Telematics I

Chapter 3
Physical Layer

- Baseband transmission over physical channels
- Limitations on data rate: Nyquist and Shannon
- Clock extraction
- Broadband versus baseband transmission
- Examples

(Acknowledgement: These slides have been taken from Prof. Karl's set of slides)

Goals of this Chapter

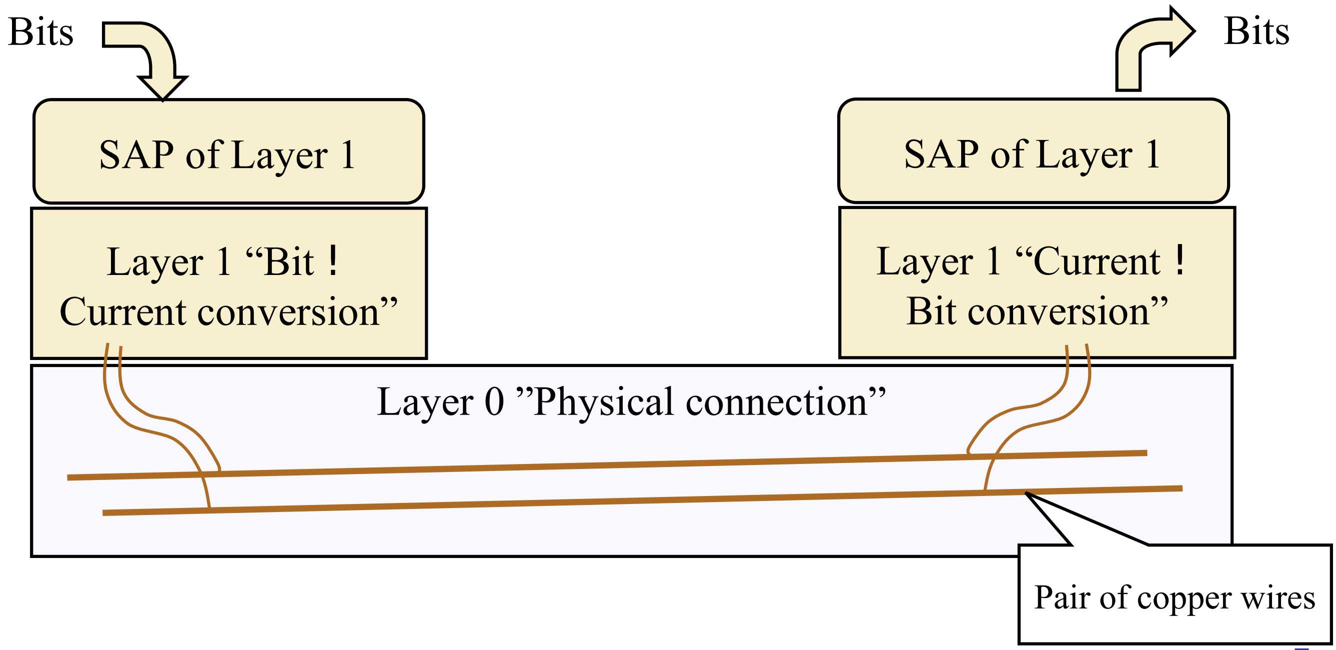
- Answer the basic question: how can data be transported over a physical medium?

- Understand the basic service provided by a physical layer
- Different ways to put "bits on the wire"
- Reasons why performance of any physical layer is limited
- Reasons for errors
- A few examples of important physical layers

- Note: This is *vastly* simplified material

Basic Service of Physical Layer: Transport Bits

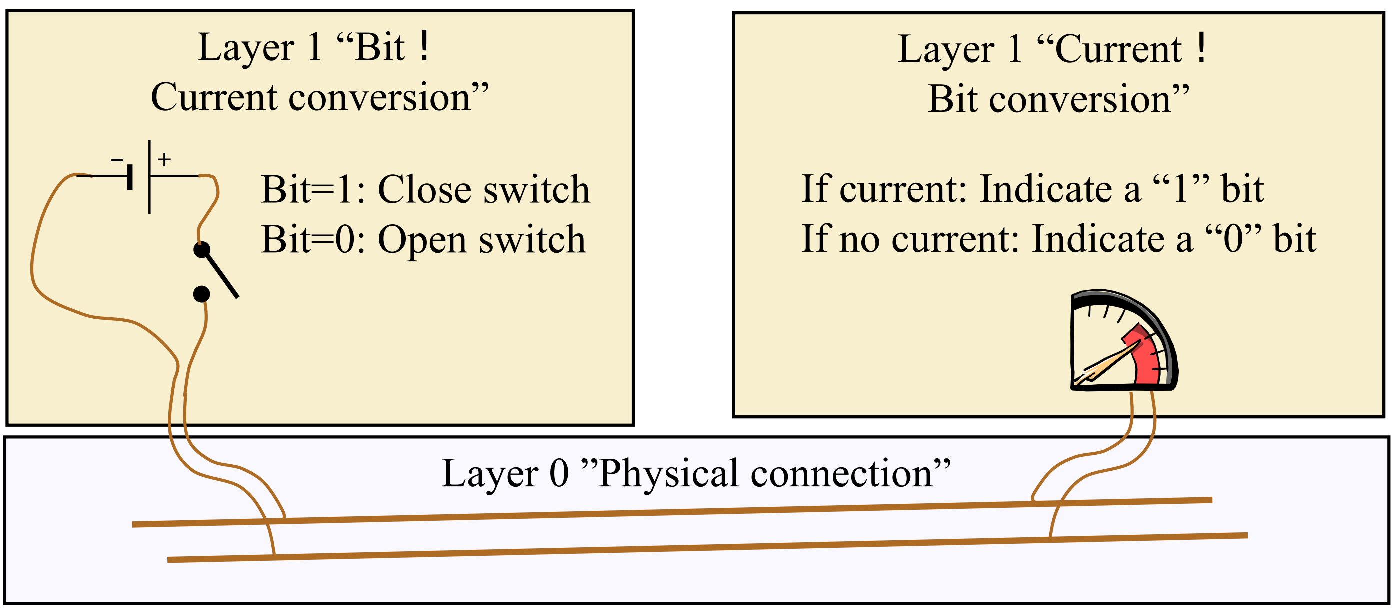
- The physical layer should enable the transport of bits between two locations A and B
- Abstraction: Bit sequence – correct, in order delivery



A Bit-to-Signal Conversion Rule

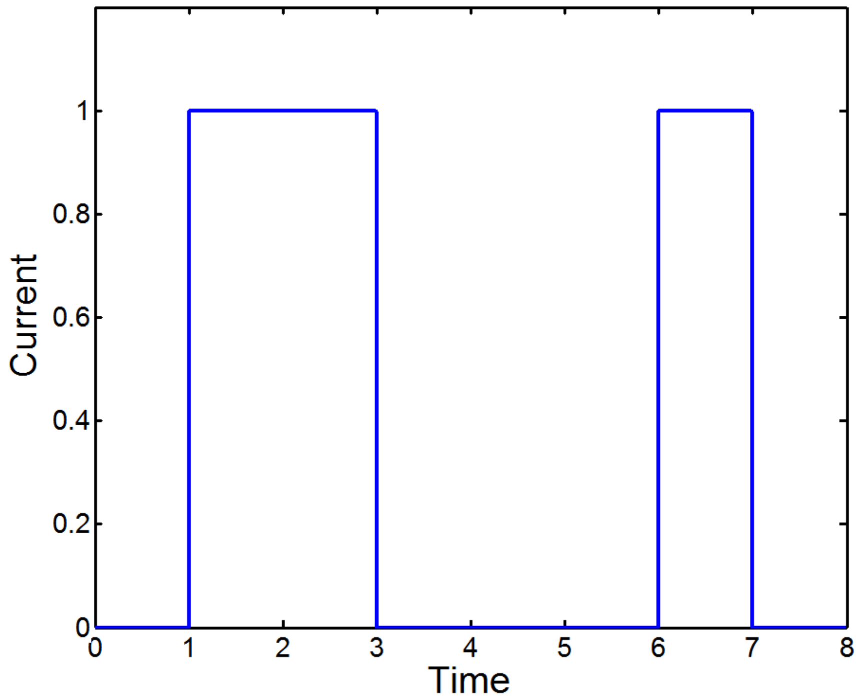
- A simple conversion rule
  - For a "1" bit, apply current to the pair of wires
  - For a "0" bit, no current

