Computational Linear Algebra in Wireless Communications

Anton Sukhinov

WAVES

Wave Equation (1D Case)

$$\partial^2 u/\partial t^2 = c^2 \cdot \partial^2 u/\partial x^2 + f.$$

- That is linear equation!
- But in what does it mean for such an equation to be linear?

Linearity of Wave Equation

Linearity means that linear superposition of two solutions is also a solution:

$$\partial^2 u/\partial t^2 = c^2 \cdot \partial^2 u/\partial x^2 + f$$

$$+$$

$$\partial^2 v/\partial t^2 = c^2 \cdot \partial^2 v/\partial x^2 + g$$

$$=$$

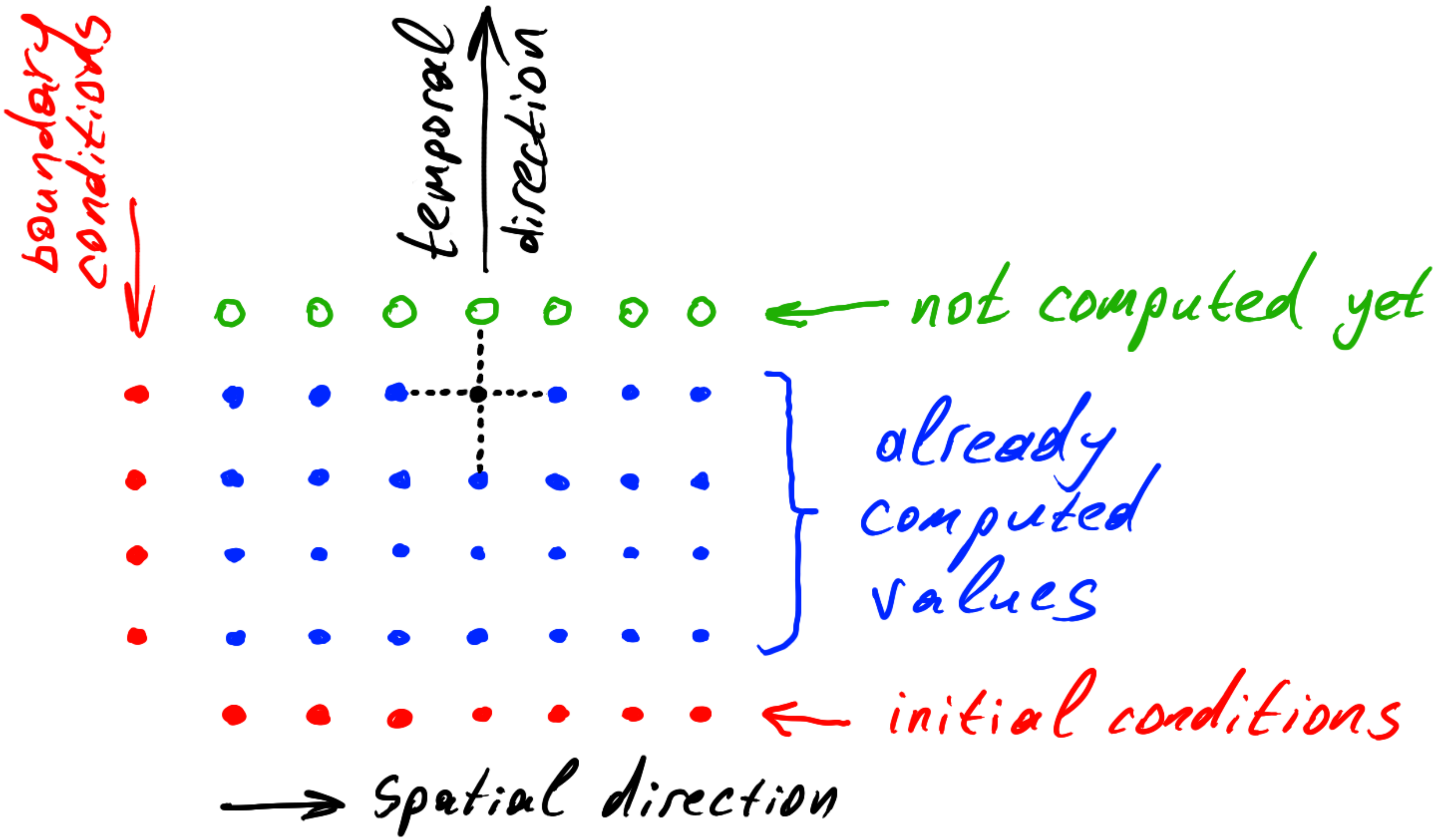
$$\partial^2 (u+v)/\partial t^2 = c^2 \cdot \partial^2 (u+v)/\partial x^2 + (f+g).$$

That in turn means that two waves can go through each other without interfering each other.

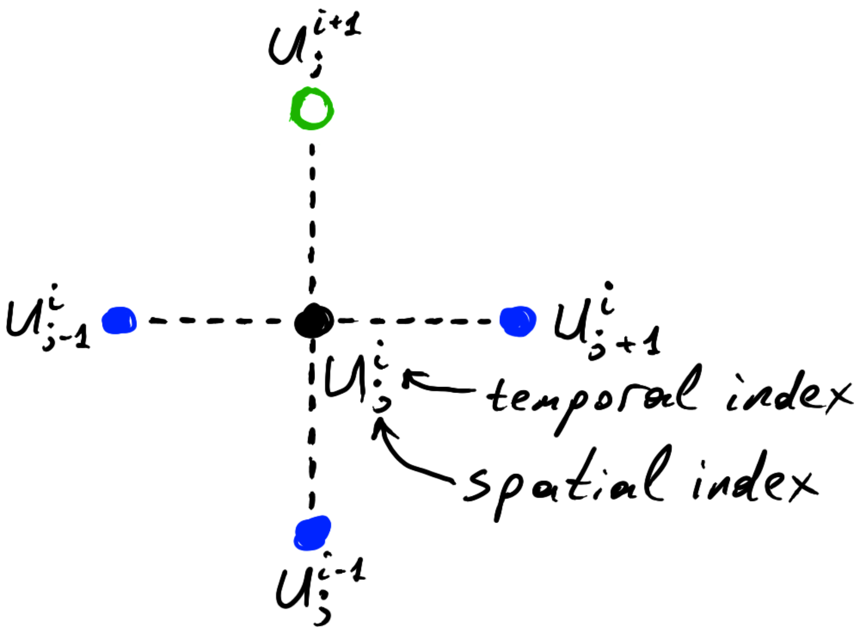
“Offset” and “Speed”

- Since wave equation has second derivative with respect to time, it sensible not only to the medium “offset” u, but also to the “speed” $\partial u/\partial t$.
- You can think of oscillation as of two values chasing each other around zero value.
- Such oscillations diffuse in space making waves.
- So, in electromagnetic wave electric and magnetic fields take place of “offset” and “speed”? No. That is a common misconception.

Simple Way of Numerical Simulation



Simple Way of Numerical Simulation



Difference scheme:

$$\frac{u_j^{i+1} - 2u_j^i + u_j^{i-1}}{\tau^2} = c^2\,\frac{u_{j+1}^i - 2u_j^i + u_{j-1}^i}{h^2} + f_j^i .$$

Stability condition:

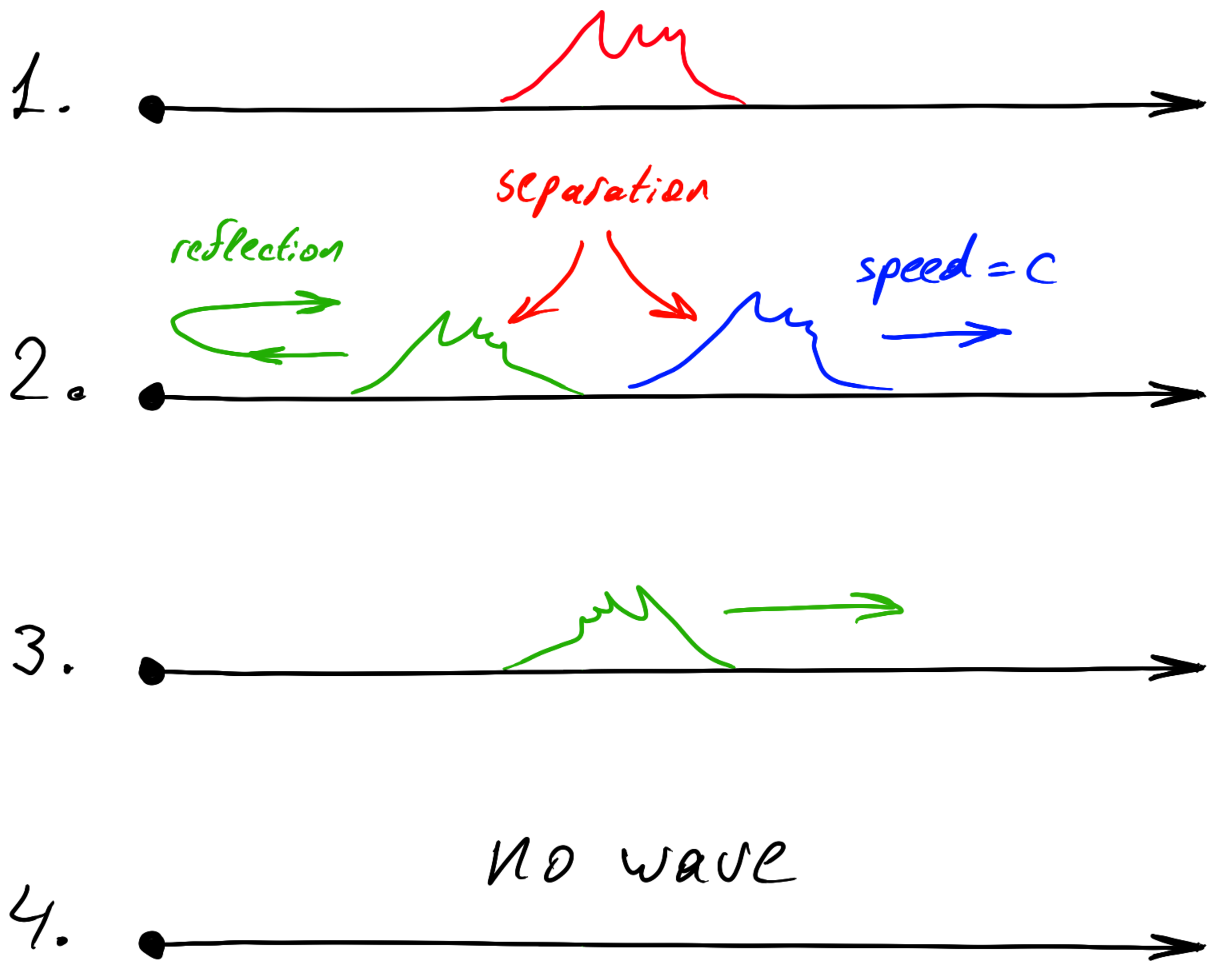
$$c \cdot \tau / h \le 1.$$

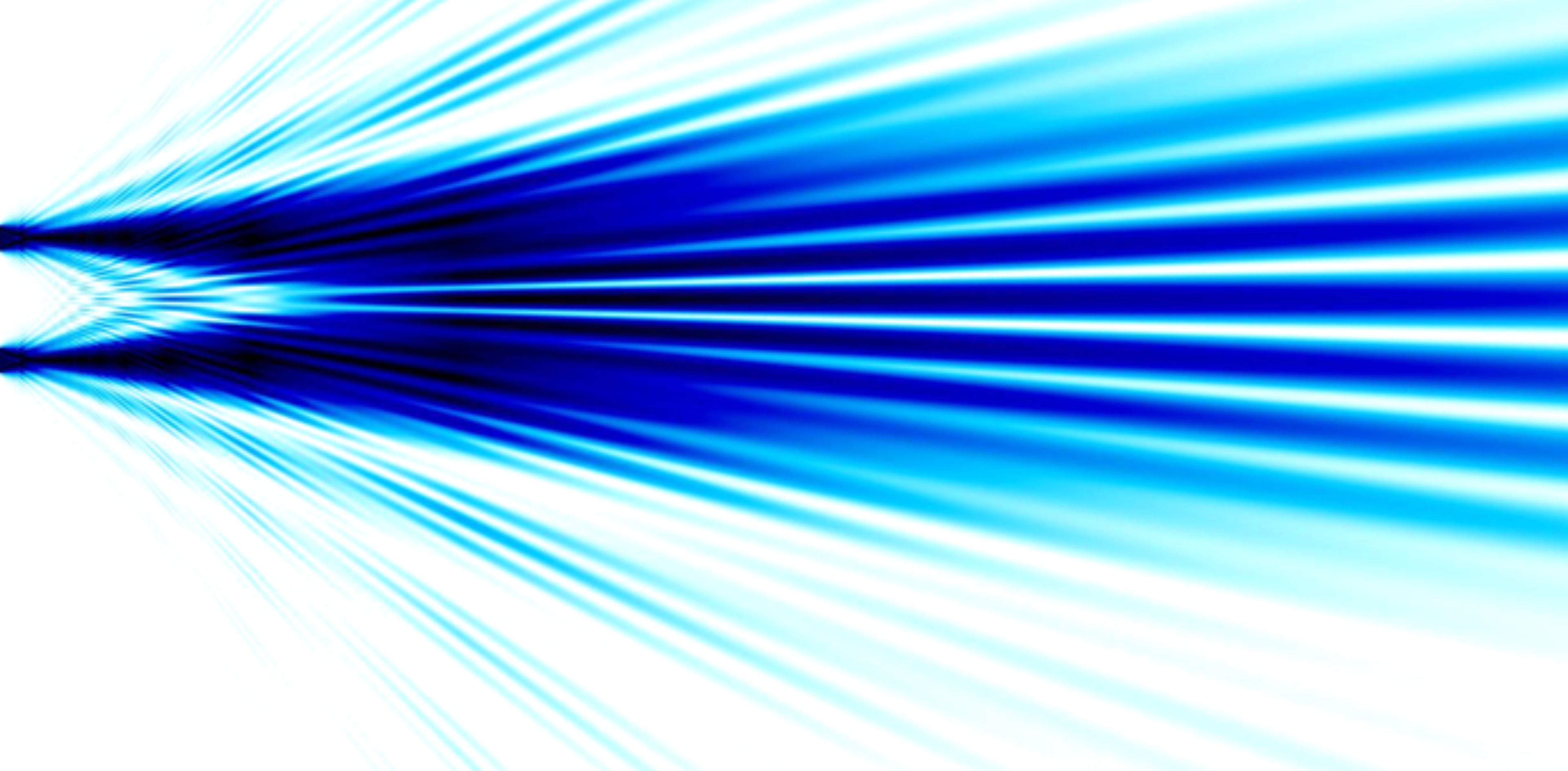
During time step τ wave should not move more than spatial step h.

Semi-Infinite String

- What will be the evolution of the situation shown at the picture?
- Since wave is sensible to both “offset” and “speed” it is not enough to provide “offset” as initial condition, “speed” is also needed.
- Let us pretend the initial “speed” is zero.

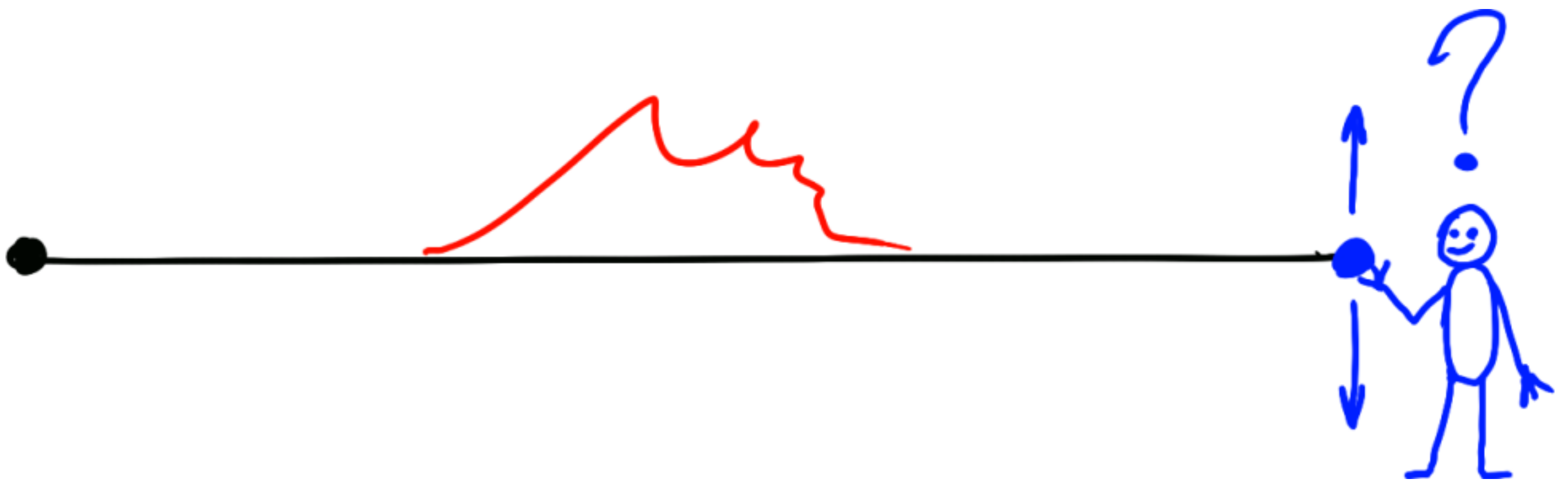
Semi-Infinite String





LINEAR TRANSMISSION EFFECTS

Creative Task

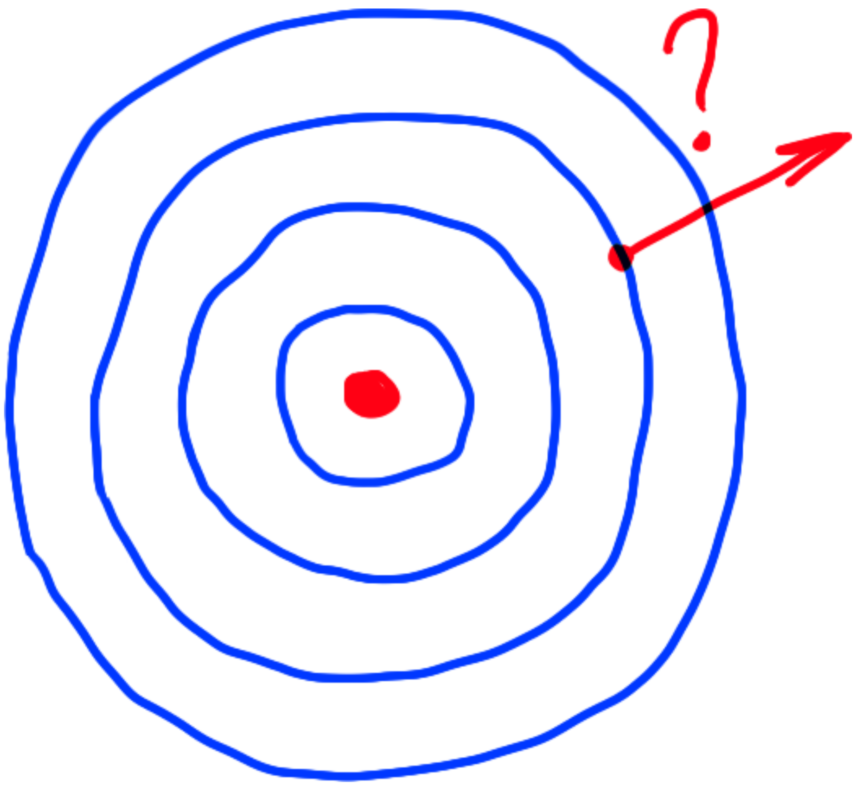


- Suppose that you can move one end of the string, and your task is to stop oscillations as fast as possible. What should you do?
- Homework: write code and prove the solution.

