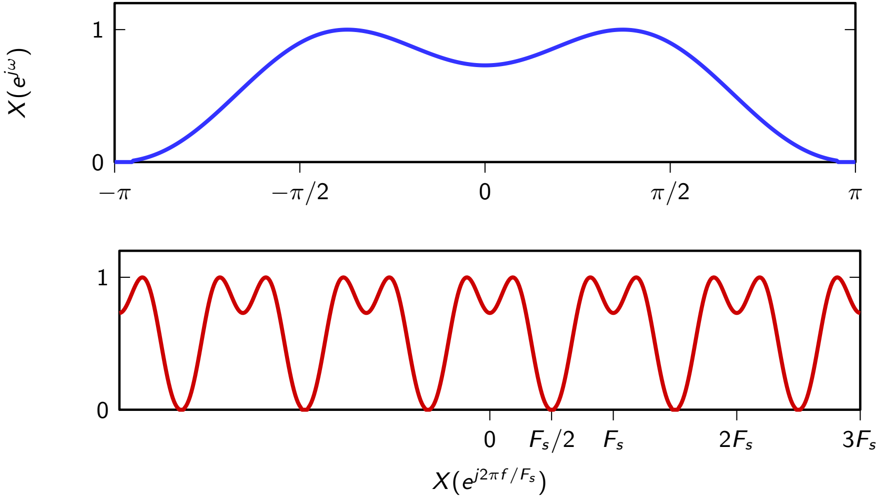
Sinc interpolation:

$$X_c(f) = \frac{1}{F_s} X(e^{j2\pi f/F_s}) \operatorname{rect}\left(\frac{f}{F_s}\right)$$