This is called "Non return to zero" NRZ



Example: Transmit Bit Pattern for Character “b”

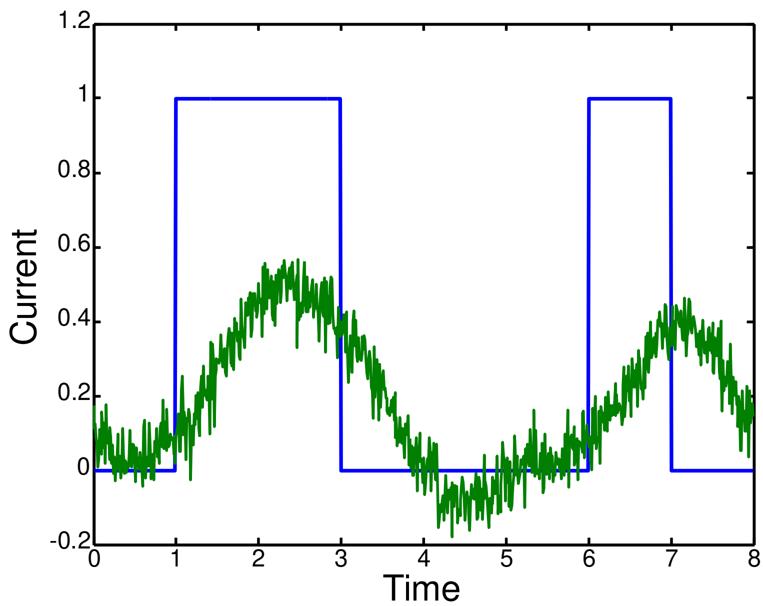
- Character “b” needs a representation as a sequence of bits
- One option: Use the ASCII code of “b”, 98, as a binary number 01100010
- Resulting current put on the wire:



Note: Abstract ***data*** is represented by physical ***signals*** – changes of a physical quantity in time or space!

What Arrives at the Receiver?

- Typical pattern at the receiver:



- What is going on here?

Note: this and the following examples are exaggerated!

Some Background: Fourier Analysis

- To understand signal propagation on a physical medium, some background is required how such signals can be analyzed/treated mathematically
- First: ***Fourier's theorem***
  Any periodic function *g(t)* (with period *T*) can be written as a (possibly infinite) sum of sine and cosine functions; the frequencies of these functions are integer multiples of the fundamental frequency $f = 1/T$.

$$g(t) = 1/2c + \sum_{n=1}^{\infty} a_n \sin(2\pi n f t) + \sum_{n=1}^{\infty} b_n \cos(2\pi n f t)$$

  Constants c, a_n, b_n are to be determined.

Fourier Analysis – Computing Coefficients

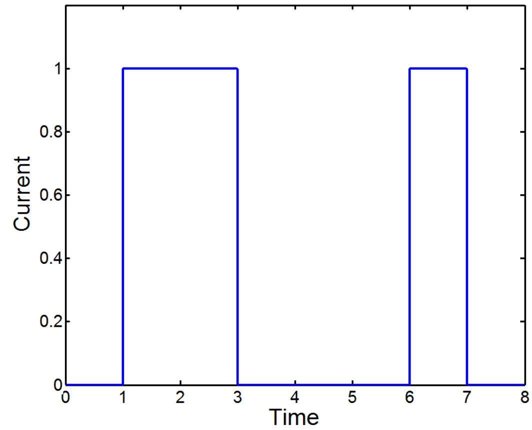
- Coefficients c, a_n, b_n in the Fourier series can be computed:
  - $$c = \frac{2}{T}\int_0^T g(t)\mathrm{d}t$$
    $$a_n = \frac{2}{T}\int_0^T g(t)\sin(2\pi n f t)\mathrm{d}t$$
    $$b_n = \frac{2}{T}\int_0^T g(t)\cos(2\pi n f t)\mathrm{d}t$$
  - Because of orthogonality of sines and cosines as basis functions
- The *n*th summary terms are called ***harmonics***
- The sum of the squares of the *n*th coefficients – $a_n^2 + b_n^2$ – is proportional to the ***energy contained in this harmonic***
  - Usually, root of it is shown - $(a_n^2 + b_n^2)^{1/2}$

TELEMATIK Rechnernetze

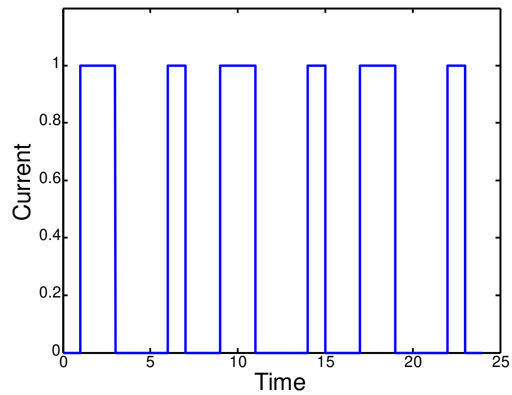
Applying Fourier Analysis to the Example

- The transmitted waveform of ‘b’ is not a periodic signal – Fourier not applicable directly

- Use a trick: Suppose waveform is repeated
infinitely often, resulting in a periodic waveform
with period 8 bit times



Repeated waveform for bit pattern ‘b’

