Statistical Theory of Quantization

Sampling the Probability Density Function

Richard C. Hendriks

TU Delft
Delft University of Technology

Sampling versus Quantization

From signal processing lectures:

- Sampling: Discretizes a function with respect to time (in general w.r.t. the ordering variable)

  The sampling theorem: describes the conditions under which the original signal can be obtained from the samples.

TU Delft

Sampling versus Quantization

This lecture: a Bridge between the signal processing and the stochastic processing lectures.

Understanding the process of quantization.

- Quantization: Discretizes the value (amplitude) at a specific sampling time.

  We will show that applying the sampling theorem within the context of quantization, we can derive conditions under which the distribution and the moments of the original unquantized data can be obtained.

TUDelft

Background

Binaural beamforming:

- Two hearing aids, one for each ear.
- Both hearing aids equipped with M_R and M_L microphones.
- Estimate target using all $M = M_R + M_L$ microphones.
- Challenges: 1) spatial cues preservation. 2) Intelligibility enhancement.
- Typically, binaural algorithms assume error free microphone signals are present at both hearing aids.

TUDelft

Interest for Statistical Theory of Quantization

Signal model:

$\mathbf{Y}_L = \mathbf{S}_L + \mathbf{N}_L$ $\qquad$ $\mathbf{Y}_R = \mathbf{S}_R + \mathbf{N}_R$

$\mathbf{Y}_L$

$\mathbf{Y}_R$

$\hat{S}_L = \mathbf{w}_L^H[\mathbf{Y_L}^T, \mathbf{Y_R}^T]^T$ $\qquad$ $\hat{S}_R = \mathbf{w}_R^H[\mathbf{Y_L}^T, \mathbf{Y_R}^T]^T$

TUDelft

Interest for Statistical Theory of Quantization

Signal model:

$\mathbf{Y}_L = \mathbf{S}_L + \mathbf{N}_L$ $\qquad$ $\mathbf{Y}_R = \mathbf{S}_R + \mathbf{N}_R$

$Q[\mathbf{Y}_L]$

$Q[\mathbf{Y}_R]$

$\hat{S}_L = \mathbf{w}_L^H[\mathbf{Y_L}^T, \mathbf{Y_R}^T]^T$ $\qquad$ $\hat{S}_R = \mathbf{w}_R^H[\mathbf{Y_L}^T, \mathbf{Y_R}^T]^T$

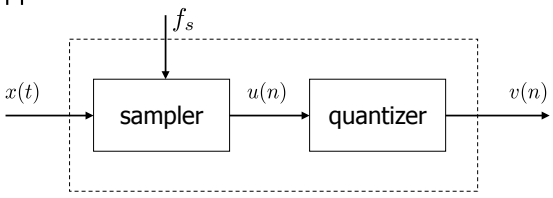
TUDelft

Background - Sampling

TU Delft

Analog-to-Digital Converter

Two-step approach:



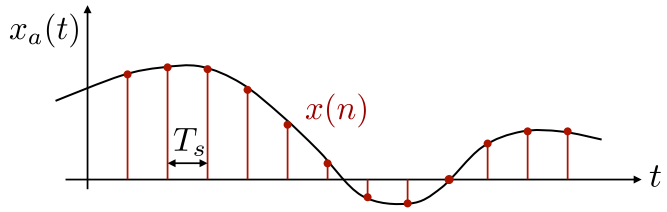
- Sampler: $u(n) = x(nT_s)$ where $T_s = 1/f_s$, the sampling period
- Quantizer: $v(n) = (Qu)(n)$, where Q is a (nonlinear) mapping from intervals of the real line (quantization cells) to reproduction levels

TU Delft

Sampling

To process a continuous-time signal using digital signal processing techniques, it is necessary to convert the signal into a sequence of numbers. This is usually done by *sampling* the analog signal, say $x_a(t)$, periodically every T_s seconds to produce the discrete-time signal $x(n)$ given by

$$x(n) = x_a(nT_s), \quad -\infty < n < \infty$$



TU Delft

Sampling

Recall that the spectrum of the discrete-time signal x is given by

$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n}$$

where the signal x can be recovered from its spectrum by the inverse Fourier transform

$$x(n) = \frac{1}{2\pi}\int_0^{2\pi} X(\omega)e^{j\omega n}d\omega$$

TU Delft

Sampling

If x_a is an aperiodic absolutely integrable signal, its Fourier transform (with $\Omega = 2\pi f$) is given by

$$X_a(\Omega) = \int_{-\infty}^{\infty} x_a(t) e^{-j\Omega t} dt$$

where the signal x_a can be recovered from its spectrum by the inverse Fourier transform

$$x_a(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_a(\Omega) e^{j\Omega t} d\Omega$$

The *angular frequency* Ω is expressed in radians per second (rad/s)

TU Delft

Sampling

Key question: What is the relation between $X(\omega)$ and $X_a(\Omega)$, and under what conditions can we recover x_a from $X(\omega)$?

