Università
della
Svizzera
italiana

Facoltà
di
scienze
informatiche

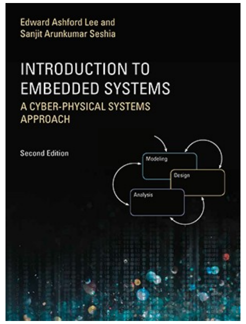
Edge Computing in the IoT

Sensors and Actuators

Alberto Ferrante (alberto.ferrante@usi.ch)

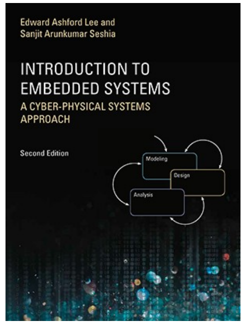
Sensors and Actuators

- Cyber-physical systems integrate computing and physical dynamics
- A **sensor** is a device that measures a physical quantity
  - The variations of the physical quantity are reflected in variations of electrical parameters (voltage/current/resistance) of the sensing element
- An **actuator** is a device that alters a physical quantity
  - An actuator is commonly driven by a voltage/current

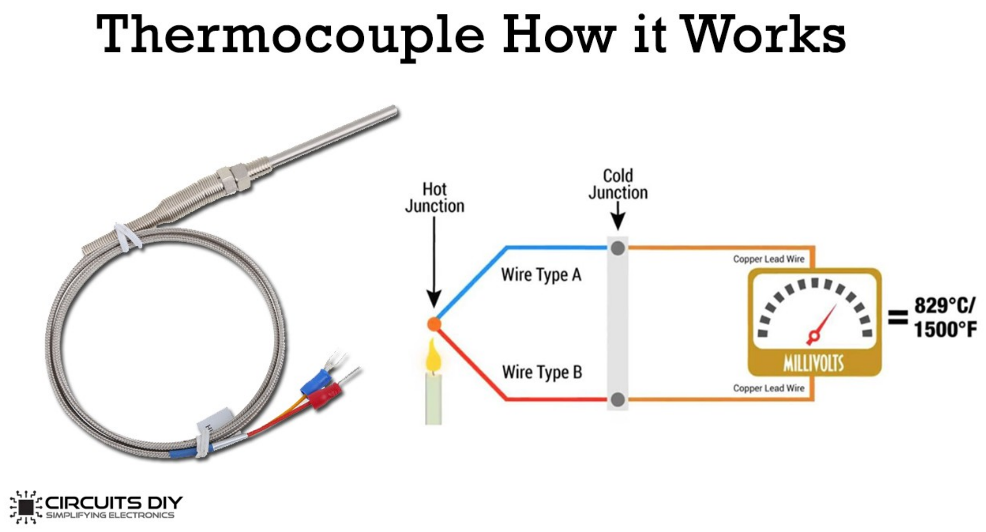


Sensors and Actuators

- Sensors and actuators are often packaged with (small) microprocessors
  - They may have network interfaces, enabling them to appear on the Internet as services
  - Or they may offer digital interfaces (e.g., I^2C) to other devices
    - In the past, the majority of devices offered only an analog interface, some are still available
    - Some devices offer both an analog and a digital interface
- Physical world functions in a multidimensional continuum of time and space. It is an **analog world**
- The world of **software is digital and strictly quantized**



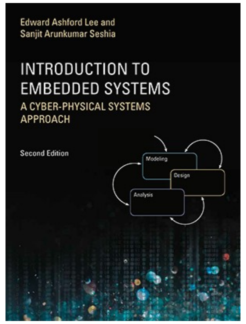
Example – Temperature Sensor



- Thermocouple: used to measure temperature
  - It provides as output different voltage levels depending on measured temperature
- Made up of two different metal wires, joined together to form a junction
  - When the junction gets hot or cold, a small amount of voltage is generated in between two junctions of two transistors

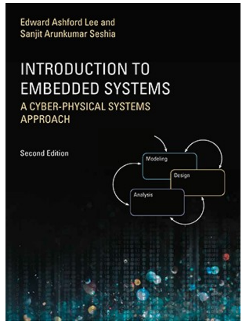
Sensors

- Analog sensors: they produce a voltage/current as output
- Digital sensors: they produce a digital number as output
  - Voltage/current converted to a number through a process called *analog to digital conversion*
    - *Limited precision, determined by the number of bits used to represent the number*



Model of Sensors

- Most sensors can be represented by a function:
  - $$f(x(t)) = ax(t) + b$$
  - *x(t)* – physical quantity to be measured
  - *a* – proportionality constant - **sensitivity**
  - *b* - bias

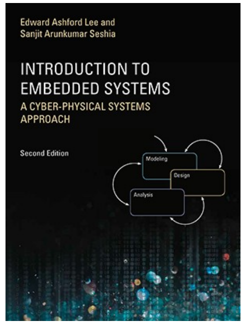


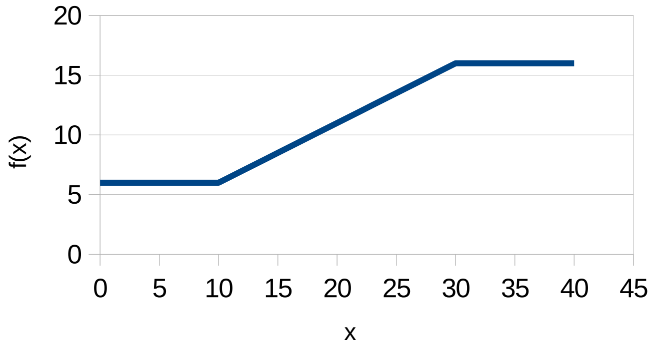
Range

- The set of values of a physical quantity that a sensor can measure is limited

$$f(x(t)) = \begin{cases} ax(t) + b & \text{if } L \leq x(t) \leq H \\ aH + b & \text{if } x(t) > H \\ aL + b & \text{if } x(t) < L, \end{cases}$$

  - *L* and *H* are the low and the high end of the **sensor range**

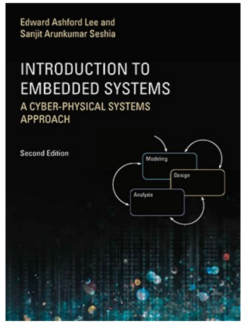




Dynamic Range

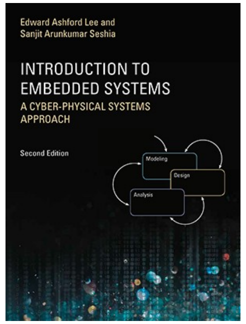
- The **precision** *p* of a sensor is the smallest absolute difference between two values of a physical quantity whose sensor readings are distinguishable
  - e.g., temperature sensor with a precision of 0.5°C
    - If three fluids have a temperature of 37.0, 37.1, and 37.2°C, respectively, the sensor may not measure any difference

- The dynamic range D is the ratio
  - $$D = \frac{H - L}{p}$$
  - Usually measured in decibels:
    - $$D_{dB} = 20 \log_{10} \left(\frac{H - L}{p} \right)$$