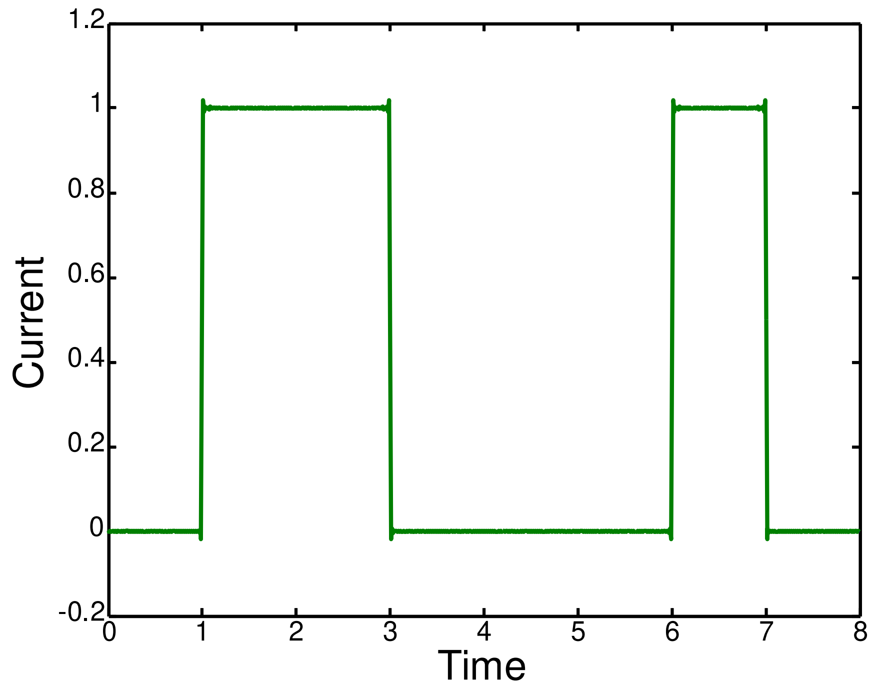


TELEMATIK Rechnernetze

Applying Fourier Analysis to the Example

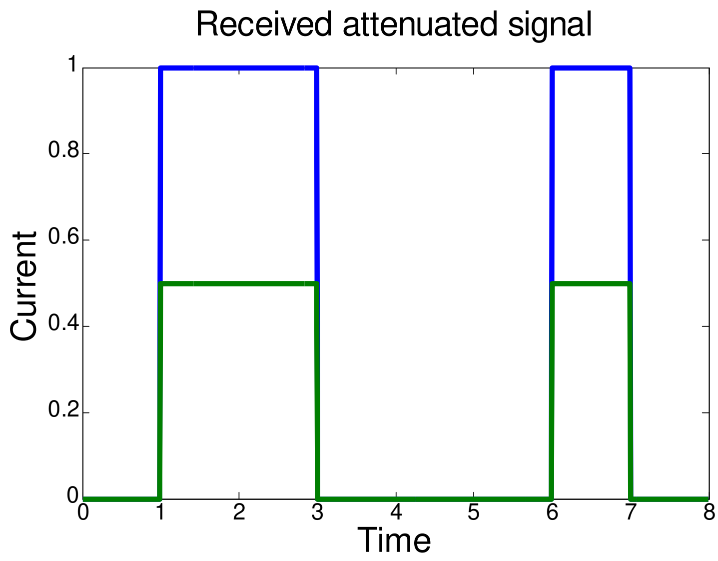
- Result of computing a_n, b_n, c and using 512 terms to represent the signal:

- Almost no discernible difference between original signal and Fourier series
  - Curves overlap



Fact 1: Signals are Attenuated in a Physical Medium

- ***Attenuation*** α : Ratio of transmitted to received power
  - $P_{\text{recv}} = P_{\text{trans}}/\alpha \Leftrightarrow \alpha = P_{\text{trans}}/P_{\text{recv}}$
  - High attenuation $\Rightarrow$ little power arrives at receiver
- Attenuation depends on
  - Actual medium
  - Distance between sender and receiver
  - … other factors
- Normalized, typically given in dB

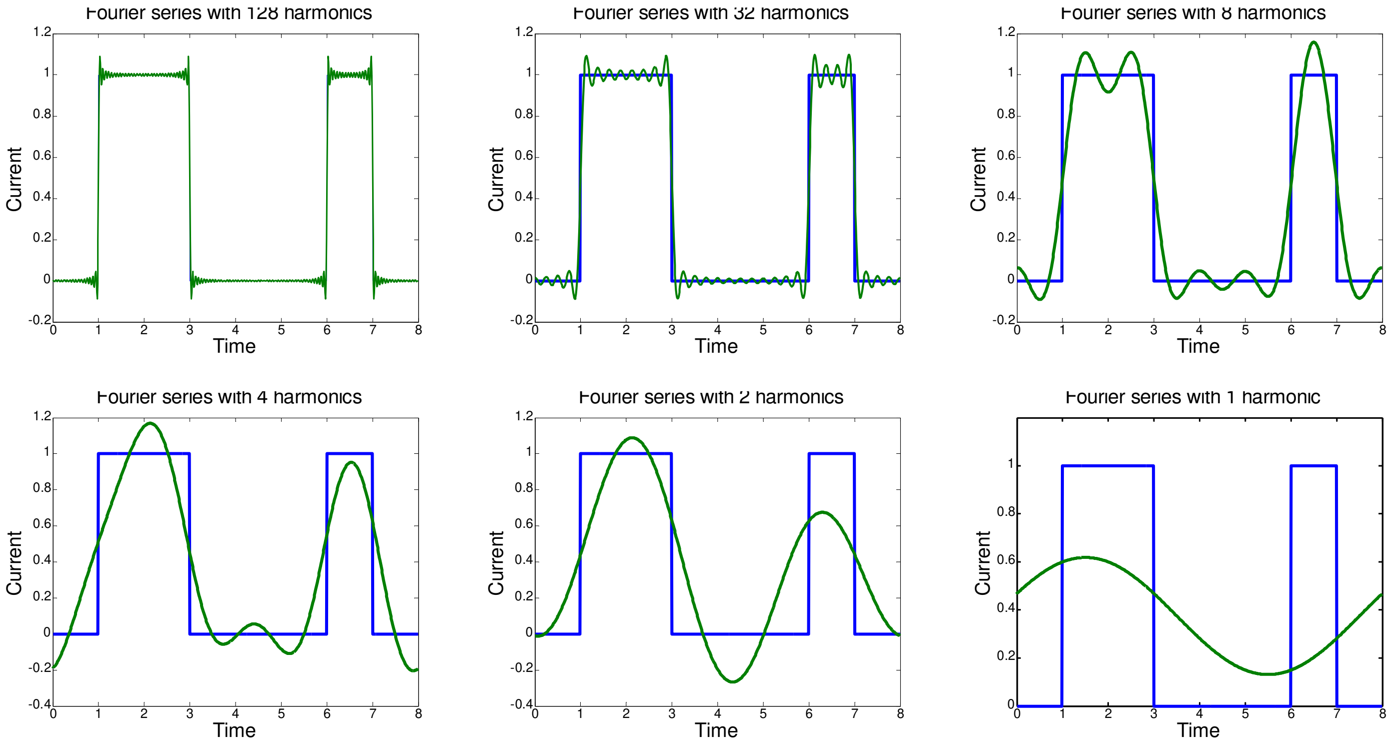


Fact 2: Not all Frequencies Pass Through a Medium

- Previous picture assumed that all frequencies travel unhindered through a physical medium
- This is not the case for real media!
- Simplified behavior: frequencies up to given upper bound f_c can pass; higher frequencies are suppressed
  - Mathematically: the Fourier series is cut off at a certain harmonic
  - High frequencies are ***attenuated*** to zero
- This frequency f_c is called the ***bandwidth*** of a physical medium (or channel)
  - Smaller f_c means fewer harmonics can pass through

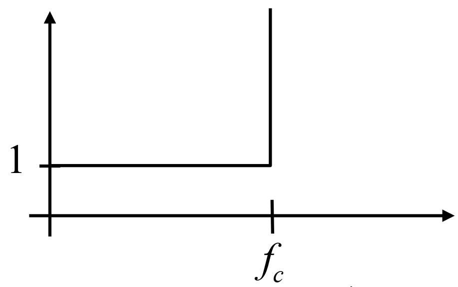
Bandwidth-Limited Medium – Example

- Result when fewer and fewer harmonics are transported

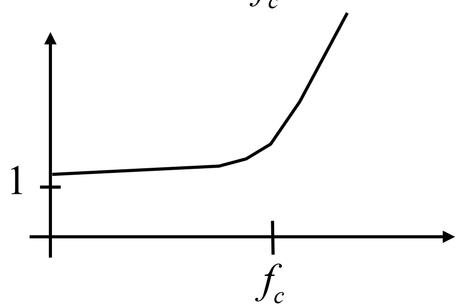


Fact 3: Attenuation Depends on Frequency

- Model just used: Cutoff
  - Attenuation is 1 below bandwidth, infinite above



- More realistic: attenuation depends on frequency
  - Attenuation close to 1 below bandwidth and increases for higher frequencies

