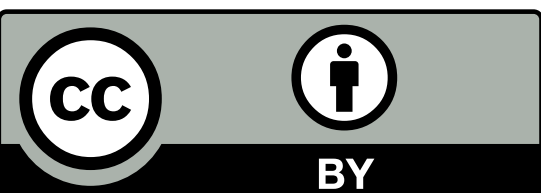


# Quantization

## Digital Audio Coding



SEBASTIEN.BOISGERAULT@MINES-PARISTECH.FR

# Scalar Quantizer

$$[\cdot] : \mathbb{R} \rightarrow \mathbb{R}$$

countable range:  $|\{[x], x \in \mathbb{R}\}| \leq |\mathbb{N}|$

idempotent mapping:  $\forall x \in \mathbb{R}, [[x]] = [x]$

**Example:** rounding functions  $\mathbb{R} \rightarrow \mathbb{Z} \subset \mathbb{R}$

$$\lfloor \cdot \rfloor, \lceil \cdot \rceil, [\cdot],$$

NUMPY: floor, ceil, round\_

# Forward/Inverse Quantizers

**forward quantizer:** mapping  $i[\cdot] : \mathbb{R} \rightarrow \mathbb{Z}$

such that  $i[x] = i[y]$  iff  $[x] = [y]$

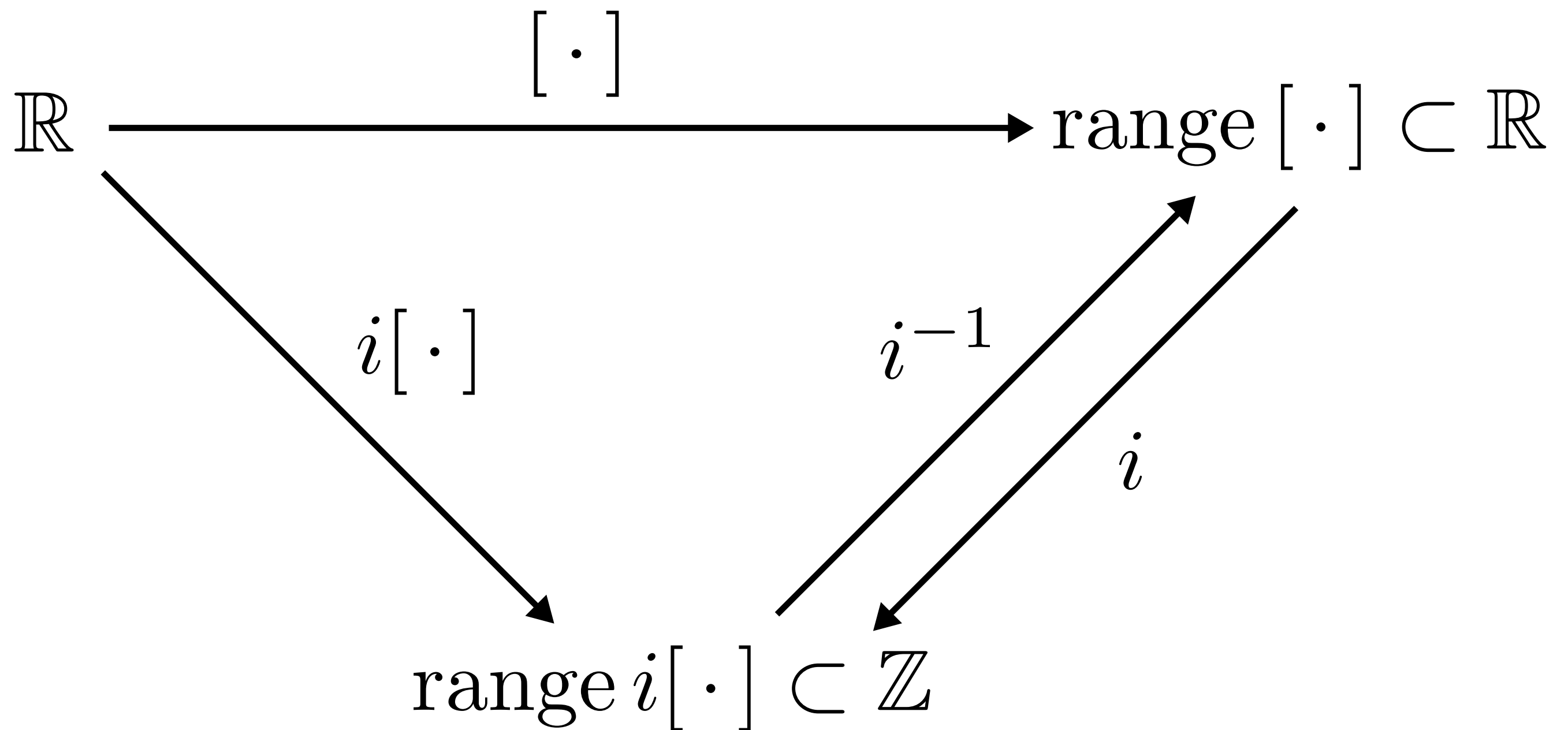
$i[\cdot] = i \circ [\cdot]$  where  $i : \text{range } [\cdot] \rightarrow \mathbb{Z}$  is into