Periodic sampling imposes a relationship between the independent variables t and n in the signals $x_a(t)$ and $x(n)$, respectively

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

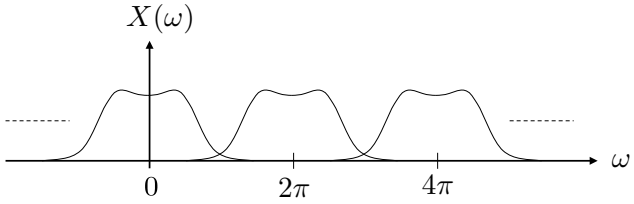
and thus between ω and Ω through Fourier transformation.

TU Delft

Sampling

Remember that it was shown that

$$X(\omega) = f_s \sum_{k=-\infty}^{\infty} X_a(\omega + k2\pi)$$

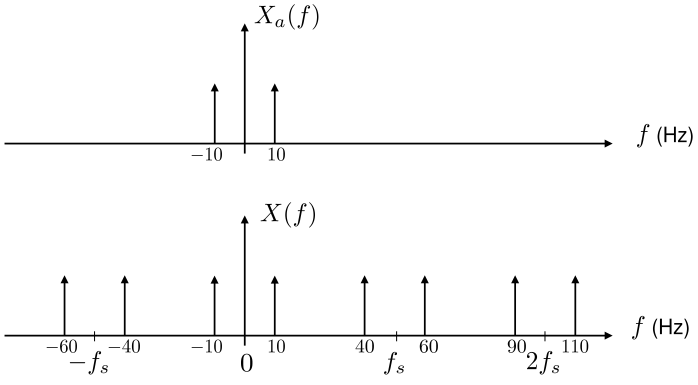


The spectrum of the discrete-time signal consists of shifted copies of the (scaled) analog spectrum

TU Delft

Sampling

Example: $x_a(t) = \cos(2\pi 10t), \quad f_s = 50$ Hz



TU Delft

Sampling

If the spectrum of the analog signal is band limited to, say B Hz, and the sampling frequency satisfies $f_s > 2B$, we have

$$X(f) = f_s X_a(f) \quad \text{for } |f| \leq \frac{f_s}{2}$$

since the periodic repetition of the spectrum of X_a does not introduce spectral overlap

In this case, we can perfectly reconstruct the original analog signal by scaling the input spectrum by f_s^{-1} and removing all spectral components $|f| > \frac{f_s}{2}$

$$X_a(f) = \begin{cases} f_s^{-1} X(f), & |f| \leq \frac{f_s}{2} \\ 0, & |f| > \frac{f_s}{2} \end{cases}$$

TU Delft

Aliasing

Aliasing occurs when the signal is sampled at a rate which is too low. For real signals, the effect can be described by folding of the frequency axis

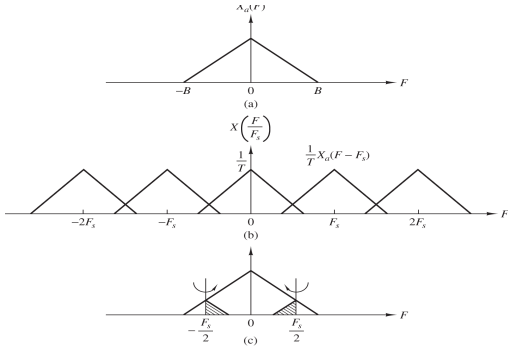
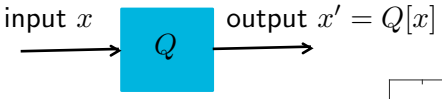


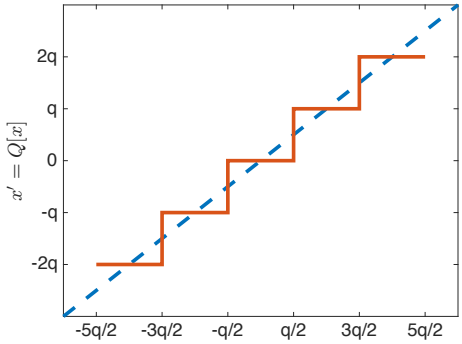
Figure 6.1.3 Illustration of aliasing around the folding frequency.