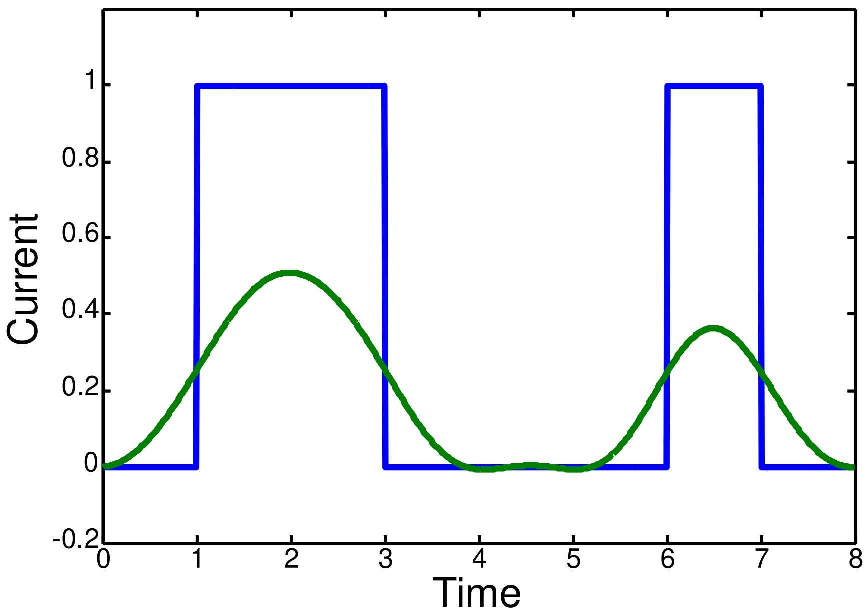


- Both are examples of ***bandwidth-limited medium / channel***

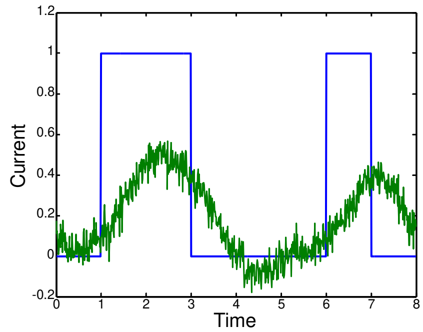
Example with Frequency-Dependent Attenuation

- Suppose attenuation is 2, 2.5, 3.333… , 5, 10, 1 for the 1st, 2nd, … harmonic

Received signal with
frequency-dependent attenuation



We have to explain this behavior:



Fact 4: Media does not only Attenuate, but also Distorts

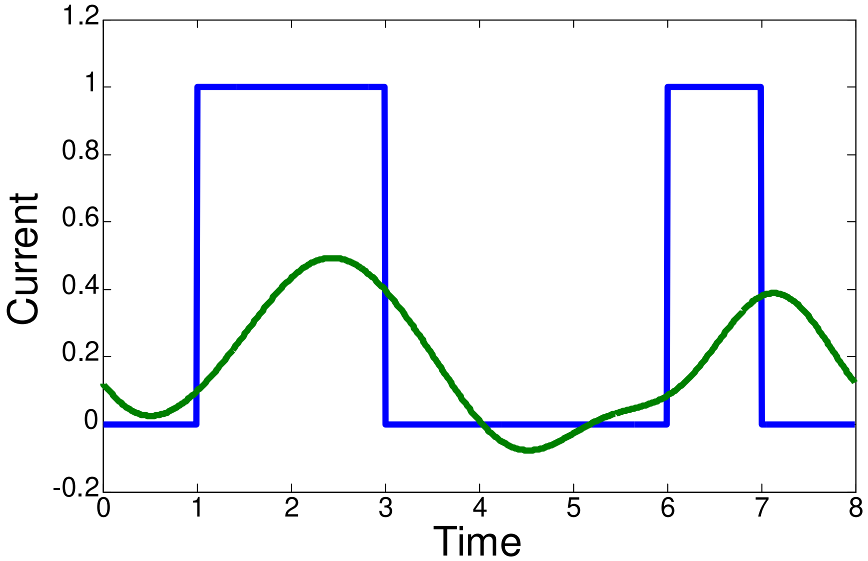
- In a physical medium other than vacuum, different frequencies have different propagation speed
  - Some wave lengths travel faster than others
- Apparent result: Waves arrive at receiver out of phase
  - Recall: a sine wave is determined by amplitude *a*, frequency *f*, and phase ϕ

$$a \sin(2\pi f t + \phi)$$

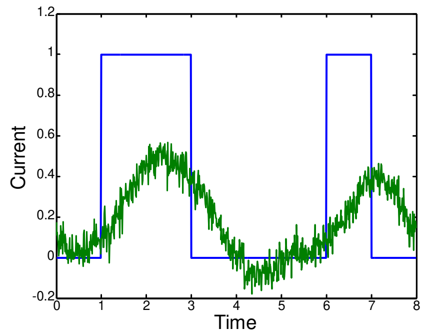
  - This is called ***distortion***
    - Sometimes also "jitter", but the term jitter will re-appear later with a different, more common definition
- Amount of phase shift depends on frequency

Example with Frequency-Dependent Attenuation and Jitter

Received signal with frequency-dependent attenuation and phase change



We have to explain this behavior:



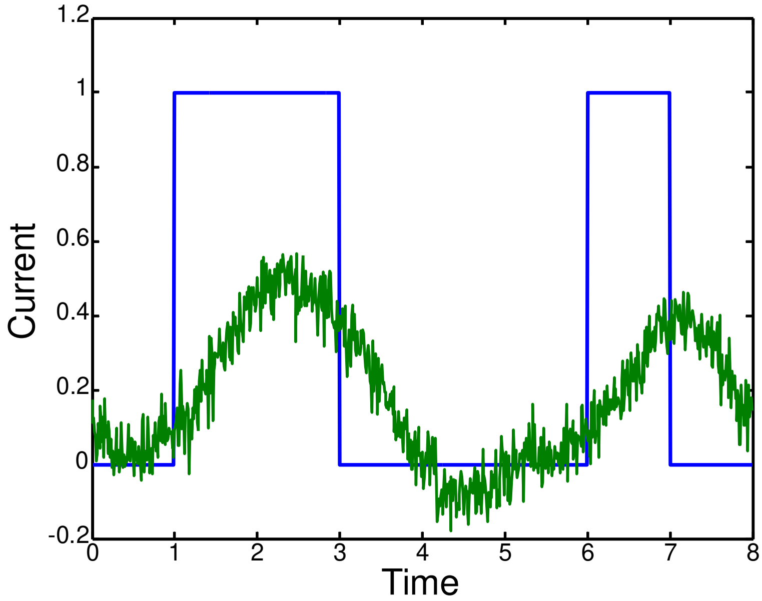
- ❑ Behavior of "real" medium already well matched!
  - ❑ What about the "wriggling"?

Fact 5: Real Media are Noisy

- ❑ A physical medium, in combination with the receiver, exhibits ***random (thermal) noise***
  - ❑ Fluctuations in the receiver circuitry, interference from nearby transmissions, etc.
- ❑ Materializes as random fluctuations around the (noise-free) received signal
  - ❑ Typical model: noise as a Gaussian random variable of zero mean, uncorrelated in time
  - ❑ More sophisticated models exist

Example w. Frequency-Dependent Attenuation, Jitter & Random Noise

- When taking all five facts into account, the received wave form can be satisfyingly explained:

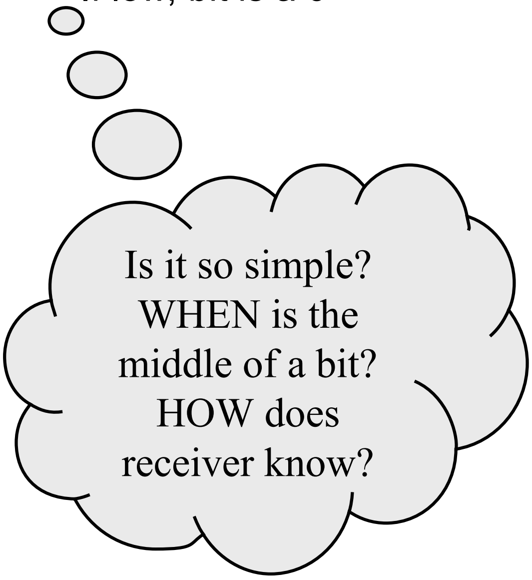


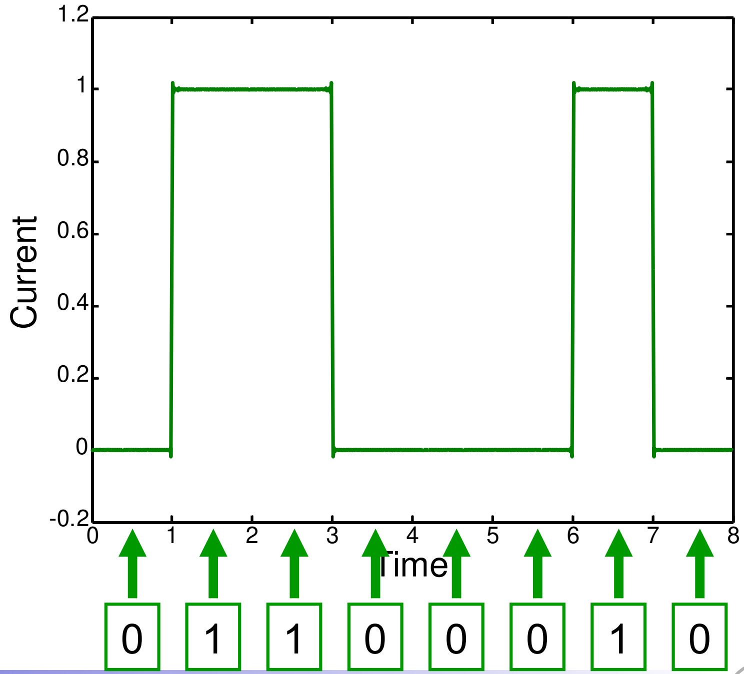
Overview

- Baseband transmission over physical channels
- ***Limitations on data rate: Nyquist and Shannon***
- Clock extraction
- Broadband versus baseband transmission
- Examples

Converting Signals to Data: Sampling

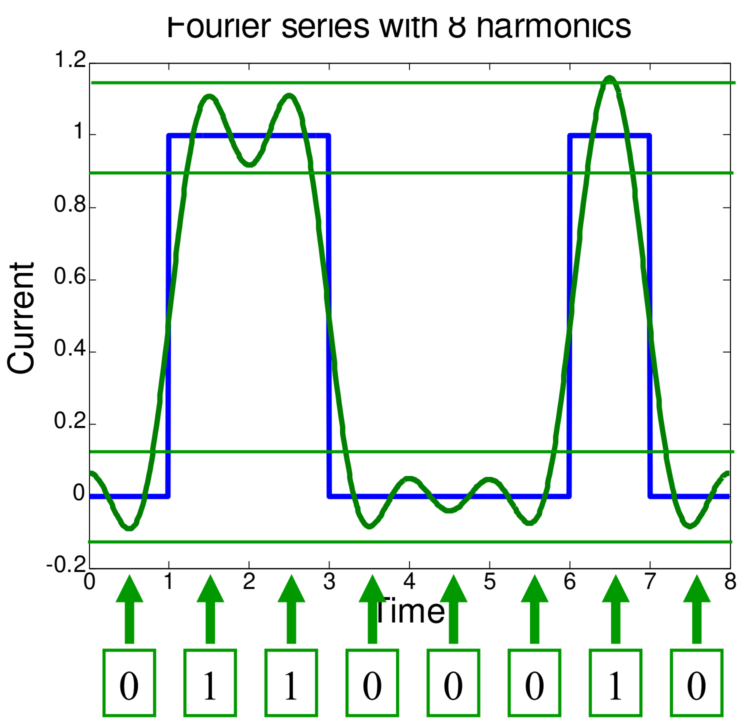
- Suppose we have a channel with "sufficient" bandwidth available
- How does a receiver convert the signal back to data?
- Simple: Look at the signal
  - If high, bit is a 1
  - If low, bit is a 0





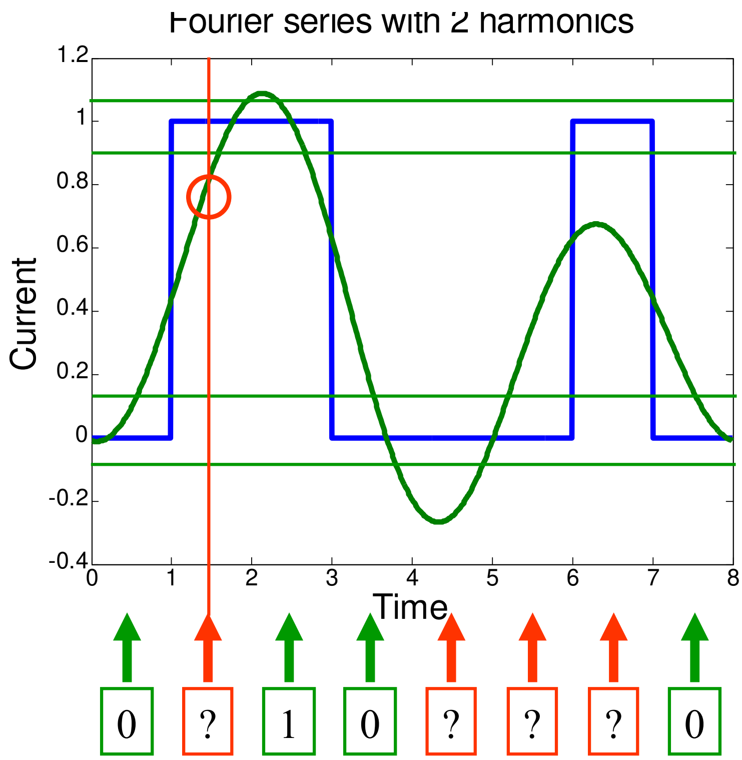
Sampling Over a Noisy or Bandwidth-Limited Channel

- In presence of noise or limited bandwidth (or both), signal will not likely be *exactly* 0 or 1
  - Or whatever 0 and 1 amounts to after attenuation
- Instead of comparing to these precise values, receiver has to use some thresholds within which a signal is declared as a 0 or a 1



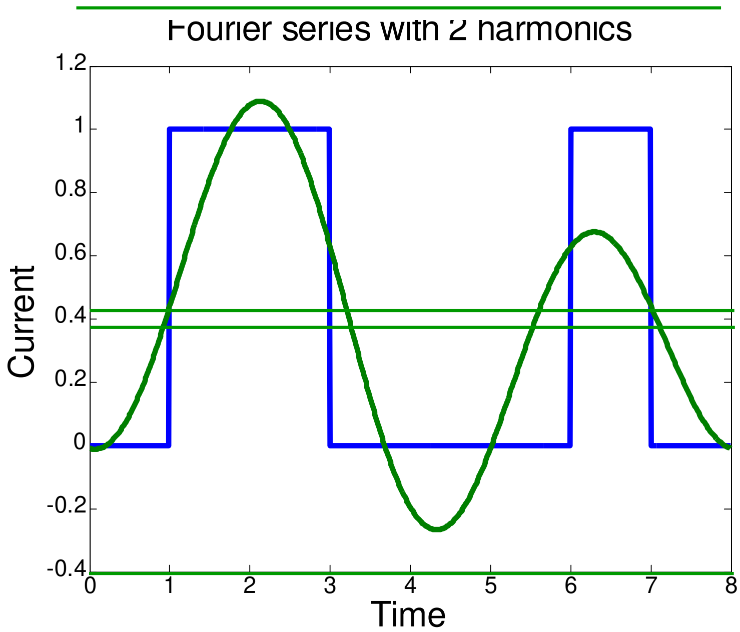
Sampling & Low Bandwidth

- What happens when little bandwidth is available?
  - Assuming same thresholds as before
- At some sampling points, the signal will be outside the thresholds!
  - No justifiable decision possible
- What are possible ways out?