Actuators

- Analog actuators: they are controlled by means of a voltage/current
- Digital actuators: they accept digital values as input
  - Internally, a *digital to analog* conversion might be needed to convert the input to a voltage/current
- Similar model to the one of sensors
- Range, dynamic range, and precision, also apply to actuators



Continuous signals

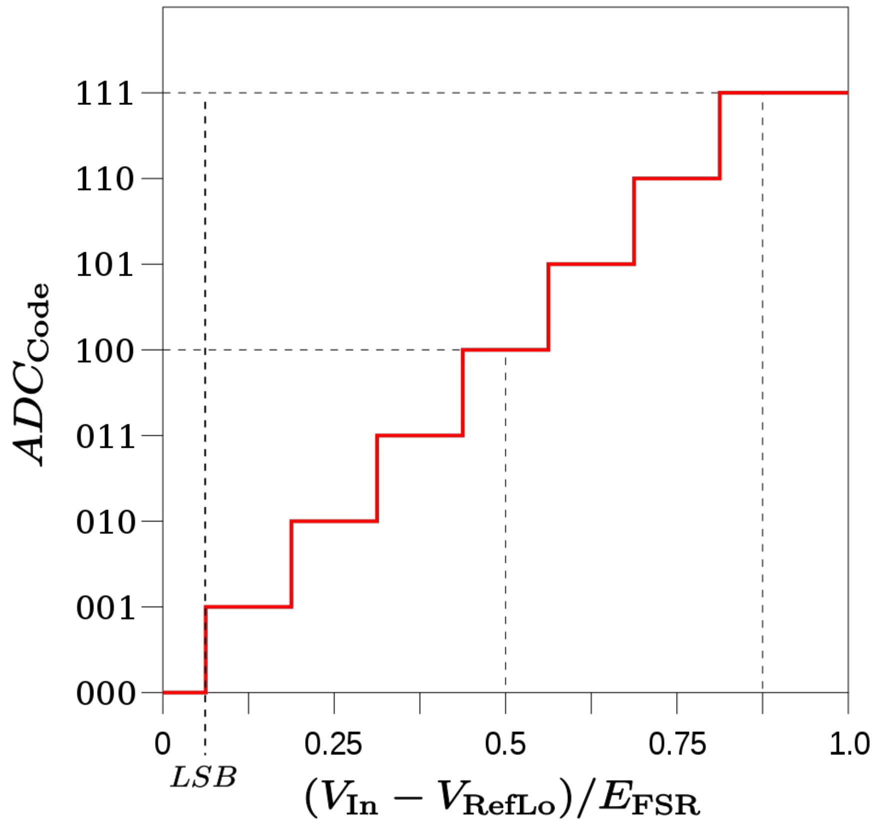
- We often need to deal with signals (e.g., sensor output) that are continuous:
  - In time: in our world, time is continuous!
  - In values: values assume any (Real) value
    - As opposed to discrete values that are finite

Digital Systems Are Discrete

- Digital systems are **discrete in time**
  - We need to make time of the signals discrete in such a way that no important information is lost
- Digital system have **finite precision**
  - We need to derive a digital value from an analog value (often a voltage)
    - We need to perform the conversion in such a way that attained precision is sufficient for our purpose

Analog to Digital Conversion

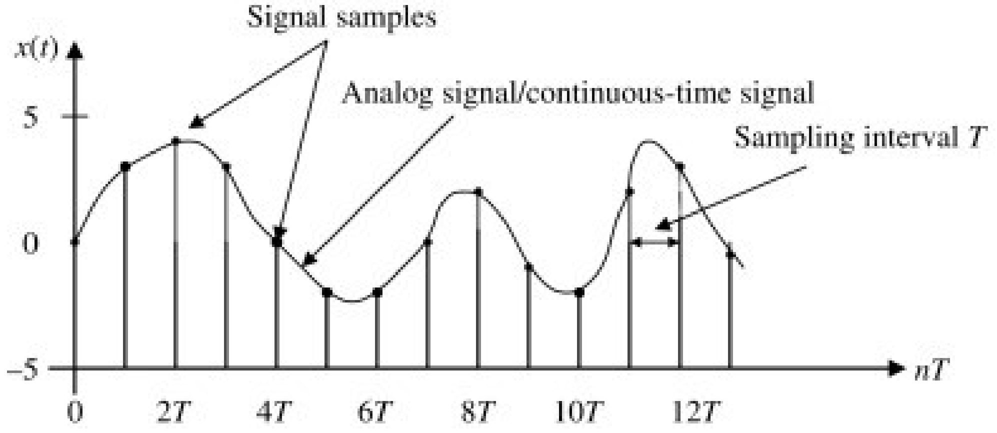
- Analog to Digital Converter = ADC
- An ADC converts a continuous-time and continuous-amplitude analog signal to a discrete-time and discrete-amplitude digital signal
- Two steps
  - **Sampling**: continuous to discrete time
  - **Quantization** of the input: continuous amplitude to discrete signal



https://en.wikipedia.org/wiki/Analog-to-digital_converter#Aliasing

Sampling

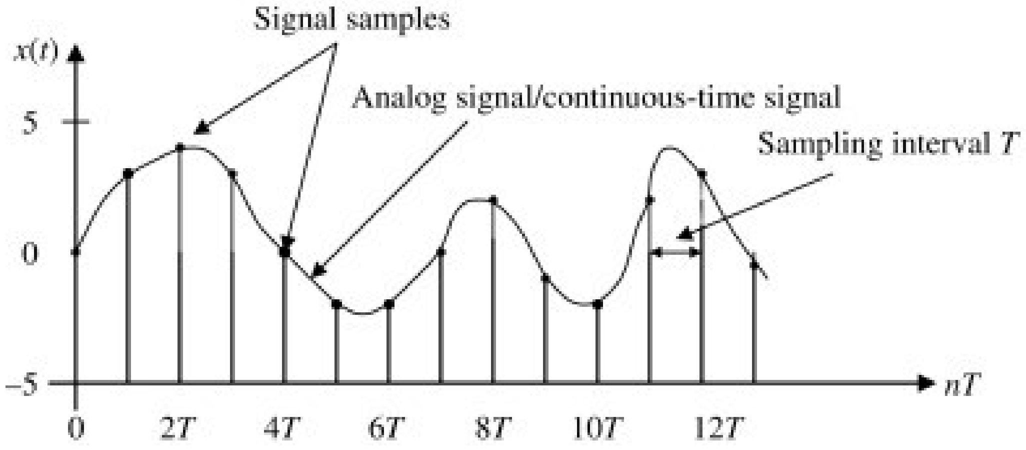
- The continuous analog signal is observed (**sampled)** at specific intervals of time
  - Conversion to a flow of digital values
- **Sampling rate**: the rate at which digital values are sampled from the analog signal
  - T=sampling period
  - Sampling rate: $f_s=1/T$
  - e.g., 10 samples per second → f_s=10Hz → T=0.1s



https://en.wikipedia.org/wiki/Analog-to-digital_converter#Aliasing
https://www.sciencedirect.com/topics/computer-science/shannon-sampling-theorem