**inverse quantizer:**  $i^{-1} : \text{range } i \subset \mathbb{Z} \rightarrow \mathbb{R}$

$i^{-1}$  is a left inverse of  $i$ .

$$\forall x \in \mathbb{R}, (i^{-1} \circ i)[x] = [x]$$

# Forward/Inverse Quantizers

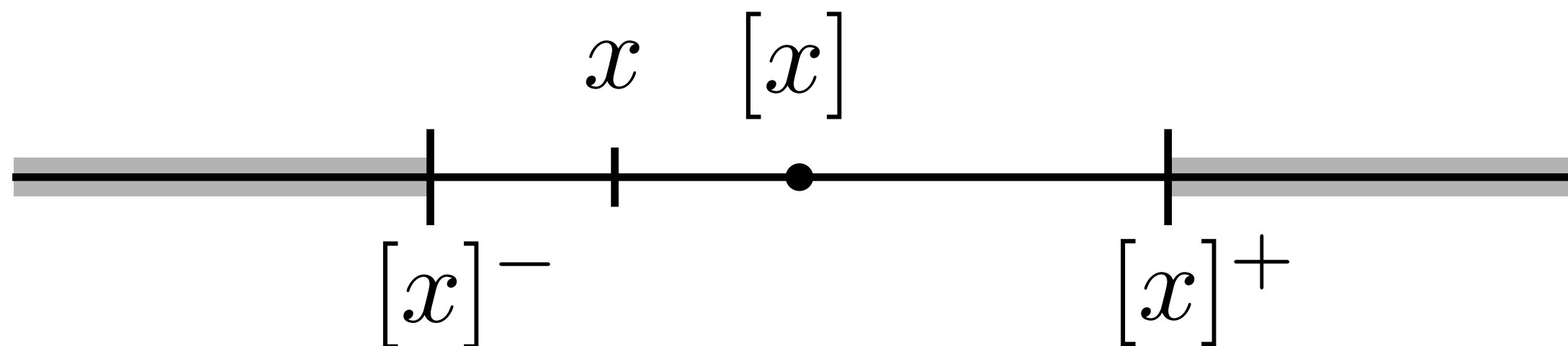


# Quantization Step

Assume that for every  $x \in \mathbb{R}$ ,

$$V_x = \{y \in \mathbb{R}, [y] = [x]\}$$

is an interval.



We define the **decision values**:

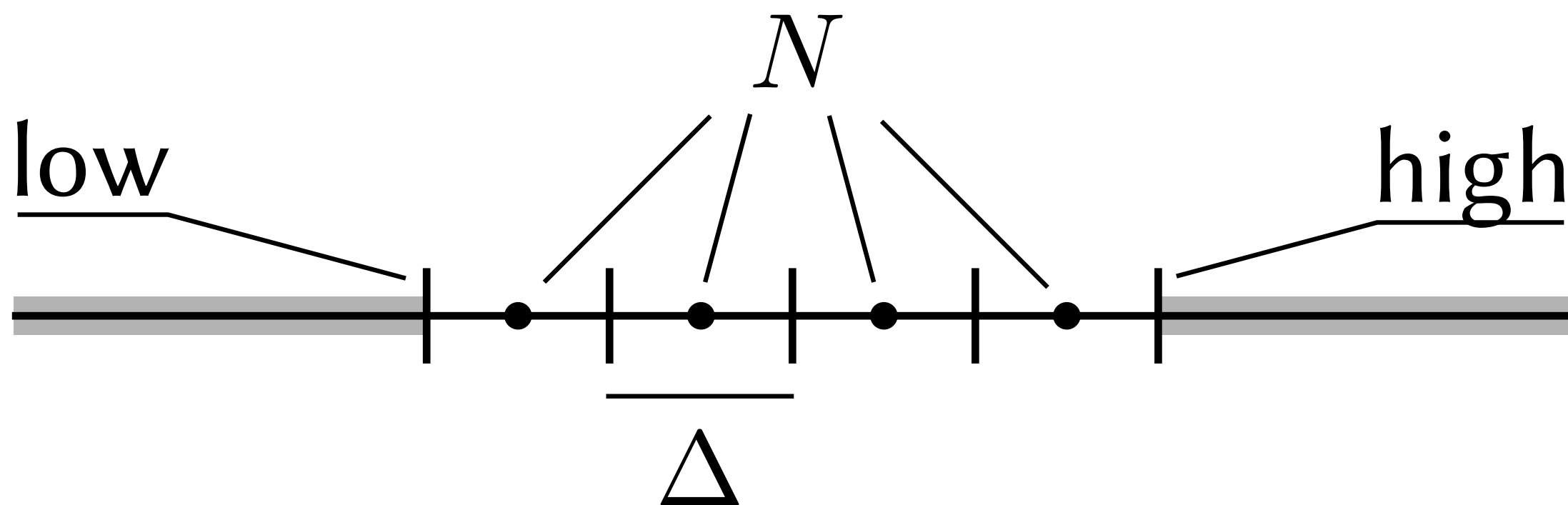
$$[x]^- = \inf V_x \quad \text{and} \quad [x]^+ = \sup V_x$$

