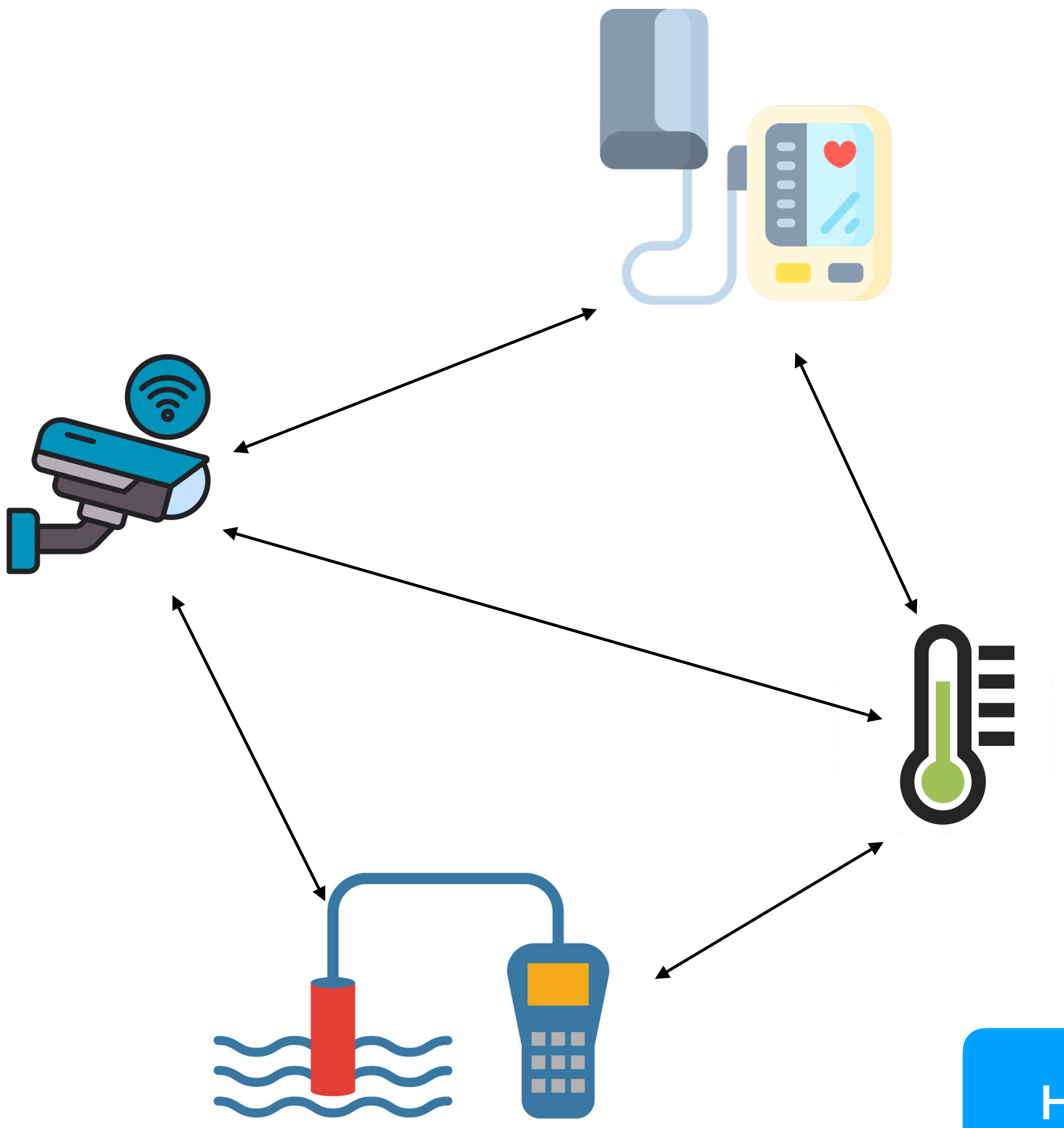
Internet of Things

5. Communication

5.1.Radio Frequency Communication (PHY)

M.Sc. Computer Science 2024-2025

Viviana Arrigoni



HOW?

- In IoT networks, smart devices collect data and work together to accomplish a common objective.
- Smart things form the **perception layer** of an IoT architecture.
- These devices have very limited computation capacities and are scattered around possibly large areas. For these reasons, they must be able to **communicate** their data (raw or processed).

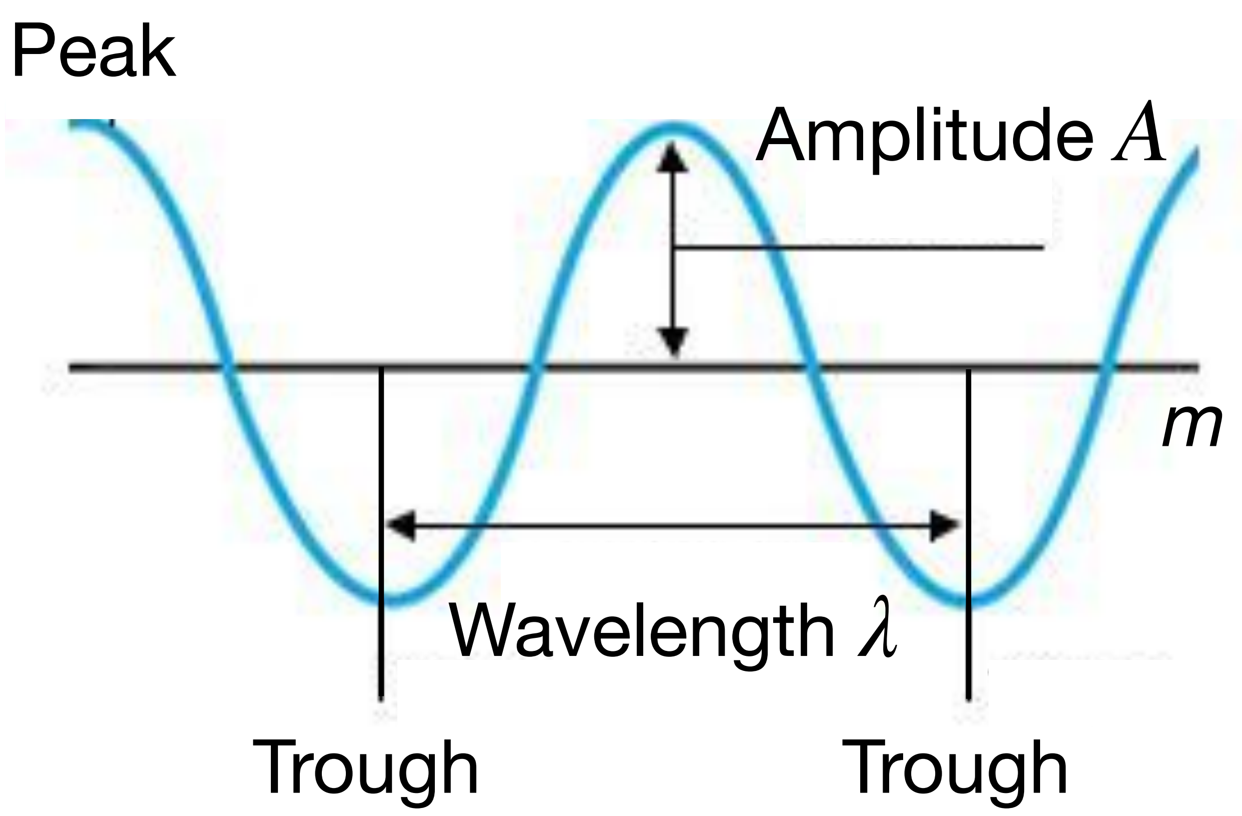
5.1.1: Electromagnetism Theory

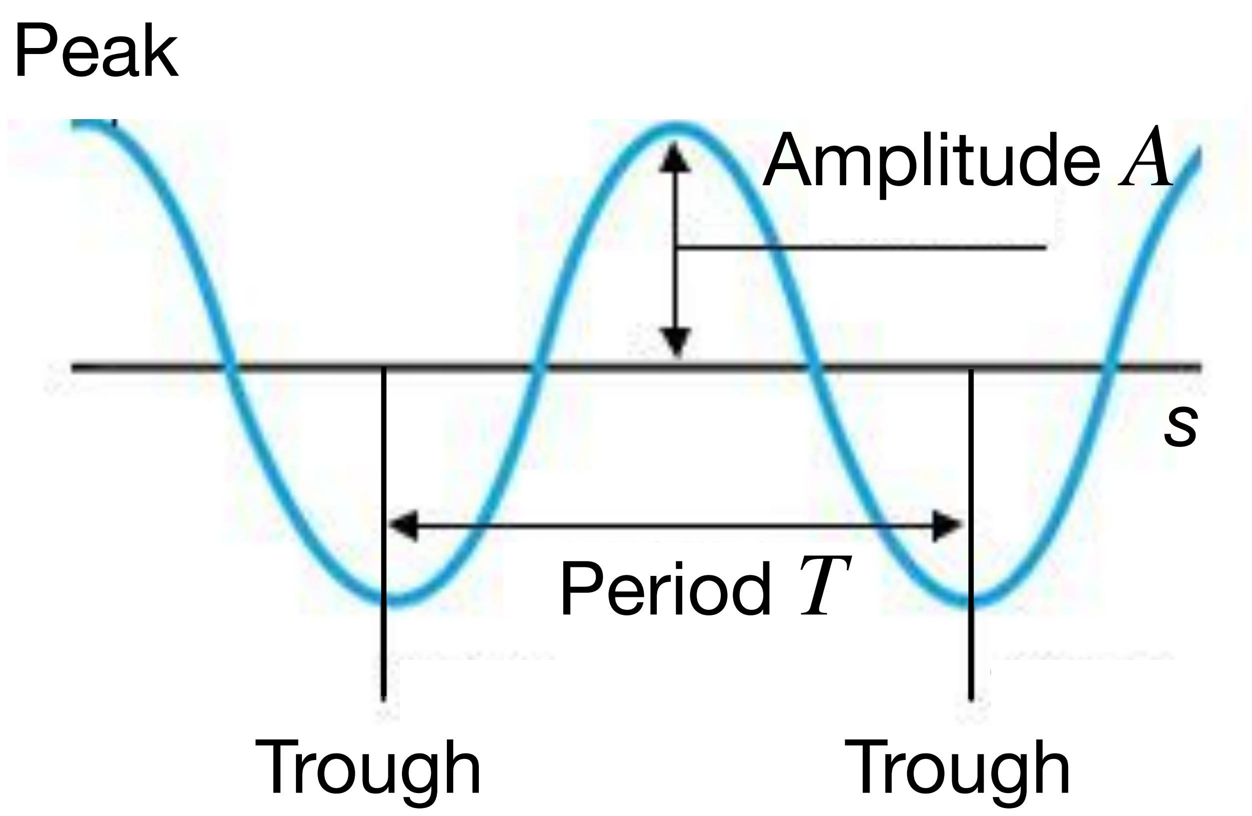
Light

...what is it?

video

- Photon: the smallest amount of energy that can be transported (*quantum* of energy). Has no mass, cannot be split, but only created or destroyed.
- Electromagnetic radiation: electromagnetic waves that carry energy distributed as “energy packets”, i.e., photons.
- Energy that can be transported by a photon: $E = hf$, where h is the Planck constant (very small, ~10^{-34}) and f is the **frequency** of the wave.





Frequency f: number of waves that pass a point in a given amount of time (hertz). It is the inverse of the period

Speed = $\lambda \times f = \dfrac{\lambda}{T}$

$\lambda = \dfrac{c}{f}$ $f = \dfrac{c}{\lambda}$ c: speed of light

Wavelength and frequency are inversely proportional!

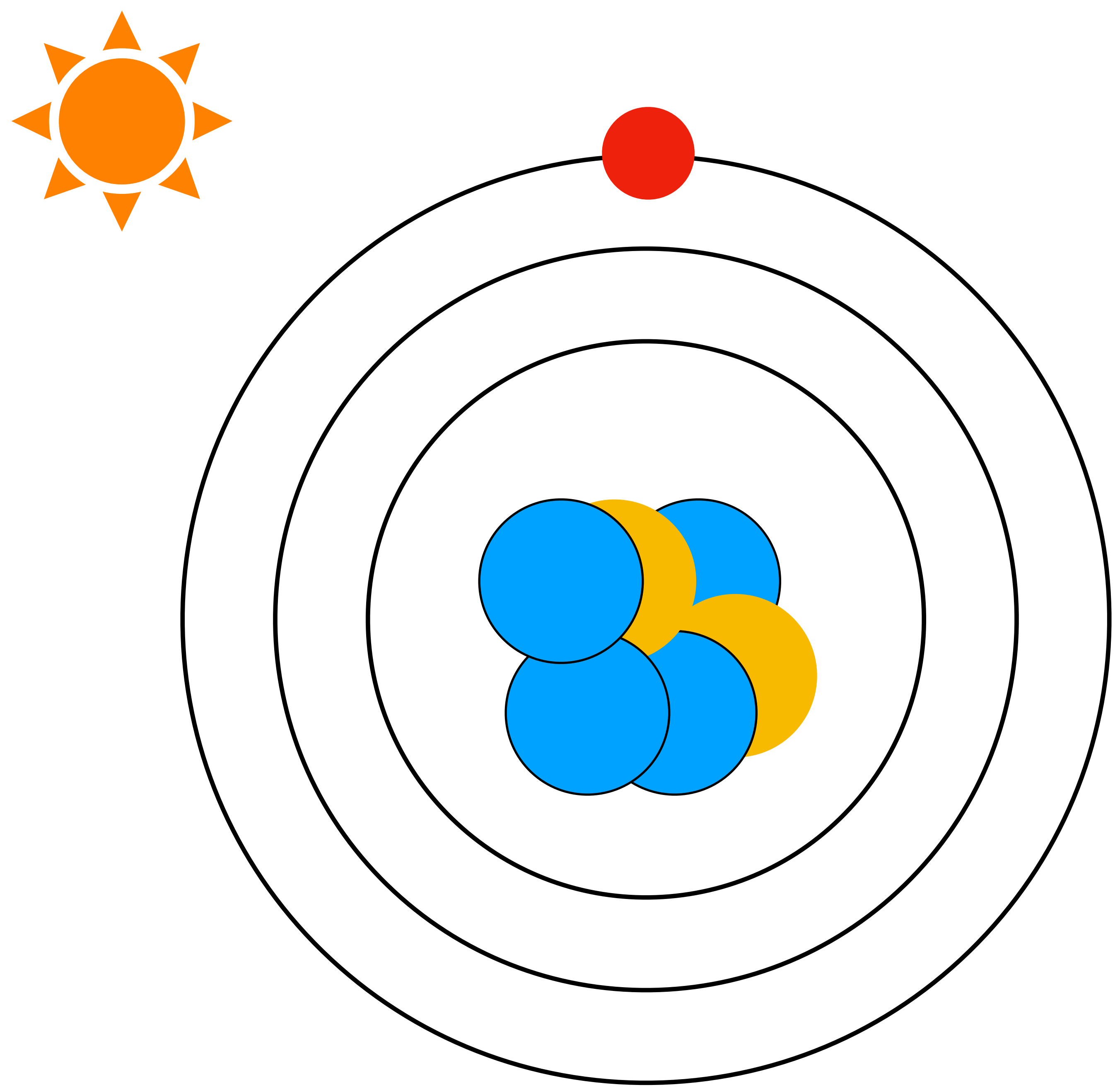
Where does the energy for electromagnetic waves come from?

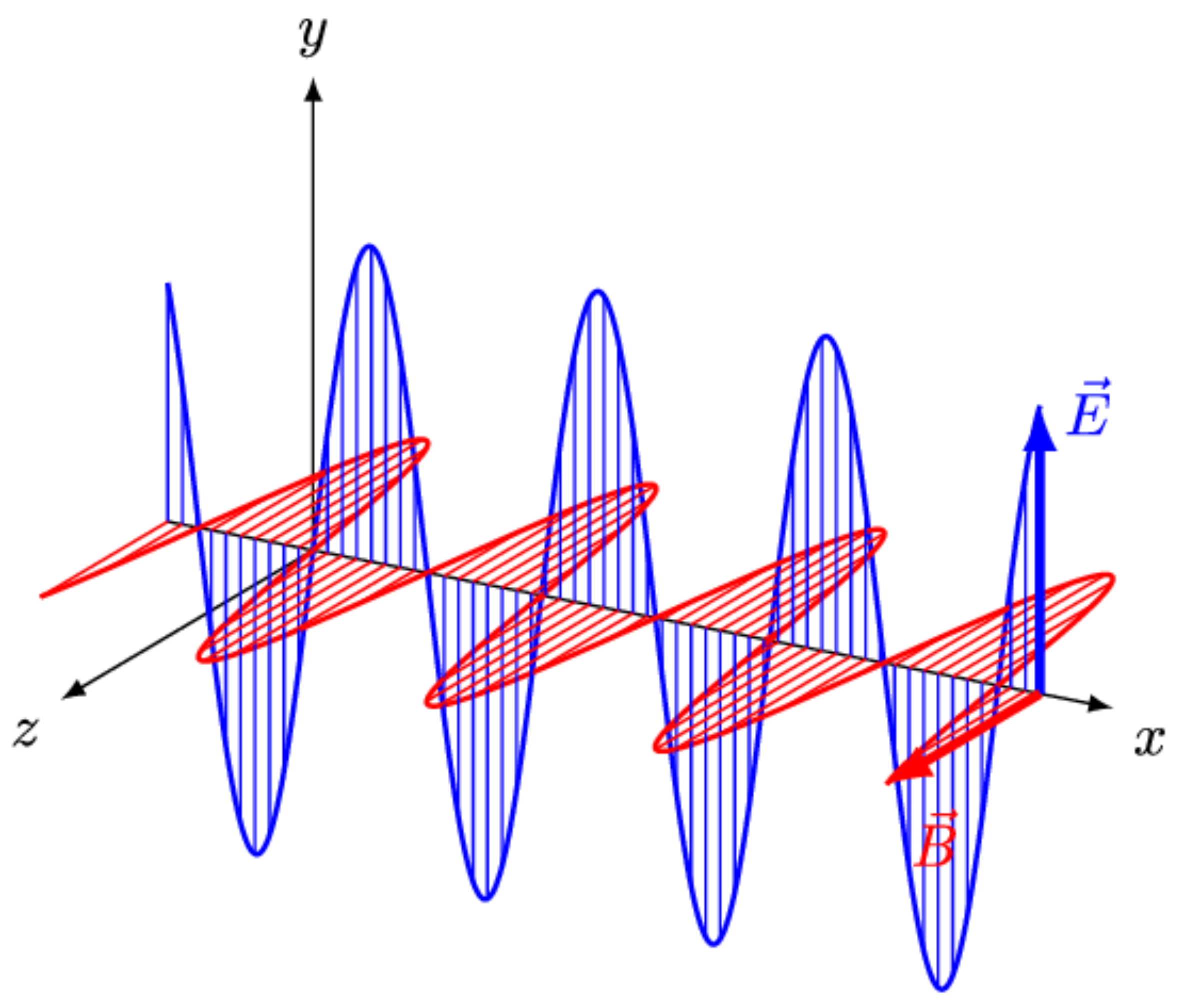
- A vast range of electromagnetic waves are created when atoms or molecules drop from a higher state of energy to a lower one: in this way, they lose energy and emit it in the form of radiation.

- An electron in an excited state (e.g., heated up) drops to a lower energy state and looses this excess energy. If heated up again, the electron can gain energy again and move to a high energy state.

- The moving charge of the electron creates an oscillating **magnetic field** $\overrightarrow{B}$, which creates an oscillating **electric field** $\overrightarrow{E}$ perpendicular to it.