Sampling

- Sampling rate
  - Depends on the characteristics of the signal and on our purposes
  - limited by the technology used for the ADC
    - Sometimes, trade-off between speed and precision



https://en.wikipedia.org/wiki/Analog-to-digital_converter#Aliasing
https://www.sciencedirect.com/topics/computer-science/shannon-sampling-theorem

Shannon's sampling theorem

- **Nyquist rate:** twice the bandwidth of a band-limited function or a band-limited channel
  - i.e., twice the highest frequency (of interest) of the signal
- <u>A faithful reproduction of the original signal is only possible if the sampling rate is higher than the Nyquist rate</u>

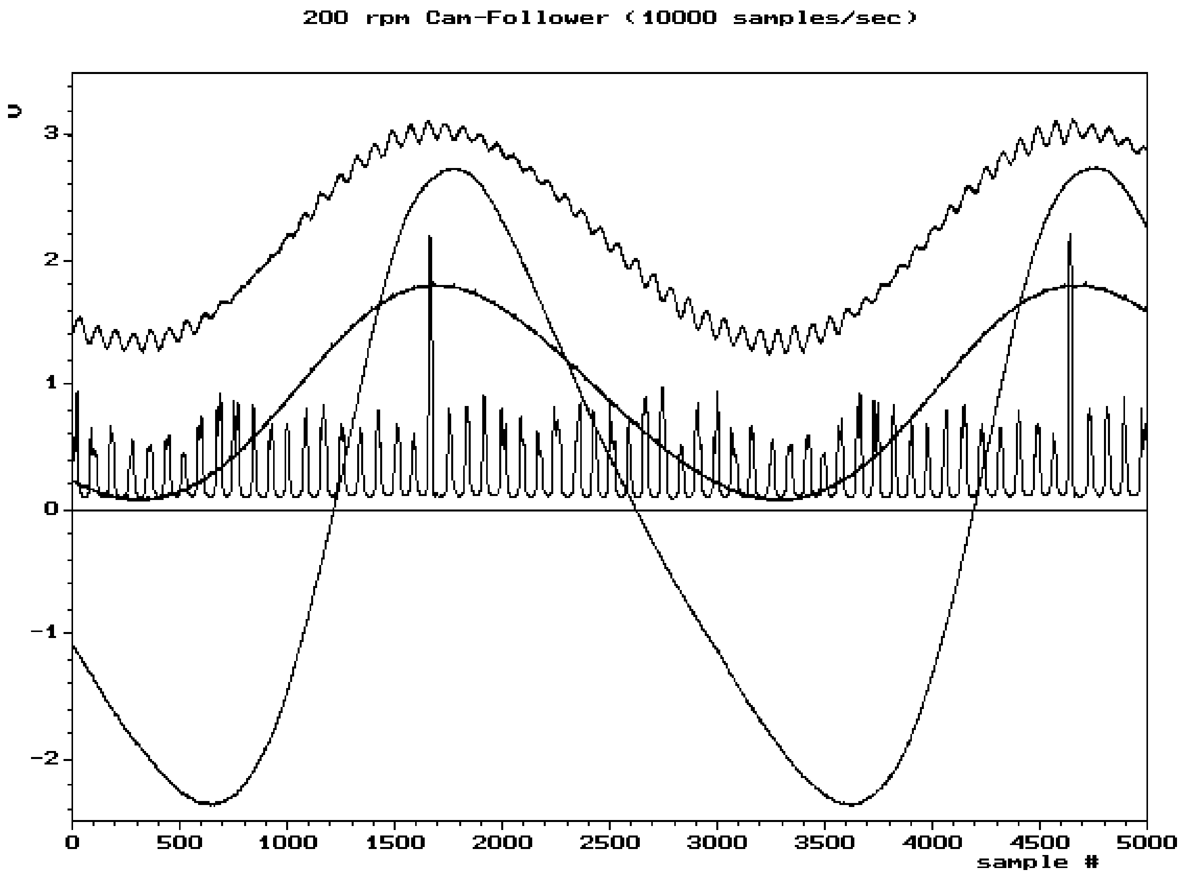
https://en.wikipedia.org/wiki/Analog-to-digital_converter#Aliasing

Shannon’s sampling theorem

- Musical instruments and voice include frequencies in different ranges
- For reproducing audio faithfully, we are only interested in those frequencies that are audible to humans (i.e., we do not need to reproduce frequencies that we cannot hear)
  - Audible frequency range: 20–20,000 Hz
  - We can consider human beings as “the channel”
- Typical sampling rate for digital audio (CD-DA coding) is 44.1kHz
  - Double of the maximum audible frequency, with some margin

https://en.wikipedia.org/wiki/Analog-to-digital_converter#Aliasing

Sampling



https://commons.wikimedia.org/wiki/File:Aliasing.gif

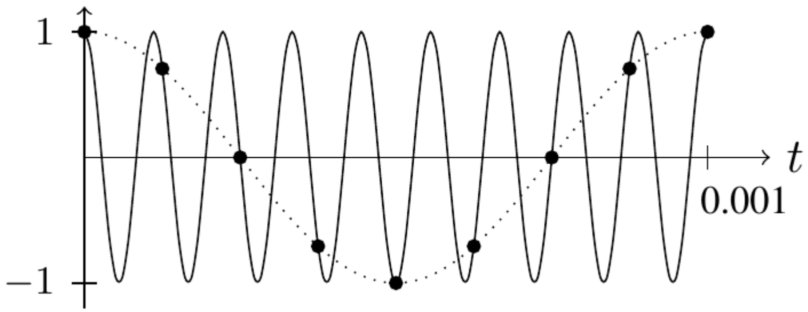
Sampling - Aliasing

- An effect that causes different signals to become indistinguishable when sampled
  - Aliases of one another
- If frequencies above half the Nyquist rate are sampled, they are incorrectly detected as lower frequencies

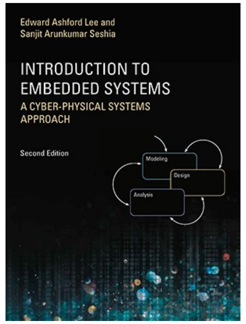
https://en.wikipedia.org/wiki/Aliasing

Sampling

- Consider a sinusoidal signal at 1kHz:
  - $$x(t) = \cos(2000\pi t)$$
- Sampling at 8kHz with no sensor distortion, the signal is
  - $$s(n) = f(x(nT)) = \cos(\pi n/4)$$
- Let us consider a second signal at 9kHz
  - $$x'(t) = \cos(18,000\pi t)$$
- The sampled signal at 8kHz is:



$$s'(n) = \cos(9\pi n/4) = \cos(\pi n/4 + 2\pi n) = \cos(\pi n/4) = s(n)$$

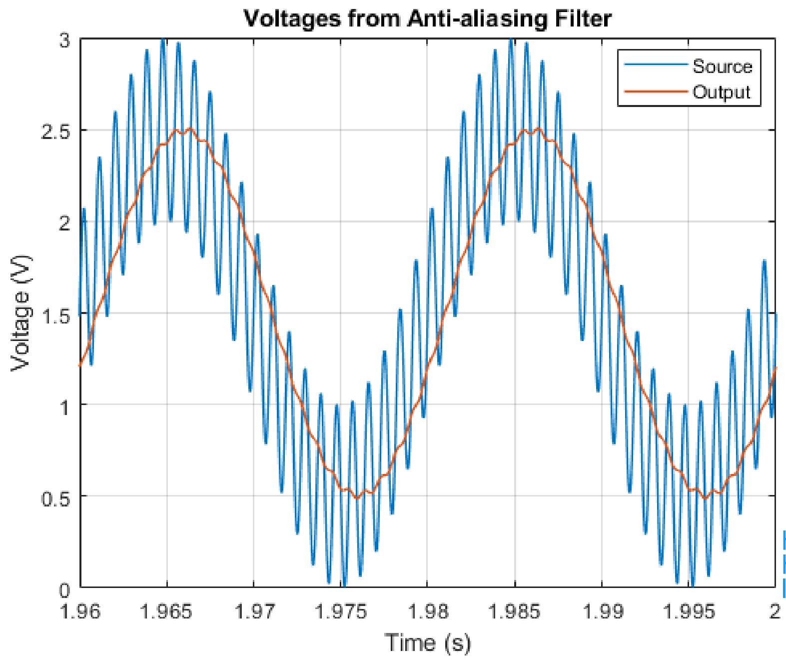


- The signal at 1kHz and the one at 9kHz look the same once sampled: **aliasing**

https://en.wikipedia.org/wiki/Analog-to-digital_converter#Aliasing

Sampling - Aliasing

- **Anti-aliasing filter**: when sampling signals that also contain higher frequencies that are not of interest, a *low-pass filter* is used to remove frequencies above half the sampling rate
- Otherwise, aliasing might occur
  - The sampled signal will contain the higher frequencies “transformed” to lower ones



https://it.mathworks.com/help/sps/ug/analog-anti-aliasing-filter.html

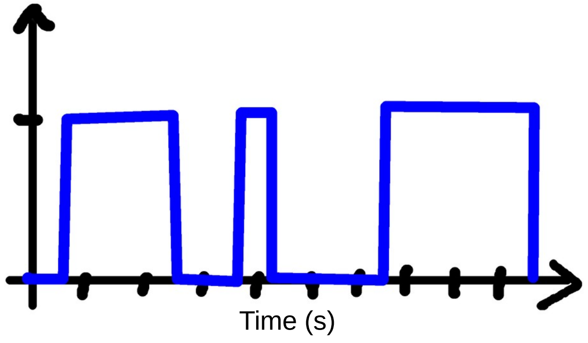
https://en.wikipedia.org/wiki/Aliasing

Why not always sampling at the highest possible frequency?

- Sampling at higher frequencies than strictly required is inefficient as it leads to an increased use of resources
  - Higher energy usage for sampling
  - Higher memory for storing data and bandwidth requirements (storage, network, …)
    - Higher energy
  - Higher computational resources for data processing
    - Higher energy
- For example:
  - Signal sampled at 100Hz, 8-bit: 100 byte/s
  - Signal sampled at 1kHz, 8-bit: 1 Kb/s
  - Signal sampled at 100kHz, 8-bit: 100 Kb/s

Sampling - Discrete events

- Usually, we do not want to reconstruct a signal, we rather want to detect certain discrete events
    - **Which is the shortest event that we would like to capture?**
    - **Do we want to capture only if an event occurred or also its length?**
- Example: design a system that detects if a light has been activated
    - Sample every 1s: it guarantees to catch the light if it is on for more than 1s. If it is on for less than 1s, there is no guarantee that it is sampled correctly
- Example: design a system that detects for how long the light remains on:
    - The system needs to detect when the light switches on and it switches off
    - Let us consider a sampling period of 1s: the error in measuring time is <2s
    - Let us consider a sampling period of 1ms: the error in measuring time is <2ms
    - What is the maximum error that we can tolerate?



Quantization

- Quantization is about converting continuous amplitude values (i.e., with infinite precision) to discrete (i.e., finite precision) values
  - It results in approximated values
  - The approximation should be compatible with the expected precision
- Analog to Digital Converters (ADCs) are used to perform sampling and quantization
  - One of the main characteristics of ADCs is the number of bits that they use for representing values
    - e.g., 4-bit ADC or 8-bit ADC

ADC Resolution

- An n-bit ADC will use n bits to represent the analog input in the allowed range by using digital numbers
  - 2^n different possible values
  - The analog signal range will be mapped to the available 2^n digital values
- Resolution of an ADC tells us how accurate our conversion is going to be

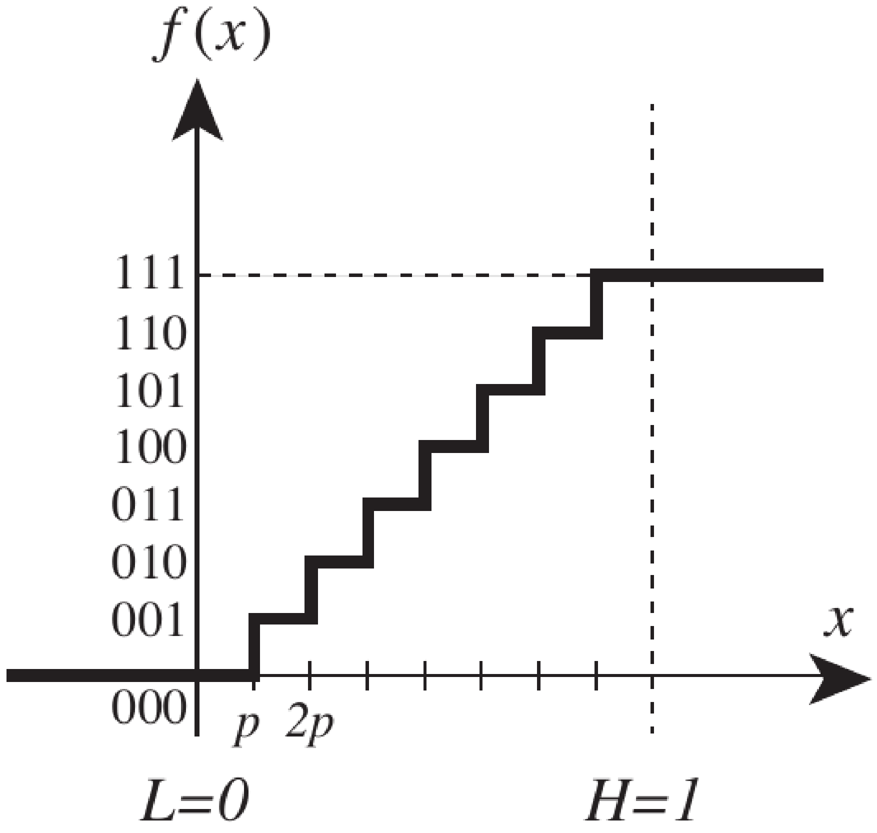
ADC Resolution

- A 4-bit ADC will be able to represent the signal by using 16 different values
  - Analog signal in the range [0..63]
  - The 4-bit ADC will map this interval to 16 different discrete values [0...15]
    - 0 is mapped to 0, 63 to 15; intermediate values are mapped to values with steps of 4