Active Noise Cancellation

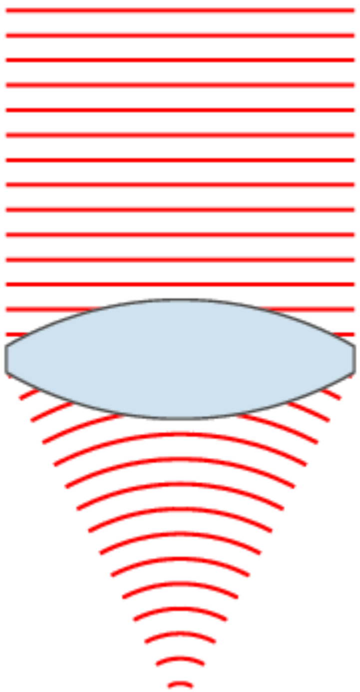
- The best wave absorber is open window: waves will go through the window and will never return.
- How about these fancy headphones with active noise cancellation? We have seen at previous slide that they cannot work…
- In fact, active noise cancellation can work:
  - Noise is cancelled in a very limited volume.
  - Wave length should me much larger than the size of that volume.

Wave Propagation Direction



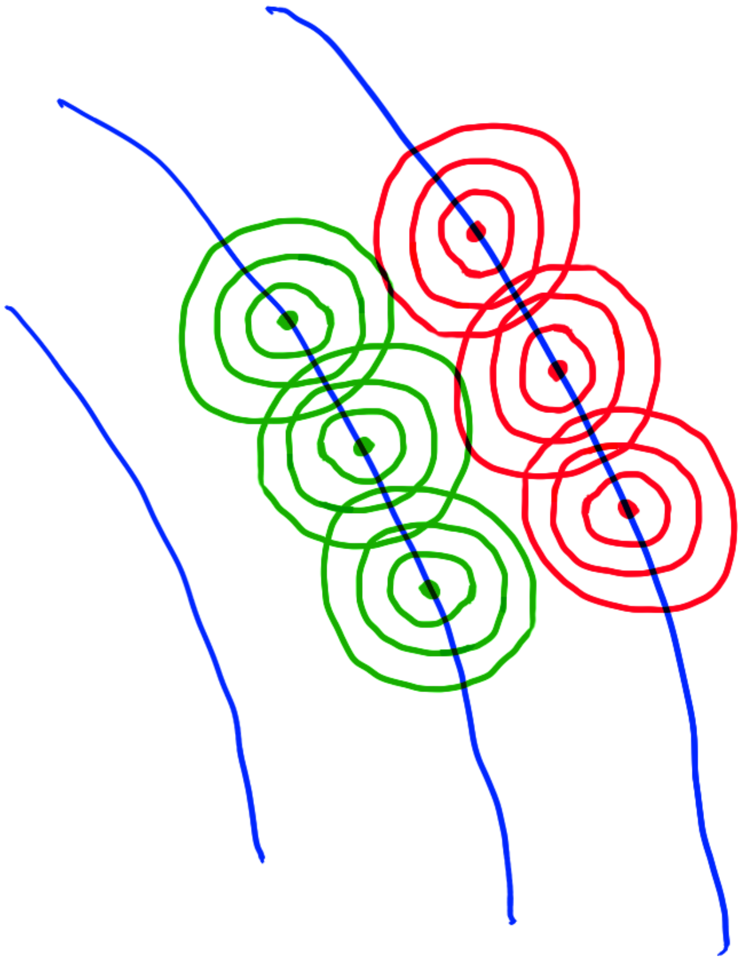
- How does the wave know where should it go?
- Does it remember the position of its source?

Wavefronts



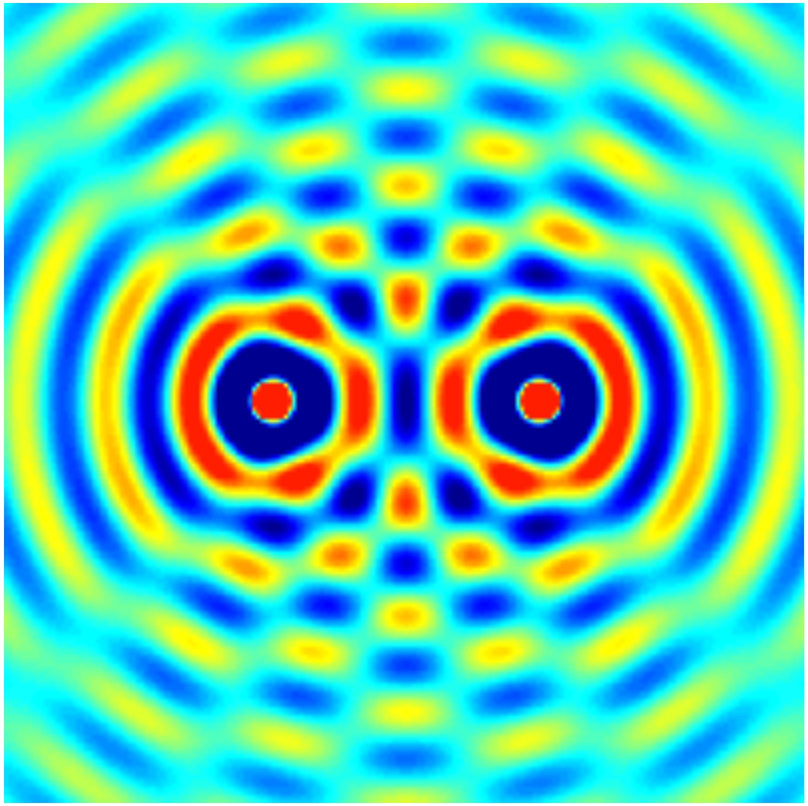
- A wavefront is the locus of points having the same phase.
- Wavefronts are defined only in simple wave propagation scenarios.
- They can be used to see how the wave is moving.
- When we follow waves on water by our eyes, we follow wavefronts, not water "particles"

Wave Itself is Its Source



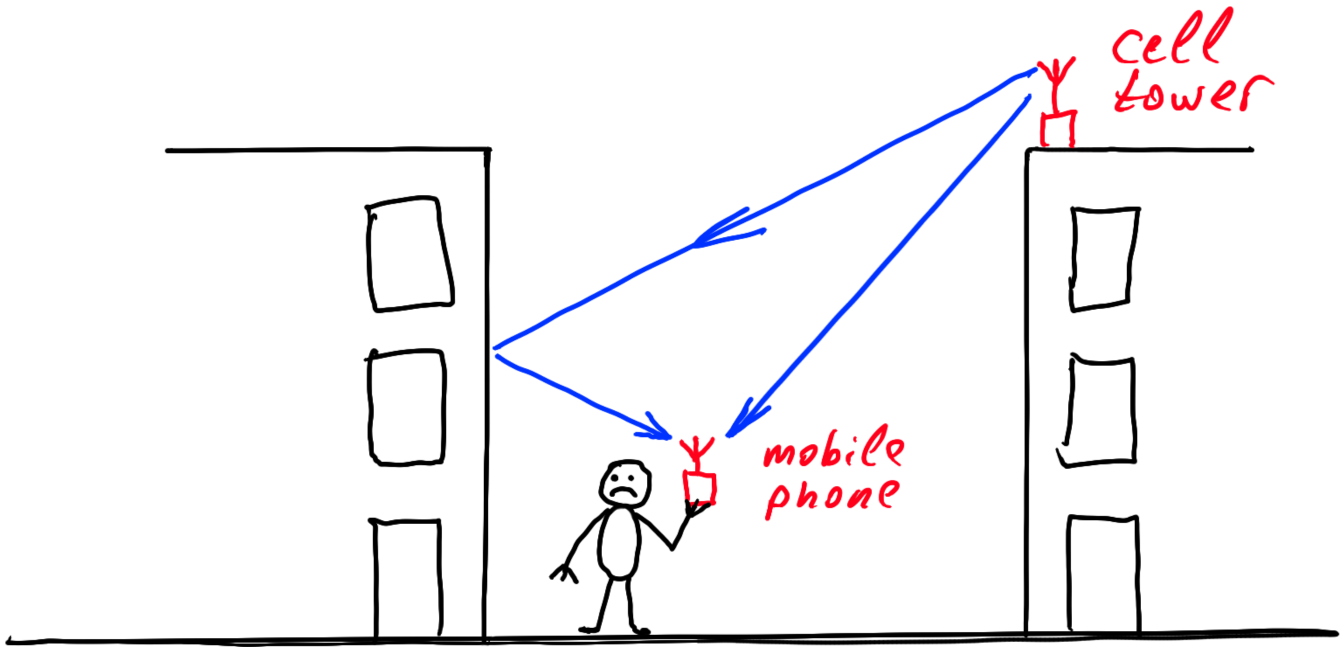
- Wave points constantly retransmit waves in all directions.
- Because of phase shift between neighboring wavefronts such retra-nsmitted waves form interference pattern which looks like moving or expanding wave

