Exercise 2 - The A2D process and Fourier Analysis

Oded Scharf

October 17, 2021

Part I
Introduction

A signal is something that conveys information. While we can think of many types of signals, both inside and outside of the biological world, in this course we'll try to describe them mathematically. Even when considering proteins or molecules going around between cells, we can think of them as packets of information transferring between an emitter and a receiver.

A signal has an independent variable x (usually representing the time-axis) that can either be continuous or discrete. Furthermore, we describe analog signals as continuous, while digital ones are discrete. Since many of our sampling and measurement devices are analog, they record a continuous signal. While their output is also analog, we use computers, which are digital, to process that information. This chain of events has to include a converter that can read an analog signal and transform it to a digital one. The course will mainly deal with discrete, digital signals and the mathematical tools we can apply to them.

Part II
Analog to digital (A2D) conversion

1 Basic process

Measurements from an experiment have to go through a process, called *analog to digital conversion* (depicted in figure 1), in order for us to analyze it in a computer. In class you also discussed The Nyquist frequency and *aliasing*, which we'll not currently mention.

1. <u>Electrode</u> - hyphenated -A device that reads the actual electrical signal,

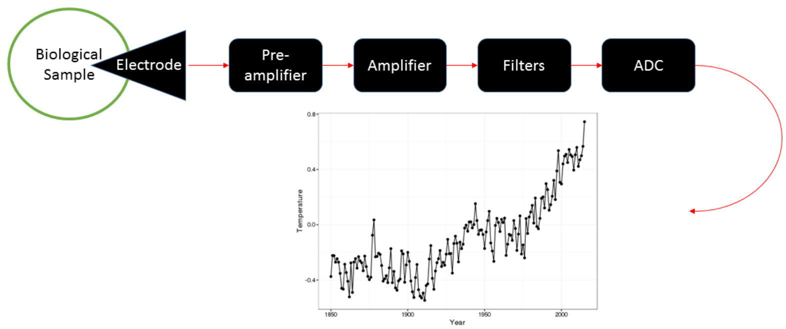

Figure 1: The complete analog to digital process of a signal.

which is completely analog, i.e. generates continuous data streams. As said in class, we wish for it to be with high input impedance and low output impedance, so it would not drastically change the samples it measures (the book explains this nicely). Its output is an electrical potential (voltage) that changes with time.

2. Pre-amp and amplifier - Serves a basic purpose of amplifying the signal, or increasing its amplitude so we could better detect it with our ADC. The amplifier might follow the filter.

3. Filters - Allow us to change the frequencies observed in the signal. We can remove high frequencies with low-pass filters, or low frequencies with high-pass filters. We can also modulate a specific band of frequencies with band-pass filters or notch filters.

4. Analog to digital converter (ADC, A2D) - A device directly connected to the computer, that sees an analog (= continuous) signal, and allocates its amplitude values to different digital (= discrete) values. An ADC that is said to have n bits actually has 2^n discrete values. Every ADC also has a sampling rate, which means that it discretizes both the amplitude and time domains. We'll discuss the implications of sampling in time soon.

2 Challenges

Signals outside of the computer are never perfect. There's always *noise* associated with every recording we'll ever take, and we'll usually try to decrease its level so that our signal to noise ratio, or SNR, is maximal. Every electronic device we put in the way of our signal generates some noise, and most of them have a specific type of noise they'll generate. For example, devices connected to a power outlet will "vibrate" at a rate of 50 Hz, since power comes from the power plant at this rate.

But noise isn't the only thing which comes between us and the perfect recording. The mentioned digitization process can bring us to saturation - a condition where the computer or measurement device allocated its maximal value to a measurement because it was "off the charts". This leads to data loss, since we're no longer aware of the "real" value of that measurement.

Another concern is data size. This might seem odd at first, but in truth recording technology has advanced in a greater pace than storage technology; it's always possible to add one more electrode, or make a camera with more sensors, but data transfer rates and storage are still quite limited. A modern imaging experiment, for example, can generates terabytes of data per minute. Storing and processing these amounts of data take time and expertise.