ADC Resolution and Sensor Precision

- 4-bit ADC
- Analog temperature sensor
  - Signal in the range [0...63]mV corresponding to a range -12, +20°C:
- 32°C range mapped to [0...63]mV range→ 0.5° for each unit variation in the signal
- A variation in the signal of 4 units (64/16) is required to obtain a change in the converted value → precision of the digital sensor: 2°C

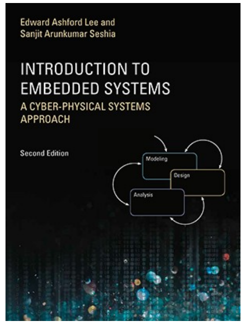
ADC Resolution



- 3-bit digital sensor (3-bit ADC) capable of measuring 0-1V:
  - Precision=$1/2^3$
  - Dynamic range:

$$D_{dB} = 20\log_{10}\left(\frac{H-L}{p}\right) \approx 18dB.$$

- Each additional bit yields approximately 6 decibels of dynamic range



ADC Quantization Error

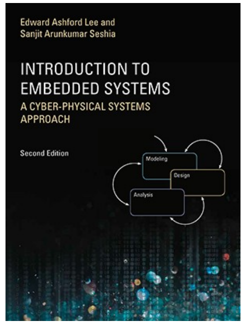
- Errors are measured in a unit called the least significant bit (LSB)
  - In an eight-bit ADC, an error of one LSB is 1/256 of the full signal range, or about 0.4%
- It is a rounding error between the analog input voltage and the output digitized value
  - The error is nonlinear and signal-dependent
    - In an ideal ADC, the quantization error is uniformly distributed between −1/2 LSB and +1/2 LSB
- Quantization error and non-linearity are intrinsic to any analog-to-digital conversion
  - All ADCs suffer from nonlinearity errors caused by their physical imperfections
- Jitter caused by non-ideal sampling clock

https://en.wikipedia.org/wiki/Analog-to-digital_converter#Aliasing

One-bit ADC: Analog Comparator

- Compares a signal value against a threshold, producing a binary digital value:

$$f(x(t)) = \begin{cases} 0 & \text{if } x(t) \leq 0 \\ 1 & \text{otherwise} \end{cases}$$



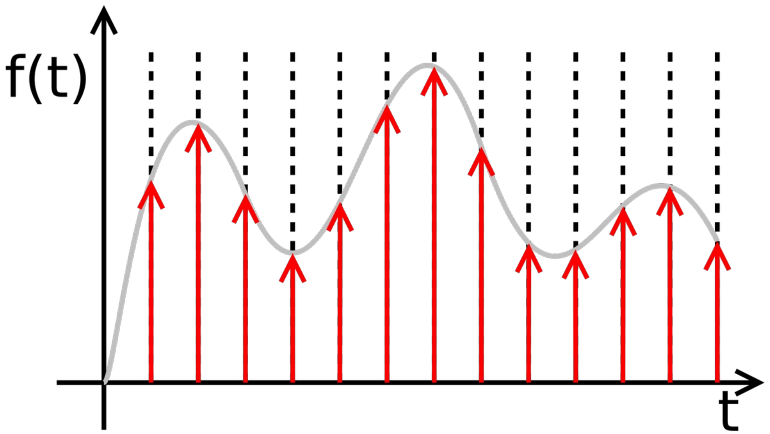
Digital to Analog Conversion

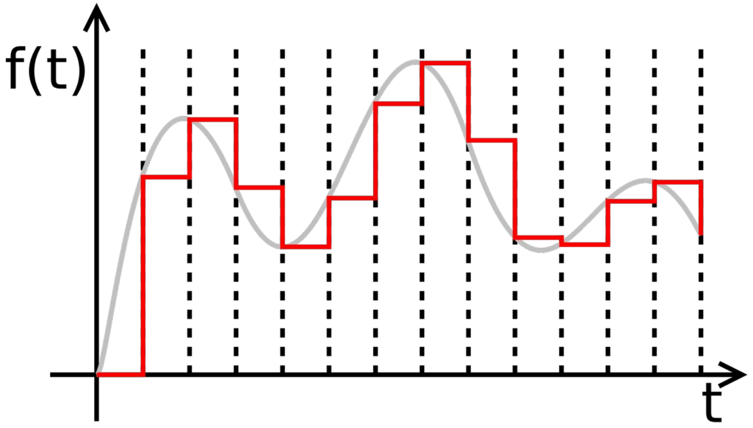
- A DAC converts an abstract finite-precision number (usually a fixed-point binary number) into a physical quantity (e.g., a voltage)
- Digital-to-analog conversion can degrade a signal, so a DAC should be specified that has insignificant errors in terms of the application

Digital to Analog Conversion

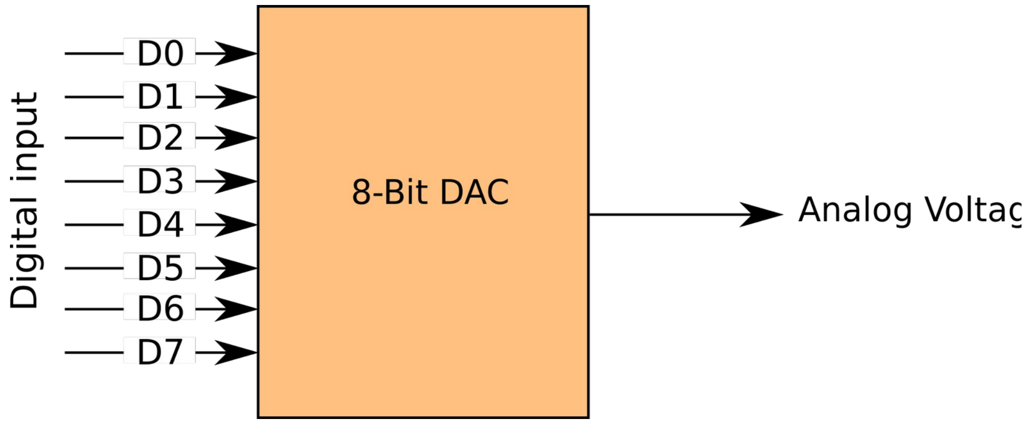
- Ideal DAC:
  - Converts the abstract numbers into a conceptual sequence of impulses that are then processed by a reconstruction filter using some form of interpolation to fill in data between the impulses

- A conventional practical DAC:
  - For example, converts the numbers into a piecewise constant function made up of a sequence of rectangular functions