and the **quantization step**:

$$\Delta(x) = [x]^+ - [x]^-$$

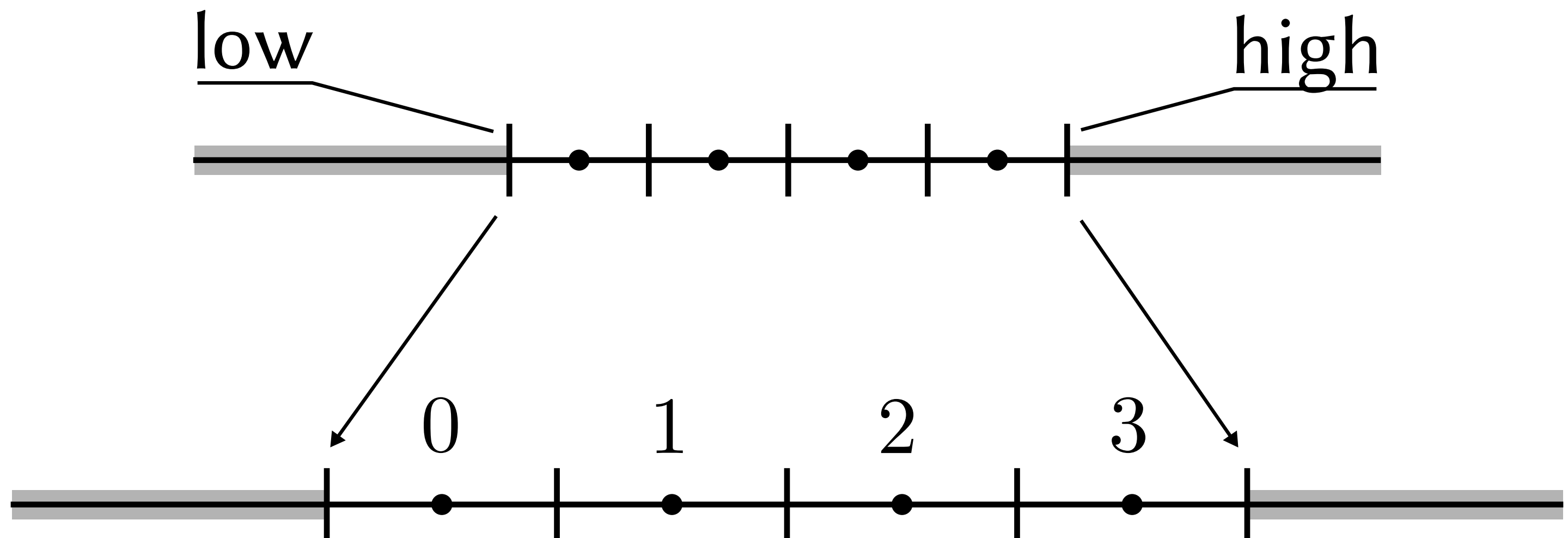
# Uniform Quantizer

```
class Uniform(Quantizer):  
    def __init__(self, low=0.0, high=1.0, N=2**8):  
        self.low, self.high = float(low), float(high)  
        self.N = N  
        self.delta = (self.high - self.low) / self.N
```