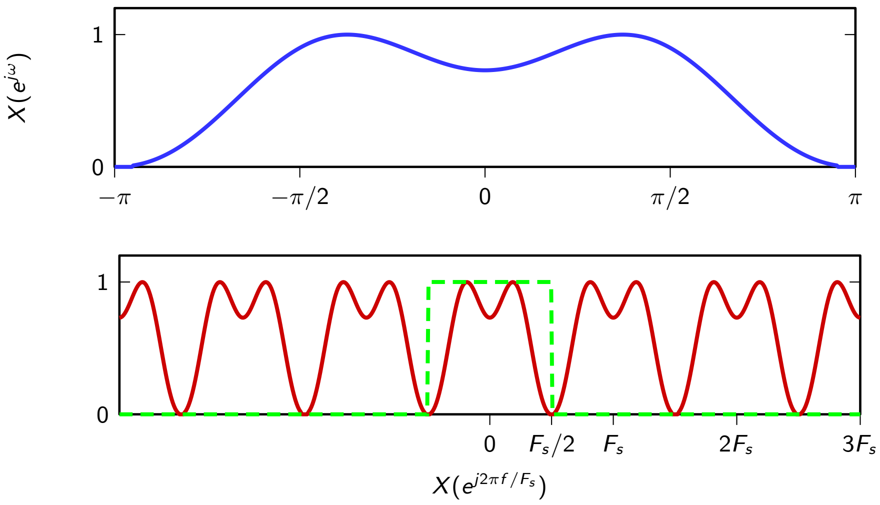
Sinc interpolation



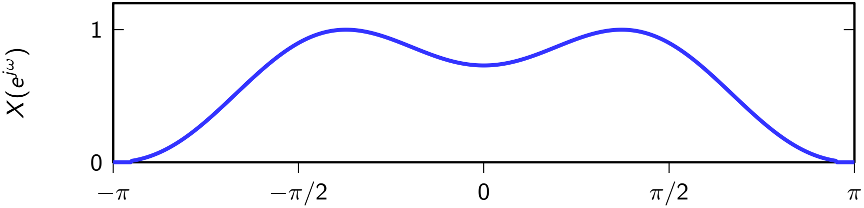
Sinc interpolation

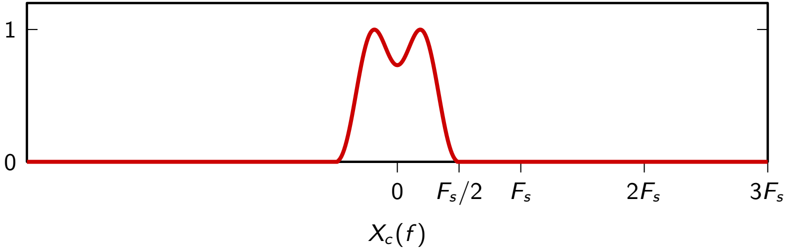


Sinc interpolation



Sinc interpolation





Oversampled D/A

In general:

$$X_c(f) = \frac{1}{F_s} X(e^{j2\pi f/F_s}) \, I\left(\frac{f}{F_s}\right)$$

The cheapest (hence most common) interpolator is the zero-order hold:

$$i_0(t) = \mathrm{rect}(t)$$

$$I_0(f) = \mathrm{sinc}(f)$$

Oversampled D/A

In general:

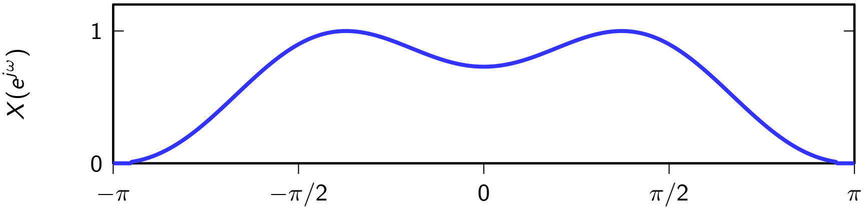
$$X_c(f) = \frac{1}{F_s} X(e^{j2\pi f/F_s})\, I\left(\frac{f}{F_s}\right)$$

The cheapest (hence most common) interpolator is the zero-order hold:

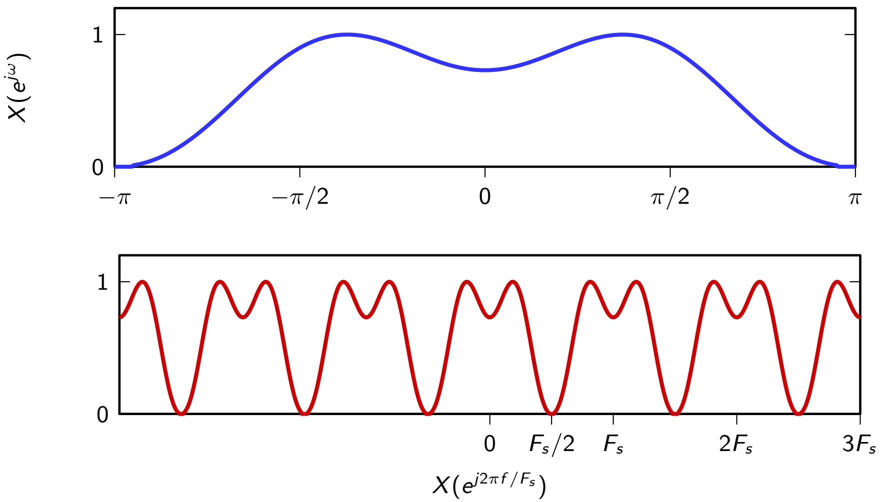
$$i_0(t) = \mathrm{rect}(t)$$

$$I_0(f) = \mathrm{sinc}(f)$$

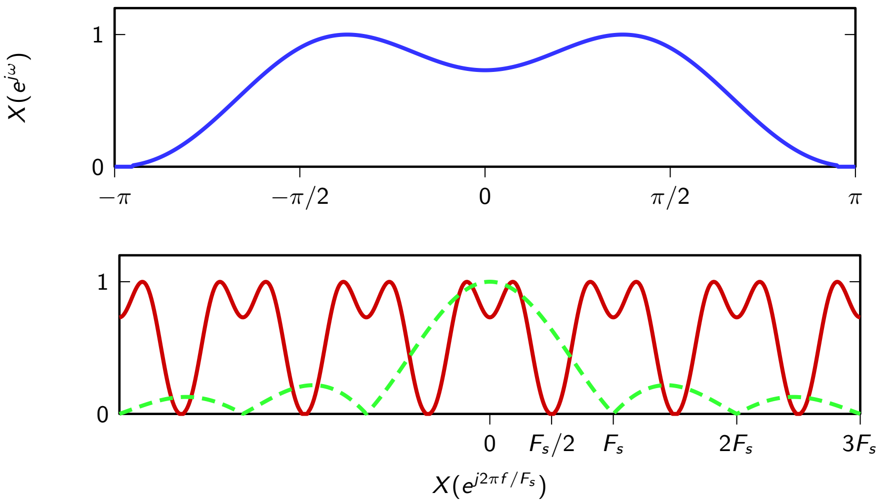
ZOH interpolation



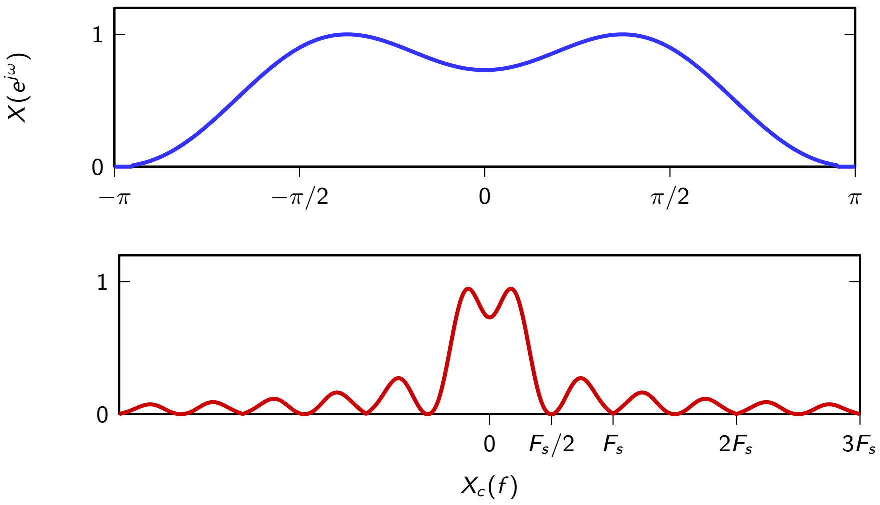
ZOH interpolation



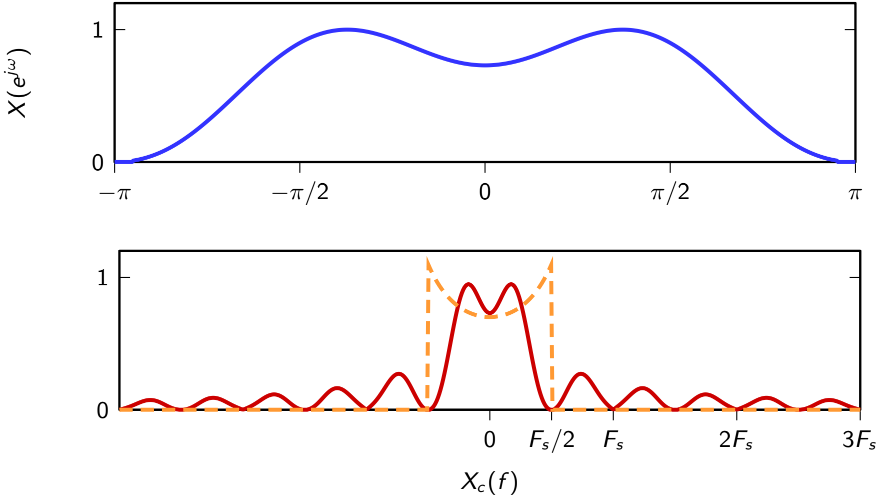
ZOH interpolation



ZOH interpolation



ZOH interpolation

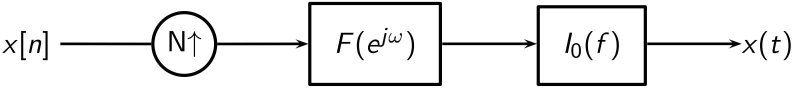


Oversampled A/D

key problems:

- we need to undo the in-band distortion in the analog domain
- we have a significant out-of-band distortion
- only advantage: minimal D/A rate

Oversampled D/A



$$F_o = NF_s$$

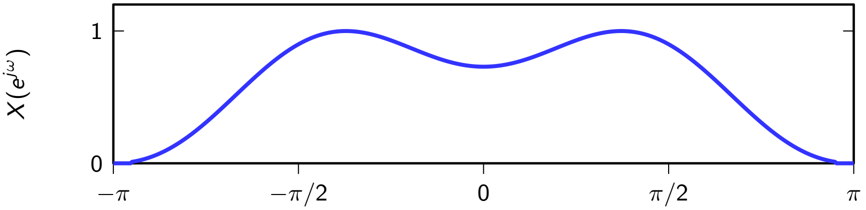
Oversampled D/A

consider a N-upsampled and interpolated version of $x[n]$:

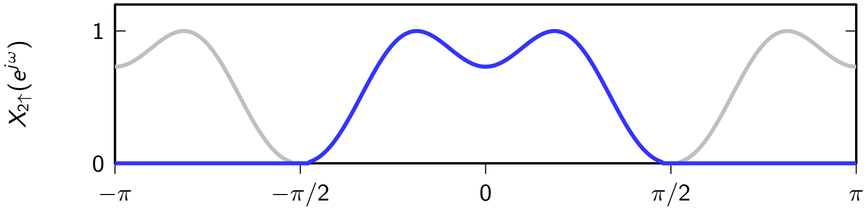
$$X_o(e^{j\omega}) = X_{N\uparrow}(e^{j\omega})\,F(e^{j\omega}) = X(e^{j\omega N})\,F(e^{j\omega})$$

- $F(e^{j\omega}) = N\ \mathrm{rect}(\omega N/(2\pi))\,C(e^{j\omega})$
- rect matches the upsampler
- $C(e^{j\omega})$ compensates for zoh in-band distortion

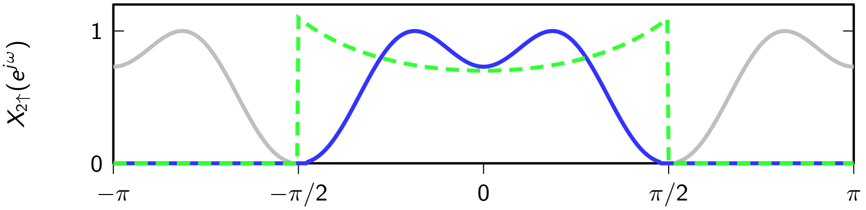
Oversampled D/A ($N = 2$)



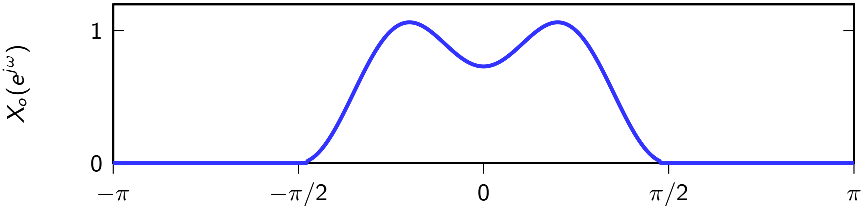
Oversampled D/A ($N = 2$)



Oversampled D/A ($N = 2$)



Oversampled D/A ($N = 2$)