Possible Way Out: Make Thresholds Wider?

- Wide thresholds would (apparently) reduce opportunity for confusion
  - E.g., +/- 0.4
- But: what happens in presence of noise?
- Wider thresholds leads to higher probability of incorrect decisions!
- ***Not good!***



Way Out 2: Increase Time for a Single Bit

- If bandwidth is limited, received signal cannot track very steep raises and falls in the signal
- Hence: give the signal more time to reach the required level for a 0 or a 1 detection.
- This means: Time for a single bit has to be extended!
  - Useable data rate is reduced!
- ***This is a fundamental limitation and cannot be circumvented***
- Formally:

  $$\text{maximum data rate} = 2H \text{ bits/s}$$

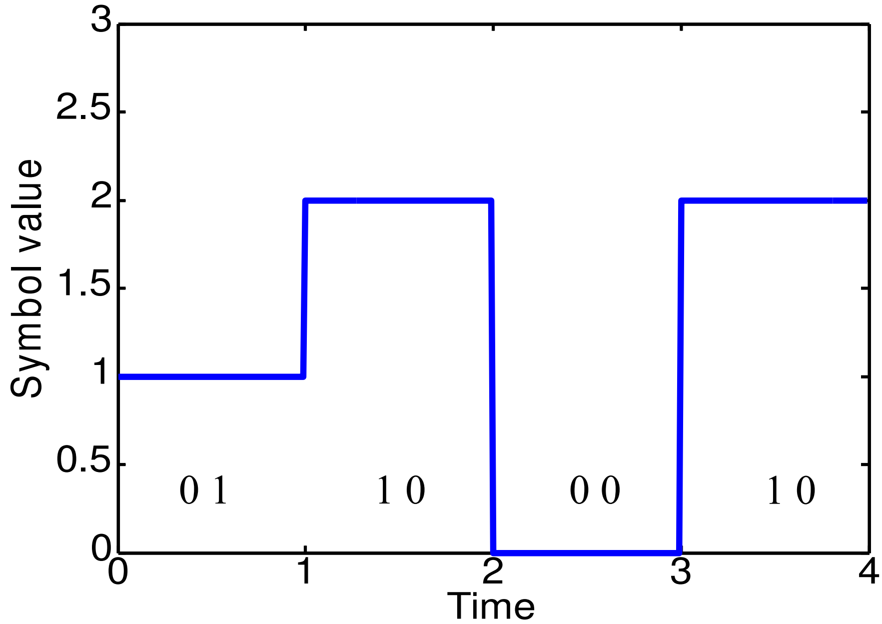
  where H is the channel bandwidth
  - Basic reason: need to sample sufficiently often

Way Out 3: Use More Than Just 0 and 1 in the Channel

- Who says we can only use 0 and 1 as possible levels for the transmitted signal?
- Suppose the transmitter can generate signals (current, voltage, …) at *four* different levels, instead of just two
- Then: to determine one of four levels, two bits are required
- Distinction:
  - "Bits" are 0 or 1, used in "higher" layers
  - "***Symbols***" can have multiple values, are transmitted over the channel
  - ***Symbol rate:*** Rate at which symbols are transmitted
    - Measured in ***baud***
  - ***Data rate:*** Rate at which physical layer processes incoming data bits
    - Measured in ***bit/s***

Way out 3: Use Four-Level Symbols to Encode Two Bits

- Example:
  - Map 00 ⇒ 0, 01 ⇒ 1, 10 ⇒ 2, 11 ⇒ 3
  - ***Symbol rate*** is then only half the ***data rate*** as each symbol encodes two bits



Data Rate with Multi-Valued Symbols – Nyquist

- Using symbols with multiple values, the data rate can be increased
- ***Nyquist formula*** summarizes:

$$\text{maximum data rate} = 2H \log_2 V \text{ bits/s}$$

where V is the number of discrete symbol values

Unlimited Data Rate with Many Symbol Levels?

- Nyquist's theorem appears to indicate that unlimited data rate can be achieved when only enough symbol levels are used
- Is this plausible?
- More and more symbol levels have to be spaced closer and closer together
- What then about noise?
  - Even small random noise would then result in one symbol being misinterpreted for another
- So not unlimited?

Shannon Limit on Achievable Data Rate

- Achievable data rate is fundamentally limited by noise
  - More precisely: by the relationship of signal strength compared to noise
  - The relatively fewer noise there is at the receiver, the easier it is for the receiver to distinguish between different symbol levels
- Relationship characterized by ***Shannon, 1948***

$$\text{maximum data rate} = H \log_2 (1 + S/N) \text{ bits/s}$$

where *S* is signal strength, *N* is noise level

- This theorem formed the basis for ***information theory***

Overview

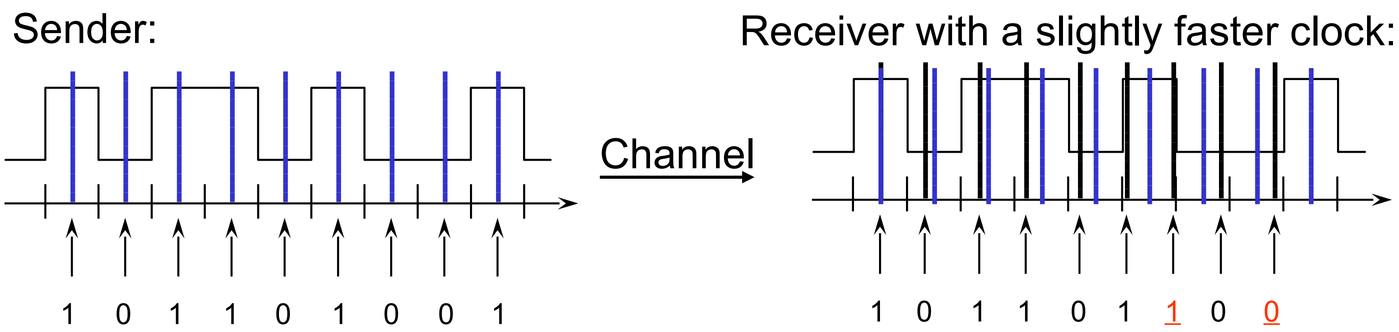
- Baseband transmission over physical channels
- Limitations on data rate: Nyquist and Shannon
- ***Clock extraction***
- Broadband versus baseband transmission
- Examples

When to Sample the Received Signal?

- How does the receiver know WHEN to check the received signal for its value?
  - One typical convention: in the middle of each symbol
  - But when does a symbol start?
    - The length of a symbol is usually known by convention via the symbol rate
- The receiver has to be ***synchronized*** with the sender at the ***bit*** level
  - The link layer will have to deal with frame synchronization
  - There is also "character" synchronization – omitted here

Overly Simplistic Bit Synchronization

- One simple option:
  - Assume that sender and receiver at some point in time are synchronized
  - That both have an internal clock that tics at every symbol step
- Usually, this does not work
  - ***Clock drift*** is major problem – two different clocks never stay in perfect synchrony
- Errors if synchronization is lost:

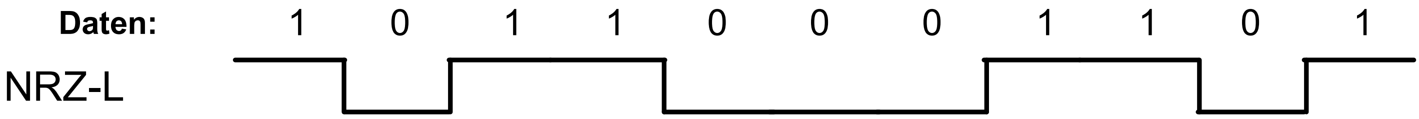


Options to Tell the Receiver When to Sample

- Relying on clock synchronization does not work
- Provide an explicit clock signal
  - Needs parallel transmission over some additional channel
  - Must be in sync with the actual data, otherwise pointless
    ⇒ Useful only for short-range communication
- Synchronize the receiver at crucial points (e.g., start of a character or of a block)
  - Otherwise, let the receiver clock run freely
  - Relies on short-term stability of clock generators (do not diverge too quickly)
- Extract clock information from the received signal itself
  - Treated next in more detail