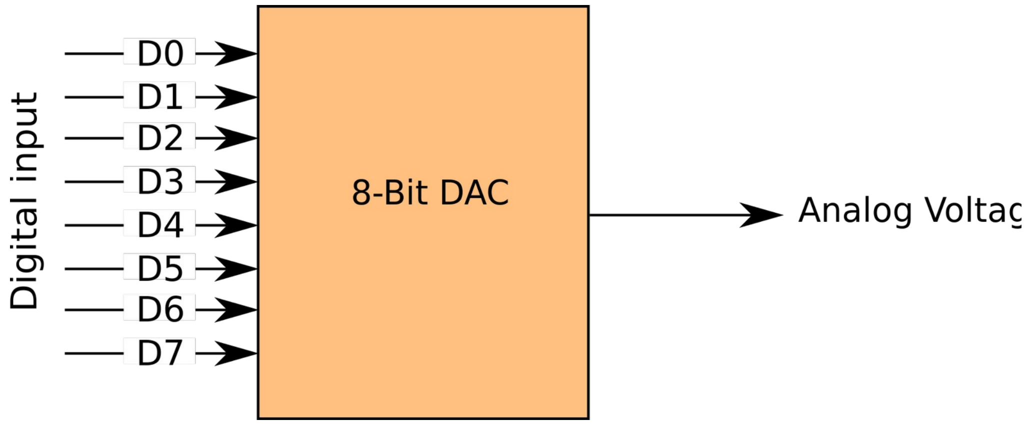


DAC - Performance



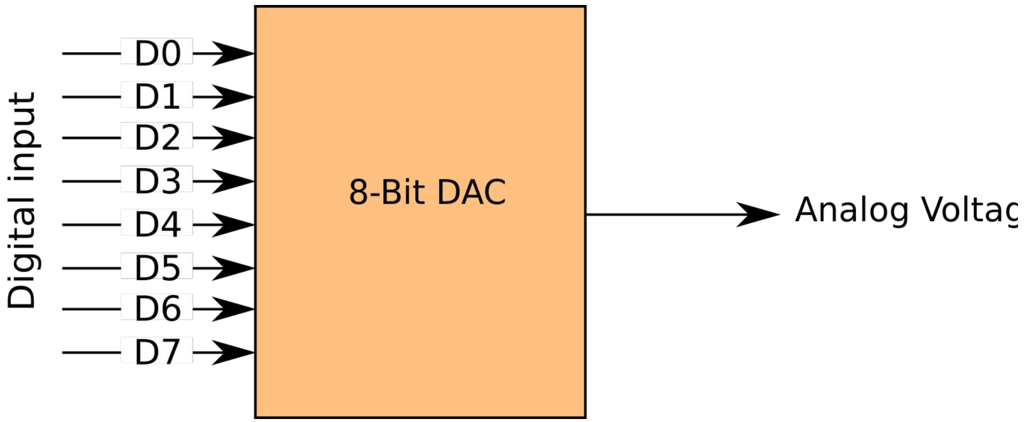
- Resolution: The number of possible output levels the DAC is designed to reproduce
  - Usually stated as the number of bits it uses:
    - For instance a 1-bit DAC is designed to reproduce 2 levels while an 8-bit DAC is designed for 256 levels
- Whenever the maximum output voltage is insufficient, we may need to adapt it by using an amplifier

DAC - Performance



- Maximum sampling rate: The maximum speed at which the DAC circuitry can operate and still produce correct output
- Monotonicity: The ability of a DAC's analog output to move only in the direction in which the digital input moves
  - i.e., if the input increases, the output doesn't dip before asserting the correct output

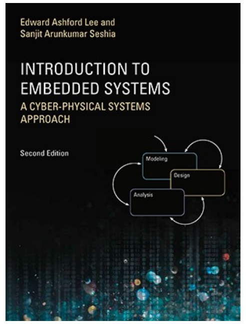
DAC - Performance



- Total harmonic distortion and noise (THD+N): A measurement of the distortion and noise introduced to the signal by the DAC
  - Expressed as a percentage of the total power of unwanted harmonic distortion and noise that accompany the desired signal
- Dynamic range: A measurement of the difference between the largest and smallest signals the DAC can reproduce
- Jitter

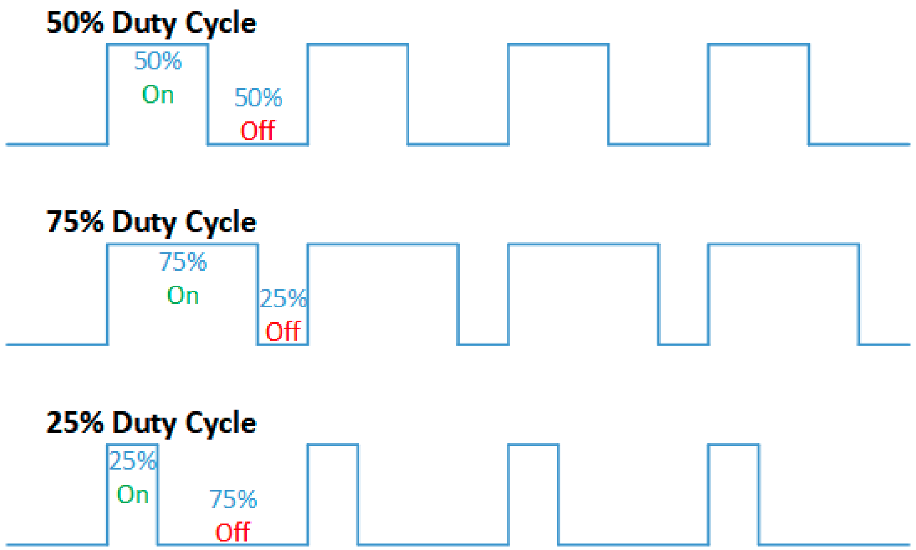
Pulse Width Modulation (PWM)

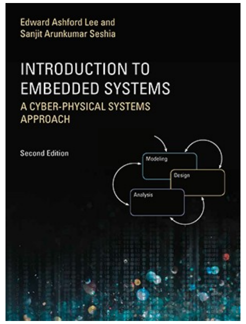
- Used to control analog devices by means of a digital output without a DAC
- The device must be able to tolerate rapid switches on and off of the power source
  - e.g. LEDs, incandescent lamps, DC motors



Pulse Width Modulation (PWM)

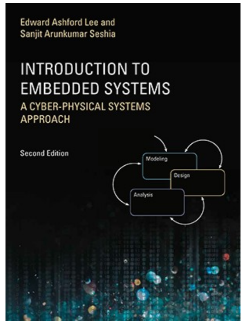
- The PWM signal has a cycle period
  - The signal is held high for a fraction of this period called **duty cycle**
  - The length of the duty cycle determines the control of the actuator
- E.g., variable-speed computer fan controllers





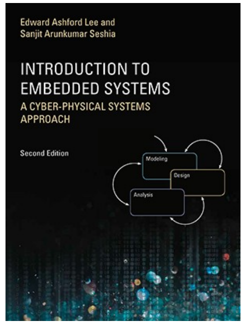
Noise

- $x'(t) = x(t) + n(t)$
- Part of the signal that we do not want:
  - we want to measure x, but we are able to measure x’
- Due to:
  - Sensor imperfections
  - Quantization
  - System and environment
    - e.g., we want to measure acceleration, but our accelerometer also measures vibrations



Noise - RMS

- Root Mean Square (RMS): square root of the average value of $n(t)^2$
  - $$N = \lim_{T\to\infty} \sqrt{\frac{1}{2T}\int_{-T}^{T}(n(\tau))^2 d\tau}$$
- Signal to Noise Ratio (SNR):
  - $$SNR_{dB} = 20\log_{10}\left(\frac{X}{N}\right)$$
    - X: RMS of the input signal x