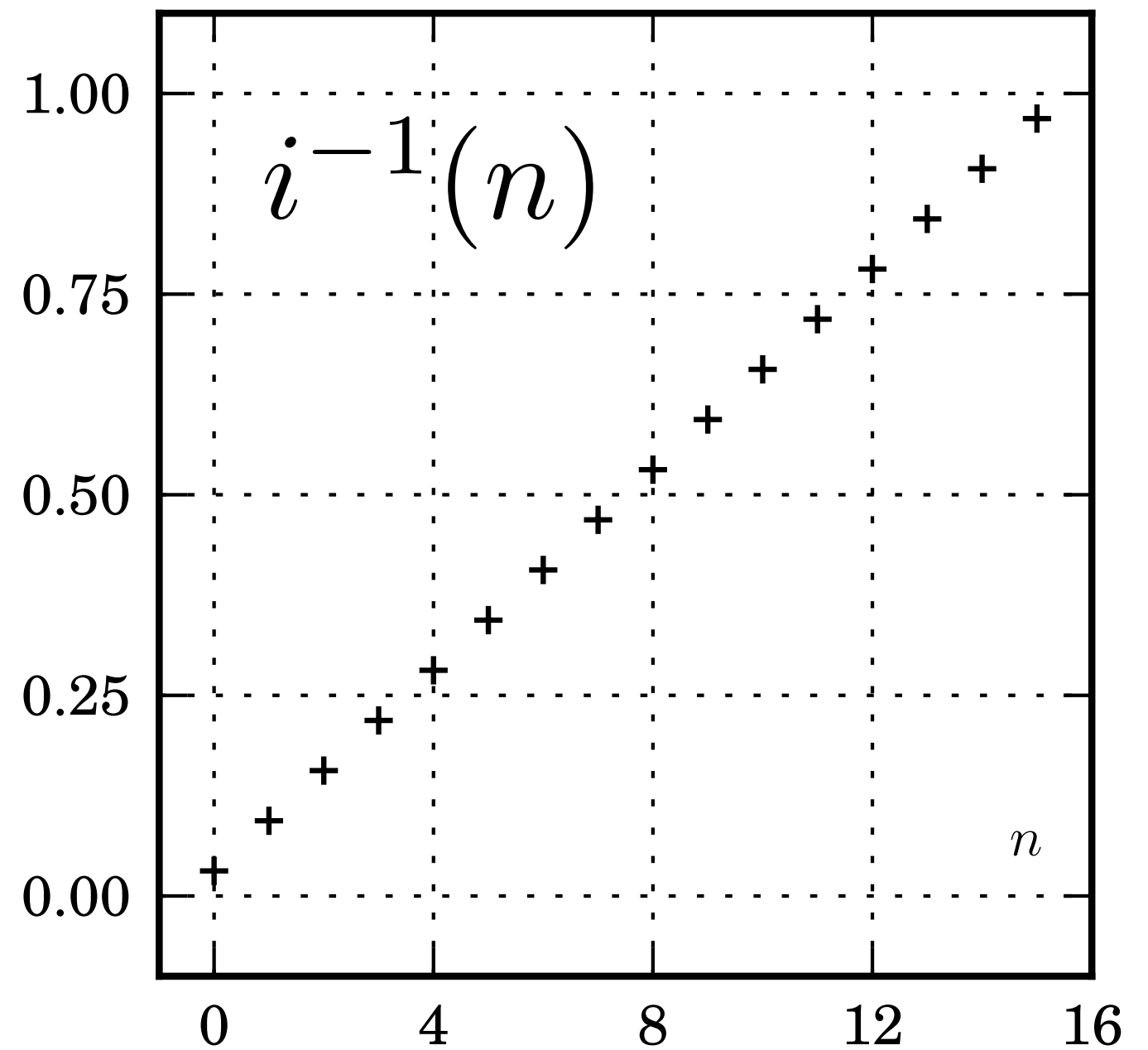
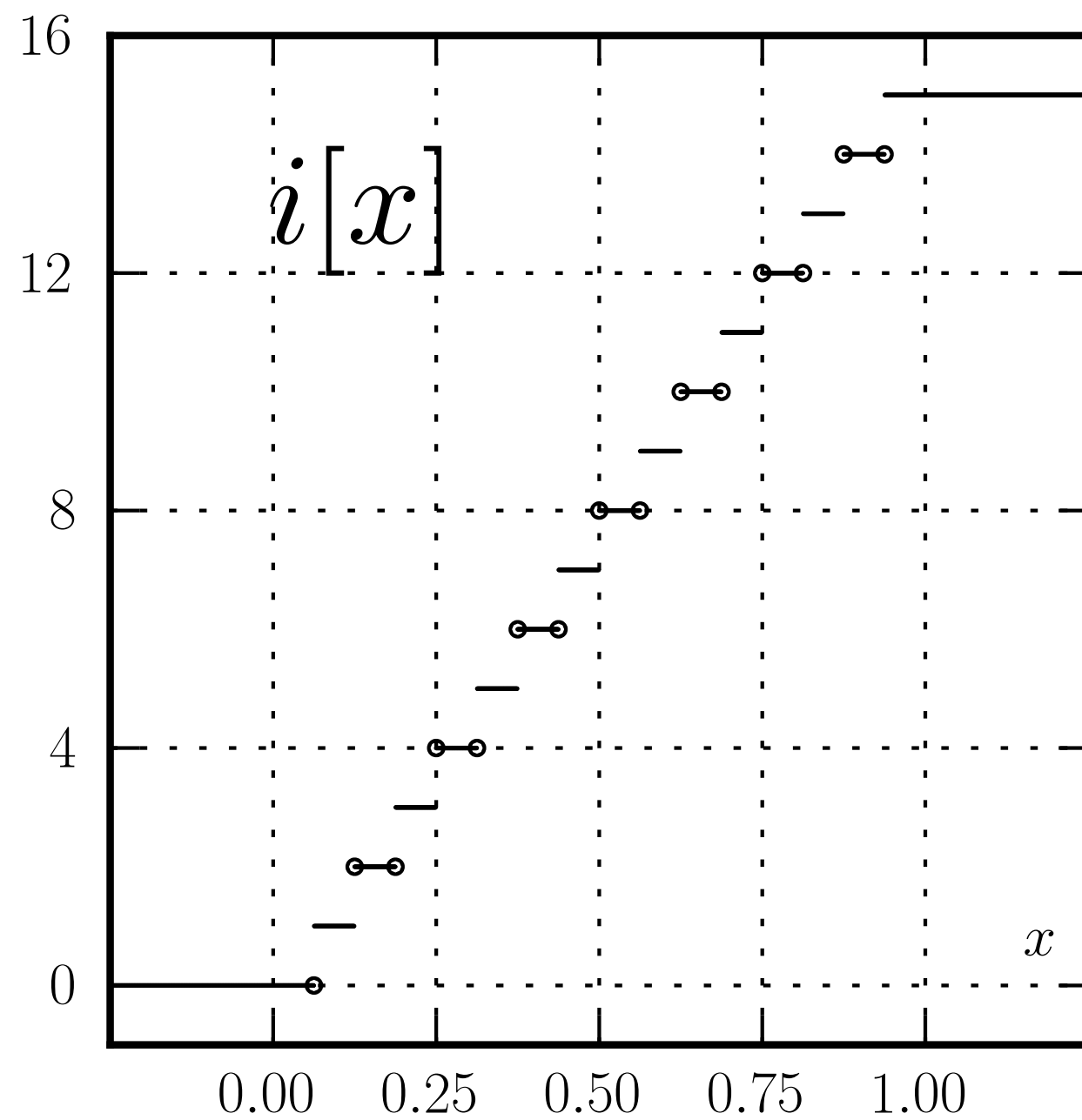