Extract Clock Information from Signal Itself – NRZ-L

- Put enough information into the data signal itself so that the receiver can know immediately when a bit starts/stops
- Would the simple 0 ⇒ low, 1 ⇒ high mapping of bit ⇒ symbol work?
- It should – after all, receiver can use 0-1-0 transitions to detect the length of a bit

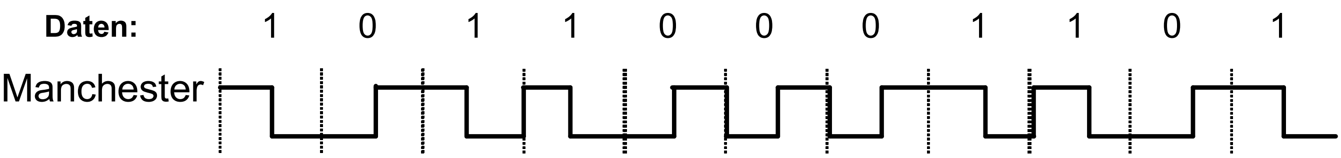


NRZ stands for "Non-Return to Zero"

- But, this scheme fails depending on bit sequences: think of long runs of 1s or 0s – receiver can loose synchronization
- Not to be able to transmit arbitrary data is not nice

Extract Clock Information From Signal Itself – Manchester

- Idea: At each bit, provide indication to receiver that this is where a bit {starts/stops/has its middle}
  - Example: Manchester encoding
  - For a 0 bit, have the symbol change in the bit middle from low to high
  - For a 1 bit, have the symbol change in the bit middle from high to low



- Ensures sufficient number of signal transitions
  - Independent of what data is transmitted!

- Drawback: needs twice the bandwidth as baudrate is twice the bitrate

Overview

- Baseband transmission over physical channels
- Limitations on data rate: Nyquist and Shannon
- Clock extraction
- ***Broadband versus baseband transmission***
- Examples

Baseband versus Broadband Transmission

- The transmission schemes described so far: ***Baseband transmission***
  - Baseband transmission directly puts the digital symbol sequences onto the wire
  - At different levels of current, voltage, … essentially, ***direct current (DC)*** is used for signaling
- Baseband transmission suffers from the problems discussed above
  - Limited bandwidth reshapes the signal at receiver
  - Attenuation and distortion depend on frequency and baseband transmissions have many different frequencies because of their wide Fourier spectrum
- Possible alternative: ***broadband transmission***

Broadband Transmission

- Idea: get rid of the wide spectrum needed for DC transmission
- Use a ***sine wave*** as a carrier for the symbols to be transmitted
  - Typically, the sine wave has high frequency
  - But only a *single* frequency!
- Pure sine waves has no information, so its shape has to be influenced according to the symbols to be transmitted
  - The carrier has to be ***modulated*** by the symbols (widening the spectrum)
- Three parameters that can be influenced
  - Amplitude *a*
  - Frequency *f*
  - Phase ϕ

$$a \sin(2\pi f t + \phi)$$

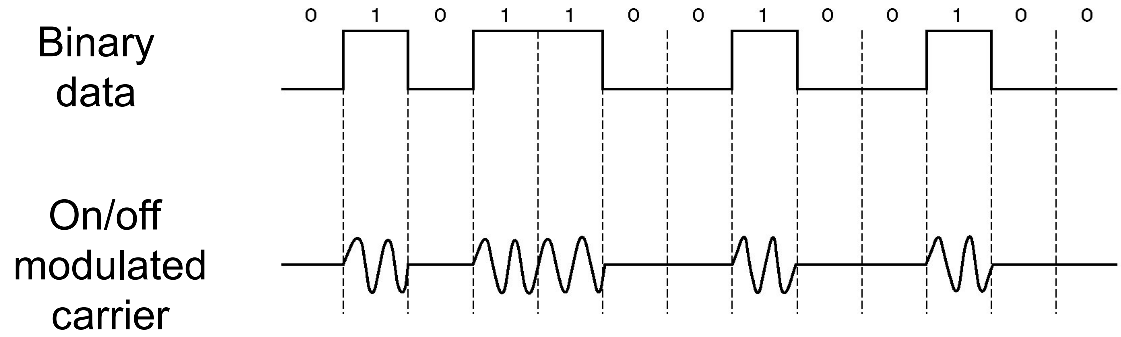
Amplitude Modulation

- Given a sine wave f(t) and a time-varying signal s(t)
  - $f(t) = a \sin(2\pi f t + \phi)$
  - Signal can be analog (i.e., a continuous function of time) or digital (i.e., a discrete function of time)
  - Signal can be e.g. the symbol levels discussed above
- The amplitude modulated sine wave $f_A(t)$ is given as:

$$f_A(t) = s(t) \sin(2\pi f t + \phi)$$

  - I.e., the amplitude is given by the signal to be transmitted
- Receiver can extract s(t) from $f_A(t)$
- Special cases:
  - s(t) is an ***analog*** signal – ***amplitude modulation***
  - s(t) is a ***digital*** signal – also called ***amplitude keying***
  - s(t) only takes 0 and 1 (or 0 and *a*) as values – ***on/off keying***

Amplitude Modulation – Example



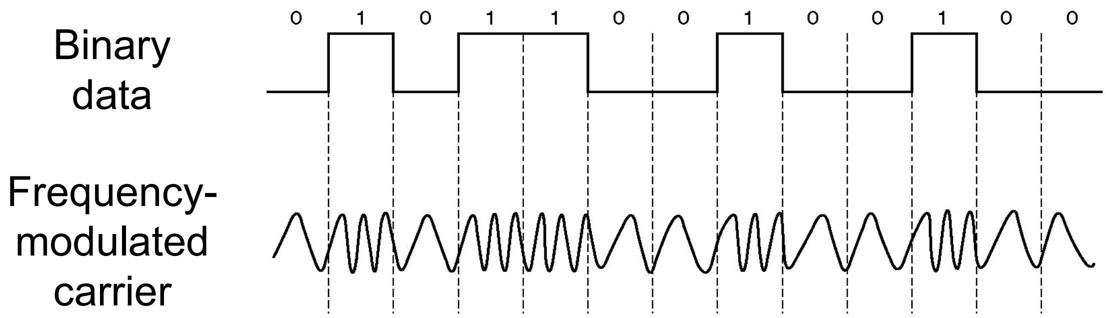
- Question:
  - How to solve bit synchronization here?
  - Is Manchester applicable?