Signal Conditioning

- Noise and the signal often have different characteristics
- We can exploit differences to filter out noise
  - Frequency selective filters
    - E.g., we measure acceleration of a vehicle: vertical accelerations due to vibrations are at higher frequencies, and we can (partly) filter them out
  - Sometimes, a simple average of digital values over a time window can mitigate high-frequency noise

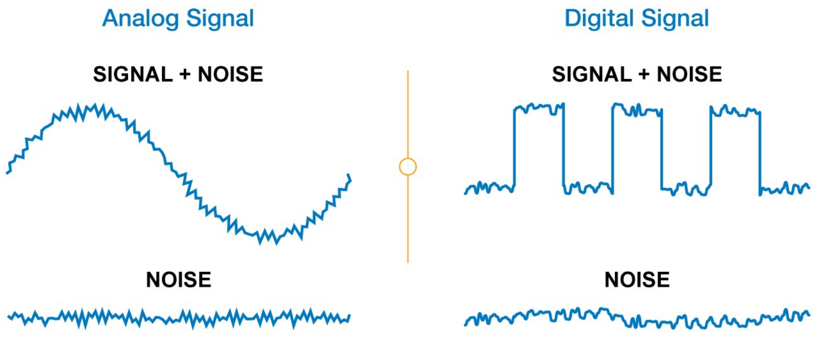
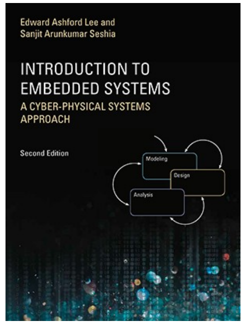


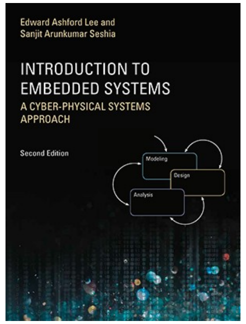
FIGURE 1. Noise in Analog and Digital Signals

https://www.predig.com/whitepaper/reducing-signal-noise-practice



Harmonic Distortion

- Typically occurs when the sensitivity of the sensor depends on the magnitude of the signal, even within operating range
- Nonlinear effect that can be modeled by powers of the considered physical quantity
  - Second harmonic distortion is a dependence on the square of the physical quantity
  - $$f(x(t)) = ax(t) + b + d_2(x(t))^2$$
  - If d_2 is small, the model approximates to the linear one



Oversampling

- Sampling frequency significantly higher than the Nyquist rate
- Improves resolution and signal-to-noise ratio
  - When oversampling by a factor of N, the dynamic range also increases by a factor of N
  - SNR increases by sqrt(N)
  - The number of samples required to get n bits of additional data precision is 2^{2n}
    - The 2^{2n} additional samples are combined to lower the SNR at the same level of a converter with n additional bits
    - E.g., to implement a 24-bit converter, we can use a 20-bit converter running at 256 times the target sampling rate
      - Combining 256 consecutive 20-bit samples can increase the SNR by a factor of 16, effectively adding 4 bits to the resolution and producing a single sample with 24-bit resolution
- Helpful in avoiding aliasing by relaxing anti-aliasing filter performance requirements

https://en.wikipedia.org/wiki/Oversampling

Oversampling

- Oversampling signals require
  - Digital filtering
  - Downsampling: the sampling rate is reduced by combining samples

https://en.wikipedia.org/wiki/Oversampling

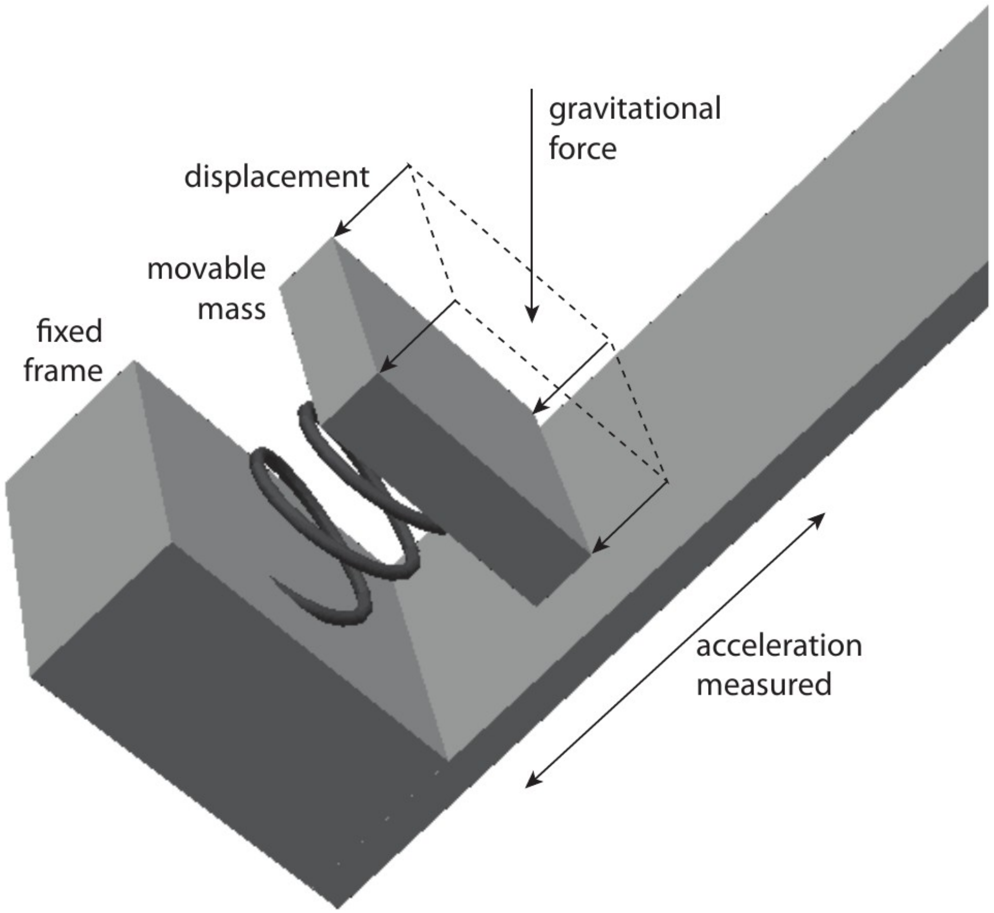
ADC in The Nordic Semiconductor nRF 52840 – SAADC

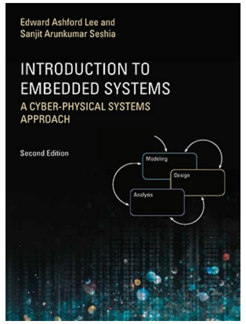
- Maximum 12-bit
  - 2^{12}=4,096 different levels
  - Can be changed to lower resolutions
    - 8/10/12-bit resolution; 14-bit resolution with oversampling
  - Samples stored as 16-bit 2’s complement
- Up to eight input channels
  - Scan mode to sample a series of channels in sequence
- Maximum sampling frequency of 200kHz
  - Continuous sampling without the need of an external timer
- Adjustable input gain

https://infocenter.nordicsemi.com/index.jsp?topic=%2Fcom.nordic.infocenter.nrf52832.ps.v1.1%2Fsaadc.html