TU Delft

What is Quantization – Uniform Quantizer



By quantization, noise $N = X - Q[X]$ is added.



TUDelft

This Lecuture:

Questions:

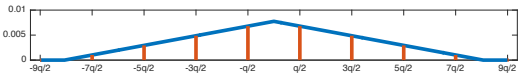
- What is the distribution $N = X - Q[X]$?
- Under which conditions can we recover the pdf of X from the pdf of $Q[X]$?
- Under which conditions can we recover the moments of X from the moments of $Q[X]$?

Even though quantizers are non-linear, these questions can be answered by applying the sampling theorem to quantization $\Rightarrow$

Area sampling of the pdf of a signal.

TUDelft

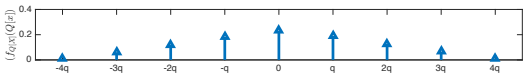
Area Sampling

Original pdf of X 

Probability of each output level equals prob. of input signal occurring within quantization band.

?

Yes, remember the sampling theorem for time domain signals.

pdf of $Q[X]$ 

This process is called area sampling.

TUDelft

Quantization = Area Sampling

Area sampling can be accomplished by

1. Convolving the input pdf $f_X(x)$ with a uniform pulse

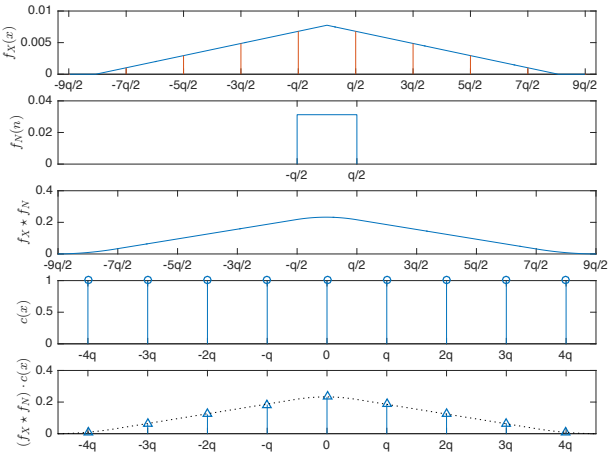
$$f_N(n) = \Pi_q(n) = \begin{cases} \frac{1}{q}, & -\frac{q}{2} < n < \frac{q}{2} \\ 0 & \text{elsewhere,} \end{cases}$$

   that is, $f_N(x) * f_X(x)$.

2. Conventional sampling of $f_N(x) * f_X(x)$.

TUDelft

Example



What is the relation between the distribution of the original data and the quantized data?

TU Delft

The Characteristic Function

The effects of quantization can easily be understood using the characteristic function.

The characteristic function is the Fourier transform of the pdf:

$$\phi_X(u) = \int_{-\infty}^{\infty} f_X(x) e^{jux} dx = E[e^{jux}].$$

- The characteristic function in quantization plays a similar role as the Fourier transform in sampling.
- Notice: If two RVs are independent, the pdf of the sum is the convolution between the two, i.e., the pdf of $W = X + V$ is given by $f_W(w) = f_X(x) * f_V(v)$.
- u is thus the "frequency".

TU Delft

The Characteristic Function

Remember area sampling:

1. Convolving the input pdf $f_X(x)$ with a uniform pulse with width q

$$f_N(n) = \Pi_q(n) = \begin{cases} \frac{1}{q}, & -\frac{q}{2} < n < \frac{q}{2} \\ 0 & \text{elsewhere,} \end{cases}$$

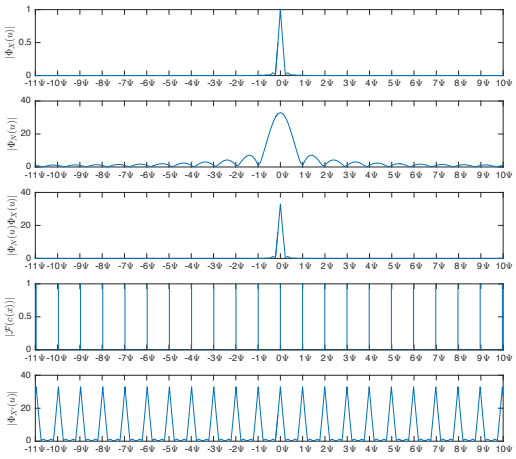
   that is, $f_N(x) * f_X(x)$.