# Uniform Quantizer

```
class Uniform(Quantizer): (continued)
    def encode(self, data):
        data = clip(self.data, self.low, self.high)
        flints = round_((self.data - self.low) / self.delta - 0.5)
        return array(flints, dtype=long)
    def decode(self, i):
        return self.low + (i + 0.5) * self.delta
```



# Uniform Quantizer



Uniform(low=0.0, high=1.0, N=2\*\*4)



# Uniform Quantizer

```
>>> uniform = Uniform(low=-1.0, high=1.0, N=2**4)
```

```
>>> uniform(1.0)
```

0.9375

```
>>> uniform(0.0)
```

0.0625 ————— mid-rise

```
>>> uniform = Uniform(low=-1.0, high=1.0, N=2**4-1)
```

```
>>> uniform(1.0)
```

0.9333333333333333

```
>>> uniform(0.0)
```

0.0 ————— mid-tread

# Random Variables

$X \in \mathbb{R}$ , density  $p$ .

$$\begin{aligned} P([X] = [x]) &= P_X \{y \in \mathbb{R}, [y] = [x]\} \\ &= \int_{[x]^-}^{[x]^+} p(y) dy \end{aligned}$$

**High Resolution assumption:**

$$\approx p(x) \times \Delta(x)$$

# Optimal Quantizer

## Criteria: Entropy

The **entropy** of  $[X]$  is maximal when all events

$$[X] = [x], \quad x \in \mathbb{R}$$

are equally likely, that is when:

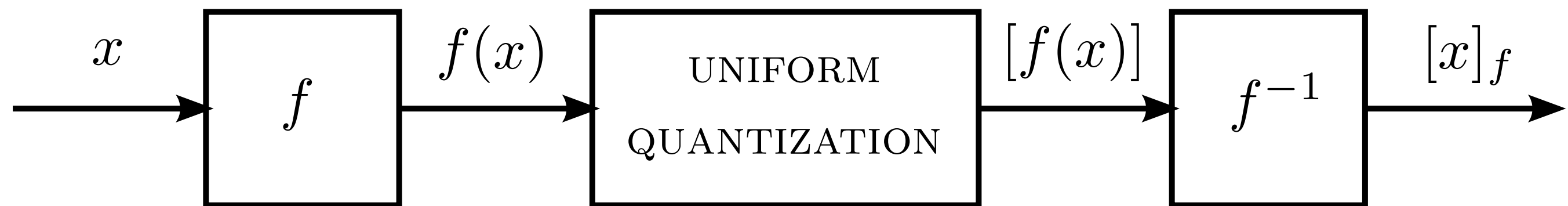
$$\Delta(x) \propto \frac{1}{p(x)}$$

# Nonlinear Quantizers

Select a **characteristic function**  $f$  and define

$$[\cdot]_f = f^{-1} \circ [\cdot] \circ f$$

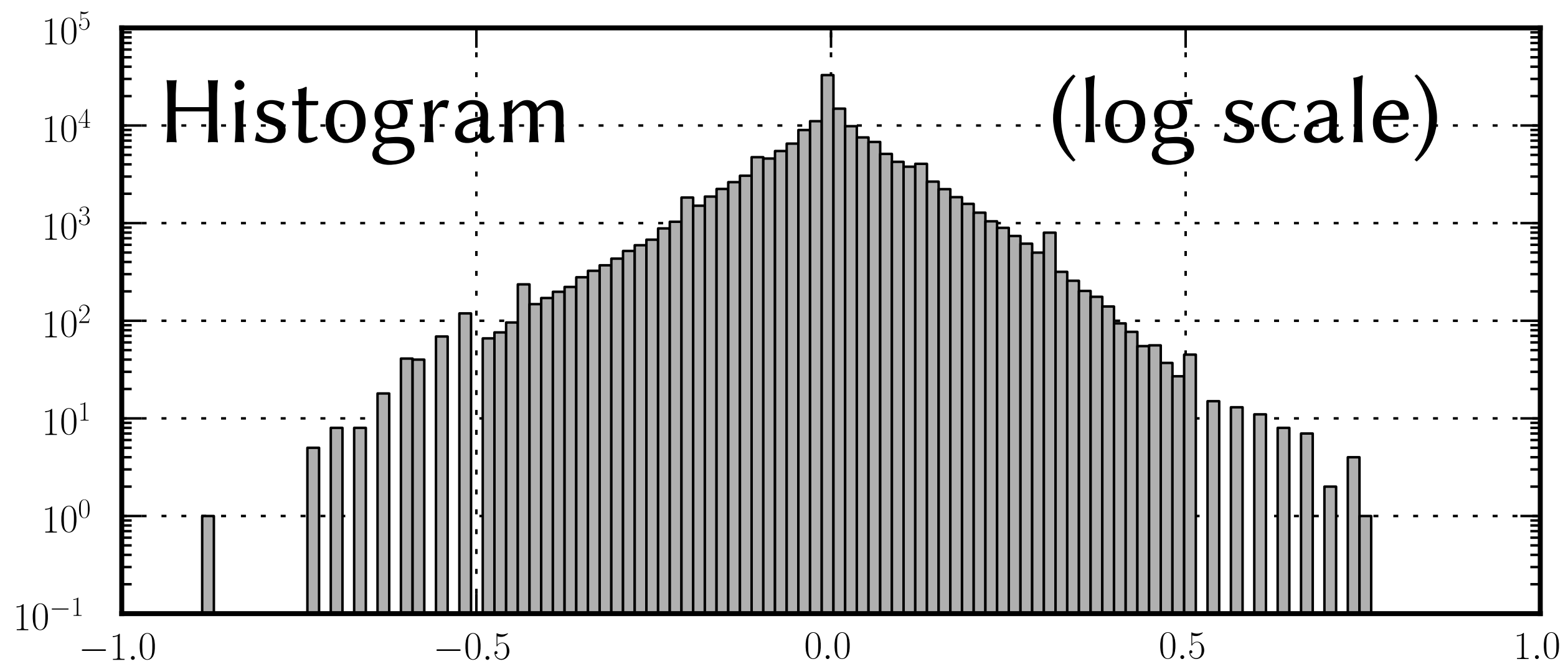
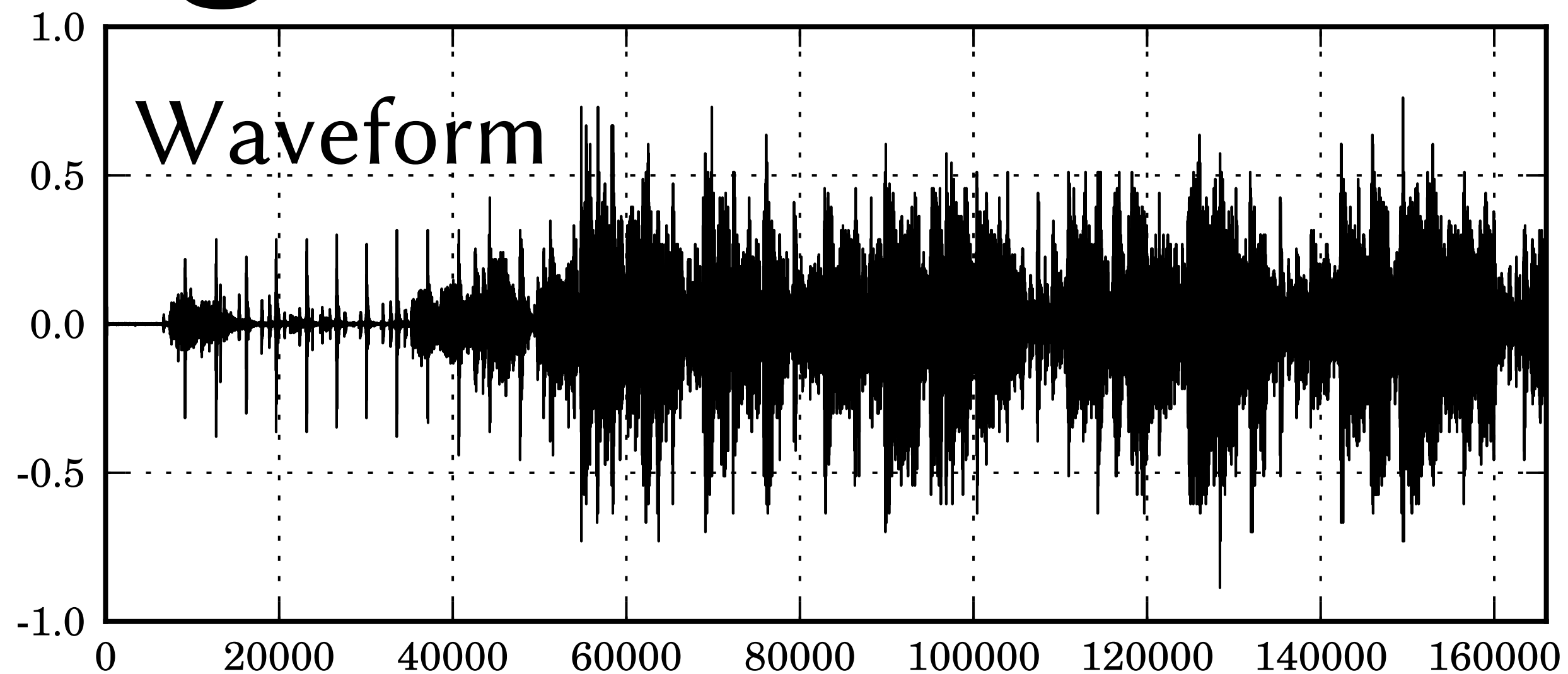
where  $[\cdot]$  is a uniform quantizer.