Common Sensors – Tilt and Acceleration

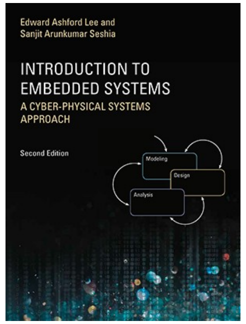
- An accelerometer is a sensor that measures the acceleration of an object
  - As observed by an observer in free fall
- It measures not only acceleration, but also gravitational force
  - It can be challenging to separate them in measures
- Accelerometers are typically implemented in silicon
  - Silicon fingers deform under gravitational pull or acceleration
  - Circuitry measures the deformation and provides a digital reading
- Often, three accelerometers are packaged together, giving a three-axis accelerometer





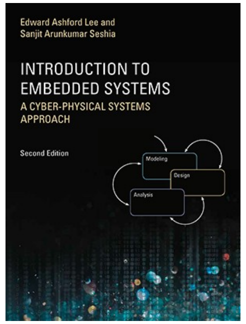
Common Sensors – Rotation

- A gyroscope measures changes in orientation
  - Mostly unaffected by gravitational field



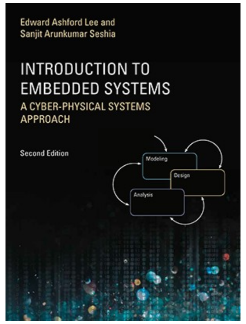
Common Sensors – IMU (Inertial Measurement Unit)

- Fusion of data from:
  - A gyroscope to measure changes in orientation
  - A 3-axes accelerometer to measure acceleration
  - A magnetometer to measure orientation (compass)
- Can be used to estimate position, often coupled with another positioning system
  - E.g., GPS outdoor and IMU for short indoor periods



Common Sensors – Position and Velocity

- Using an accelerometer is not a good idea, unless for very short periods of time and starting from known position and speed
- IMUs can do better but they still work well for limited periods of time
- Otherwise:
  - Anemometer
  - Satellite positioning (GPS, Galileo, Glonass, …)
  - Wifi / Bluetooth / UltraWideBand

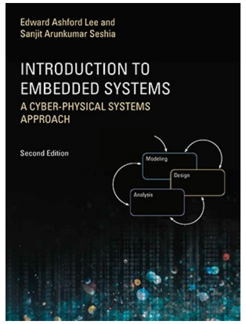


Common Sensors – Other sensors

- Temperature and humidity
- Sound
- Luminosity
- Presence
- Distance
- Chemicals
- ...

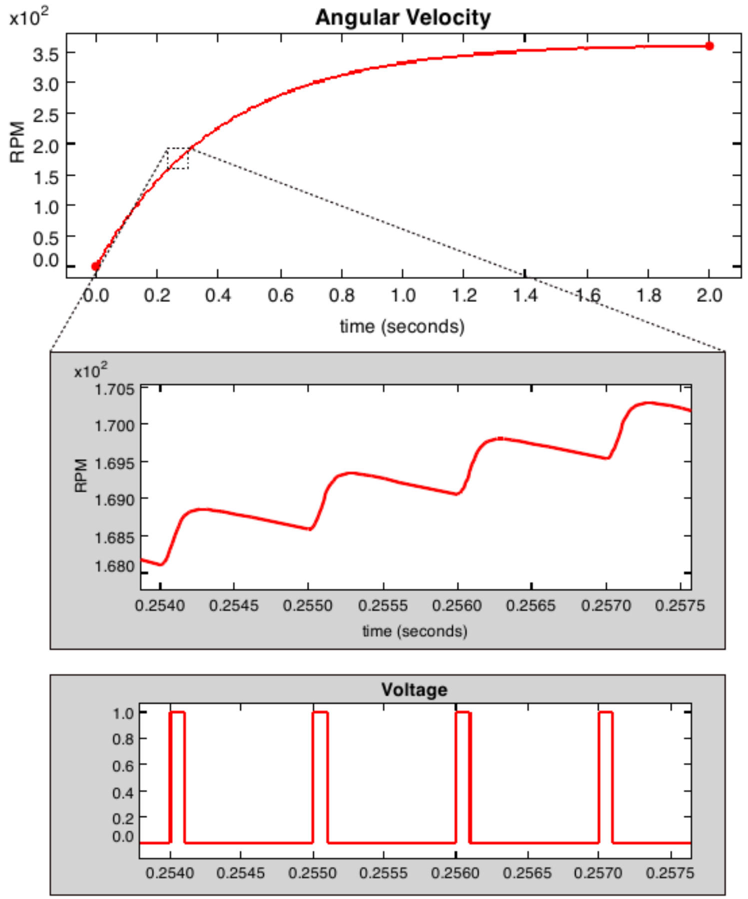
Common Actuators – LEDs

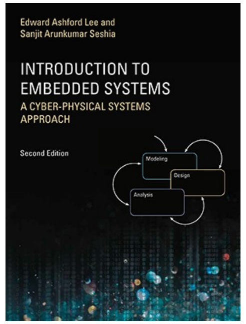
- Light Emitting Diodes are small devices that emit light
- They can be driven directly by digital outputs



Common Actuators – Motors

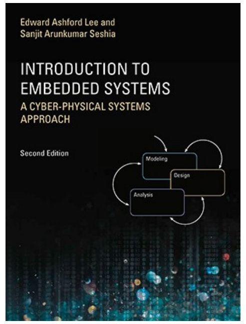
- A motor applies a torque (angular force) to a load proportional to the current through the motor windings
- The motor requires an analog input
  - We can use a DAC, but DACs often cannot deliver enough power
    - A power amplifier is required
      - Expensive, bulky, and inefficient
  - PWM
    - We use feedback on the speed to adjust the duty cycle





Common Actuators – Motors

- A **servo motor** is any motor coupled with a feedback sensor to facilitate positioning
- **Stepper motors** move from point to point in fixed increments



Common Actuators – Relay

- A relay is an electrically operated switch
- A digital output can be used to drive a relay and, thus, to open or close a circuit
  - e.g., switch on or off a lamp that requires 220V AC
- The correct type, depending on the required power to be handled, needs to be selected

Practical Notes – Calibration

- Sensor and actuators often need to be calibrated before use
- Calibration requires setting some coefficients that correlate the values reported by the sensor (voltage) to the physical quantity being measured:
  - E.g., When our temperature sensor provides 0.1V, which is the correspondent measured temperature?
- Calibration requires reference points
  - E.g., temperature precisely set, or predetermined CO_2 values
  - The coefficients are then computed by using the (linear) model of the sensor

Practical Notes – Sensor/Actuator Installation

- Sensor/actuator installation needs to be planned carefully
- It might influence how the device works and the obtained results
  - Compensation might be required if installation is not “optimal”
- For example, measuring the temperature of a room might provide different results depending on where we measure it
- Measured humidity of soil might depend on the specific location where we place the sensor
  - If the sensor is in a spot that is far away/close from/to water sources, it might provide humidity levels that are lower/higher than expected
  - If the sensor is in a spot that is often lighted by Sun, it might provide humidity levels that are lower than expected