Interference



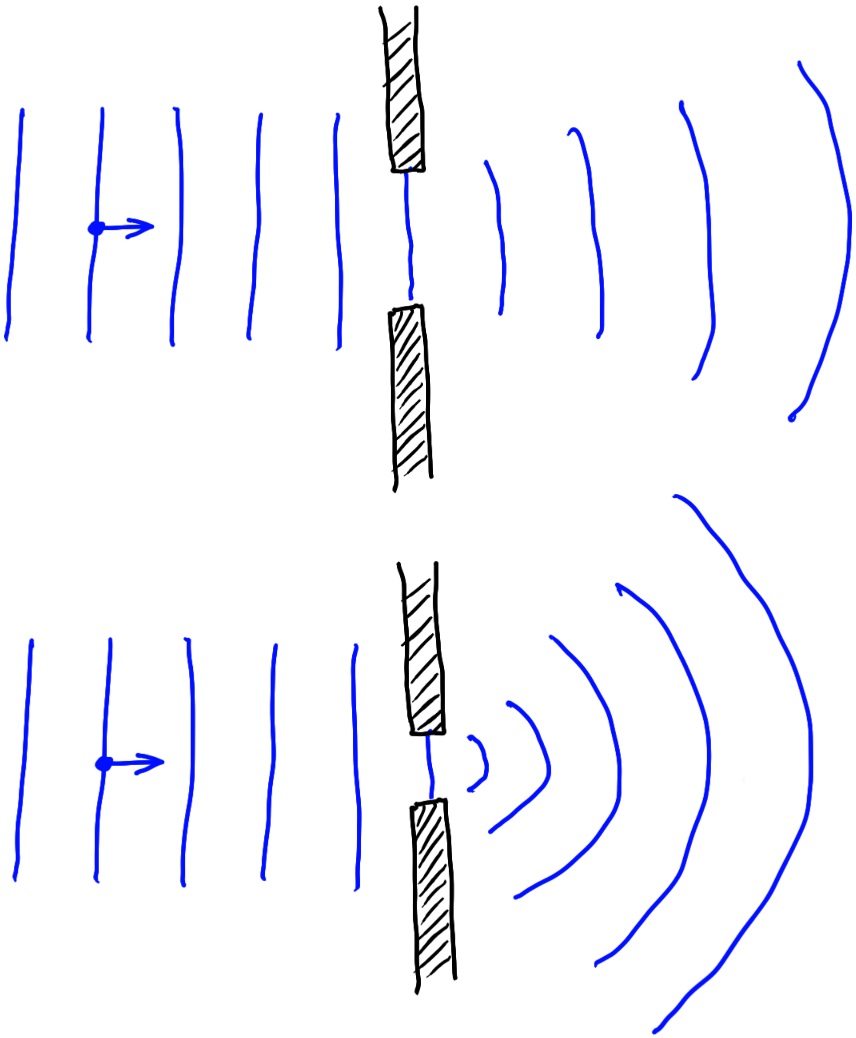
- Interference pattern is formed when coherent waves of the same frequency overlap each other.
- Waves are coherent when they have same frequen-cies and constant phase difference.
- Non-coherent wave do not form stable interfe-rence patterns.

Interference – Good or Bad?



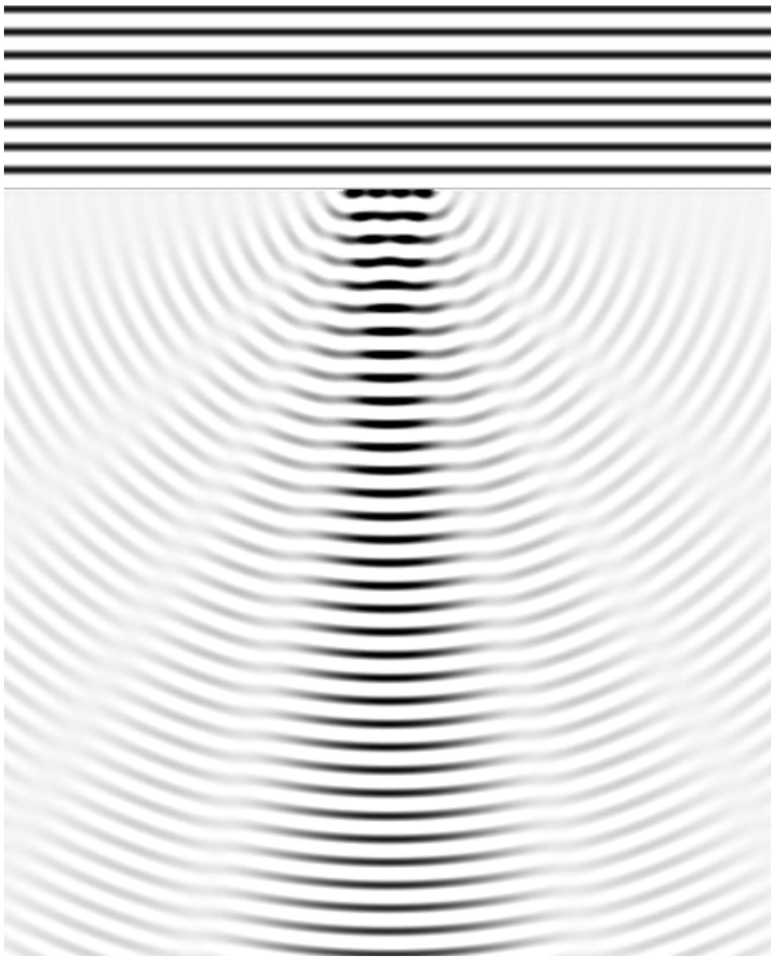
- Due to interference you can get zero signal strength even in direct visibility of cell tower (destructive interference).
- Typical size of interference patterns is similar to wavelength. Wavelength for 2GHz signal is 15 centimeters.
- Transmitters can use multiple antennas to increase bandwidth by means of constructive interference.

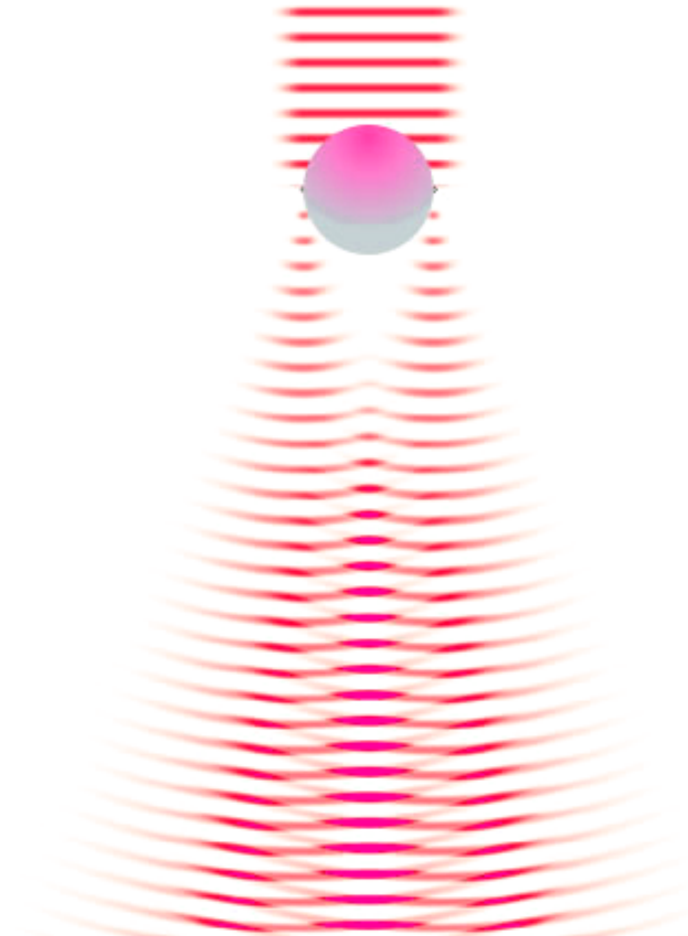
Diffraction



- Diffraction occurs when a wave encounters an obstacle.
- Diffraction can be expla-ined by interference of the wave with itself.

Diffraction and Interference Always Appear Together





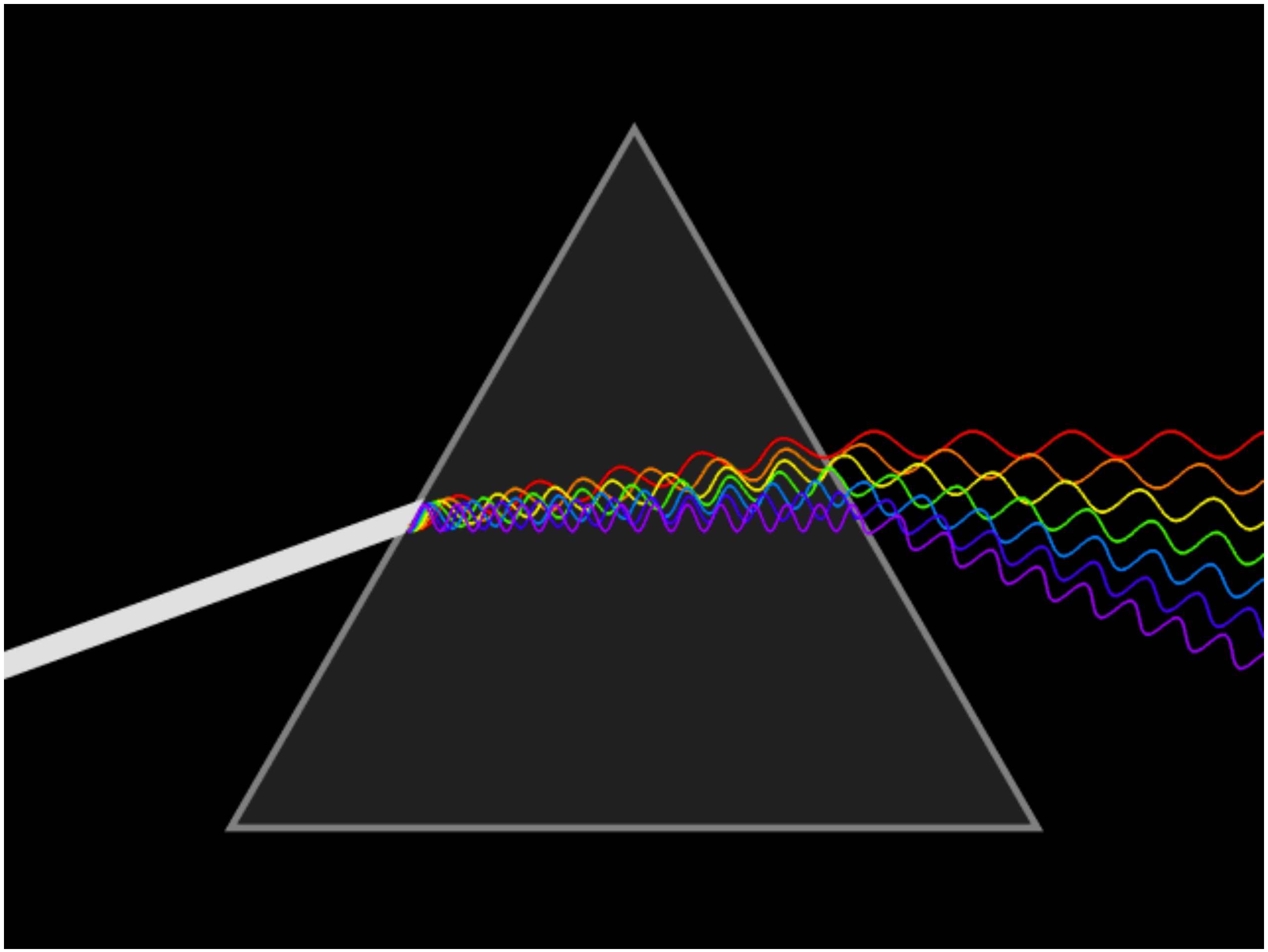
Diffraction – Good or Bad?