Frequency Modulation

- The frequency-modulated sine wave $f_F(t)$ is given by

$$f_F(t) = a \sin(2\pi s(t)t + \phi)$$

- Modulation/keying terminology like for AM
- Example



Note: s(t) has an additive constant in this example to avoid having frequency zero

Phase Modulation

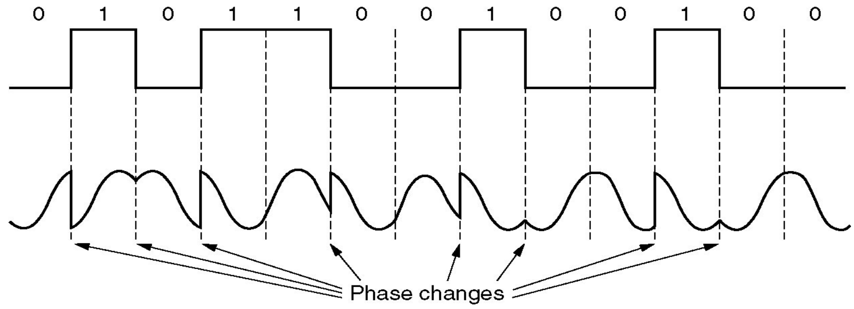
- Similarly, a phase modulated carrier is given by

$$f_P(t) = a \sin(2\pi f t + s(t))$$

  - Modulation/keying terminology again similar
- Example:

Binary data

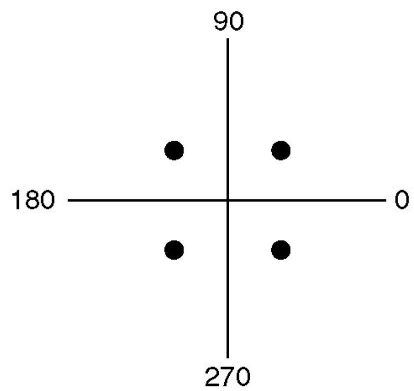
Phase-modulated carrier



- Here, s(t) is chosen such that there are phase changes when the binary data changes
  - Typical example for ***differential coding***
- Other possibilities: 0 ⇒ no phase shift, 1 ⇒ phase shift, or vice versa

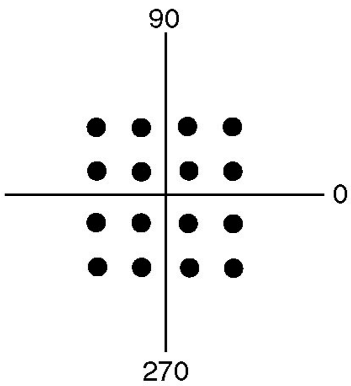
Phase Modulation With High Multiple Values per Symbol

- A receiver can usually quite well distinguish phase shifts
- Hence: Use phase shifts of 0, $\pi/2$, π, $3/2\ \pi$ to encode two bits per symbol
  - Even better: Use $\pi/4$, $3/4\pi$, $5/4\pi$, $7/4\pi$ phase shifts for each bit
  - Why better? Clock extraction!
  - Result: Data rate is twice the symbol rate
- Technique is called Quadrature Phase Shift Keying (QPSK)
- Visualization as constellation diagram:



Combinations of Different Modulations

- Amplitude, frequency, and phase modulations can be fruitfully combined
- Example: 16-QAM (Quadrature Amplitude Modulation)
  - Use 16 different combinations of phase change and amplitude for each symbol
  - Per symbol, $2^4 = 16$ bits are encoded and transmitted in one step
  - Constellation diagram:



Interlude: Digital vs. Analog Signals

- A sender has two principal options what types of signals to generate
  - It can choose from a finite set of different signals – ***digital transmission***
  - There is an infinite set of possible signals – ***analog transmission***
- Simplest example: Signal corresponds to current/voltage level on the wire
  - In the digital case, there are finitely many voltage levels to choose from
  - In the analog case, any voltage is legal
- More complicated example: finite/infinitely many sinus functions
  - In both cases, the resulting ***wave forms in the medium*** can well be continuous functions of time!
- Advantage of digital signals: There is a principal chance that the receiver can precisely reconstruct the transmitted signal

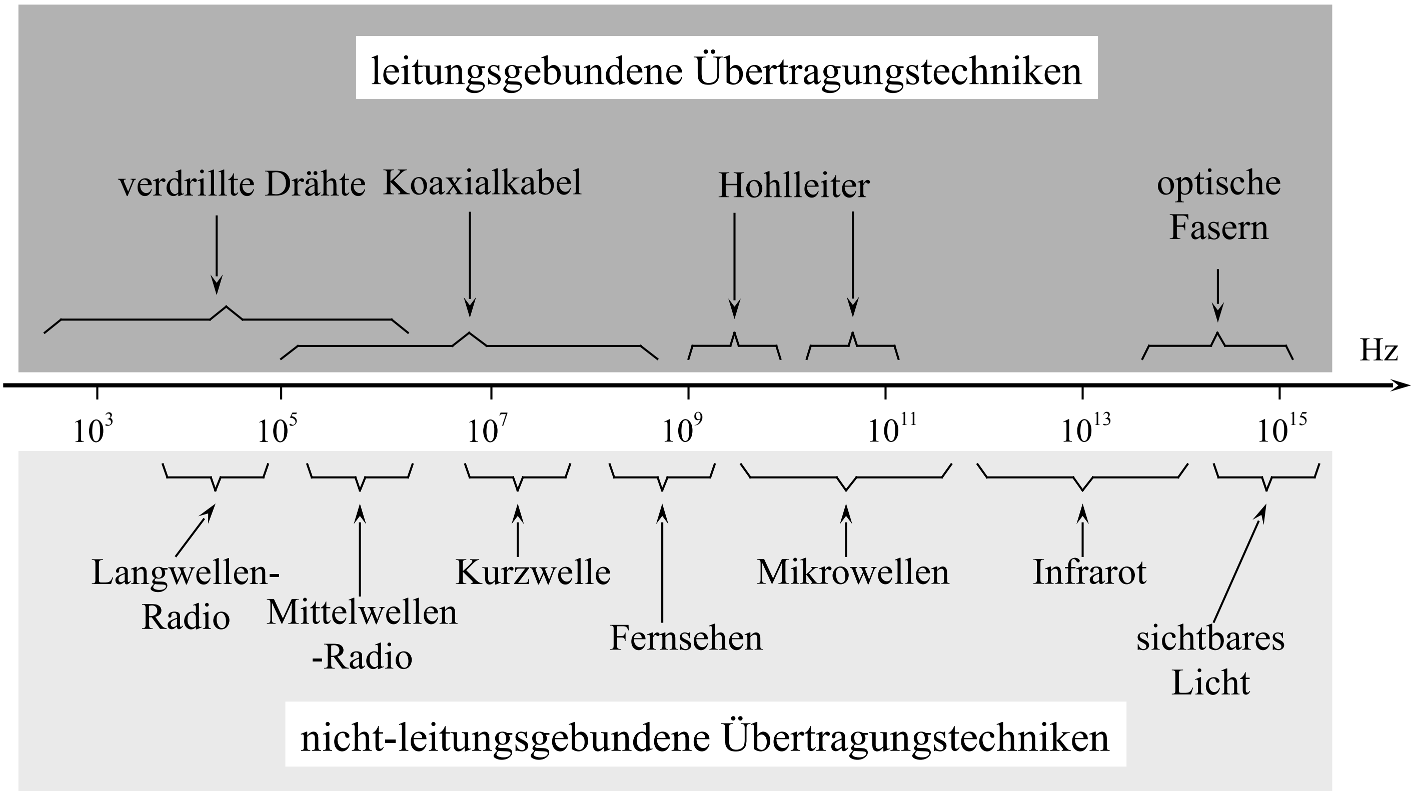
Overview

- Baseband transmission over physical channels
- Limitations on data rate: Nyquist and Shannon
- Clock extraction
- Broadband versus baseband transmission
- ***Examples***

Example Physical Layers

- Guided transmission media
  - Copper wire – twisted pair
  - Copper wire – coaxial cable
  - Fiber optics
- Wireless transmission
  - Radio transmission
  - Microwave transmission
  - Infrared
  - Lightwave

Electromagnetic Spectrum and Media



Conclusion

- The physical layer is responsible for turning a logical sequence of bits into a physical signal that can propagate through space
- Many different forms of physical signals are possible
- Signals are limited by their propagation in a physical medium (limited bandwidth, attenuation, dispersion) and by noise
- Bits can be combined into multi-valued symbols for transmission
  - Gives rise to the difference in data rate and baud rate
- Baseband transmission is fraught with problems, partially overcome by modulating a signal onto a carrier (broadband transmission)