Lecture 9: Discretization and Quantization of Continuous Signals - Kotelnikov Theorem

Dosbayev Zhandos Makhsutuly, senior lecturer
E-mail: zh.dosbayev@satbayev.university

Outline

Discretisation

Nyquist-Shannon Sampling Theorem

Quantization

Kotelnikov's Theorem and its Applications

Lecture 9: Discretization and Quantization of Continuous Signals - Kotelnikov Theorem

Welcome to Lecture 9, where we delve into the fascinating world of discretization and quantization of continuous signals, with a special focus on the groundbreaking Kotelnikov Theorem. This lecture serves as a crucial bridge between the realms of continuous and discrete signal processing, providing the foundation for modern digital communication systems and signal analysis techniques.

Throughout this lecture, we'll explore how continuous signals are converted into discrete forms, the implications of this process, and the theoretical underpinnings that govern these transformations. We'll also examine the practical considerations and challenges that arise when implementing these concepts in real-world applications. By the end of this lecture, you'll have a comprehensive understanding of the principles that enable our digital world to interact with and represent continuous physical phenomena.

Introduction to Discretisation and Quantisation

Discretisation and quantisation are fundamental processes in digital signal processing that bridge the gap between continuous, real-world signals and their digital representations. Discretisation, also known as sampling, involves converting a continuous-time signal into a discrete-time signal by taking measurements at specific time intervals. This process allows us to represent an infinite number of signal values with a finite set of samples.

Quantisation, on the other hand, deals with the amplitude domain of the signal. It involves approximating the continuous range of amplitude values with a finite set of discrete levels. This process is essential for storing and processing signals in digital systems, where we must represent values using a limited number of bits.

1. **Continuous Signal**

   The original, unprocessed signal with infinite resolution in both time and amplitude.

2. **Discretisation (Sampling)**

   Converting the continuous-time signal into a sequence of discrete-time samples.

3. **Quantisation**

   Approximating the continuous amplitude values with a finite set of discrete levels.

4. **Digital Signal**

   The final representation, discrete in both time and amplitude, suitable for digital processing.

Sampling of Continuous-Time Signals

Sampling is the process of converting a continuous-time signal into a discrete-time signal by taking measurements at regular intervals. This process is fundamental to digital signal processing and forms the basis for analog-to-digital conversion. The sampling process can be mathematically represented as the multiplication of the continuous-time signal $x(t)$ with a train of impulses, resulting in a sequence of discrete samples $x[n]$.

The sampling rate, or sampling frequency (f_s), is a critical parameter in this process. It determines how frequently samples are taken from the continuous signal. The choice of sampling rate has profound implications for the quality and accuracy of the digital representation. If the sampling rate is too low, we risk losing important information about the original signal, leading to phenomena such as aliasing.

1

Continuous Signal

The original analog signal $x(t)$ with infinite time resolution.

2

Sampling Process

Multiplication with impulse train at intervals of $T_s = 1/f_s$.

3

Discrete Samples

Resulting sequence of samples $x[n]$ at discrete time points.

4

Digital Representation

The sampled signal ready for digital processing or storage.

Nyquist-Shannon Sampling Theorem

The Nyquist-Shannon sampling theorem is a cornerstone principle in signal processing that establishes the conditions for perfect reconstruction of a continuous-time signal from its samples. The theorem states that for a bandlimited signal with maximum frequency component fmax, the sampling rate fs must be at least twice the highest frequency component (fs ≥ 2fmax) to avoid aliasing and ensure perfect reconstruction.

This minimum required sampling rate, 2fmax, is known as the Nyquist rate. When a signal is sampled at or above this rate, all the information in the original continuous-time signal is preserved in the discrete samples. This principle is crucial in various applications, from audio and video processing to data acquisition systems and telecommunications.

Theorem Statement

A bandlimited signal can be perfectly reconstructed from its samples if the sampling rate is greater than twice the maximum frequency component in the signal.

Nyquist Rate

The minimum sampling rate required for perfect reconstruction, equal to 2fmax.

Applications

Widely used in audio sampling (44.1 kHz for CD-quality audio), video capture, and various data acquisition systems.

Aliasing and its Consequences

Aliasing is a distortion phenomenon that occurs when a signal is sampled at a rate lower than the Nyquist rate. When aliasing happens, high-frequency components of the original signal are misinterpreted as lower frequencies in the sampled signal, leading to a loss of information and potential distortion in the reconstructed signal.

The consequences of aliasing can be severe, particularly in applications where signal fidelity is crucial. In audio processing, aliasing can introduce unwanted artefacts and distortions that degrade sound quality. In image processing, it can lead to visual artefacts such as moiré patterns. To prevent aliasing, anti-aliasing filters are often employed before sampling to remove frequency components above the Nyquist frequency.

Causes of Aliasing

Sampling rate below the Nyquist rate, presence of frequencies above half the sampling rate in the input signal.

Effects

Misinterpretation of high frequencies as lower frequencies, loss of information, distortion in reconstructed signal.

Prevention

Use of anti-aliasing filters, increasing sampling rate, proper signal conditioning before sampling.

Detection

Spectrum analysis, visual inspection of reconstructed signal, listening tests in audio applications.

Quantisation of Discrete-Time Signals

Quantisation is the process of mapping a large set of input values to a smaller set of output values. In the context of discrete-time signals, quantisation involves approximating the continuous amplitude values of sampled signals with a finite set of discrete levels. This process is essential for representing signals in digital systems, where we must use a limited number of bits to represent each sample.