2. Conventional sampling of $f_N(x) * f_X(x)$.

The operation $f_N(x) * f_X(x)$ thus implies that we add two random variables, and thus, that we add uniform noise to the (unquantized) data X.

TUDelft

Example – Using the CF



Sampling (convolving in CF domain by Fourier transform of $c(x)$) leads to repetitions with frequency $\Psi = 2\pi/q$.

CF of quantizer output: $\Phi_{X'}(u) = \sum_{l=-\infty}^{\infty} \Phi_X(u + l\Psi) sinc\left(\frac{q(u+l\Psi)}{2}\right)$

"sampling frequency"

TUDelft

Quantization Theorem I

What are the conditions to be able to get back the original pdf or CF of X?

This is analogous to the sampling theorem:

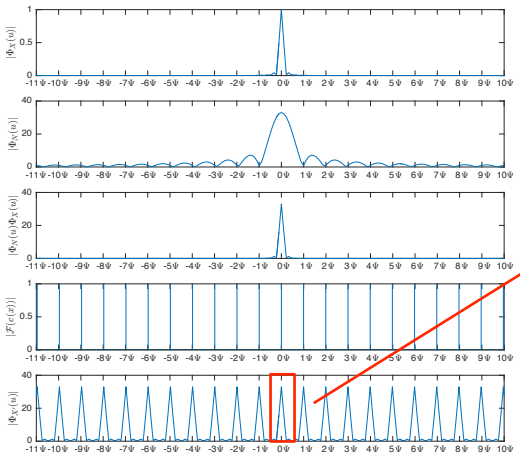
Quantization Theorem I: If the CF of X is bandlimited, such that

$$\Phi_X(u) = 0 \quad \text{for } |u| > \frac{\pi}{q} = \frac{\Psi}{2},$$

then the CF and the pdf of X can be obtained from the CF and pdf of X', respectively.

TU Delft

Example – Using the CF



Take central replica. Notice that this gives back the convolution between f_X and f_N.

TU Delft

Analogy Sampling and Quantization

What if $\Phi_X(u)$ is not band limited?

With sampling we would apply an anti-aliasing filter before sampling. With quantization we can do something similar.

Define a "filter" with response $\Phi_V(u)$ such that $\Phi_X(u)$ becomes band-limited. We then get $\Phi_{X+V}(u) = \Phi_X(u)\Phi_V(u)$. Notice that this means that we add an independent variable V to X and that

$$f_{X+V} = f_X * f_V.$$

Variable V acts as a dither signal.

TU Delft

The Moments of the Quantizer Output

The moments of a RV can be calculated using the CF:

$$E[X'^k] = \frac{1}{j^k} \frac{d^k \Phi_{X'}(u)}{du^k}\bigg|_{u=0}.$$

TU Delft

The Moments of the Quantizer Output

When QTI holds:
The moments of the quantized signal x' and $x+n$ correspond exactly, thus,

$$E[(X')^k] = E[(X+N)^k]$$

with N uniform noise.
However, QTI is a much stricter condition (on the pdfs). If we are only interested in the moments we can use a relaxed condition.

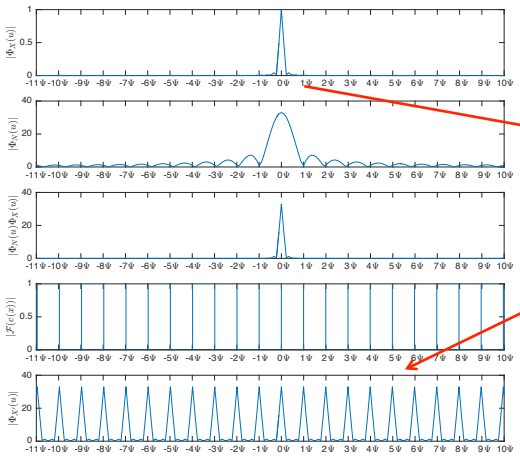
Quantization Theorem II: If the CF of X is bandlimited, such that

$$\Phi_X(u) = 0 \quad \text{for } |u| > \tfrac{2\pi}{q} - \varepsilon = \Psi - \varepsilon,$$

with positive and arbitrarily small ε, then the moments of X can be calculated from the moments of X'.