Oversampled D/A

interpolate $x_o[n]$ with $F_o = NF_s$:

$$X_o(f) = \frac{1}{F_o} X_o(e^{j2\pi f/F_o}) \, I_0\left(\frac{f}{F_o}\right)$$

$$= \frac{1}{F_o} \left[X(e^{j\omega N}) \, F(e^{j\omega})\right]_{\omega=2\pi f/F_o} I_0\left(\frac{f}{F_o}\right)$$

$$= \frac{N}{F_o} \left[X(e^{j\omega N}) \, \mathrm{rect}\left(\frac{\omega N}{2\pi}\right) \, C(e^{j\omega})\right]_{\omega=2\pi f/F_o} I_0\left(\frac{f}{F_o}\right)$$

$$= \frac{1}{F_s} X(e^{j2\pi f/F_s}) \, \mathrm{rect}\left(\frac{f}{F_s}\right) \, C(e^{j2\pi f/F_o}) \, I_0\left(\frac{f}{F_o}\right)$$

$$= X(f) \quad \text{for } |f| < F_s/2$$

Oversampled D/A

interpolate $x_o[n]$ with $F_o = NF_s$:

$$X_o(f) = \frac{1}{F_o} X_o(e^{j2\pi f/F_o})\, I_0\left(\frac{f}{F_o}\right)$$

$$= \frac{1}{F_o} \left[X(e^{j\omega N})\, F(e^{j\omega})\right]_{\omega=2\pi f/F_o} I_0\left(\frac{f}{F_o}\right)$$

$$= \frac{N}{F_o} \left[X(e^{j\omega N}) \operatorname{rect}\left(\frac{\omega N}{2\pi}\right) C(e^{j\omega})\right]_{\omega=2\pi f/F_o} I_0\left(\frac{f}{F_o}\right)$$

$$= \frac{1}{F_s} X(e^{j2\pi f/F_s}) \operatorname{rect}\left(\frac{f}{F_s}\right) C(e^{j2\pi f/F_o})\, I_0\left(\frac{f}{F_o}\right)$$

$$= X(f) \quad \text{for } |f| < F_s/2$$

Oversampled D/A

interpolate $x_o[n]$ with $F_o = NF_s$:

$$\begin{aligned}
X_o(f) &= \frac{1}{F_o} X_o(e^{j2\pi f/F_o})\, I_0\left(\frac{f}{F_o}\right) \\
&= \frac{1}{F_o} \left[X(e^{j\omega N})\, F(e^{j\omega})\right]_{\omega=2\pi f/F_o} I_0\left(\frac{f}{F_o}\right) \\
&= \frac{N}{F_o} \left[X(e^{j\omega N})\ \mathrm{rect}\left(\frac{\omega N}{2\pi}\right)\ C(e^{j\omega})\right]_{\omega=2\pi f/F_o} I_0\left(\frac{f}{F_o}\right) \\
&= \frac{1}{F_s} X(e^{j2\pi f/F_s})\ \mathrm{rect}\left(\frac{f}{F_s}\right)\ C(e^{j2\pi f/F_o})\, I_0\left(\frac{f}{F_o}\right) \\
&= X(f) \quad \text{for } |f| < F_s/2
\end{aligned}$$

Oversampled D/A

interpolate $x_o[n]$ with $F_o = NF_s$:

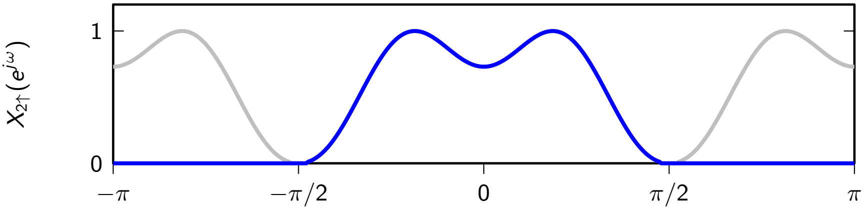
$$
\begin{aligned}
X_o(f) &= \frac{1}{F_o} X_o(e^{j2\pi f/F_o})\, I_0\left(\frac{f}{F_o}\right) \\
&= \frac{1}{F_o} \left[X(e^{j\omega N})\, F(e^{j\omega})\right]_{\omega=2\pi f/F_o} I_0\left(\frac{f}{F_o}\right) \\
&= \frac{N}{F_o} \left[X(e^{j\omega N})\ \mathrm{rect}\left(\frac{\omega N}{2\pi}\right)\ C(e^{j\omega})\right]_{\omega=2\pi f/F_o} I_0\left(\frac{f}{F_o}\right) \\
&= \frac{1}{F_s} X(e^{j2\pi f/F_s})\ \mathrm{rect}\left(\frac{f}{F_s}\right)\ C(e^{j2\pi f/F_o})\, I_0\left(\frac{f}{F_o}\right) \\
&= X(f) \quad \text{for } |f| < F_s/2
\end{aligned}
$$

Oversampled D/A

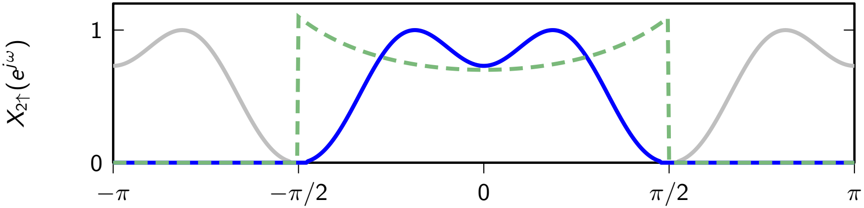
interpolate $x_o[n]$ with $F_o = NF_s$:

$$
\begin{aligned}
X_o(f) &= \frac{1}{F_o} X_o(e^{j2\pi f/F_o})\, I_0\left(\frac{f}{F_o}\right) \\
&= \frac{1}{F_o} \left[X(e^{j\omega N})\, F(e^{j\omega})\right]_{\omega = 2\pi f/F_o} I_0\left(\frac{f}{F_o}\right) \\
&= \frac{N}{F_o} \left[X(e^{j\omega N}) \operatorname{rect}\left(\frac{\omega N}{2\pi}\right) C(e^{j\omega})\right]_{\omega = 2\pi f/F_o} I_0\left(\frac{f}{F_o}\right) \\
&= \frac{1}{F_s} X(e^{j2\pi f/F_s}) \operatorname{rect}\left(\frac{f}{F_s}\right) C(e^{j2\pi f/F_o})\, I_0\left(\frac{f}{F_o}\right) \\
&= X(f) \quad \text{for } |f| < F_s/2
\end{aligned}
$$

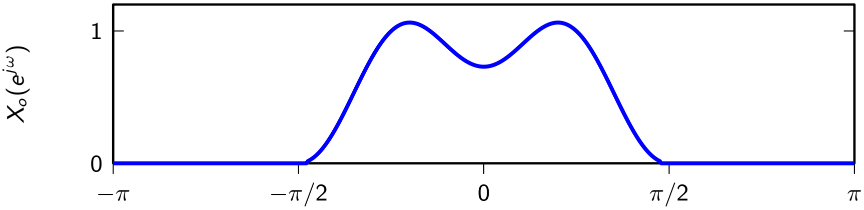
Oversampled D/A, using a ZOH ($N = 2$)



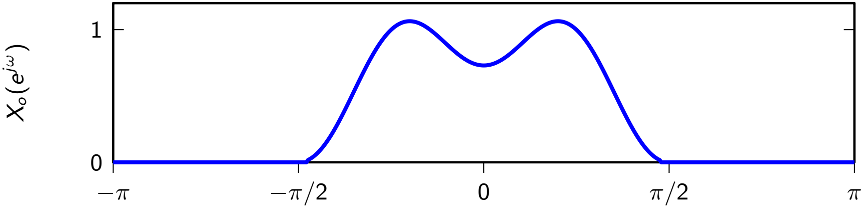
Oversampled D/A, using a ZOH ($N = 2$)