- In radio engineering diffraction is mostly good, since it allows having signal reception even without direct visibility of the transmitter.
- Nevertheless photographers hate diffraction, since it reduces sharpness of their images.
- In microscopy diffraction limits the size of objects that can be viewed in optical microscope.

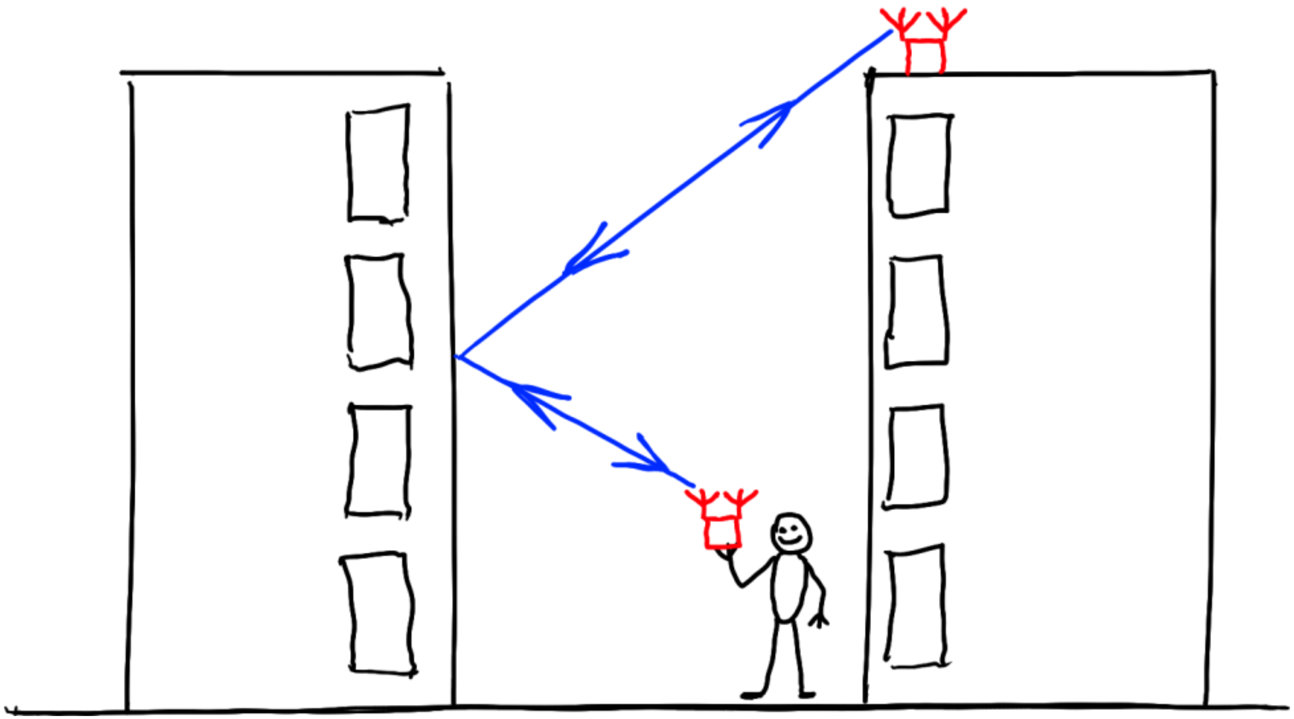
Reciprocity

- Reciprocity of waves is their ability to travel the same path in both directions.
- If you "reverse in time" the wave propagation process, then you still have correct solution to the wave equation. That is because it has second derivative on time.
- Reciprocity is important feature of waves transferred in linear media.
- Lossy medium is not reciprocal.

Reciprocity Illustration



Reciprocity – Good or Bad?

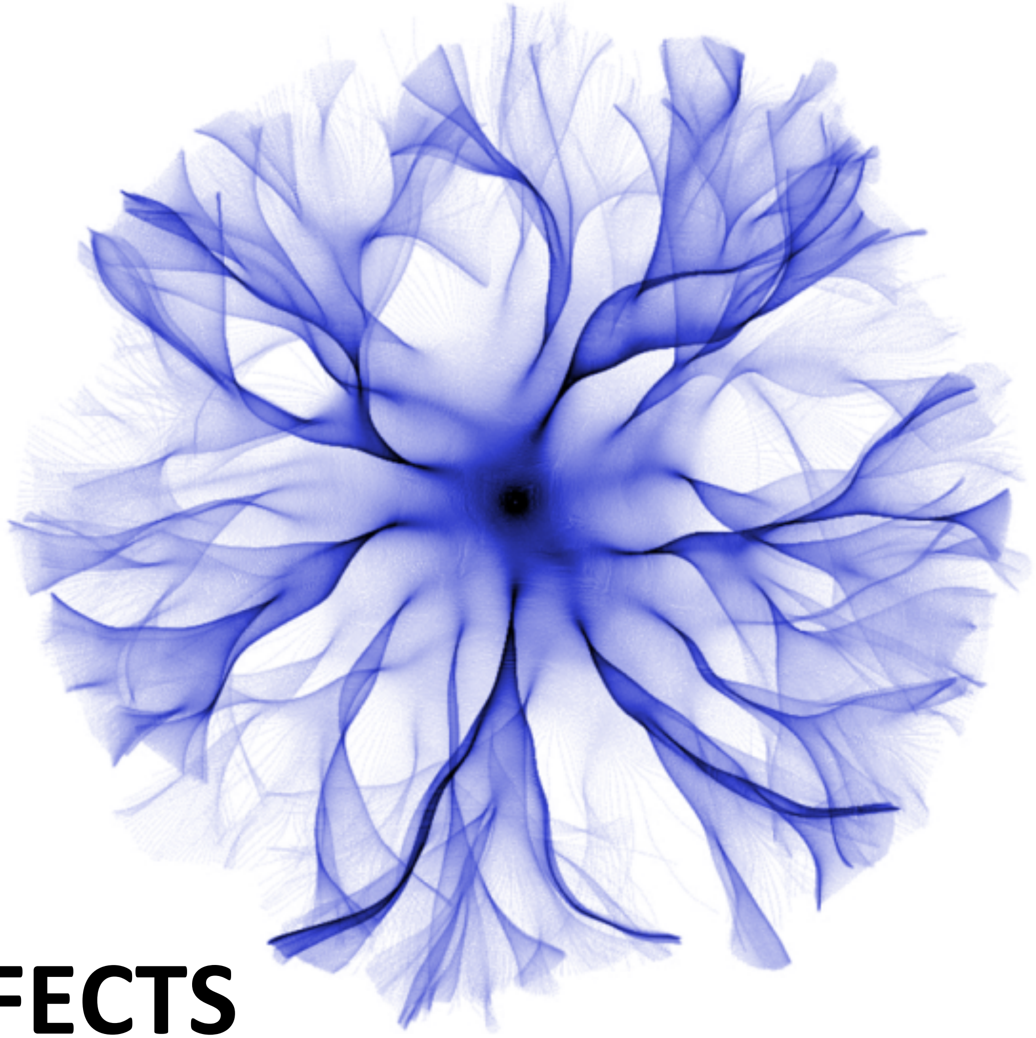


Reciprocity allows efficient double-sided communication by transmitting the signal in reverse direction to received signal (multiple antennas needed).

Spectrum as a Way to Understand Waves

Although you can transmit signal of arbitrary shape (for example, rectangular waves instead of sinusoid), the best way to understand waves is to think of any signal as of sum of sinusoidal waves:

- Different frequencies have different diffraction and interference patterns. So, harmonics of your transmitted signal may be received with different attenuation and phase offsets.
- Dispersion: different frequencies may travel with different speeds.
- Shapes of transmit and receive antennas and corresponding electronic circuits are optimized for specific ranges of frequencies.
- Usage of radio waves is licensed by frequency ranges.



NON-LINEAR EFFECTS

Non-Linear Effects

- Lossy media.
- Intermodulation.
- Nonlinear signal amplification.

Rusty Bolt Effect

- Oxide layer between different metals may act as a diode.
- This leads to frequencies.
- This is an example of passive intermodulation.