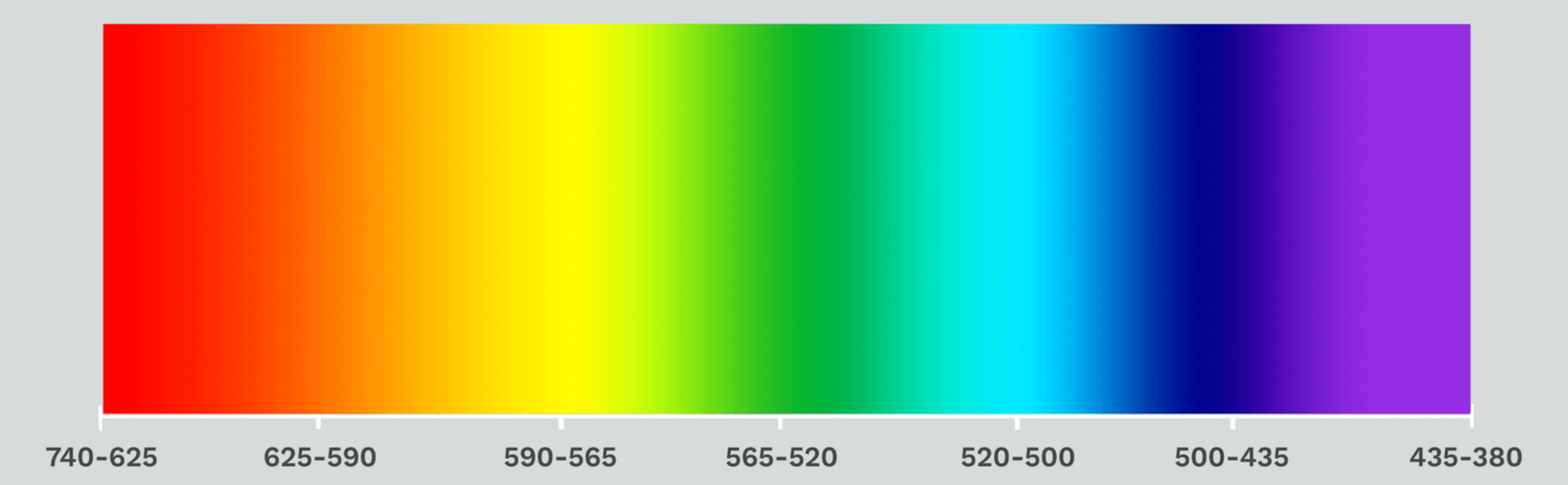
- Curiosity: all bodies at a temperature higher than the absolute zero (-273,15°C) emit electromagnetic radiations! (Planck's Law)



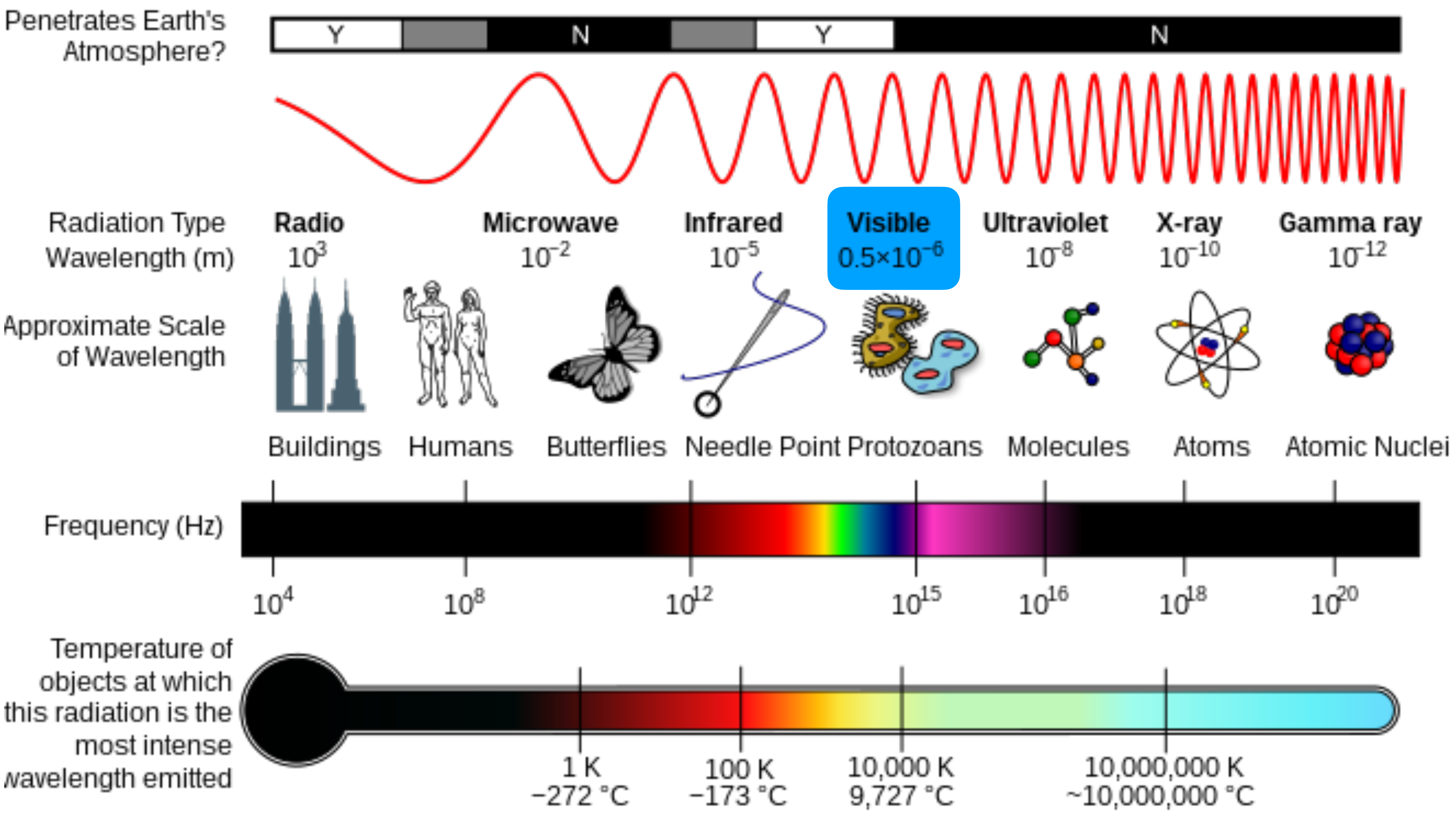


- Electric waves’ wavelengths and frequencies come in wide a spectrum.

- Visible light spectrum is the segment of the electromagnetic spectrum that the human eye can view (380 to 740 nanometers wavelength [10^{-9} meters]- size similar to that of bacteria. 400-780 terahertz [10^{12} cycles per second]).



Wavelength



Red light is cooler than blue light!

Can we use EM waves to communicate?

- Yes. We can encode information by sending an EM source and varying its properties (e.g., brightness, colour, etc.)

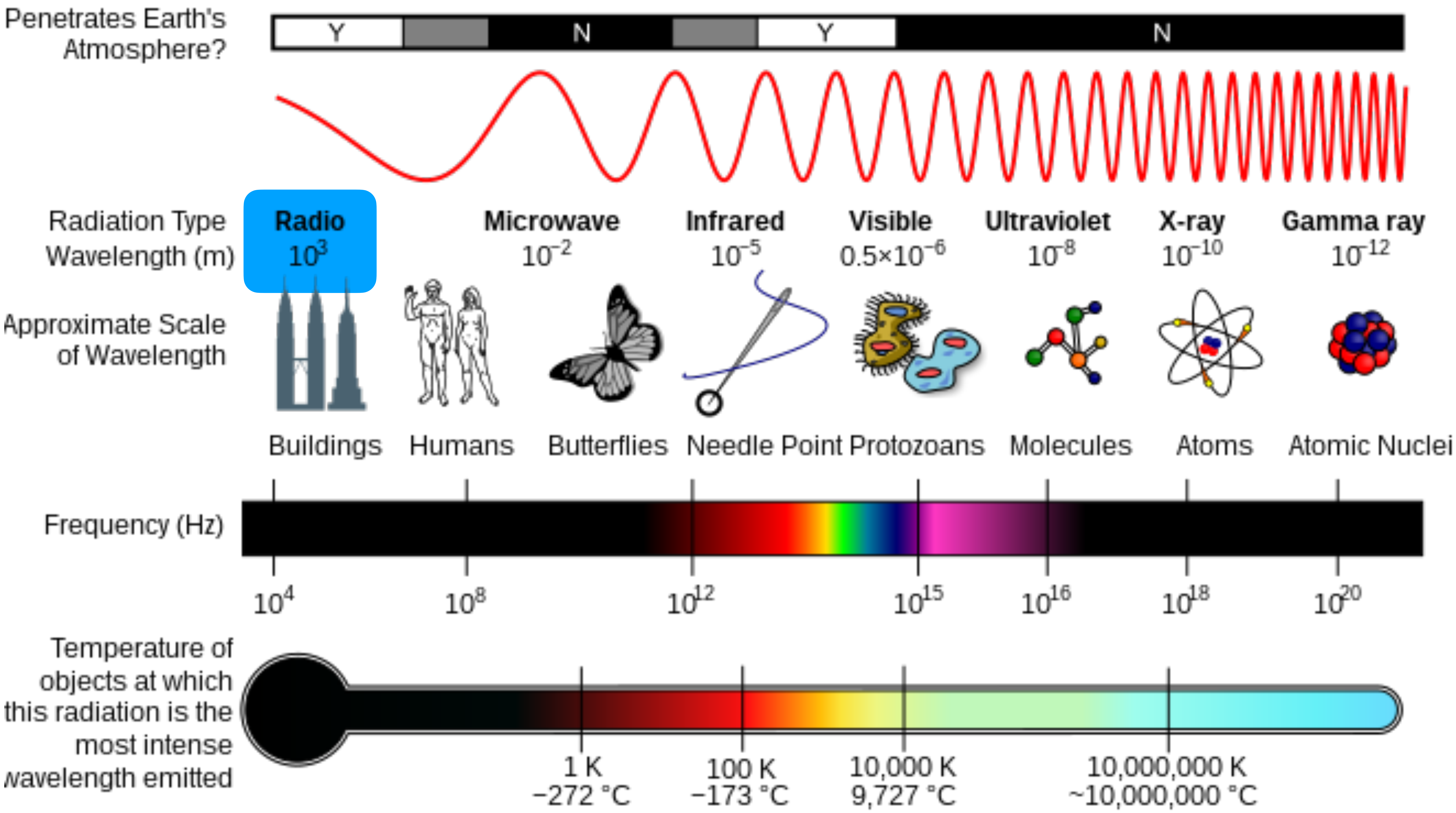
- What frequency should we use?

EM radiation acts differently at different wavelengths

- Higher frequencies are more dangerous to humans.

- High and low frequencies tend to go through objects (IR, visible and UV tend to be dissipated by air, can be reflected or refracted by objects).

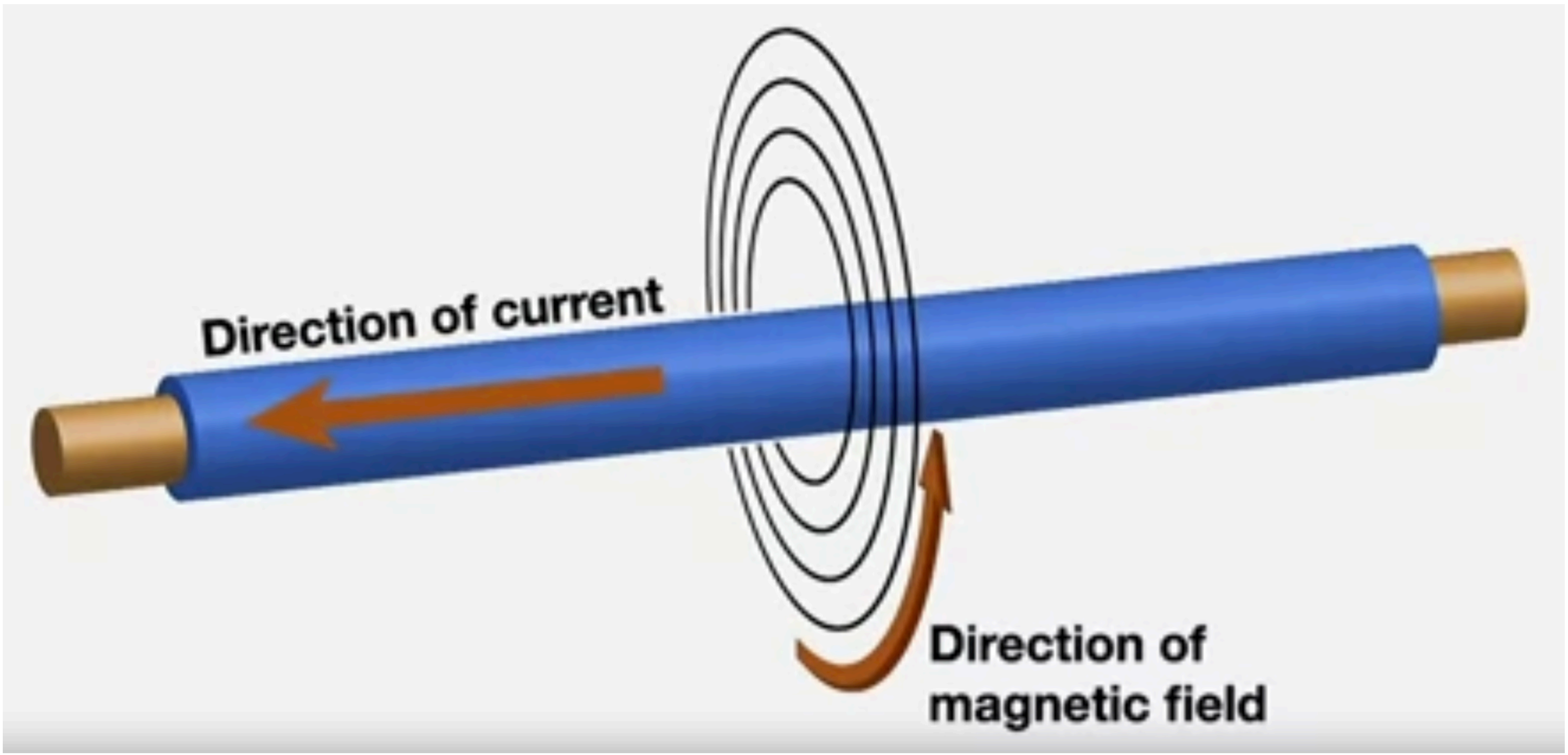
- Lower and middle frequencies are easier to make and control with circuits. Higher frequencies are more difficult and expensive to make: need to module high frequencies very precisely.

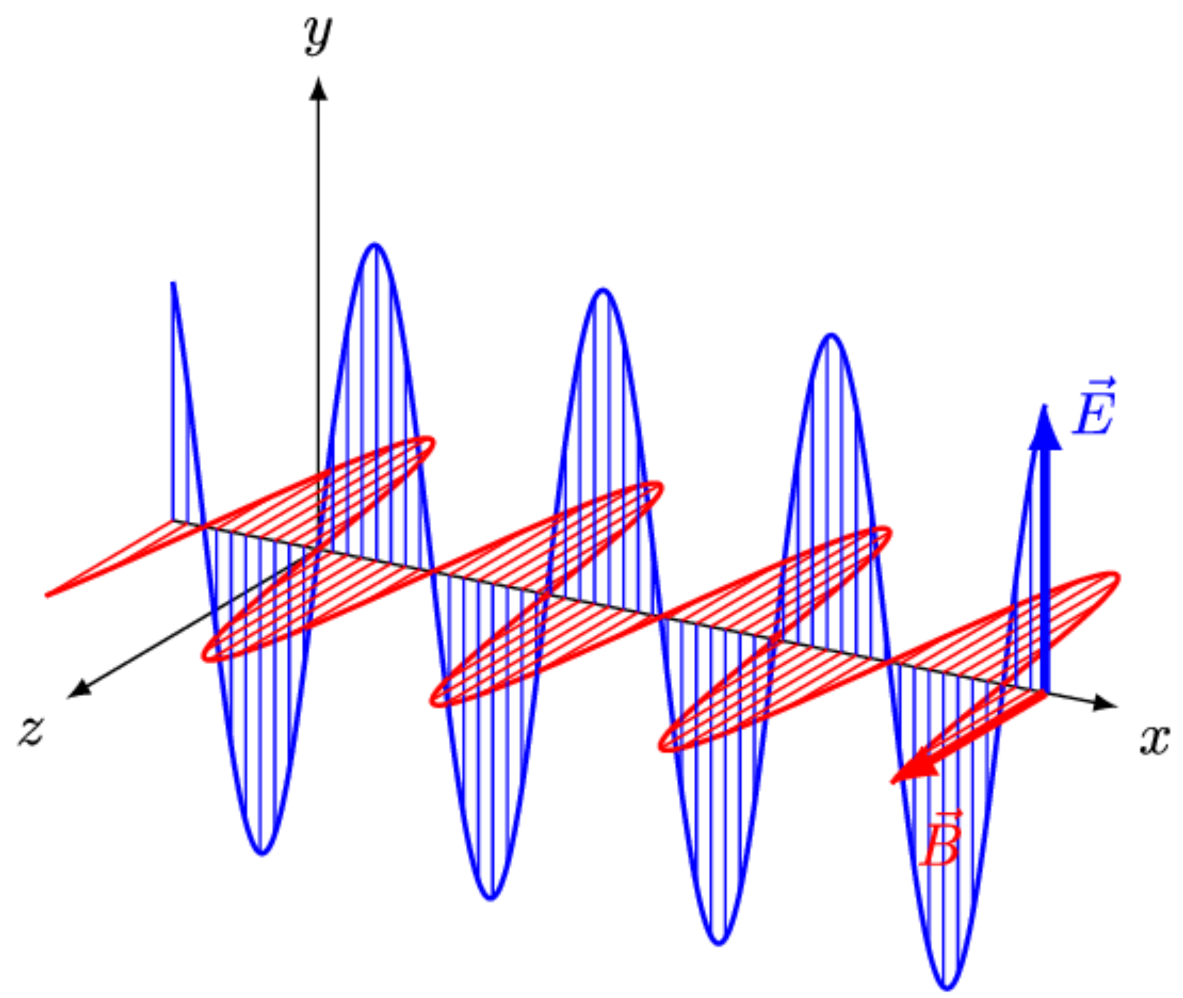


Radio frequency: 3kHz to 300 GHz. Easy to generate, good propagation characteristics, relatively safe for humans

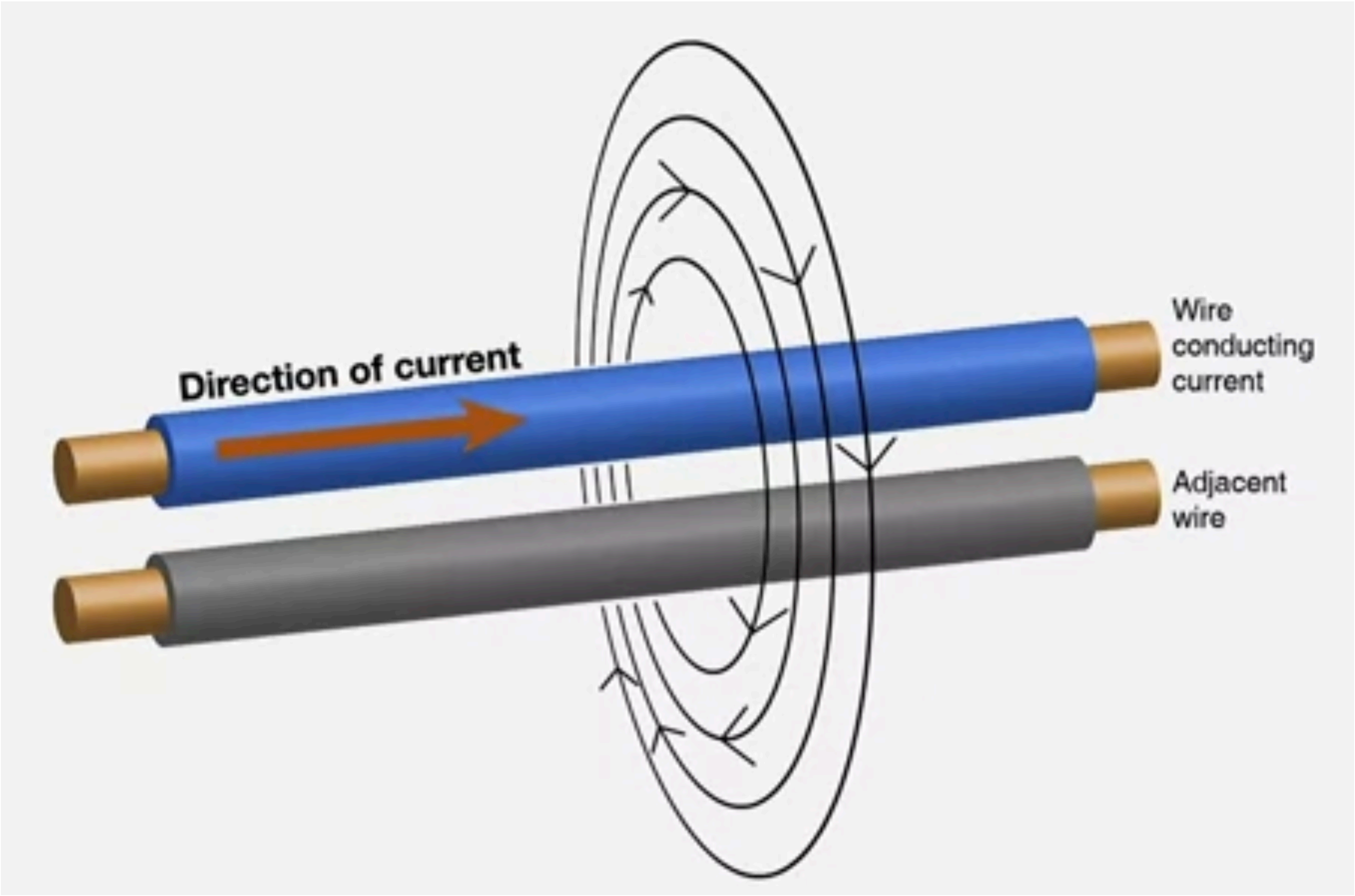
How to generate EM waves?