3 Examples

Question Assuming an ADC with 12 bits, and a system voltage of 5 Volts, what is the digital value of a 1.83 V signal? What is the system's resolution?

Answer We can write a simple equation to solve this question:

$$\begin{aligned} \frac{\text{Total number of ADC bins}}{\text{Maximal system voltage}} &= \frac{\text{Digital value}}{\text{Analog voltage measured}} \\ \frac{2^{12}}{5\text{V}} &= \frac{x}{1.83\text{V}} \qquad (1) \\ \Rightarrow x &= 1499.136 \approx 1499 \end{aligned}$$

As we can see, there's a rounding error in this process. The system's resolution is the amount of volts a single bin takes:

$$\frac{5\text{V}}{2^{12}} \approx 1.22\text{mV} \qquad (2)$$

If we divide this number by the amplification of our A2D system, we'll receive the actual resolution our measurement device has (in *real* voltage units, that correspond to the biological sample).

Another way to display the digitization process is with images - the image *bit depth* is the exact same idea as the bit depth of an ADC. Take a look at figures 2-5 which show the same image with increasing bit depth.

Figure 2: Image with a bit depth of 1, corresponding to black or white colors.

Figure 3: Image with a bit depth of 2, or four levels of gray.

Figure 4: Image with a bit depth of 5.

Figure 5: Image with a bit depth of 8.

Part III
Introduction to the Frequency Domain

4 Sampling in Time

As we previously mentioned, the ADC samples a continuous, i.e. analog, signal, into bins in two dimensions - the first (and more obvious) is the amplitude domain, i.e. the values of the measurement are discretized. But it also digitizes the time domain by sampling a continuous signal into time bins.

Many "natural" signals are continuous. For example, the membrane voltage of a neuron has a continuous value which we can measure with an electrode. The current and voltage on the electrode are also continuous (=analog) values, since the can have any values they wish to - the wires that they're made up with aren't digital. But as soon as we wish to observe these values we have to sample them. This process "asks" the wire at a given point in time "what is your current value?", and the answer is logged in the ADC. After a hopefully-short period of time the ADC will ask the wire again "what is your current value?", and will again log the answer. In the period between these two "questions" the wire could've changed its voltage value and the ADC wouldn't have known. This is why we must sample much faster than the rate at which we expect the measurement value to change.

To put the previous paragraph in a more mathematical fashion, we'll define *sampling interval* T_s and the *sampling frequency* $F_s = f_s = 1/T_s$ as the two important quantities that define the act of sampling. The resulting signal is only defined in the specific time points which were sampled. In other words, starting with an analog signal $x_{analog} = x(n)$, where n can mean time, or any other independent variable, we'd like to sample this signal into our digital, discrete, world. Mathematically, we represent discrete signals as a sequence of numbers x, in which the nth number is denoted $x[n]$. The square brackets are indicative that this is a discrete, rather than continuous, set. This set is formally written as:

$$x = x[n], \; -\infty < n < \infty \tag{3}$$

where n is an integer. How can we connect the analog and digital signal, how can we "sample" the analog data in the mathematical sense? This definition helps:

$$x[n] = x_{\text{analog}}(nT_s), \; -\infty < n < \infty \tag{4}$$

Note that $x[n]$ can mean two things - either entire sequence, or just the nth sample of it. Figure 6 shows a typical signal train. While the x-axis looks