Nonlinear System – Some Math

Harmonic wave of frequency f:

$u(t)=a\sin(2\pi\cdot f\cdot t)$.

System with quadratic nonlinearity:

$g(u)=k\cdot u+\varepsilon\cdot u^2$.

Output signal:

$g(u(t))=k\cdot a\cdot\sin(2\pi\cdot f\cdot t)+$

$+\varepsilon a^2/2\ (1-\cos(2\pi\cdot 2f\cdot t))$.

So, double frequency ($2f$) appeared at the output. The relative power of intermodulated harmonic is higher when signal amplitude a is high. It can also be shown that combination of two waves will lead to appearance of sum and difference frequency (do it yourself).

Non-Linearity – Good or Bad?

- Nonlinearities in electromagnetic wave transmission are bad for us, since it is very hard to design compensation algorithms.
- The same can be told about sound processing: linear defects can be alleviated by equalizer, but non-linear effects are source of nasty sound distortions at high sound volume.
- There are 3 main ways of dealing with nonlinearities:
  - designing quality (highly linear) hardware and keeping it in good condition (tuning, cleaning, etc);
  - reducing signal power;
  - just ignoring non-linear effects.

INFORMATION TRANSMISSION

Channel Capacity

Shannon–Hartley theorem:

$C = B \log_2 \left(1 + S/N\right)$.

Here:

- C — theoretical channel capacity (bits/sec);
- B — bandwidth (hertz);
- S — signal power (watts);
- N — noise power (watts).

In case of wideband signal:

$C = \int_{f_{min}}^{f_{max}} \log_2 \left(1 + S(f)/N(f)\right) df$.

Channel Capacity

- Increasing of signal power S is not very productive way of capacity increasing.
- It is better to increase the bandwidth B.
- From hardware point of view it is easier to increase bandwidth for high-frequency signals.
- High-frequency allows hardware miniaturization.

Creative question

- Let us transmit pure sine wave, and turn it on and off according to Morse code.
- When the wave is on, it has zero bandwidth ($B=0$). When the wave is off, it has zero energy ($S=0$). So in both cases we do not transfer information according to Shannon–Hartley theorem.
- Nevertheless we can receive this wave and decode Morse code. So, information is transferred.
- What is wrong with our reasoning?

Modulation

Harmonic wave with amplitude a, frequency f, and phase p:

$u(t,a,f,p)=a\sin(2\pi\cdot f\cdot t+p)$.

In order to transfer information we should change some wave properties:

- a — amplitude modulation (AM);
- f — frequency modulation (FM);
- p — phase modulation (PM).

The modulated wave is called carrier, because it carries our information.

Amplitude modulation

- AM can be digital or analog.
- Our example with Morse code is in fact digital amplitude modulation.
- In digital modulation we use few fixed levels of modulated parameter.
- In analog modulation the parameter is continuous.

AM Bandwidth

Suppose we want to modify amplitude by some sinusoidal function:

$a(t) = b \sin(2\pi \cdot g \cdot t)$,

then (when phase is fixed it can be considered zero):

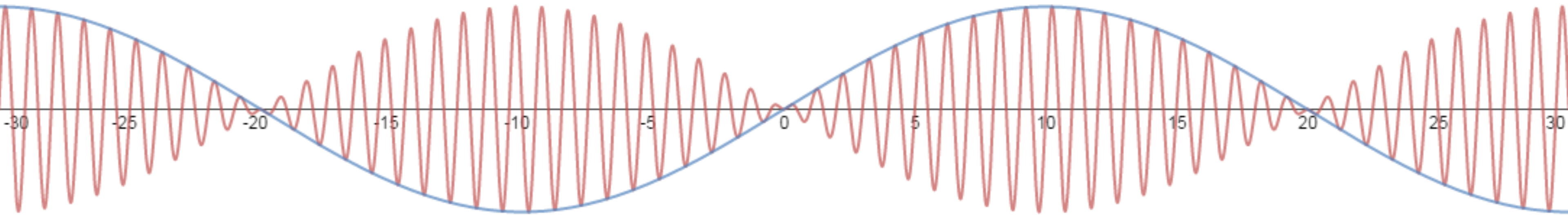
$u(t, a(t), f) =$

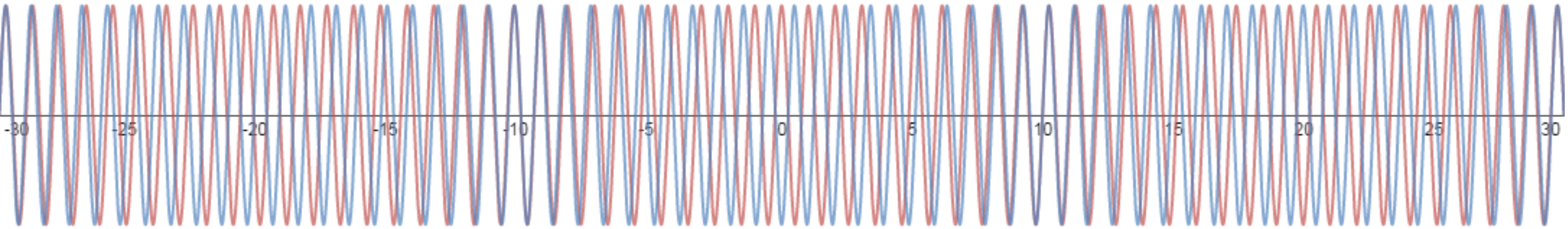
$= b \sin(2\pi \cdot g \cdot t) \sin(2\pi \cdot f \cdot t) =$

$= b/2\ (\cos(2\pi \cdot (f - g) \cdot t) - \cos(2\pi \cdot (f + g) \cdot t)\ )$.

So, changing amplitude of frequency f with frequency g will require bandwidth $[f - g, f + g]$.

AM as a Sum of Two Waves

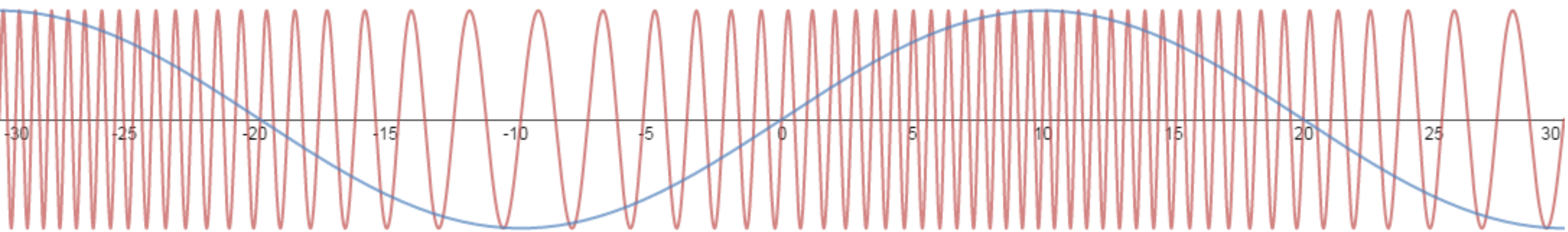




Required bandwidth

- We have seen that change of signal amplitude requires wider bandwidth than zero-width band of pure unmodulated wave.
- It can be shown that change of any signal parameter (amplitude, frequency, phase) will widen the band according to the frequency of parameter changing.

Frequency Modulation

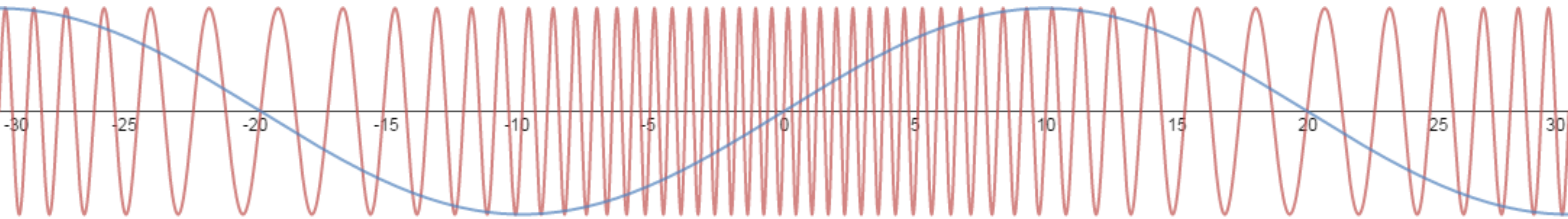


- We change carrier frequency (red) according to information to be transferred (blue).
- Analog FM radio has higher quality than AM radio simply because FM has much higher bandwidth.

Digital Frequency Modulation?

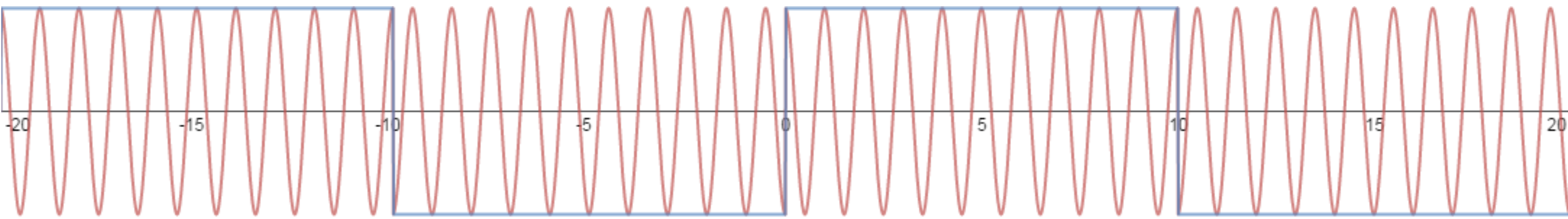
- Let us suppose that we switching carrier between 2 different frequencies in order to transmit our binary data.
- We can think of this process as of two carriers alternately switching their amplitudes between zero and non-zero.
- We can transfer more information if will switch them independently (not alternately).
- So, digital frequency modulation is a special case of amplitude modulation with multiple carriers.