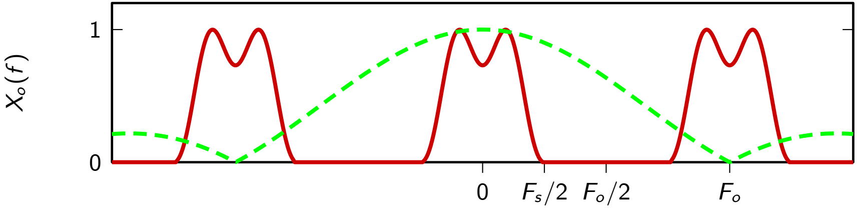


Oversampled D/A, using a ZOH ($N = 2$)

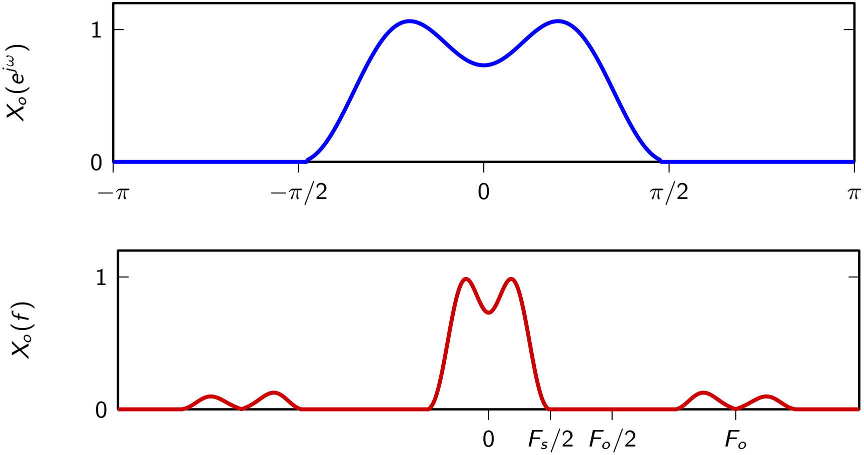


Oversampled D/A, using a ZOH ($N = 2$)

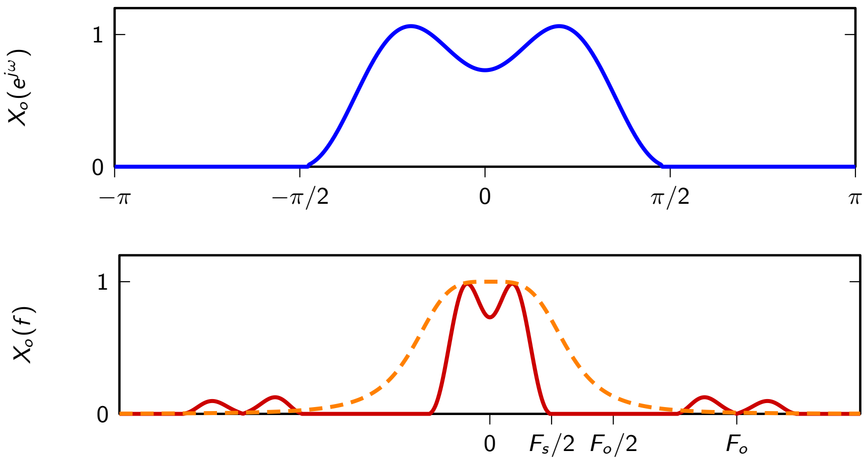




Oversampled D/A, using a ZOH ($N = 2$)



Oversampled D/A, using a ZOH ($N = 2$)



Oversampled A/D

key points:

- we can pre-compensate the in-band distortion in the digital domain
- we can interpolate with a cheap ZOH
- the higher the upsampling, the cheaper the analog lowpass needed to eliminate out-of-band distortion
- only price: higher D/A rate