Under the high resolution assumption:

$$\Delta_f(x) \propto \frac{1}{f'(x)} \quad \left( \Delta_f(x) = \frac{\Delta_{[\cdot]}}{f'(x)} \right)$$

# "legende.au" (NeXT)



# Nonlinear Quantizer

## Maximal Entropy

If we model the distribution of values with:

$$p(x) \propto \exp(-a|x|)$$

the optimal quantizer satisfies:

$$\Delta(x) \propto \exp(a|x|) \text{ and } f'(x) \propto \exp(-a|x|)$$

Setting  $f(0) = 0$  and  $f'(0) = a$  yields:

$$f(x) = \operatorname{sgn}(x)(1 - e^{-a|x|})$$

$$f^{-1}(x) = -\frac{\operatorname{sgn}(x)}{a} \log(1 - |x|)$$

# Nonlinear Quantizer

Setting  $f(0) = 0$  and  $f'(0) = a$  yields:

$$f(\mathbb{R}) = (-1, 1)$$

Given implementations **f** and **f\_inv** of the characteristic function and of its inverse, we define the optimal 8-bit quantizer with:

```
uniform = Uniform(low=-1.0, high=1.0, N=2**8-1)  
quantizer = NonLinear(f, f_inv, uniform)
```

# Noise and SNR

Given a random value  $X$  and a quantizer  $[\cdot]$ , the quantizer **noise**  $B$  is defined by:

$$[X] = X + B$$

and the **signal-to-noise ratio (SNR)** by:

$$\text{SNR}^2 = \frac{\mathbb{E} X^2}{\mathbb{E} B^2}$$

or in **decibels** by:

$$\text{SNR [dB]} = 10 \log_{10} \text{SNR}^2$$



# SNR: Number of Bits

Consider a  $n$ -bit nonlinear quantizer with

$$f([-1, 1]) = [-1, 1]$$

The high-resolution assumption yields

$$\mathbb{E} B^2 \propto \mathbb{E} \Delta(X)^2 \quad \text{and} \quad \Delta(x) = \frac{2^{-n+1}}{f'(x)}$$

and as a consequence

$$\text{SNR [dB]} \approx 6.0 \times n + c(f)$$

# Optimal Quantizer

## Criteria: SNR

The best characteristic function is solution of:

$$\min_{f'} \int_{-1}^1 \frac{1}{f'(x)^2} p(x) dx$$

subject to  $f(1) - f(-1) = 2$

Solution:  $f'(x) \propto p(x)^{1/3}$

(Reminder:  $f'(x) \propto p(x)$  optimal for the entropy.)