- Need some way to generate EM waves in a controllable way (i.e., very specific frequencies, with certain patterns to encode data)
- **Electricity in a wire produces an EM field!**





Magnetic field induces current flow



- By taking another wire and placing it next to the first one when electric current is flowing, the magnetic field created by the flowing current induces current on the second wire, too.

- In wired networks, this can be bad (information is carried by electric pulse, can cause **crosstalks** - totally want to avoid it).

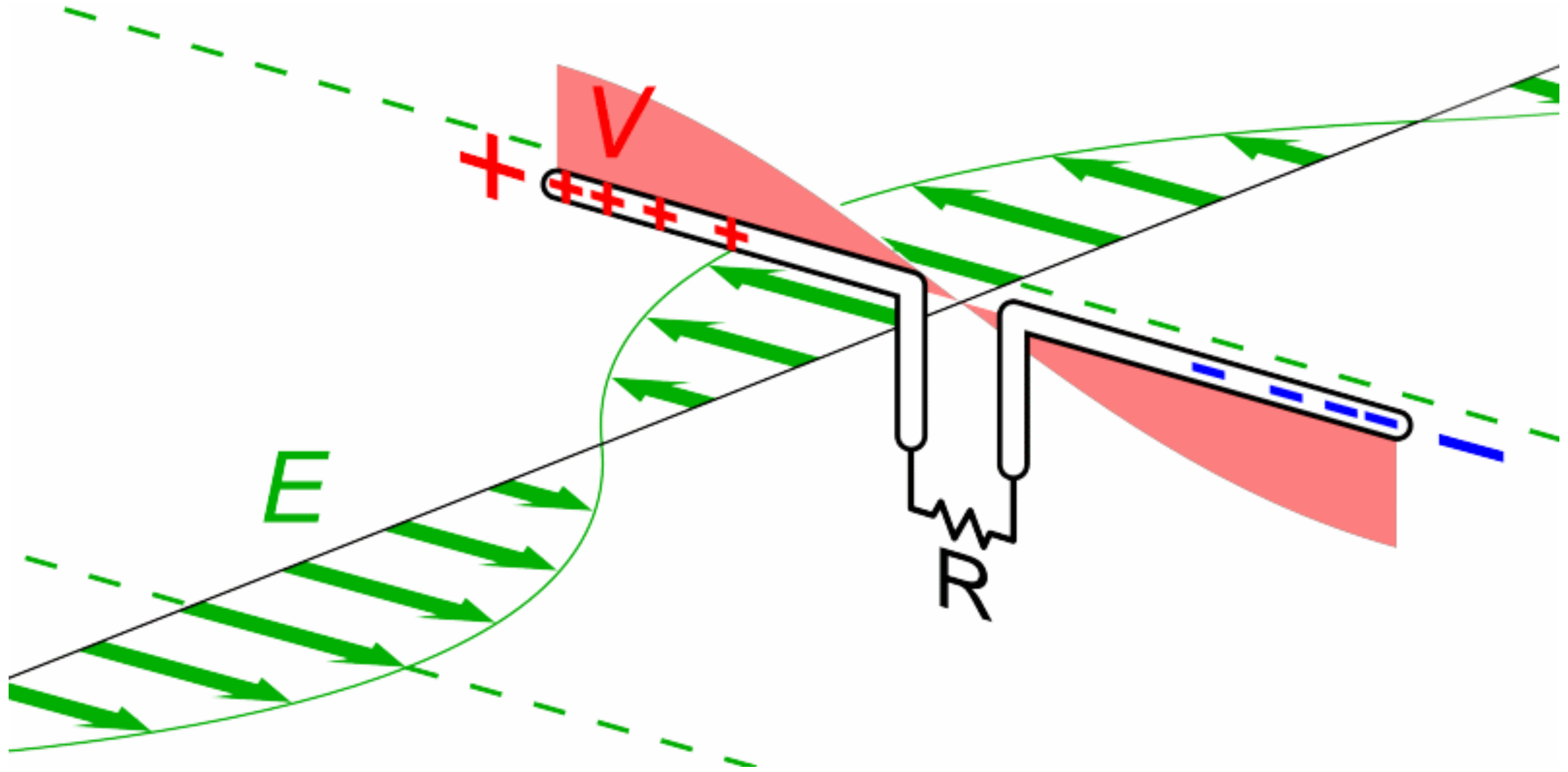
- In wireless networks, this can be good! We modulate communication in one wire and it appears in the other one.

How do we structure these wires?

- To make electricity flow in a wire, you need to have two points with a voltage difference.

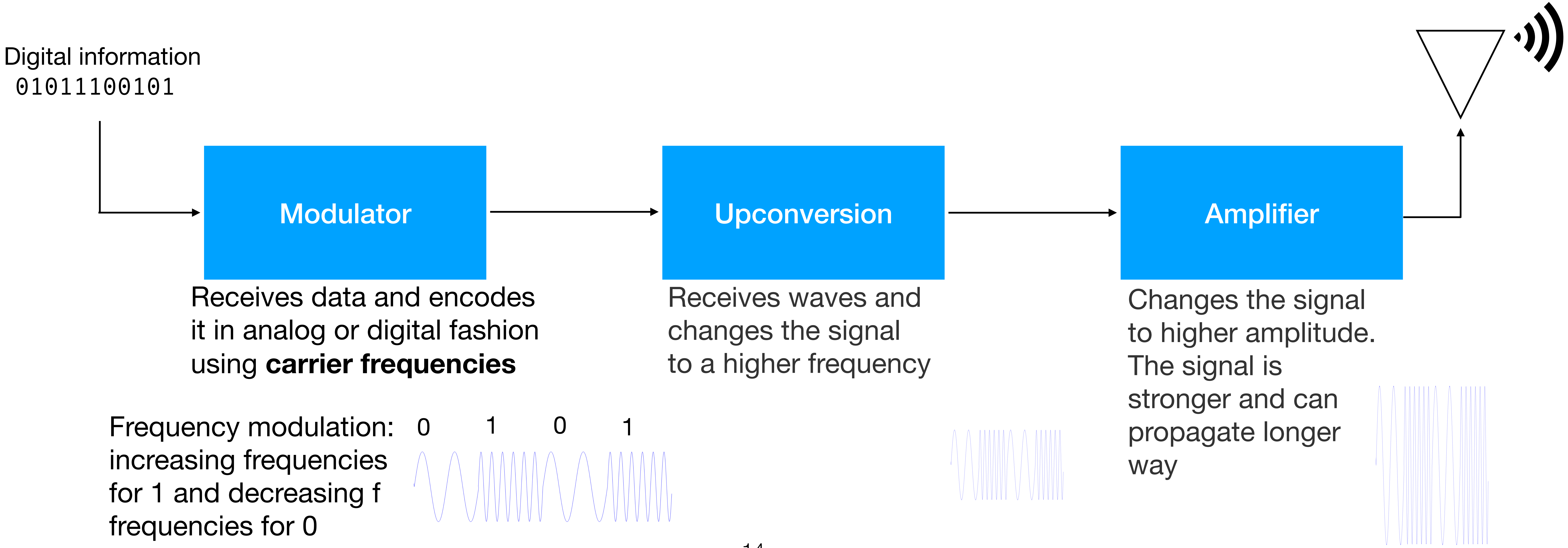
- Particular antenna design: **dipole** (bend wires to improve signal strength) YouTube link

- RF Circuits: Circuits where the electricity flowing through them is **modulated** at radio frequencies (we will learn about signal modulation later).



RF Transmitter

- An RF Transmitter is a RF circuit that takes data as input (sequence of 0s and 1s) and produces a fluctuating magnetic field as a output that propagates out and can be received by a receiver further away.



Why do we need to up convert the signal?

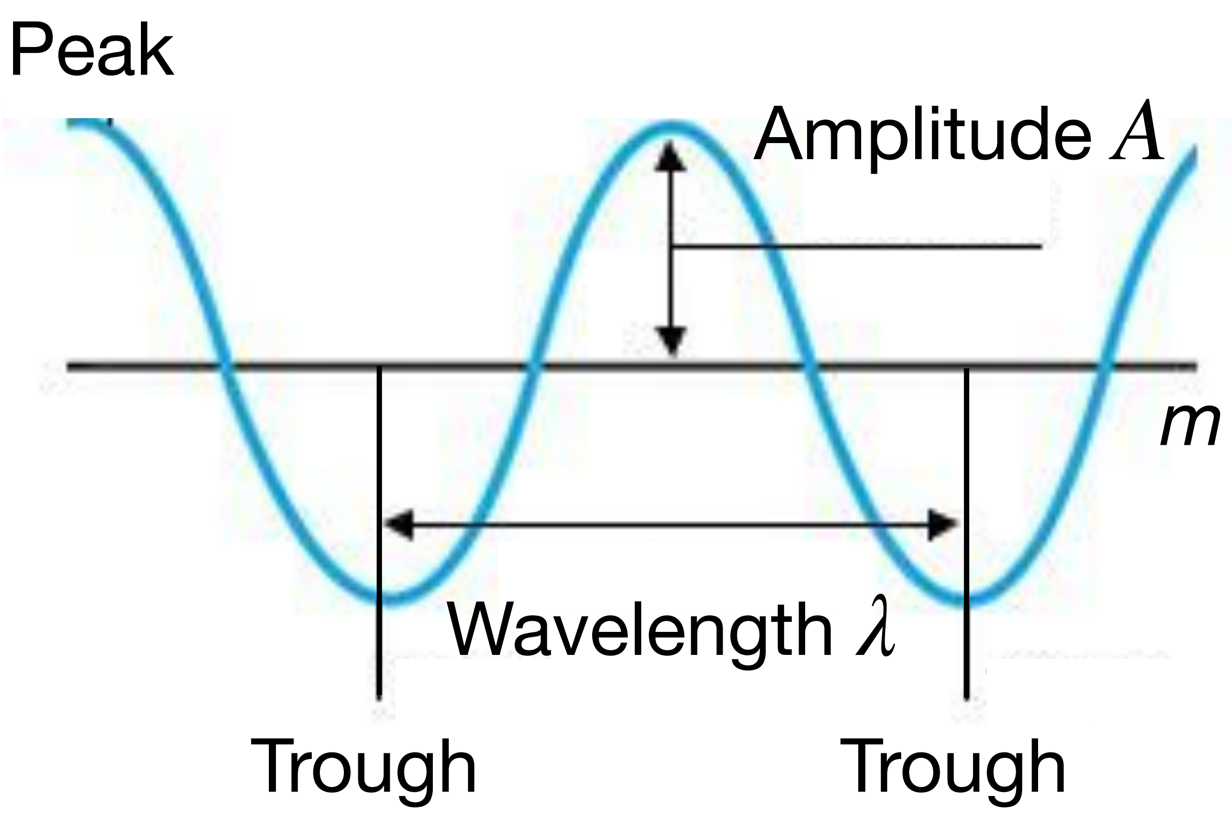
- As we have seen from the video, the length of a dipole, L, needs to be equal to half the wavelength of the signal, i.e., $L = \frac{\lambda}{2}$.
  - Low frequency means large wavelength waves, and require huge antennas. For instance, 100 KHz signal needs an antenna ~3 km long, while a 2 GHz (Wi-Fi) signal needs an antenna ~15 cm long.
- The waves after upconversion are still radio waves.
- Why don’t we directly modulate digital signal into high frequency waves?
  - It is possible to directly generate high-frequency signals from digital sources, but:
  - High-frequency digital-to-analog converters (DACs) and digital mixers are very expensive and power-hungry.
  - Generating high-frequency signals directly from digital circuits requires ultrafast transistors, which are difficult to manufacture.

Why higher wave amplitudes mean stronger signals?

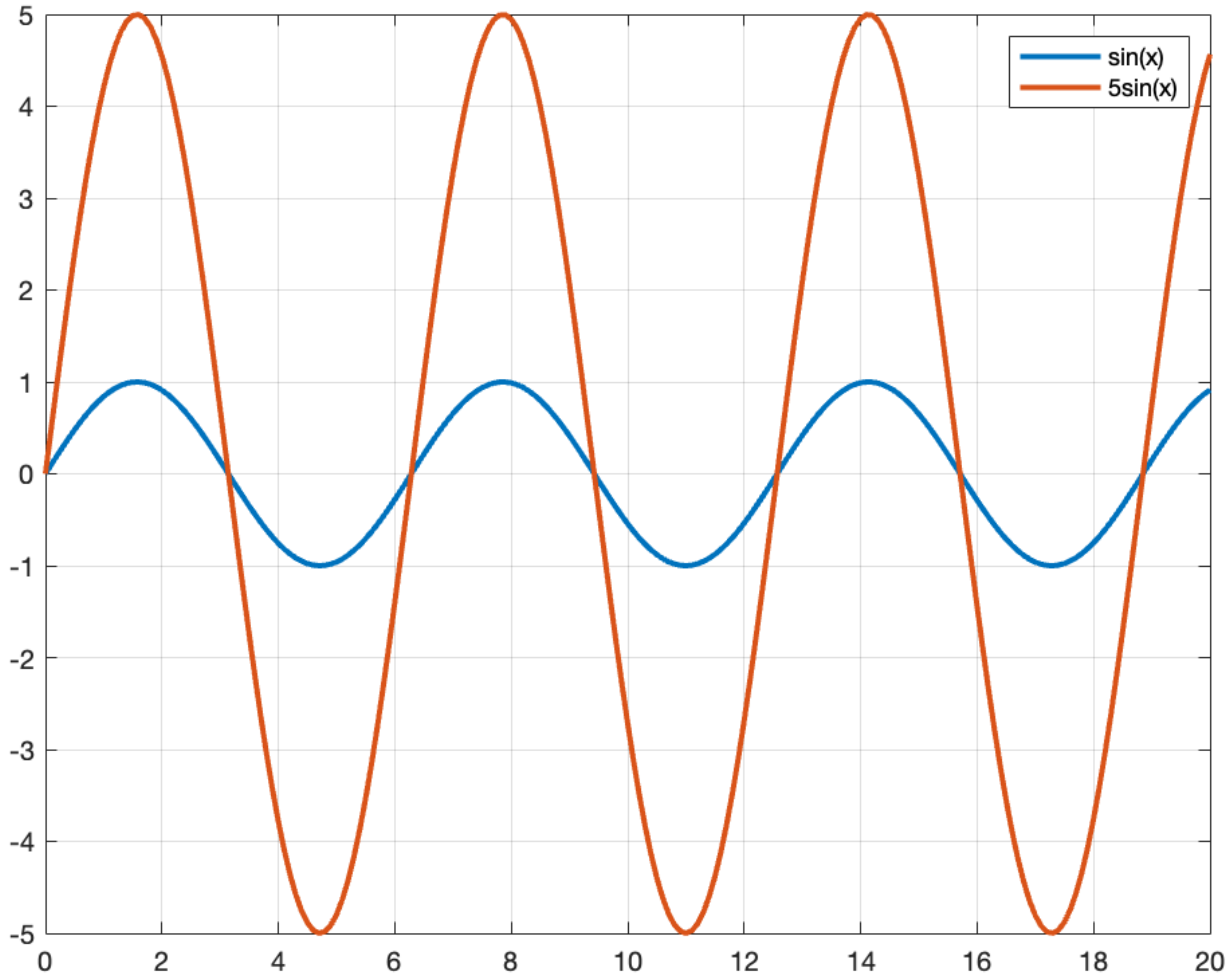
- We have seen that objects that generate or radiate electromagnetic waves create an electromagnetic field which carries energy through space.

- The amplitude of an electromagnetic wave is **the peak value of the electric field $\overrightarrow{E}$ and magnetic field $\overrightarrow{B}$,** i.e., the maximum strength of the fields.

- The energy carried by an electromagnetic wave is proportional to its amplitude squared.

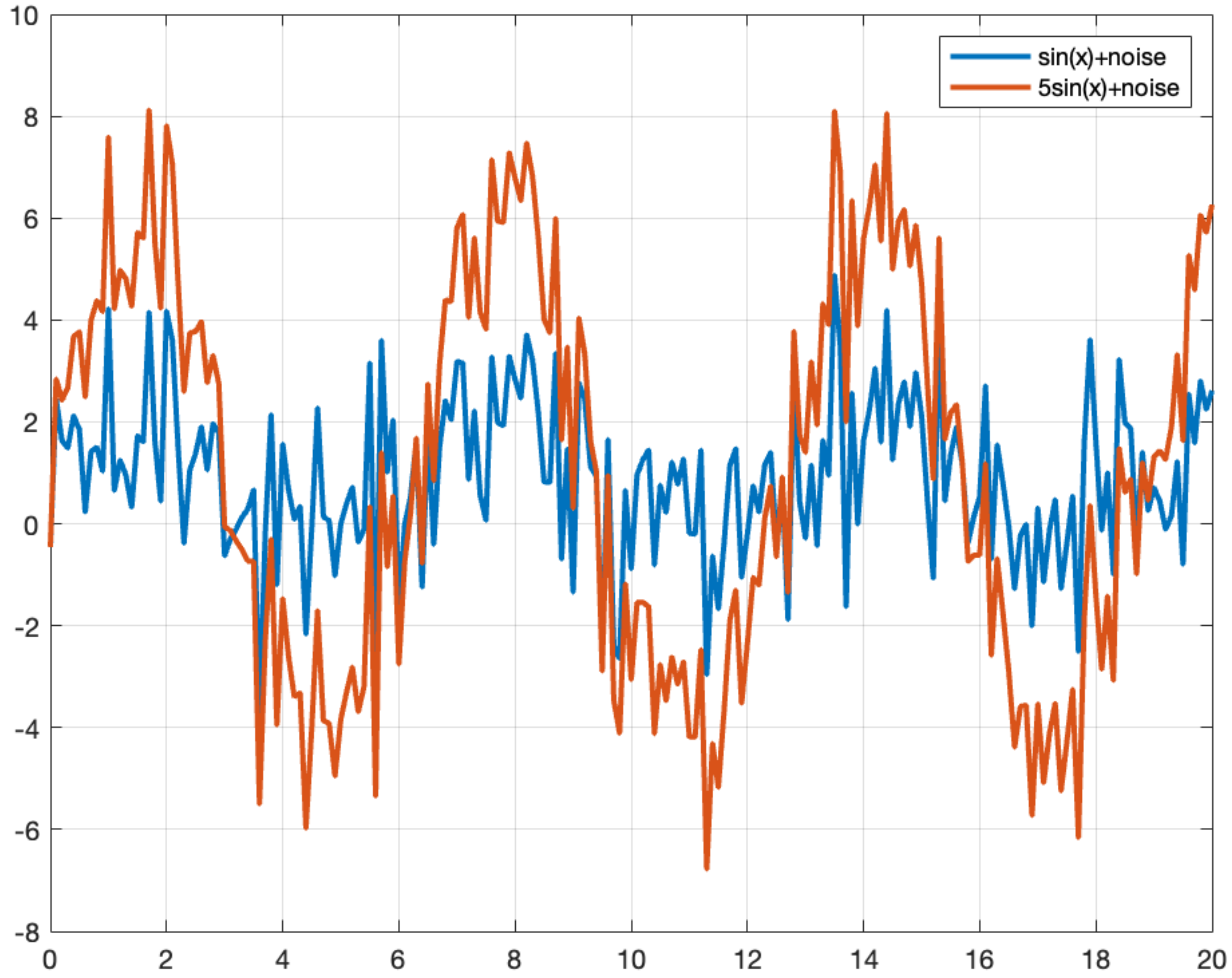


Why higher wave amplitudes mean stronger signals?