TUDelft

Consequence of QTII



If CF is "zero" from $\Psi - \epsilon$, moments can be obtained from $\Phi_{X'}$.

TUDelft

The Moments of the Quantizer Output

The moments of a RV can be calculated using the CF:

$$E[X'^k] = \frac{1}{j^k} \frac{d^k \Phi_{X'}(u)}{du^k}\bigg|_{u=0}.$$

This means that if $\Phi_X(u) = 0$ for $|u| > \frac{2\pi}{q} - \epsilon$, obviously, (derivatives of) the higher order copies do not overlap at $u = 0$.

The moments of X can still be calculated from the moments of X' (the aliasing terms then have no influence).

TU Delft

The Moments of the Quantizer Output

Let $G_X(u) = \Phi_X(u) sinc\left(\frac{qu}{2}\right)$ (the zeroth copy in the CF domain).

A weaker condition: If

$$\frac{d^k G_X(u)}{du^k}\bigg|_{u=\frac{2\pi k}{q}=\Psi k} = 0.$$

for $\forall k$ except $k = 0$,

then,

$$E[X'^k] = \frac{1}{j^k} \frac{d^k G_X(u)}{du^k}\bigg|_{u=0},$$

and moments of X' can completely be described in terms of the moments of X

TU Delft

The Distribution of the Quantizer Noise

Let $\Pi_q(n)$ be defined as

$$\Pi_q(n) = \begin{cases} \frac{1}{q} & -\frac{q}{2} \leq n < \frac{q}{2} \\ 0 & \text{otherwise.} \end{cases}$$

and $W_q(n) = \sum_{k=-\infty}^{\infty} \delta(n - kq)$. The quantizer noise pdf is then given by

$$\begin{aligned} p_N(n) &= \sum_{k=-\infty}^{\infty} p_X(kq+n), \quad \text{for} \quad -\frac{q}{2} \leq n < \frac{q}{2} \\ &= q\Pi_q(n) \sum_{k=-\infty}^{\infty} p_X(kq+n) = q\Pi_q(n)[W_q * p_X](n) \end{aligned}$$

selects part in range $-\frac{q}{2} \leq n < \frac{q}{2}$

Pulse train that generates shifted copies of p_X.

TUDelft

The CF of the Quantizer Noise

The quantizer noise pdf:

$$p_N(n) = q\Pi_q(n)[W_q * p_X](n)$$

CF due to area sampling (addition of uniform noise).

$$\begin{aligned} \Phi_N(u) &= sinc\left(\frac{qu}{2}\right) * [W_{2\pi/q}(u)\Phi_X(u)] \\ &= \sum_{k=-\infty}^{\infty} \Phi_X\left(\frac{2\pi k}{q}\right) sinc\left(\frac{q(u - k\Psi)}{2}\right) \end{aligned}$$

TUDelft

The CF of the Quantizer Noise

$$\Phi_N(u) = \sum_{k=-\infty}^{\infty} \Phi_X\left(\frac{2\pi k}{q}\right) sinc\left(\frac{q(u-k\Psi)}{2}\right)$$

The error will be uniformly distributed if $\Phi_N(u)$ reduces to a single sinc function. This happens under QTI.

More general: The error (in an undithered system) is uniform, if and only if

$$\Phi_X(u)|_{u=2\pi k/q} = 0,$$

$\forall k$ except $k = 0$.

TU Delft

Subtractive Dither

Let v denote a dither signal independent from the input x, that is, the quantizer input is $x + v$ and the output of the system is $y = Q(x + v) - v$. The total error is $n = Q(x + v) - (x + v)$.
The quantizer noise pdf of a subtractive dithered signal is then given by

$$p_N(n) = q\Pi_q(n)[W_q * p_X * p_v](n)$$

The CF of the subtractively dithered quantizer noise is

$$\Phi_N(u) = sinc\left(\frac{qu}{2}\right) * [W_{2\pi/q}(u)\Phi_X(u)\Phi_v(u)]$$

TU Delft

Subtractive Dither

The CF of the subtractively dithered quantizer noise is

$$\Phi_N(u) \quad = \quad sinc\left(\frac{qu}{2}\right) * [W_{2\pi/q}(u)\Phi_X(u)\Phi_v(u)]$$

The error in a subtractively dithered quantization system will be uniformly distributed and statistically independent of the input for any input distribution if and only if

$$\Phi_V(u)|_{u=2\pi k/q} = 0,$$

$\forall k$ except $k = 0$.

TU Delft