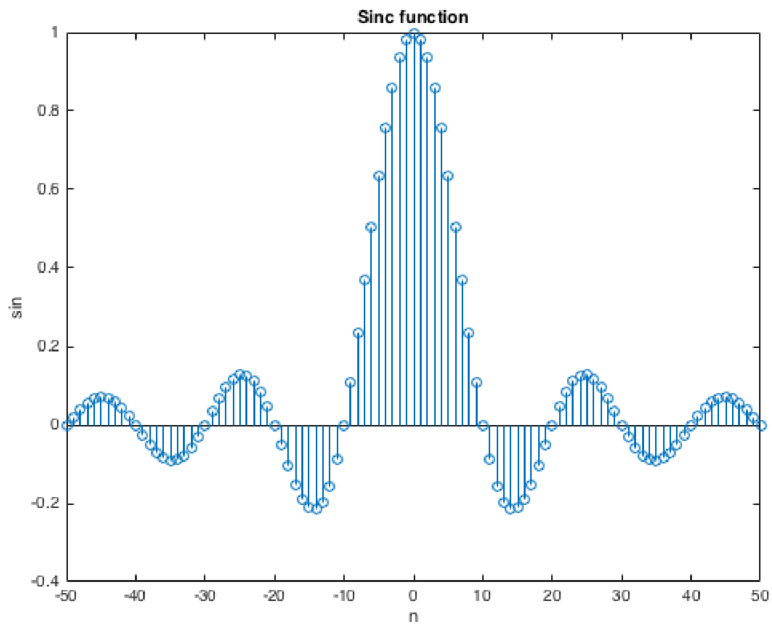

Figure 6: A basic discrete (digital) signal.

continuous, mind that the function is defined only when x, or n, is an integer. The function isn't 0 between these values - it's just not defined.

While equation 4 shows the final result of the sampling, i.e. it has $x[n]$ before the equality sign, it still doesn't explain *how* we got that $x[n]$. To that end, we'll have to introduce *Dirac's Delta Function*, which will show us the mathematically-correct way to sample a continuous function $x(n)$ into a discrete one. In the signal processing world, it's also called the *unit sample sequence*, or *unit impulse*.

$$\delta[n] = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases} \quad \Bigg| \quad \delta[n-k] = \begin{cases} 0, & n \neq k \\ 1, & n = k \end{cases} \tag{5}$$

The impulse function can be used to represent any arbitrary value $x[n]$ of a sequence:

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k] \tag{6}$$

Here's $x[k]$ is either continuous or a discrete signal $x(k)$. Our delta function can handle both cases. The process of creating multiple deltas and integrating them with a continuous signal to receive a discrete signal is sometimes called a "Dirac's comb".

5 Spectrum

The spectrum of a signal shows the frequencies contained in that signal. Fourier Analysis, which we'll study in detail later, states that every signal can be decomposed to sine and cosine waves. When these waves are added up, each with its own amplitude, phase and frequency, they'll recreate perfectly the original signal. The spectrum of a signal will show the result of Fourier Analysis done to a signal.

When looking at simple signals the conclusion of the Fourier Analysis is trivial. A sine wave signal is clearly periodic, and its spectrum should indicate that it has one frequency. Later on we'll analyze more complex signals and see how they can also be considered periodic, and what are the periods of the sines and cosines that compose them.

A basic example of what a spectrum is, and how its calculated, can be seen in the Moodle.

6 Aliasing

As discussed in class, there's a potential issue when sampling a signal. If we sample a periodic signal with a low f_s, we might lose information about its periodicity. This idea is known as *aliasing*. The Nyquist theorem states that if we wish to preserve information of the frequency of the recorded signal, we have to sample them at least at twice the maximal frequency which they contain. For example, if we have a signal composed of two sine waves $x(n) = \sin(2\pi f_1) + \sin(2\pi f_2)$, where $f_1 = 10$ Hz and $f_2 = 22$ Hz, we'll have to sample the signal at least at $f_s = 2 * f_2 = 44$ Hz for it to retain its frequency properties. When we'll discuss Fourier Analysis in depth you'll be able to see how the frequencies in a signal are distorted when sampled below the Nyquist limit.

For example, if we sample a 10 Hz signal with $f_s = 21$ Hz, the main component we'll notice is an added 1 Hz component on top of the signal, since that's the remainder when subtracting the 10 Hz signal from f_s. You'll explore this phenomenon in your HW.

The `aliasing.ipynb` Jupyter notebook availble in the Moodle can shed more light on this topic.