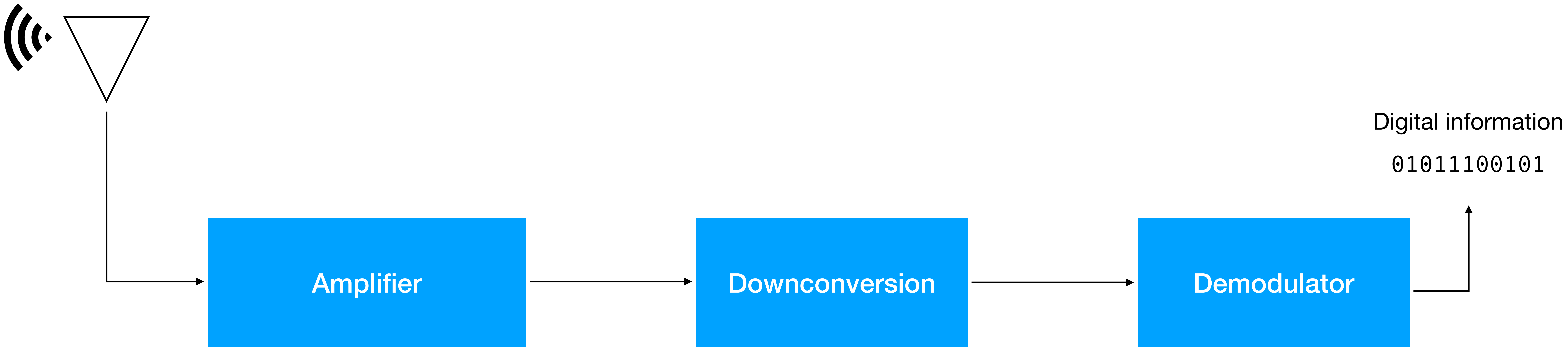
Two waves with equal frequency/wavelength but different amplitudes

Why higher wave amplitudes mean stronger signals?



The same waves with Gaussian noise.
It is still easy to see peaks and troughs of the red wave, while it is much more difficult for the blue one.

RF Receiver



Amplifier

- Changes the signal to higher amplitude.
  - As the signal travelled, it got degraded and we need to amplify it back.
- Amplification allows to reconstruct the signal by filtering out noise

Downconversion

Uses filters to filter out extra noise. Outputs a signal with reduced frequency.

Demodulator

Converts the signal into a digital sequence

We call transceiver an antenna that can transmit and receive signals

What Frequencies to use?

- Radio frequencies span from 3 kHz to 300 GHz (bandwidth)
- When two phase of the same frequency overlap, their amplitudes combine according to the principle of **superposition**:
  - If the waves are **in phase** (peaks align with peaks and troughs with troughs), they add up, leading to **constructive** interference (stronger signal).
  - If the waves are **out of phase** (one wave's peak aligns with the other's trough), they cancel each other out, leading to **destructive** inferences (nullified signal).
- For this reason:
  - Different protocols use different frequencies.
  - Multiple antennas are usually deployed
  - Adaptive modulation

UNITED STATES FREQUENCY ALLOCATIONS

THE RADIO SPECTRUM

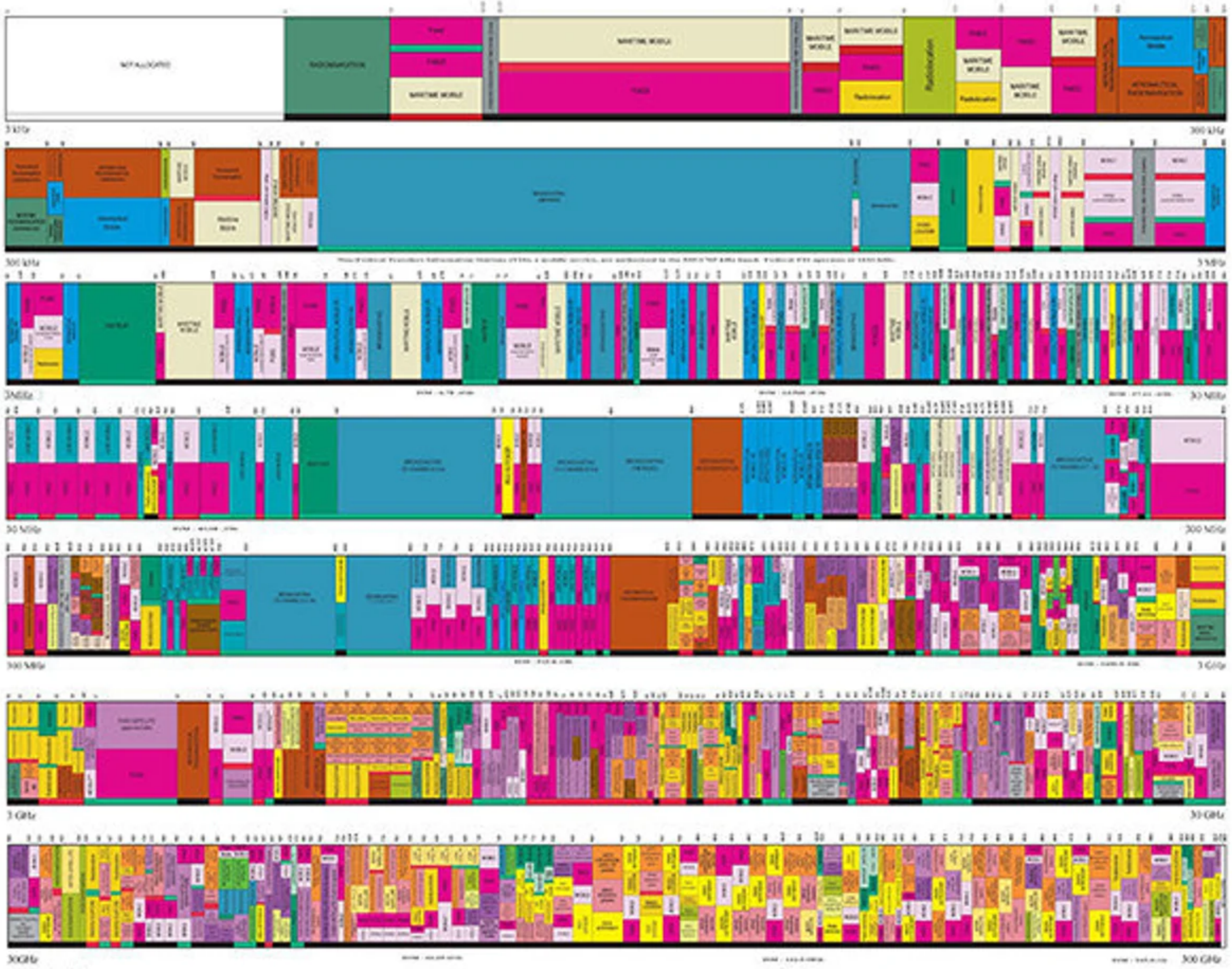
RADIO SERVICES COLOR LEGEND

ACTIVITY CODE

ALLOCATION USAGE DESIGNATION

U.S. DEPARTMENT OF COMMERCE
National Telecommunications and Information Administration
Office of Spectrum Management

August 2011



5.1.2 Antenna design

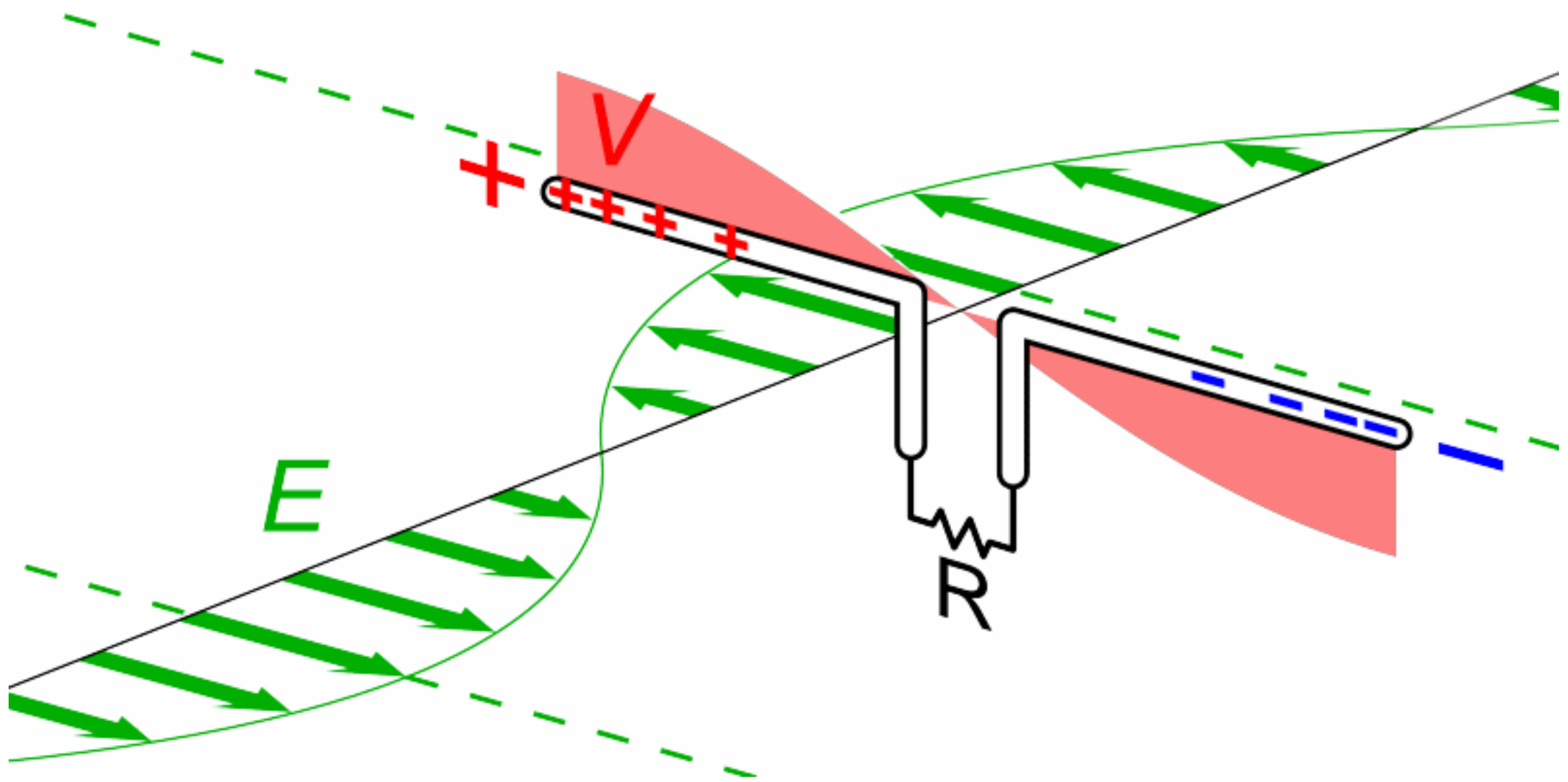
Antenna design

- In real deployments, it is imperative to carefully choose what antennas to deploy, how they are oriented and how they are tuned.

Polarisation of EM Waves

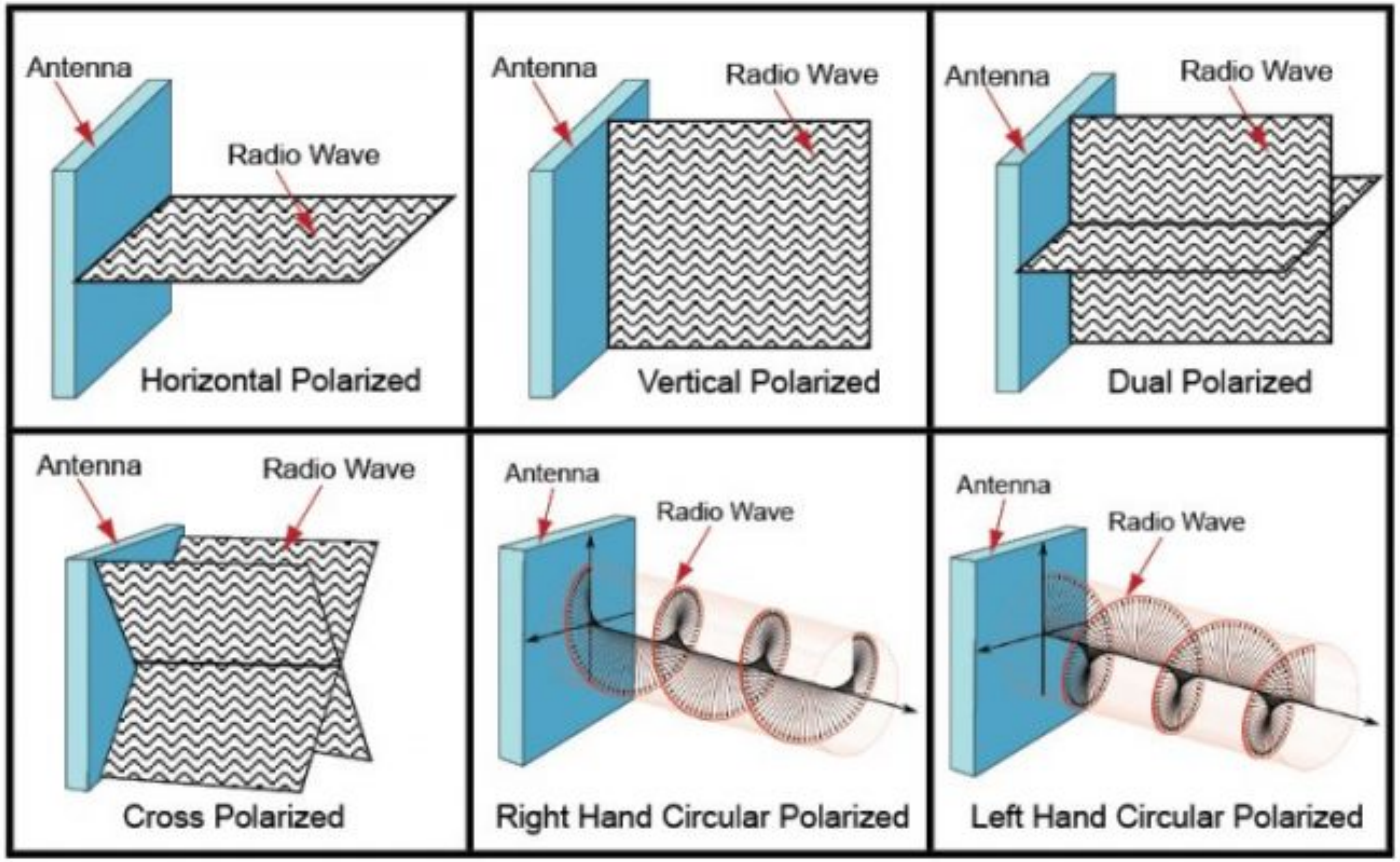
- We have seen that antennas create an EM field that is propagated through a direction

- The field (i.e., the signal) has a certain orientation, i.e., it is **polarised** in a direction.



- In order to be able to receive the signal, the orientation of the receiver antenna must match that of the transmitter.

Polarization options



- In order for the receiver to be able to receive the signal, it must be able to match the pattern created by the transmitter and must be oriented according to the transmitter’s orientation.
- This is not always possible or easy, antennas spread around are usually multi proprietary.
- Transmitters/Receivers often have more than one antennas with different orientations.

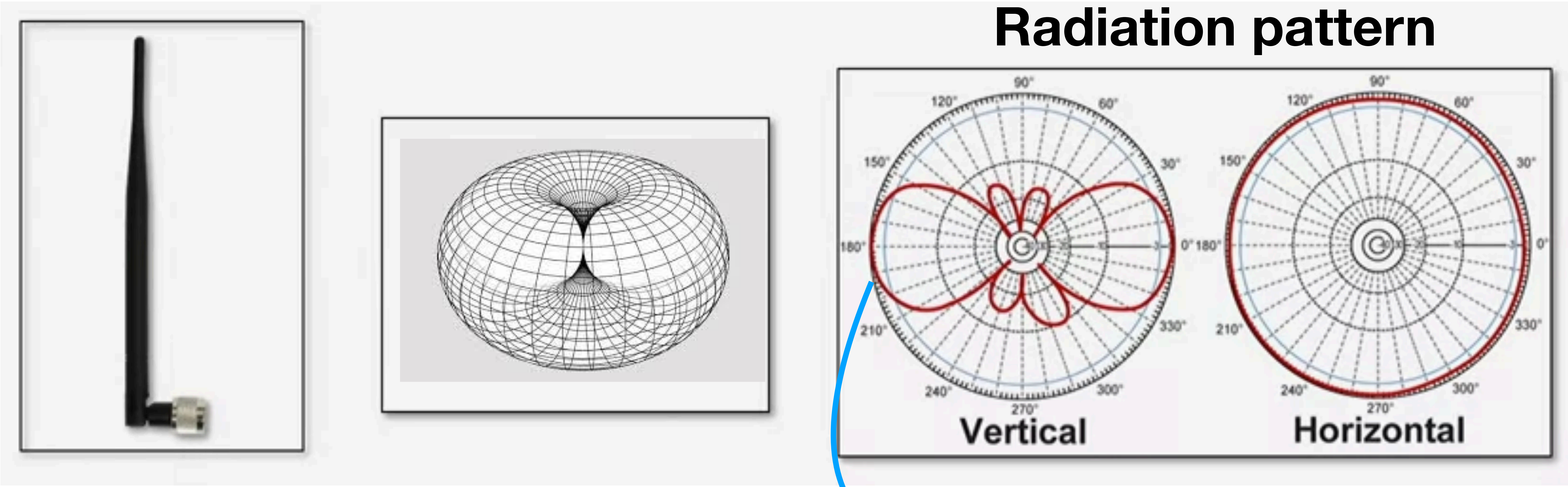
Radiation patterns

- We can classify antennas also depending on the radiation patterns, similarly to light sources.
- The radiation pattern is the graphical representation of the radiation strength of the antenna as a function of space.
  - Defines the variation of the **power** radiated by an antenna as a function of the **direction** away from the antenna.