The quantisation process introduces an error, known as quantisation noise, which is the difference between the original continuous value and its quantised approximation. The severity of this error depends on the number of quantisation levels used, which is typically determined by the bit depth of the digital system. Higher bit depths allow for more quantisation levels, resulting in finer amplitude resolution and lower quantisation noise.

1. **Continuous Amplitude**

   The original sampled values with infinite amplitude resolution.

2. **Quantisation Process**

   Mapping continuous values to nearest discrete levels.

3. **Discrete Levels**

   The finite set of amplitude values used to represent the signal.

4. **Quantised Signal**

   The final digital representation with discrete time and amplitude.

Uniform Quantisation and its Effects

Uniform quantisation is the simplest and most common form of quantisation, where the range of input values is divided into equal-sized intervals. Each interval is represented by a single quantisation level, typically the midpoint of the interval. This approach is straightforward to implement and analyse, making it popular in many digital systems.

The effects of uniform quantisation depend largely on the number of quantisation levels used. As the number of levels increases (higher bit depth), the quantisation error decreases, and the signal-to-quantisation-noise ratio (SQNR) improves. However, uniform quantisation may not be optimal for signals with non-uniform amplitude distributions, as it allocates the same resolution to all amplitude ranges, regardless of their probability of occurrence.

Bit Depth	Number of Levels	SQNR (dB)
8-bit	256	~49.9
16-bit	65,536	~98.1
24-bit	16,777,216	~146.2
32-bit float	~4.3 billion	~192

Non-uniform Quantisation and Logarithmic Encoding

Non-uniform quantisation is an alternative approach that allocates quantisation levels unevenly across the amplitude range. This technique is particularly useful for signals with non-uniform amplitude distributions, such as speech or music, where smaller amplitudes occur more frequently than larger ones. By allocating more quantisation levels to the more probable amplitude ranges, non-uniform quantisation can achieve better overall signal quality for a given number of bits.

Logarithmic encoding is a common form of non-uniform quantisation where the quantisation levels are spaced logarithmically rather than linearly. This approach provides finer resolution for smaller amplitudes and coarser resolution for larger amplitudes, which often matches the characteristics of human perception. Common logarithmic encoding schemes include μ-law and A-law encoding, widely used in telecommunications.

Improved SNR

Non-uniform quantisation typically offers better signal-to-noise ratio for signals with non-uniform amplitude distributions.

Perceptual Matching

Logarithmic encoding often aligns better with human auditory perception, providing better perceived quality.

Compression

Non-uniform quantisation can achieve effective signal compression by allocating bits more efficiently.

Telecommunications

Widely used in telephony and audio transmission systems to optimise signal quality and bandwidth usage.

Kotelnikov's Theorem and its Applications

Kotelnikov's theorem, also known as the Sampling theorem or the WKS (Whittaker–Kotelnikov–Shannon) sampling theorem, is a fundamental principle in signal processing. Developed independently by Vladimir Kotelnikov in 1933, it predates the better-known Shannon formulation. The theorem states that a bandlimited continuous-time signal can be perfectly reconstructed from a set of samples taken at a rate of at least twice the highest frequency component in the signal.

This theorem has far-reaching applications in various fields of engineering and science. It forms the theoretical basis for digital communications, enabling the faithful transmission and reconstruction of analog signals in digital form. In audio and video processing, it guides the choice of sampling rates to ensure high-quality digitisation. The theorem also plays a crucial role in data acquisition systems, where it helps determine the minimum sampling rate required to capture all relevant information from physical phenomena.

1 Digital Communications

Kotelnikov's theorem underpins the design of digital communication systems, ensuring that analog signals can be accurately represented and transmitted in digital form.

2 Signal Processing

It provides the foundation for various signal processing techniques, including filtering, modulation, and spectral analysis in the digital domain.

3 Data Acquisition

The theorem guides the design of data acquisition systems, helping to determine appropriate sampling rates for capturing physical phenomena accurately.

4 Compression Techniques

Understanding the minimum sampling rate required for perfect reconstruction also informs the development of efficient data compression algorithms.

Practical Considerations in Discretisation and Quantisation

While the theoretical foundations of discretisation and quantisation are well-established, their practical implementation involves various considerations and challenges. In real-world applications, engineers must balance factors such as signal fidelity, system complexity, power consumption, and cost. For instance, while higher sampling rates and bit depths can improve signal quality, they also increase data storage requirements and processing complexity.

Anti-aliasing filters play a crucial role in practical systems, ensuring that signals are properly bandlimited before sampling. The design of these filters involves trade-offs between stopband attenuation, passband ripple, and filter complexity. Additionally, the choice between uniform and non-uniform quantisation depends on the specific application and signal characteristics. In many cases, a combination of techniques, such as oversampling followed by noise shaping and decimation, is used to achieve optimal performance.

Hardware Limitations

Practical systems are constrained by the capabilities of analog-to-digital converters (ADCs) and digital-to-analog converters (DACs), including their speed, resolution, and noise characteristics.

System Design Trade-offs

Engineers must balance factors such as signal quality, power consumption, cost, and complexity when designing discretisation and quantisation systems.

Advanced Techniques

Oversampling, noise shaping, and adaptive quantisation are often employed to overcome limitations and improve overall system performance in specific applications.