# Logarithmic Quantizers

Consider the probability distribution:

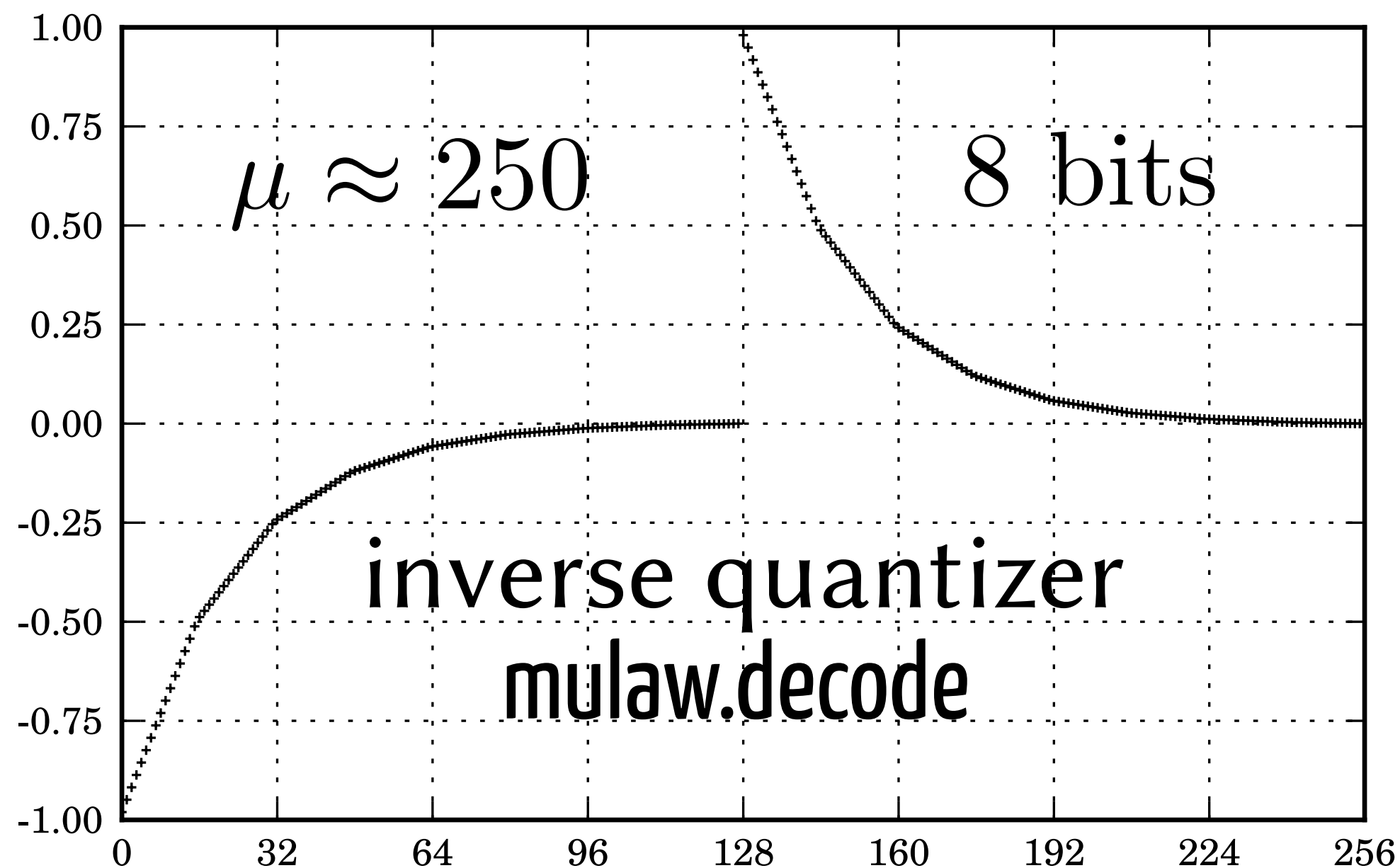
$$p(x) \propto \begin{cases} \frac{1}{1 + \mu|x|} & \text{if } |x| \leq 1.0, \\ 0 & \text{otherwise.} \end{cases}$$

The optimal quantizer w.r.t. entropy such that  $f([-1, 1]) = [-1, 1]$  is defined by:

$$f(x) = \operatorname{sgn}(x) \frac{\log(1 + \mu|x|)}{\log(1 + \mu)}$$

# $\mu$ -law

Implements a piecewise approximation of  $f$ .  
Part of the G.711 (ITU-T) standard.



Used in Sun/NeXT AU file format.