Omnidirectional

Directional

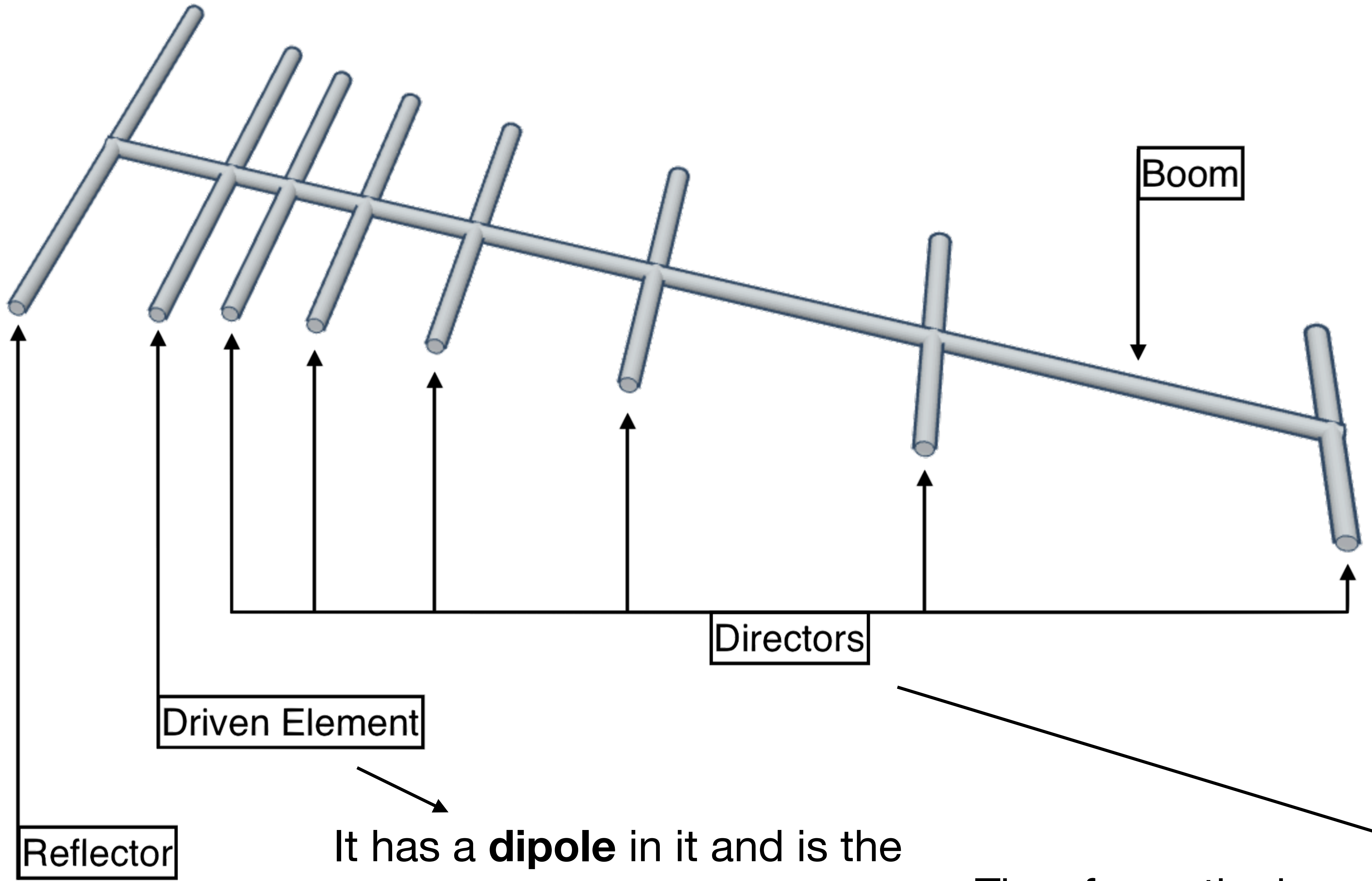
Omnidirectional antenna



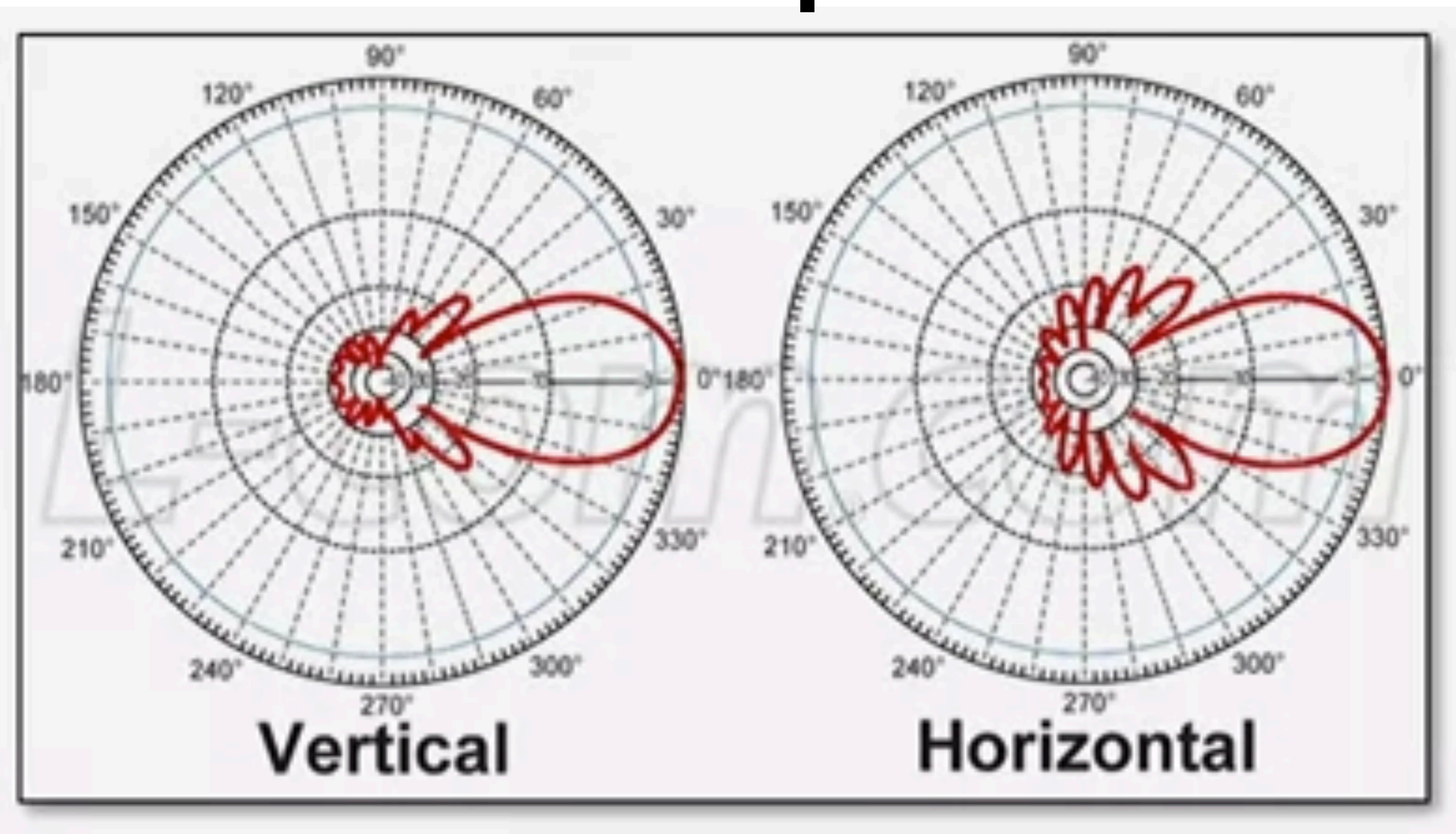
Lobes: show the direction where the radiation is maximal

- Doughnut (or torus) shaped, radiates equally in all directions of the plane cutting it horizontally (power decreases above and below that plane).
- Widely used in Access Points (APs) and Network Interface Cards (NICs)

Directional antenna (Yagi)



Radiation pattern



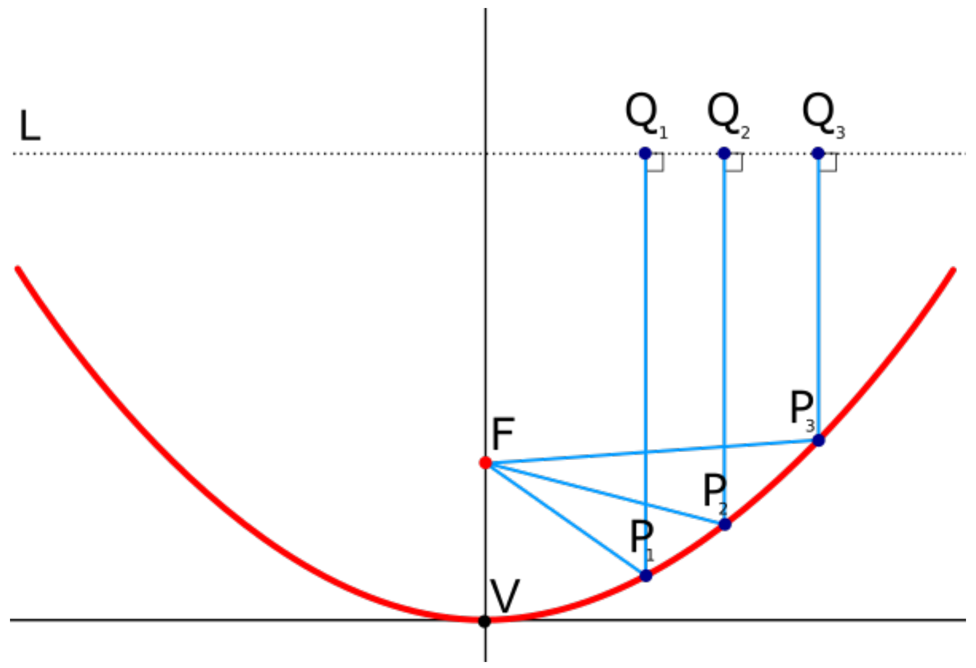
It has a **dipole** in it and is the element generating the signal to be transmitted

They focus the beam out in a certain direction. They receive and re-radiate the radio waves from the driven element and modify the driven element's radiation pattern. About 5% shorter than the driven element.

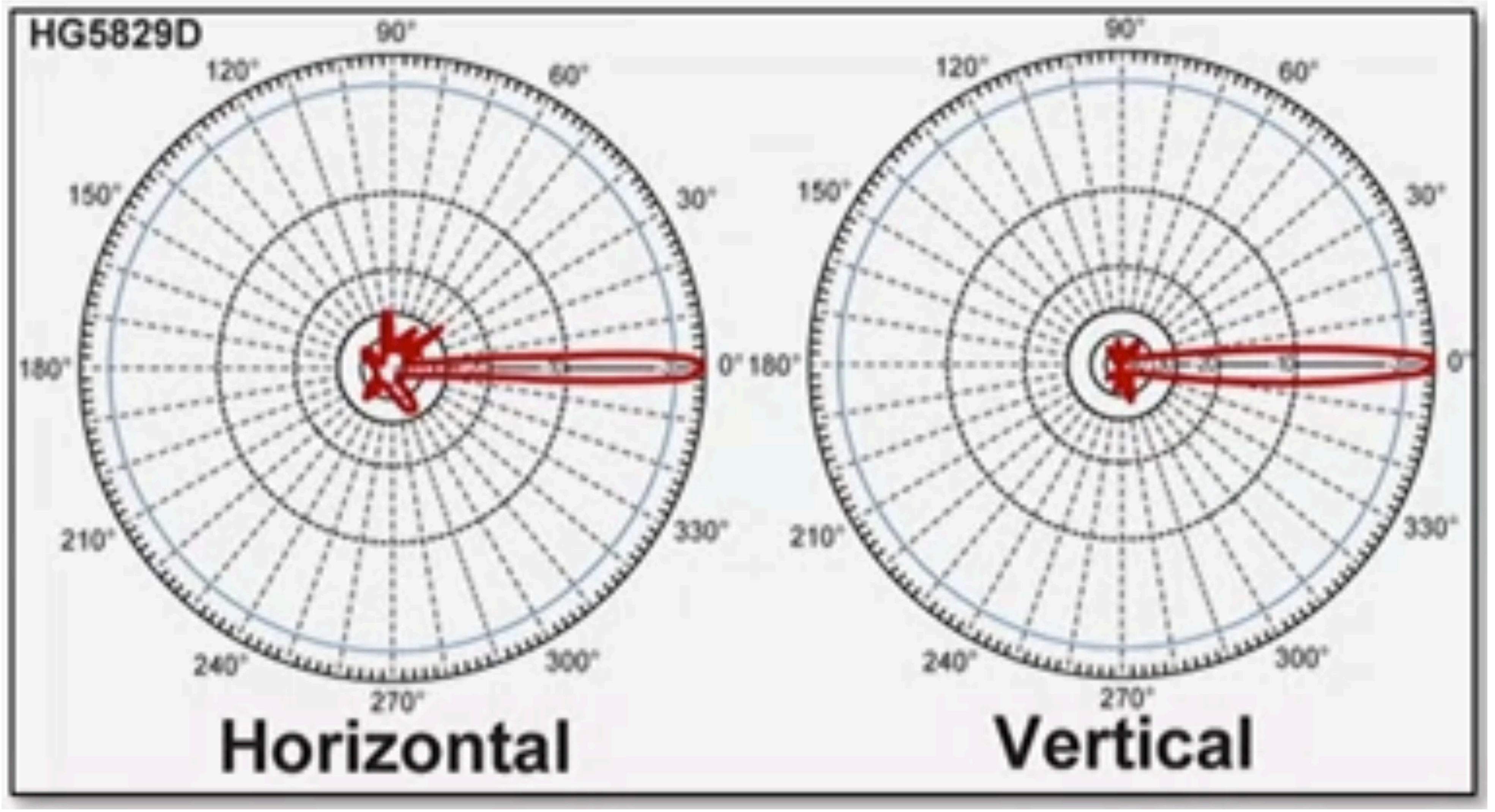
Adds gain to the signal by reflecting the energy produced at the driven element and sending it in a single direction. Usually there is a single reflector. It is about 5% longer than the driven element.

Spacing of the wires is a function of the frequencies of the transmitting signals.

Directional antenna (parabolic)



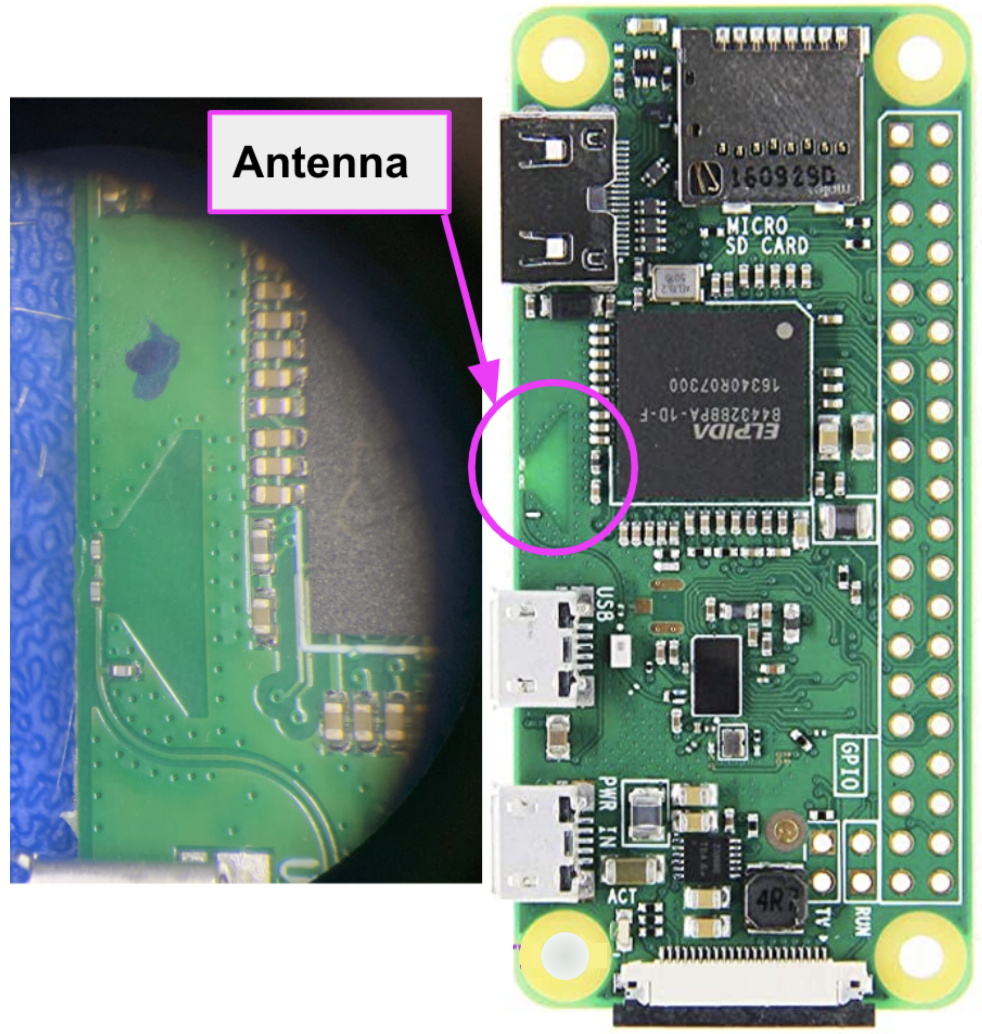
Radiation pattern



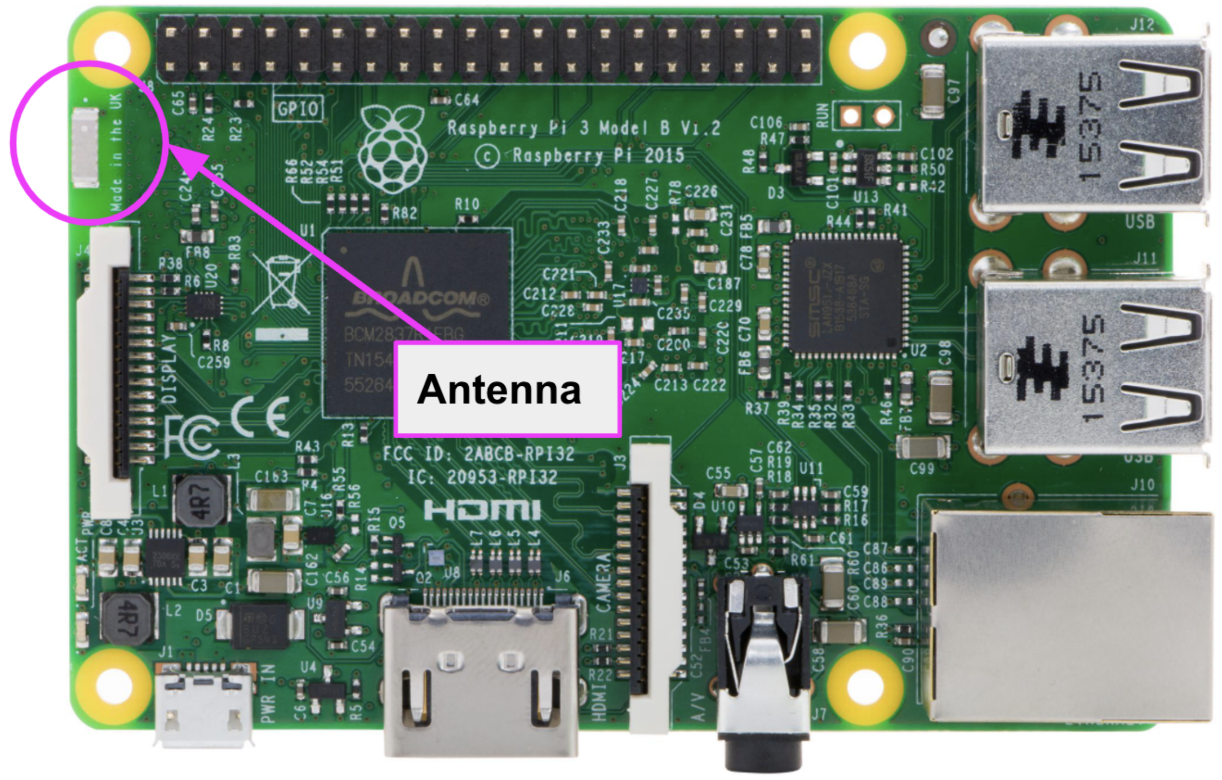
- Has a metal parabolic reflector with an antenna suspended at the focus.
- Uses properties of **elliptic paraboloids**: the parabolic reflector transforms an incoming plane wave travelling along the axis into a spherical wave converging toward the focus. Conversely, a spherical wave generated by a point source placed in the focus is reflected into a plane wave.