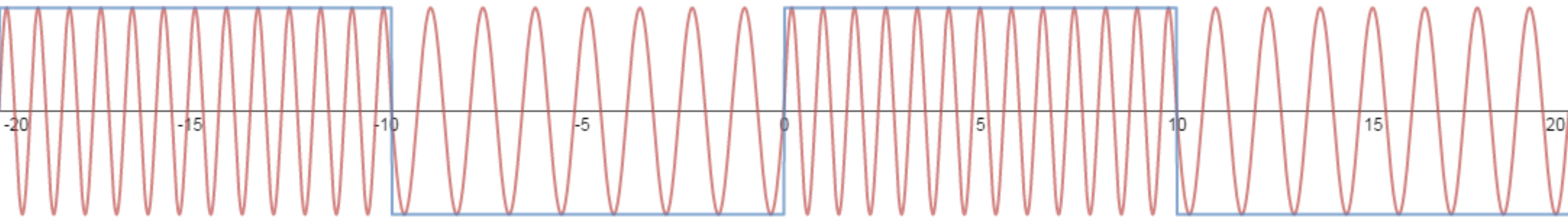
Phase Modulation



In case of modulation by sinusoidal signal phase modulation is similar to frequency modulation, except frequency of phase-modulated signal is higher where derivative of information signal is high, not its value.

Phase Modulation vs Frequency Modulation for Binary Signal



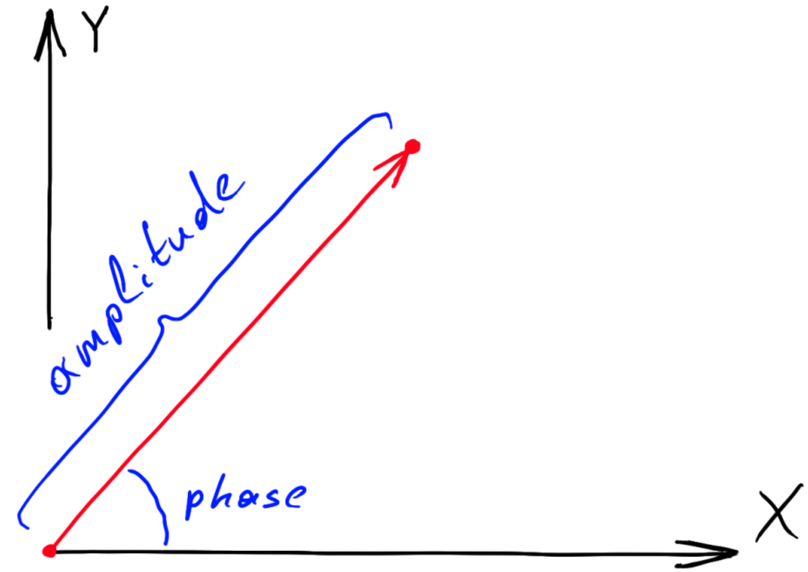


In order to decode phase-modulated signal receiver and transmitter should be phase-synchronized. So, such transmission is called synchronous.

Combined Phase-Amplitude Digital Modulation

- We already know that there is no such thing as digital frequency modulation (we can consider multiple carriers instead).
- For each carrier we have 2 parameters to change: amplitude and phase.
- But these parameters are not independent: phase shift by 180° is the same as negative amplitude.

Complex amplitude



- We can think of amplitude and phase as of length and angle of some planar vector.
- In this case coordinates x and y of that vector are independent.
- These coordinates can be packed into complex number $c = a \cdot \cos(p) + i \cdot a \cdot \sin(p)$, which is called complex amplitude.

Euler's formula

Remember original harmonic wave:

$u(t,a,f,p)=a\sin(2\pi\cdot f\cdot t+p)$.

Euler’s formula

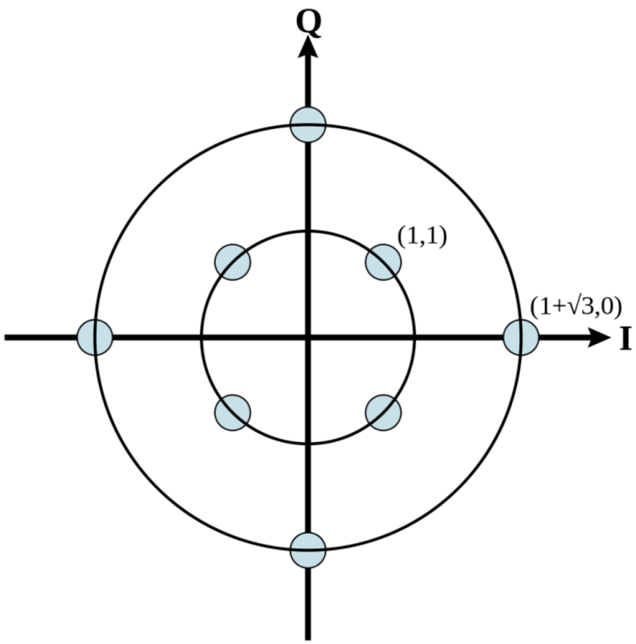
$e\uparrow i\cdot x=\cos x+i\cdot\sin x$

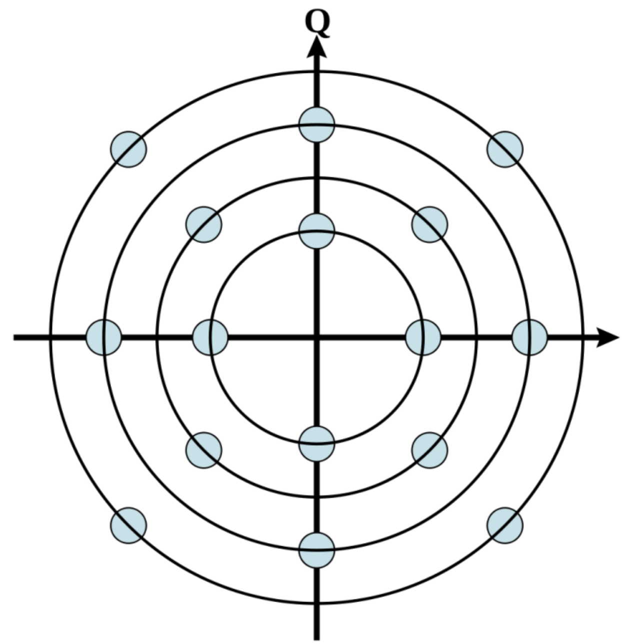
allows us to easy deal with complex amplitude:

$u(t,c,f)=\mathrm{Im}(c\cdot e\uparrow 2\pi i\cdot f\cdot t\ )$,

where $c=a\cdot\cos(p)+i\cdot a\cdot\sin(p)=a\cdot e\uparrow i\cdot p$.

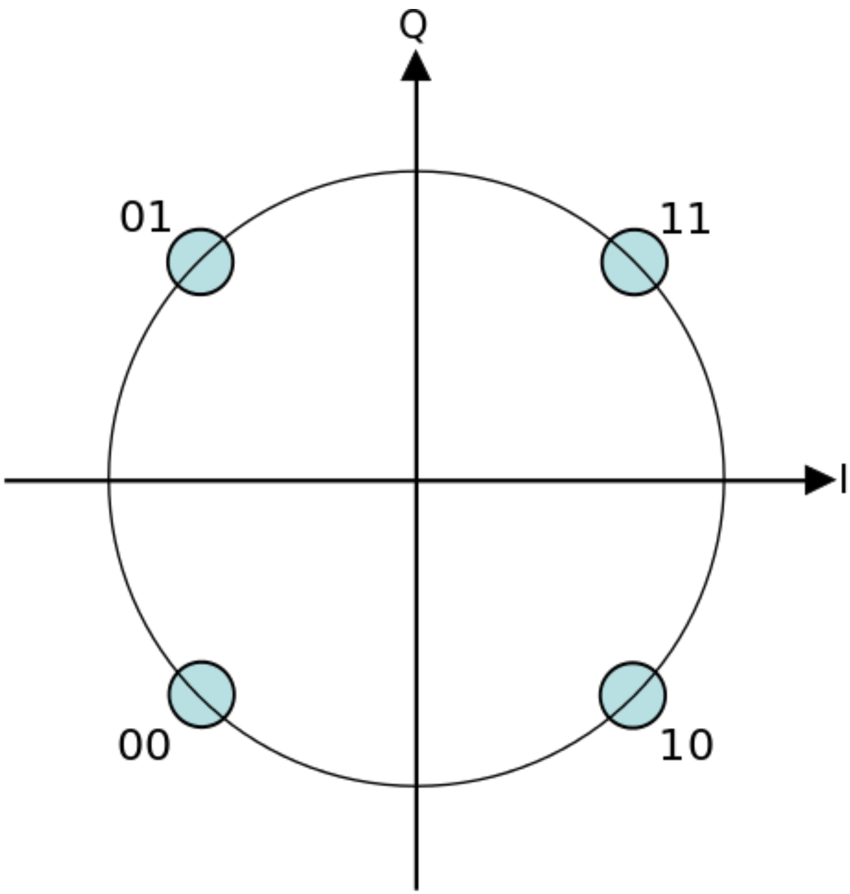
QAM Modulation





- Peek some fixed set of complex amplitude points and encode your information by switching carrier between these points.
- Such set is called QAM constellation.
- Each point encodes several bits and called QAM symbol.
- The more points are packed in QAM symbol the faster the information will be transferred, but symbols with many points are sensible to noise. So, balance is needed.

QPSK Constellation



- QPSK is the most com-mon example of QAM mo-du-la-tion.
- All points have the same distance from zero, so QPSK can be considered as pure phase modu-lati-on with 4 fixed phases.
- Alternatively you can think of QPSK as having 2 phases and 2 amplitudes.
- Each QPSK symbol carry 2 bits of information.