Antennas for IoT Devices

- PCB Antennas (Printed Circuit Board Antennas) are antennas printed directly into the board of an embedded system - reduced dimensions and maintenance costs.
- Can be directional or omnidirectional



PCB trace antenna
Raspberry Pi

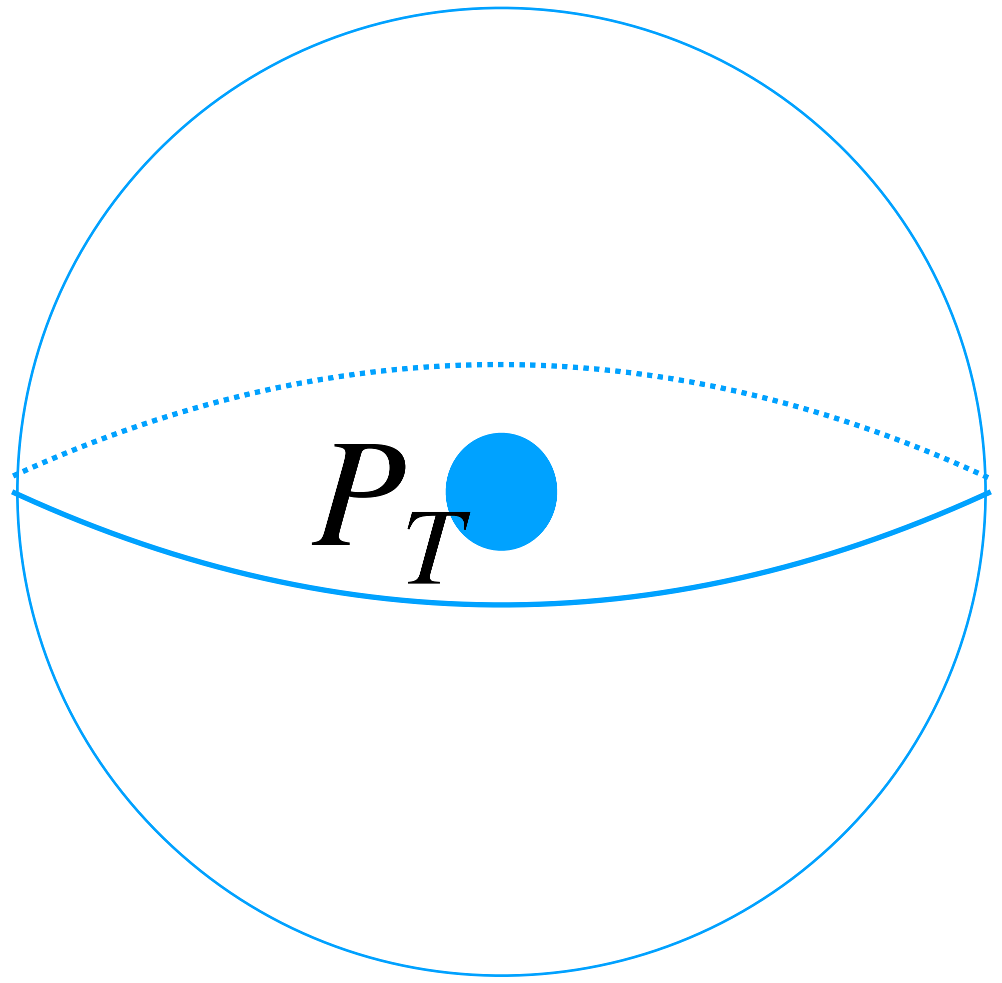


Chip antenna
Raspberry Pi 3

- Chip antennas are surface-mounted directly onto the PCB.
- Mostly omnidirectional

Antenna metrics - gain (1)

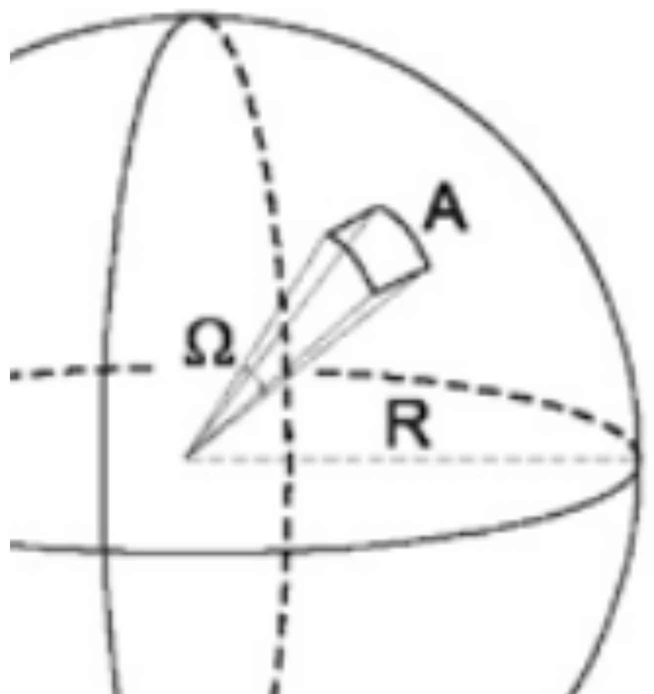
- The gain G of an antenna is a measure of how well the antenna directs its radiated power in a specific direction compared to an **isotropic** antenna.
- An isotropic antenna is an ideal antenna that radiates the signal equally in all directions.
  - This means that the gain is defined for each possible direction around the antenna, as antennas emit EM waves with different energy in different directions.
  - For this reason, we usually talk about the **maximum gain** of an antenna, which is the gain in the direction when the antenna radiates the most power (main lobe).



Isotropic antenna

Antenna metrics - gain (2)

- $G = \eta \cdot D$
  - η is the antenna's radiation **efficiency.** $\eta = \frac{P_{radiated}}{P_{input}} \times 100\,\%$, where $P_{radiated}$ is the power actually emitted by the antenna as RF energy, and P_{input} is the total power supplied to the antenna.
  - D is the **directivity**, that is the ability of an antenna to focus radiation in one direction. $D = \frac{U}{U_{iso}}$, where U is the radiation intensity of the antenna in a direction and U_{iso} is the radiation intensity of an isotropic antenna.
    - The radiation intensity is the amount of power radiated per unit solid angle.



Antenna metrics - gain (3)

- It is an adimensional metric and is usually expressed in Decibels (dB).
  - Decibels are a logarithmic unit used to measure **ratios** for several physical quantities.
  - For example, power levels are often measured in dBmW (decibel relative to 1 milliwatt):

    $$P(\text{dBmW}) = 10\log_{10}\left(\frac{P}{\text{1mW}}\right)$$

    In this case, a power of 0 dBmW means 1 mW. A power of 20 dBmW means 100 mW. A power of -10 dBmW means 0.1 mW.

Antenna metrics - gain (3)

- It is an adimensional metric and is usually expressed in Decibels (dB).
  - Decibels are a logarithmic unit used to measure **ratios** for several physical quantities.
  - So, for an antenna with gain G, we can say that the gain in decibels is $G(\text{dBi}) = 10\log_{10}(G)$
    The "i" in dBi means that the ratio is with respect to an **i**sotropic antenna.

Antenna metrics - gain (4)

- We can also express the antenna gain in terms of its **effective aperture**, or **effective area** A_e.
  - It is a constant factor of the antenna measuring how much signal power it is able to capture (larger effective area -> more signal power captured -> better reception).
  - $G = \dfrac{4\pi A_e}{\lambda^2}$, where λ is the wavelength of the signal.
  - In a **transmitting** antenna, the gain describes how well the antenna converts input power into radio waves headed in a specified direction.
  - In a **receiving** antenna, the gain describes how well the antenna converts radio waves arriving from a specified direction into electrical power.

Antenna metrics - bandwidth (1)

- The **Bandwidth** of an antenna is the range of frequencies over which the antenna can effectively transmit or receive signals.

Bandwidth of a channel = difference between the highest frequency and the lowest frequency that can be transmitted through the channel = size of the frequency range.

- **Wideband** antennas can transmit more data by using a larger portion of the frequency spectrum (more information can be transmitted simultaneously).
- A **narrowband** antenna can only transmit data within a small part of the frequency spectrum (limited amount of data can be sent).
  - Are large bandwidth antennas better than low bandwidth antennas?
    - It very much depends on the application

Antenna metrics - bandwidth (2)

Amphenol RF
ST0428-31-002-A

- **Wide range** cellular outdoor antenna, supporting 2G/…/5G.
- Bandwidth: [617 MHz, 5.925 GHz].
- Good if your system has multiple devices transmitting/receiving at a wide range of frequencies.

- But in some cases, you only need for instance Bluetooth communication, an antenna with a **small bandwidth** around 2.4 GHz is enough.
  - Narrowband antennas are less prone to interferences with other devices and antennas.
  - Nevertheless, frequencies can change over time and regions, so the bandwidth should not be too narrow.

Directionality

OMNIDIRECTIONAL

- Used to send EM signals to receivers whose location is not fixed.
- Easier to install
- Prone to interferences
- Lower gain and directivity

DIRECTIONAL

- Used to send EM signals to fixed receivers (satellite TVs, space craft)
- Harder to install
- Less interferences
- Higher gain and directivity

5.1.3 Signal Propagation

Wireless Channels

1. Transmit antenna

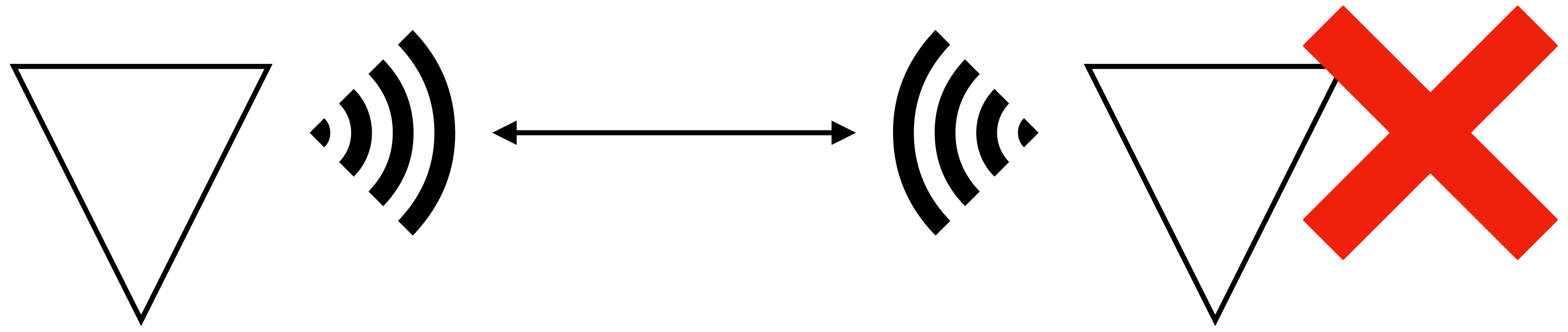
3. Receive antenna

2. Wireless signal propagation

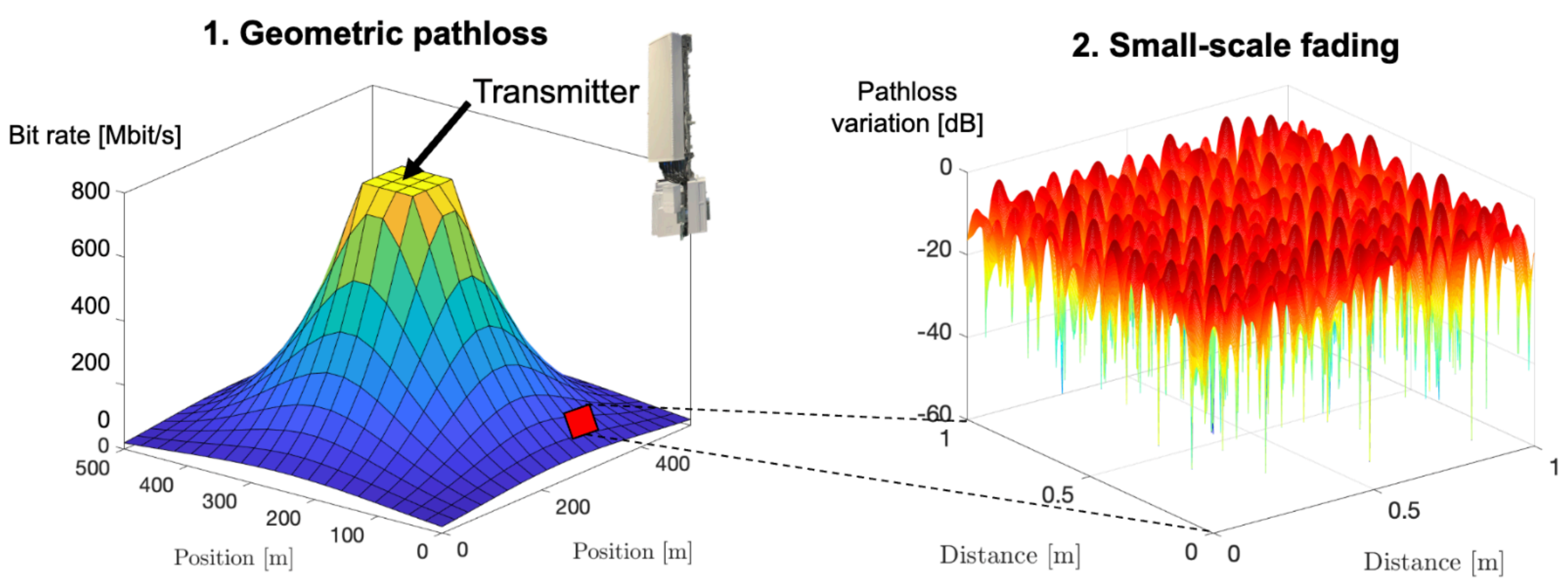
They all determine the channel quality!

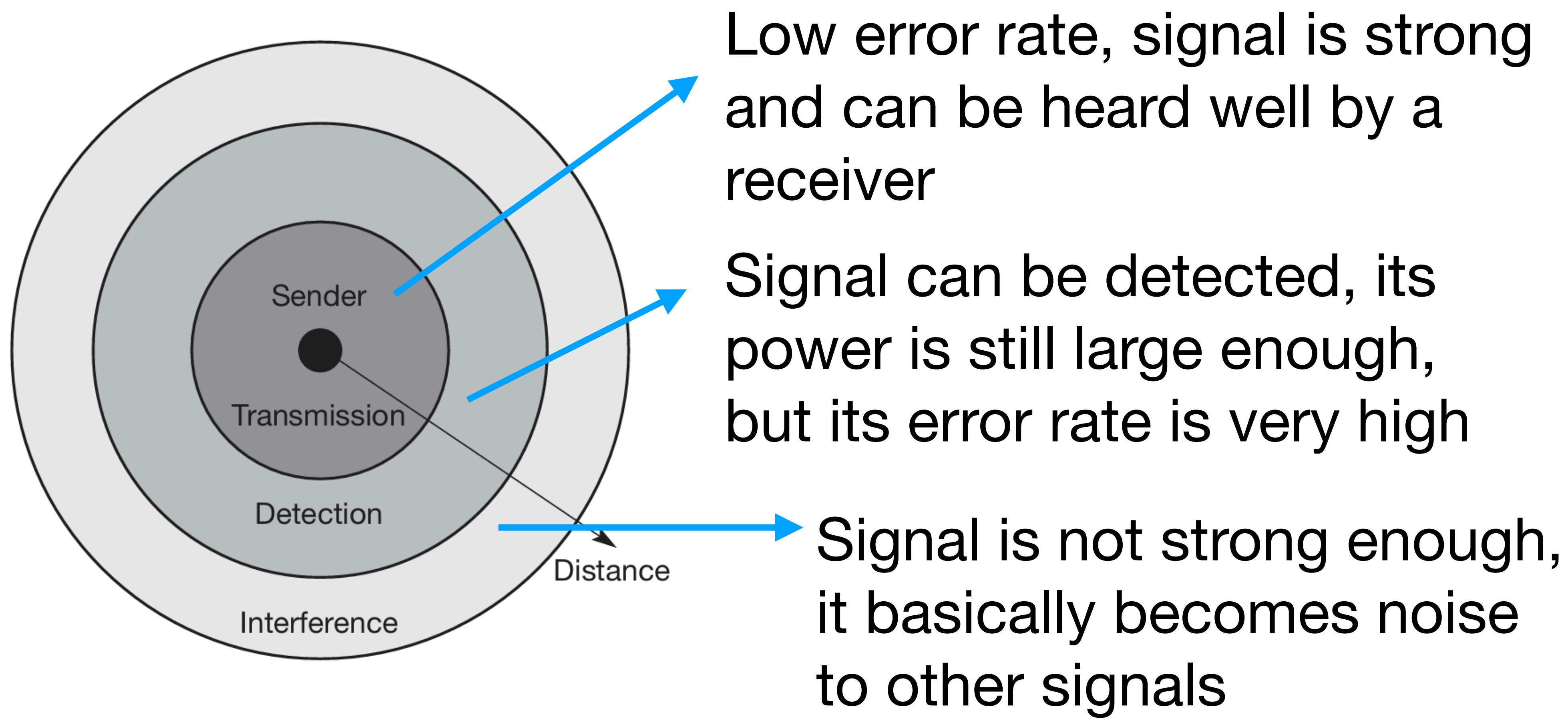
Signal Propagation

- What happens to radio waves between antennas?
- The signal does not travel in a straight line.



- Radio waves lose energy as they propagate (**path loss**) by going through air.



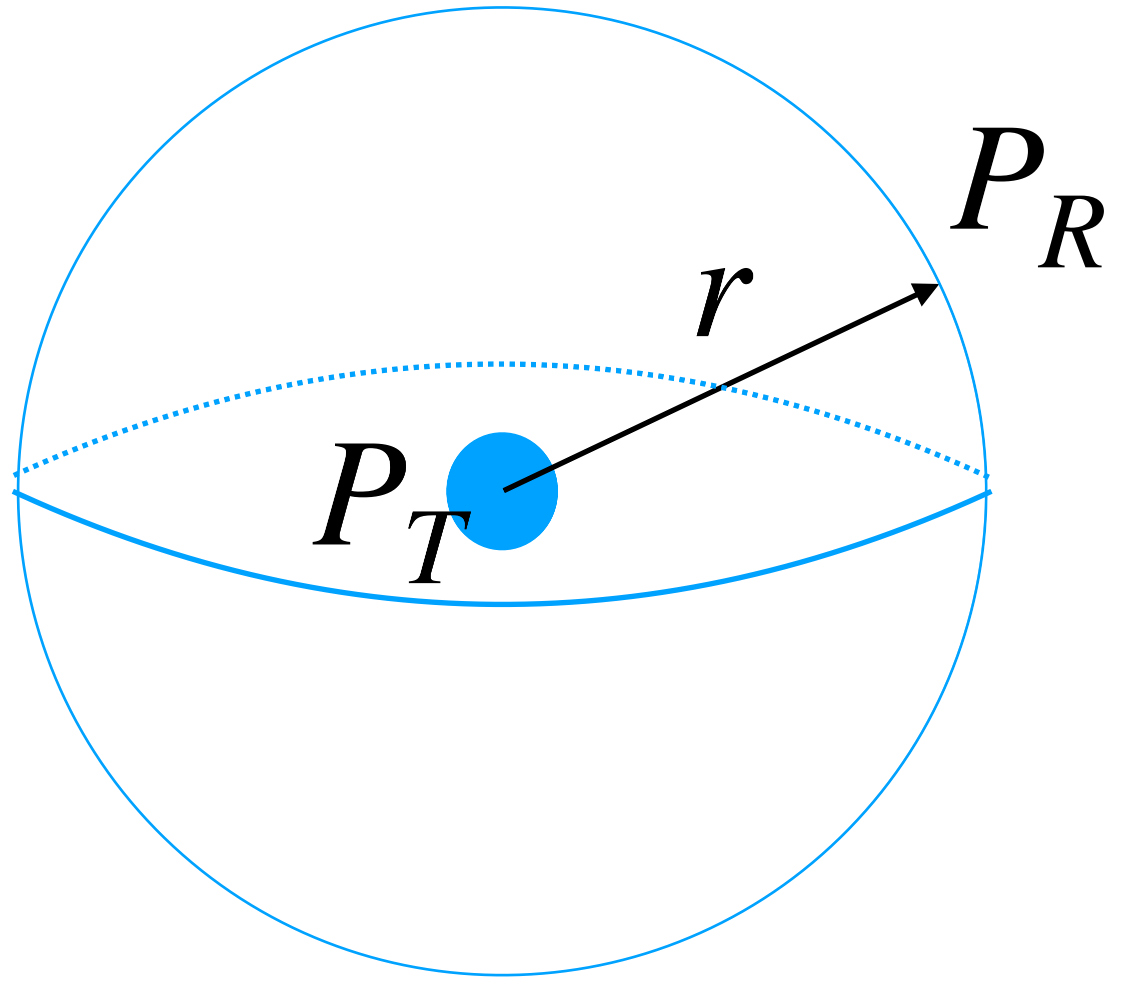


- Radio waves get **deflected** or **reflected** when they hit objects.
- They bounce off things and then recombine (**multipath** or **fading**).

Free space propagation (1)

- Let's considering a simplified model, the **free space propagation model**.
- Consider an ideal point source of EM waves which emanates power P_T equally in all directions (i.e., an **isotropic** antenna).
- Energy emanated at the source expands out in a spherical way. A receiver on the sphere at distance r receives power:

  $P_R = P_T \frac{A_e}{4\pi r^2} = \frac{P_T}{L}$, where $L = \frac{4\pi r^2}{A_e}$ is the **free space pathloss**. Notice that this depends on the effective area of the receiver and that signal decreases proportionally **to the square of the distance** to the sender.



  - Suppose that the power transmitted from the transmitter is 100 Watt. If the effective area of the receiver is 1/100 of the surface of the area around the transmitter, the power received by the receiver is 1 Watt.
- We loose a lot of power even in free space ideal scenario

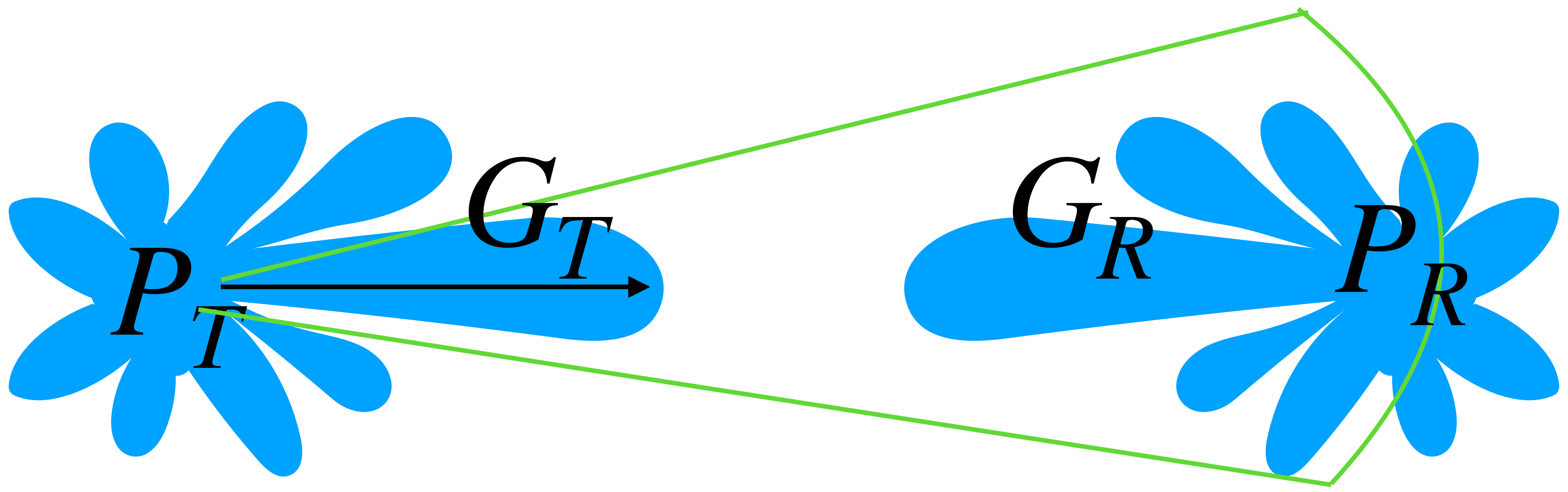
Free space propagation (2)

- We can express the free space propagation formula also in terms of the gains of the antennas.

- If the transmitter and receiver antennas have gain G_T, G_R respectively, then the power at the receiver is $\boxed{P_R = P_T G_T G_R \left(\frac{\lambda}{4\pi r}\right)^2,}$

  Friis transmission formula

  λ= wavelength.

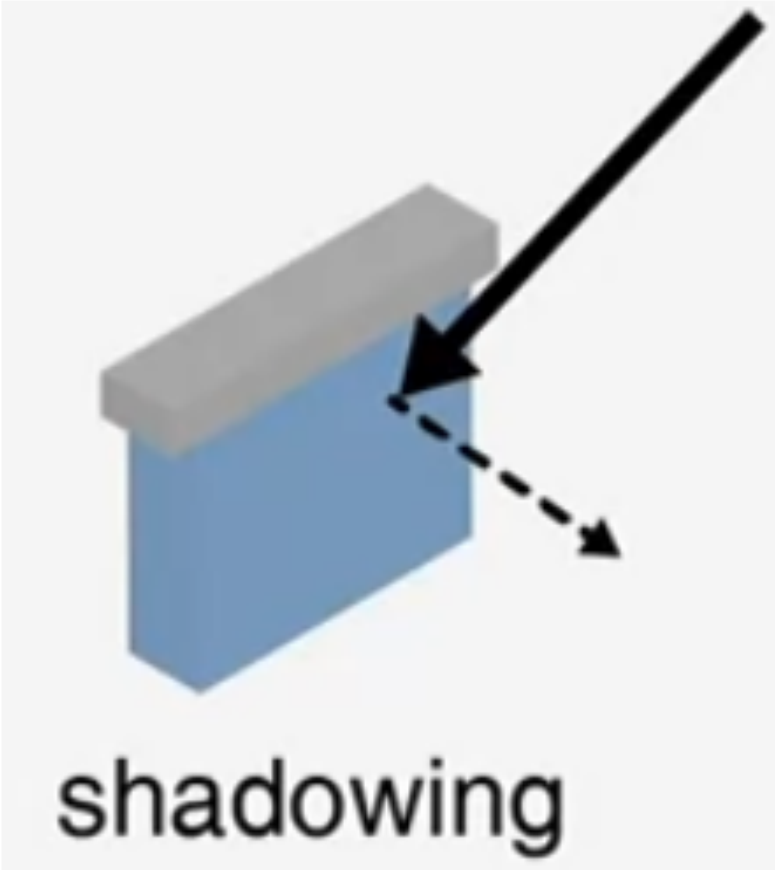


- The free space propagation model is useful for gaining intuition on how much signal strength you should expect between two points in a perfect **idealized** environment.

Path Loss Effects

- In reality, the space between the transmitter and the receiver is not free.
- EM waves **interact with objects**. This can be bad (signal degradation) or good (can leverage some effects to improve the signal)
- The interactions can produce different kinds of effects.

Path Loss Effects: shadowing



- Shadowing is the effect that causes the fluctuation of a signal's power due to objects obstructing the propagation path.

- Shadowing losses depend on the frequency of the EM wave: EM Waves can penetrate through various surfaces but at the cost of loss in power i.e **signal attenuation**.

  - The losses depend on the type of the surface and frequency of the signal.

Path Loss Effects: reflection

- Reflection: similarly to light hitting a mirror, a signal can be reflected when hitting an object.

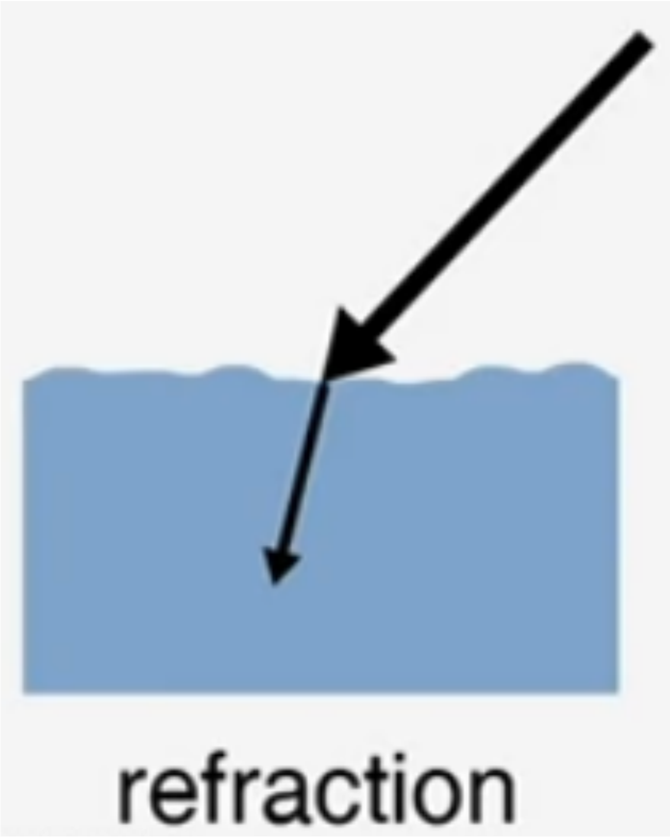
- This can be good: as we have seen, in directional antennas reflectors are used to amplify the signal and focus it in the desired direction.

- Reflection effects depend on the means that the signal hits. Some materials have good reflection properties (e.g., mirrors, some metals).

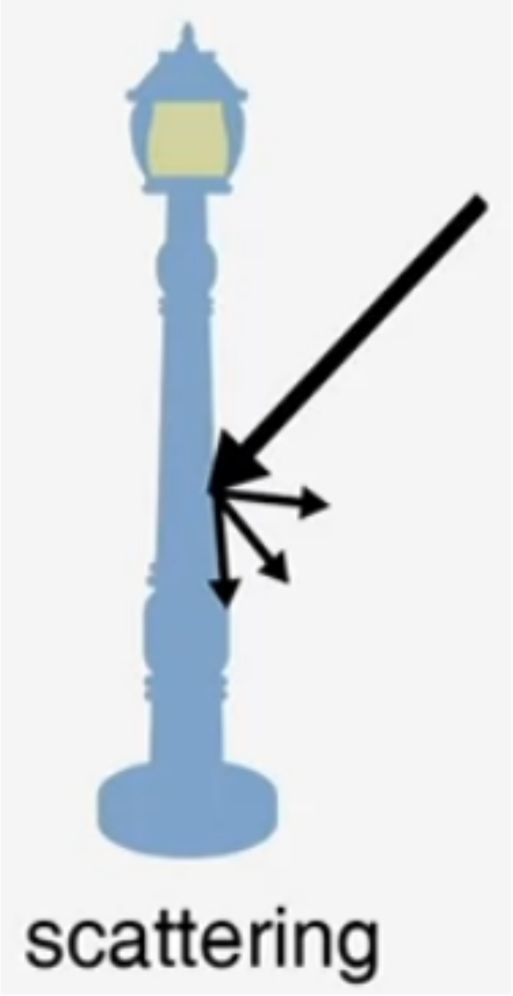
Path Loss Effects: refraction

- Refraction is the change in direction of a wave passing from one medium to another.

- As light going through water is bent, EM waves can change their direction after passing through some means, e.g., glass or drywall.



Path Loss Effects: scattering

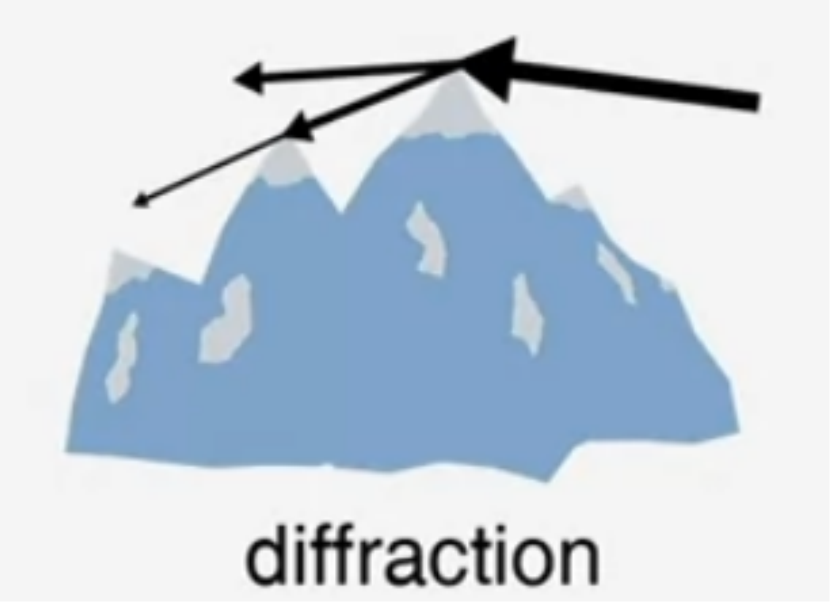


- Scattering is the reflection of the signal in multiples directions when hitting an uneven surface.
- While some material (e.g., mirrors) are very good at reflecting EM waves, some others are not.

  For example, think of light hitting a white wall. The wall appears brighter and light does not bounce back to your eyes perfectly, but it is rather scattered around the wall.

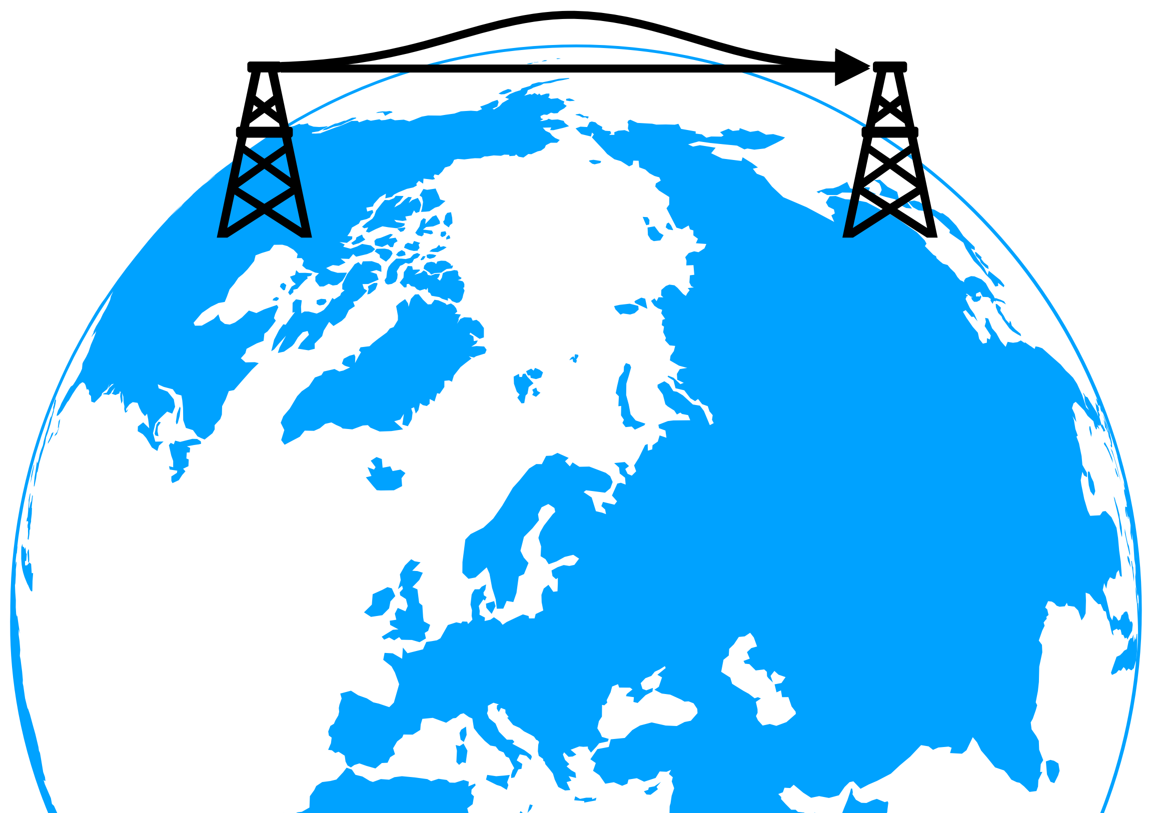
Path Loss Effects: diffraction

- Diffraction is the interference or bending of waves around obstacles or through an aperture into the region.



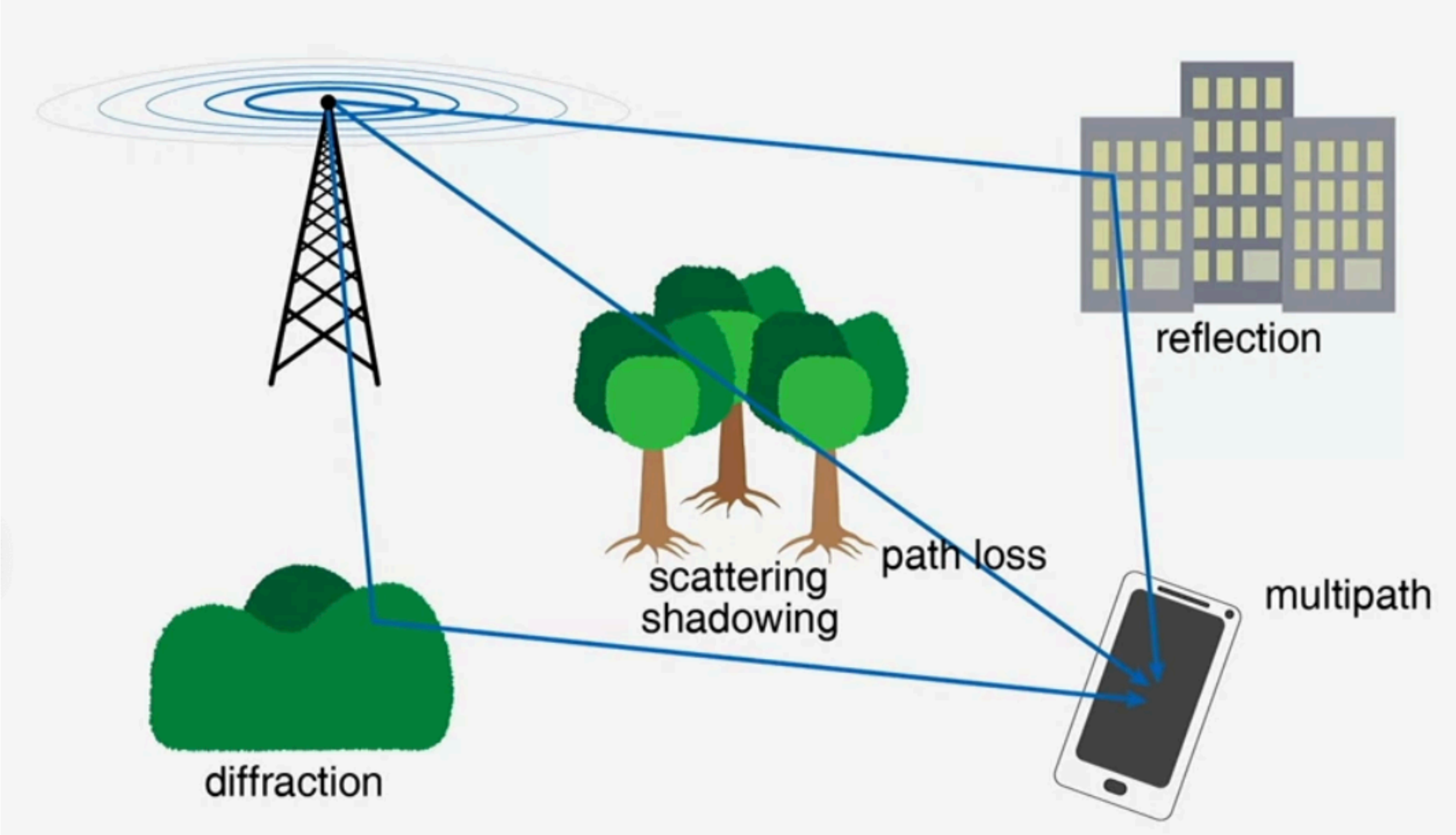
- Does not happen much with high frequency signals (e.g., light).