Linear Distortion of Symbols

- The unique feature of QAM symbols is that they are distorted the same way the carrier is distorted (we talk about linear distortions here).
- If wave is delayed by some time amount then each transmitted symbol is delayed too. This delay can be interpreted as phase shift both in carrier and QAM symbols (rotation of symbol constellations around zero).
- The amplitude of each symbol is changed the same way the carrier amplitude is changed. So, symbol constellations are scaled.
- The carrier frequency can not be changed in linear transmission media. This corresponds to the absence of frequency information in QAM symbols.
- So we can forget about carrying waves and imagine that symbols are directly transmitted and received as complex numbers.

Linear Channel

- We come to the simple and fundamental concept of linear transmission media and QAM modulation: linear channel.
- The main idea:
  - Transmitter antenna transmits stream of complex values (symbols).
  - These numbers flying in the air, bounce of walls etc, and finally come into the receiver antenna.
  - Most of distortions can be combined into single complex value which slowly changes in time.

Linear Channel

The final formula:

$$y_i = H_i \cdot x_i + n_i .$$

Here:

- $x_i \in \mathbb{C}$ — i-th transmitted symbol;
- $y_i \in \mathbb{C}$ — i-th received symbol;
- $H_i \in \mathbb{C}$ — channel state when the i-th symbol was transmitted;
- $n_i \in \mathbb{C}$ — noise of i-th received symbol.

Channel H_i and noise n_i are usually normalized so that noise has unit power:

$$E(n_i^2) = 1.$$

The channel M slowly changes in time:

$$H_{i+1} \approx H_i .$$

It should be noted that a large time can pass from the transmission of symbol to receiving of this symbol. But that time is not important for most of the tasks.

Reciprocity Model

Symbol transmitted from A to B:

$$y_i^A = H_i^{AB} \cdot x_i^A + n^A .$$

Symbol transmitted from B to A at the same time:

$$y_i^B = H_i^{BA} \cdot x_i^B + n^B .$$

Reciprocity equality:

$$H_i^{AB} = H_i^{BA} .$$

Pilot Symbols

- You may wonder how can transmitted information be decoded if each symbol is randomly rotated and scaled?
- These distortions changing slowly in time, so we can transmit thousands of symbols with mostly the same distortions.
- Transmitter sends symbols in packets and adds into each packet few special symbols with predefined values. These special symbols are called pilot symbols.
- Receiver knows that each packet has pilot symbols in certain positions, and can use them to correctly decode phase and amplitude of other received symbols.

Symbol Rate

- Each QAM symbol is transmitted by several periods of carrier wave (typically from 10 to 100 periods).
- The frequency of symbol transmission is called symbol rate.
- So, in noiseless case transmission rate is the symbol rate multiplied by number of bits per symbol.

Noise

- Main sources of noise at high frequencies:
  - Thermal electric noise in receiver and transmitter amplifiers.
  - Interference from other devices working on the same frequencies.
  - Non-linear intermodulation from the devices working at different frequencies.
- Electrical equipment and atmospheric processes do not generate noise in gigahertz spectrum used for modern high-speed wireless communications.

Methods of Dealing with Noise

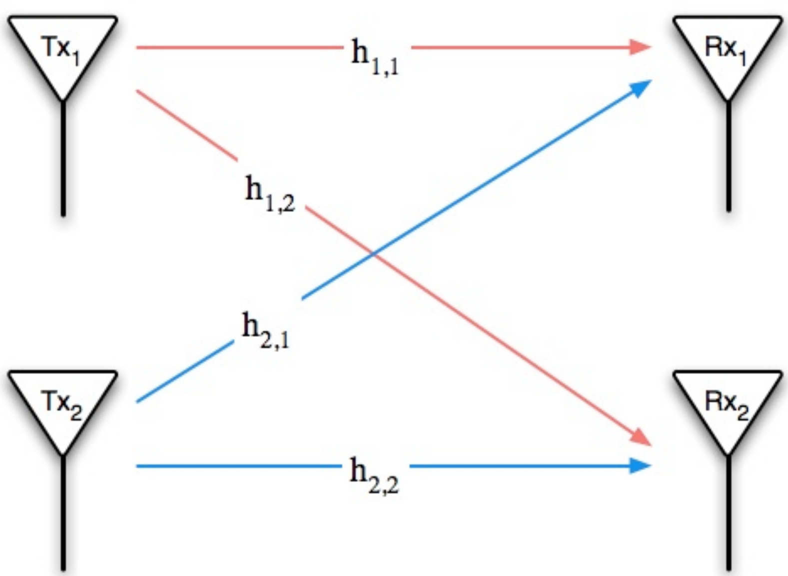
- Use more transmission power.
- Use smart antenna arrays capable of sending signal at the direction of receiver.
- Transmit less bits per symbol, so different constellation points are separated by larger energy gaps.
- Reduce symbol rate, so each symbol will be longer and have more energy.
- Use forward error correction, so small amount of errors can be fixed without retransmission.
- Use automated repeat requests for retransmitting damaged parts of data.

MULTI-ANTENNA COMMUNICATIONS

Why Multiple Antennas?

- Transmitter can use multiple antennas to send signal exactly where the receiver is (remember reciprocity illustration). Thus, more signal power will go to the receiver (larger signal-to-noise ratio).
- Receiver can use multiple antennas to listen to the direction the signal comes from. Thus more signal will be received and more noise will be filtered.
- But the most important thing is that multiple antennas in both receiver and transmitter allow opening parallel virtual wireless channels at the same frequency. That is called MIMO — multiple input and multiple output transmission medium.

2×2 MIMO System



If transmitter and receiver have 2 antennas each then we have total 4 channels, which can be packed into 2×2 matrix $\boldsymbol{H}$:

$$y = \boldsymbol{H} \cdot x + n \, .$$

Here x is vector of transmitted symbols (one per antenna), and y is vector of received symbols, n is vector of noise (normalized so that each component has unit variance).

Simple MIMO Solution

- In MIMO system transmitter can transmit different symbols from each antenna, these symbols become mixed in the air, and such mixes arrive to the receiver (different mixes to different antennas).
- Can the receiver decode the symbols were sent?
- May be, if matrix ***H*** can be inverted:

$$\hat{x} = \boldsymbol{H}^{-1} \cdot y,$$

Noise Gain

Let us consider the noise:

$$\hat{x} = \boldsymbol{H}^{-1} \cdot y = \boldsymbol{H}^{-1} (\boldsymbol{H} \cdot x + n) = x + \boldsymbol{H}^{-1} \cdot n .$$

So, sent symbol vector is almost recovered, except that noise vector is multiplied by the inverted channel matrix.

Should we worry about that?

Yes! Channel matrix usually has small norm, so the inverted matrix has huge norm, noise amplified, and sent data cannot be recovered.

Precoding

In order to reduce noise gain transmitter and receiver should cooperate. Transmitter can do some linear calculations with symbols to be transmitted:

$$x_t = \boldsymbol{P}\cdot x,$$

where $\boldsymbol{P}$ is called precoding matrix. Additional limitation to precoding matrix is that it should not increase transmitted power:

$$\mathrm{tr}\left(\boldsymbol{P}\boldsymbol{P}^H\right) = 1.$$

The receiver should then inverse both the channel and precoder:

$$x' = \boldsymbol{P}^{-1}\boldsymbol{H}^{-1}\cdot y = \boldsymbol{P}^{-1}\boldsymbol{H}^{-1}\left(\boldsymbol{H}\boldsymbol{P}\cdot x + n\right) =$$

$$= x + \boldsymbol{P}^{-1}\boldsymbol{H}^{-1}\cdot n.$$

Optimal Solution

The idea is that combined matrix $\boldsymbol{P}\uparrow -1\ \boldsymbol{H}\uparrow -1$ should have small norm.

Surprisingly, the optimal solution is given by SVD decomposition of channel matrix:

$$(\boldsymbol{U},\boldsymbol{\Sigma},\boldsymbol{V})=\mathrm{SVD}(\boldsymbol{H}),$$

where

- $\boldsymbol{V}$ is used as optimal precoder,
- $\boldsymbol{\Sigma}$ are inverse noise values of virtual MIMO channels,
- $\boldsymbol{U}\uparrow H$ is optimal decoder.

Power Allocation Problem

- Suppose that we have several independent channels with noise powers N_i, $i=1\ldots m$.
- We have transmitter with power S.
- We want to distribute power among m channels such that

$$\sum_{i=1}^{m} \log_2 \left(1+ S_i / N_i \right) \longrightarrow \max,$$
$$\sum_{i=1}^{m} S_i = S.$$

Water-Filling Power Allocation

The solution to the power allocation problem is given by a so-called water-filling algorithm:

$S_i = \max(\mu - N_i, 0)$,

Where μ is selected to satisfy power constraint:

$\sum_{i=1}^{m} S_i = S$.

Water-Filling for SVD Precoding

In case of SVD precoding virtual channels are independent, with noise powers $N_i = 1/\boldsymbol{\Sigma}_i^2$.

Plugging that into water-filling we get optimal power allocation matrix

$\boldsymbol{S} = \mathrm{diag}(\sqrt{S_1}, \ldots, \sqrt{S_m})$.

This matrix can be used in addition to precoder:

$x_t = \boldsymbol{SP} \cdot x$

Comments on MIMO Precoding

- Water-filling algorithm may select less number of channels than the rank of matrix $\boldsymbol{H}$ if some channels are “bad”.
- Such bad channels correspond to small singular values of matrix $\boldsymbol{H}$.
- In the worst case MIMO system will degrade to single-channel system, but nevertheless will perform better than SISO system.
- We have omitted here some peculiarities such as adaptive selection of symbol rate and QAM constellation for each virtual channel.