- Low frequency signals can bend over buildings, trees, and even around the curvature of the Earth.
  That is why we are able to receive AM radio signals around the curve of the Earth with low enough frequencies



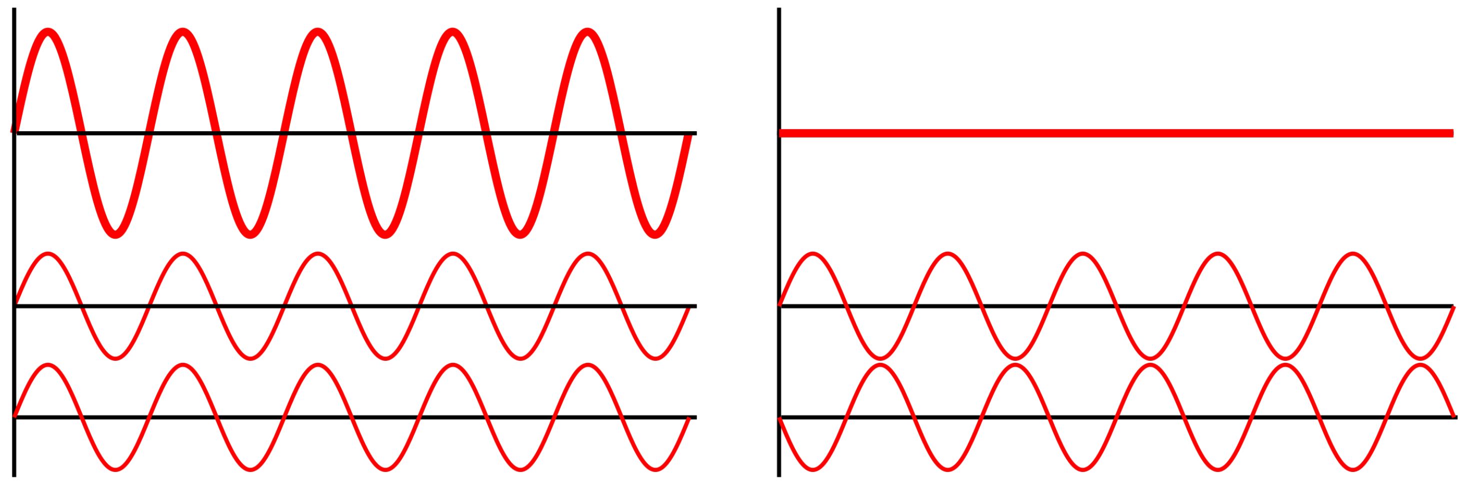
Path Loss Effects

- In real life, the way signals propagate very much depends on the environment.
- Signal transmitted by an antenna will travel around, be scattered, diffracted, reflected before reaching an IoT device.
- The signal received is a **multi-path**, composed of several signals that are then added together.



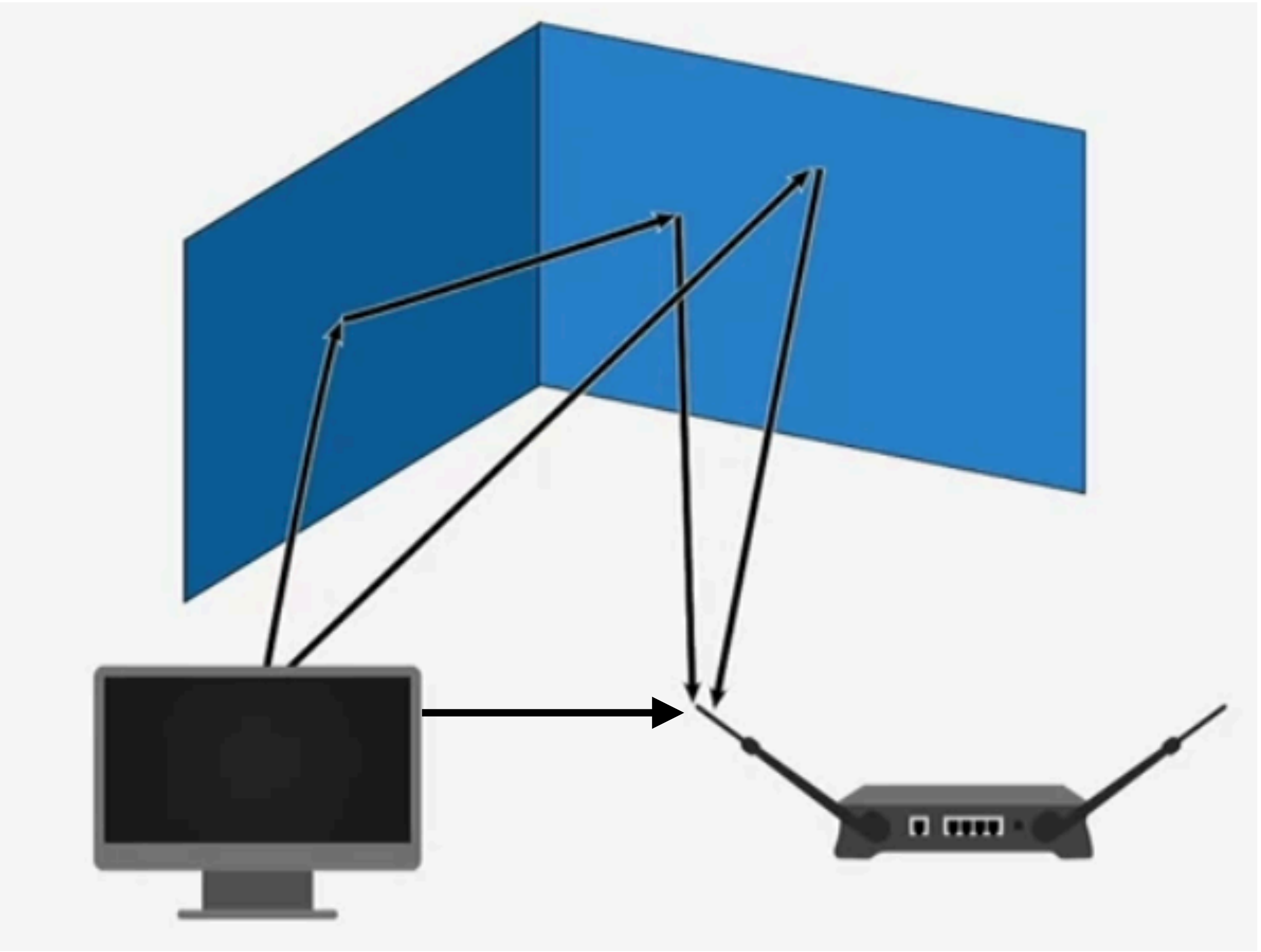
Wireless signals propagate like waves (1)

- Matching waves construct, opposite wavelengths destruct



- When two waves interfere, the resultant wave may have greater amplitude (***constructive interference***) or lower amplitude (***destructive interference***) if the two waves are in phase or out of phase, respectively.

- When you have a transmitter and a receiver, the receiver does not only get the signal from the direct path from the transmitter (i.e., **line of sight**), but also all the signal bouncing around coming from the transmitter and reaching the receiver (multi path).



Wireless signals propagate like waves (2)

- Frequency of the EM waves can have massive effect on how they propagate.
  - An IoT designer has to make decisions on what protocol to use, i.e., on what frequency to use between their devices.

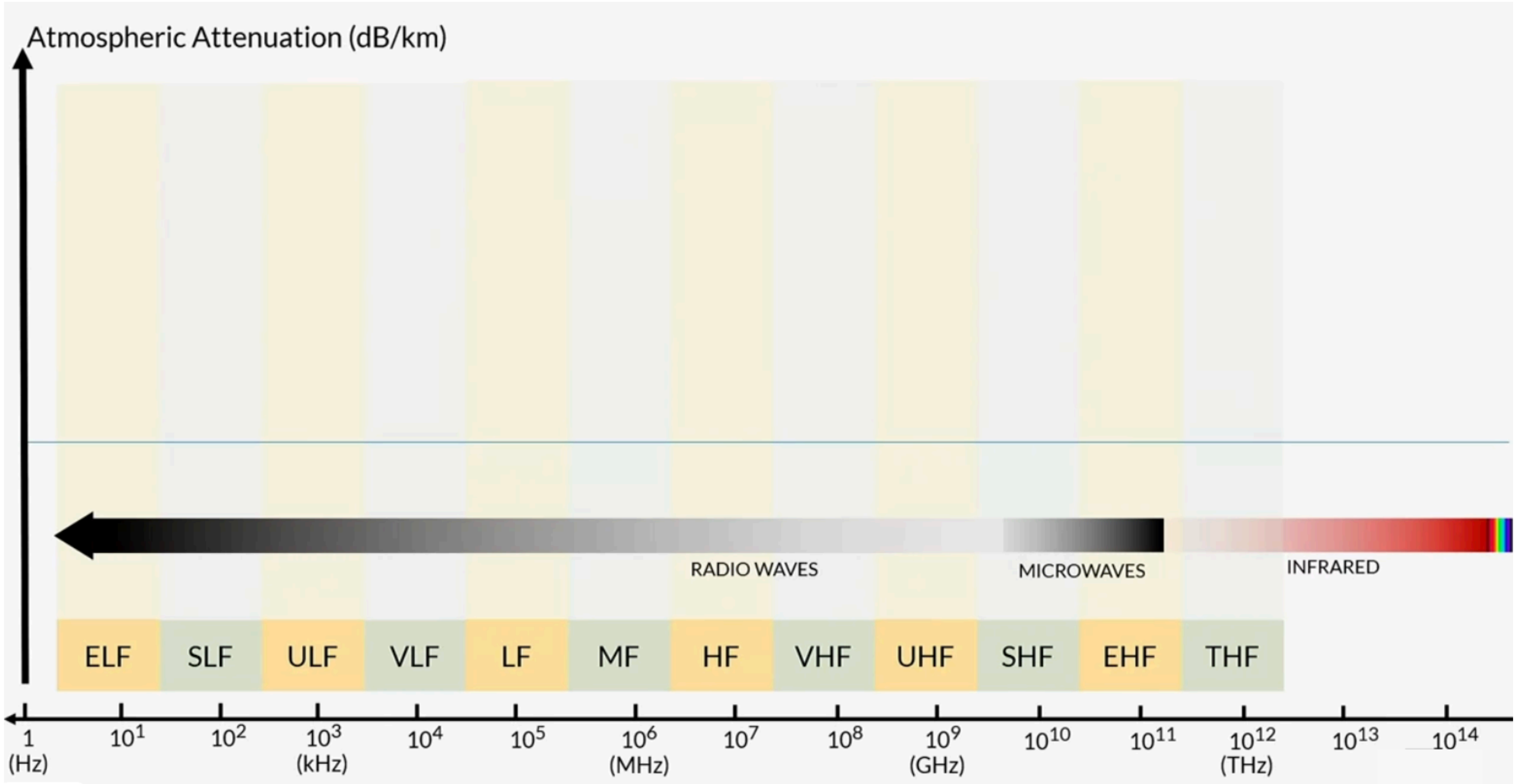
- Low Frequencies/Long wavelengths:

✅ Diffract around objects, can follow curvature of the Earth
✅ Can penetrate through objects
❌ Require bigger antennas

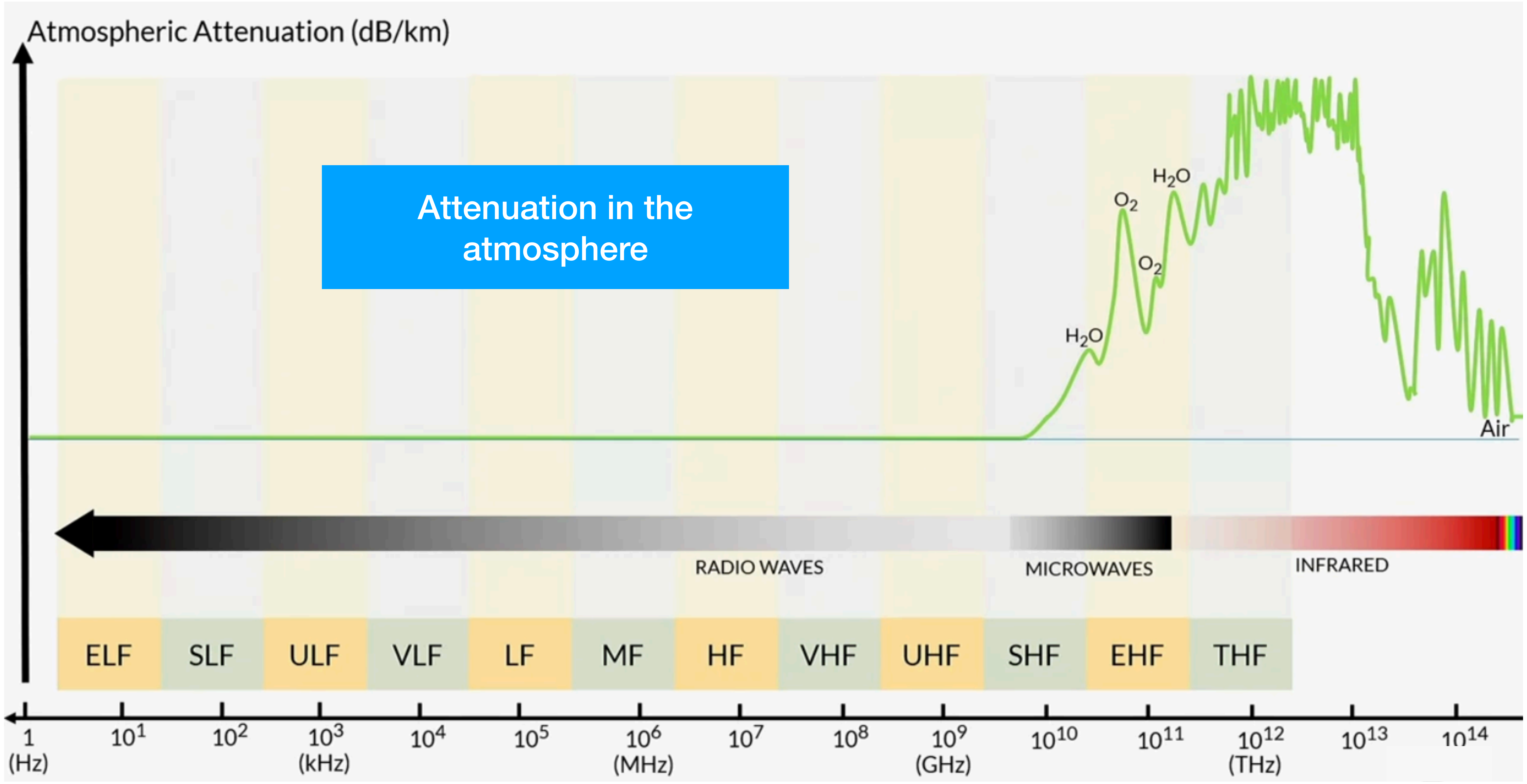
- High frequencies/short wavelengths:

✅ act more like light, can be more easily focused (parabolic antennas)
❌ attenuate more in air; increased scattering by objects and water (rain)

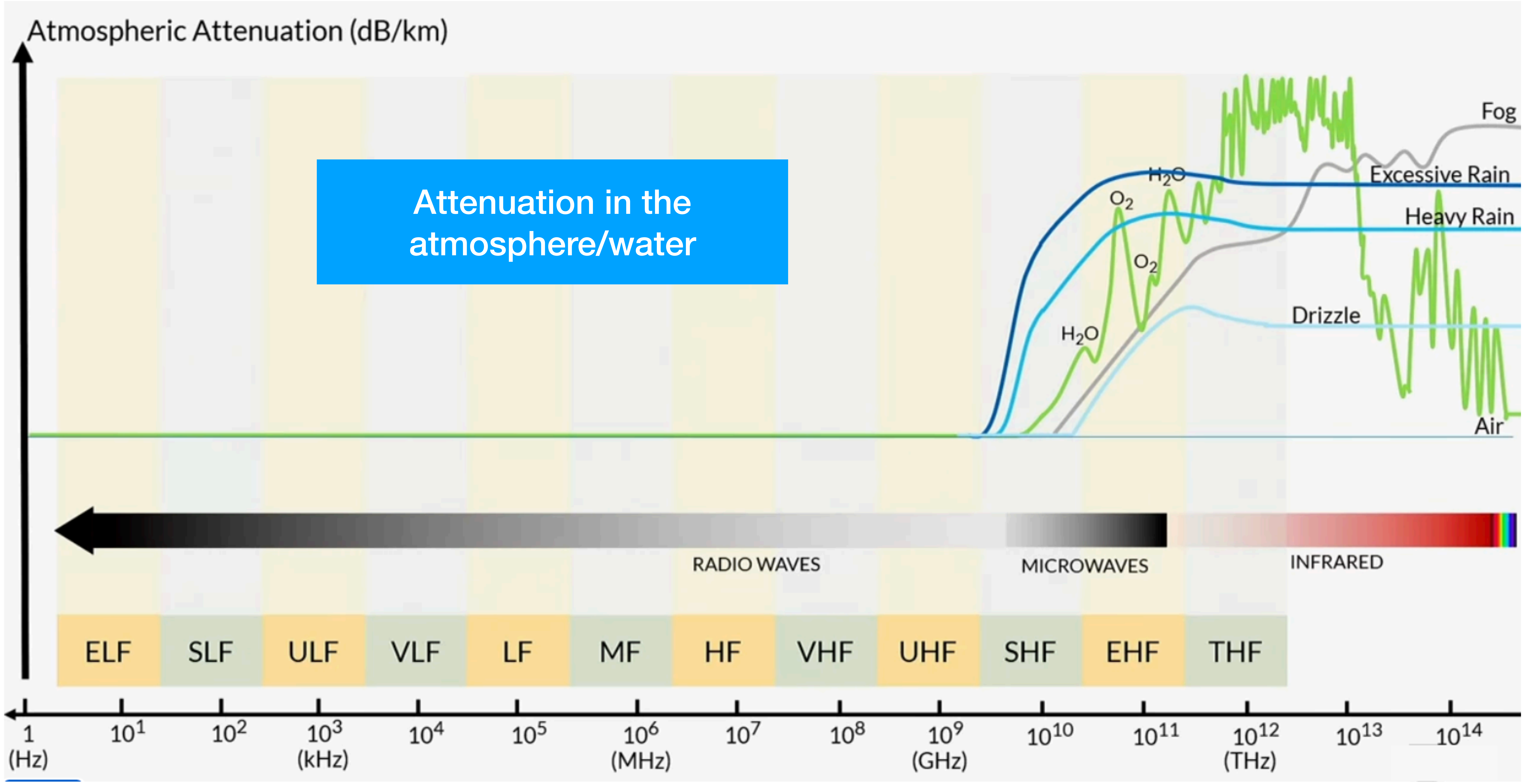
5.1.4. More about attenuation



- Ranges of low frequency waves are further split and have different names.
- Their interaction with air/water is very changeling.

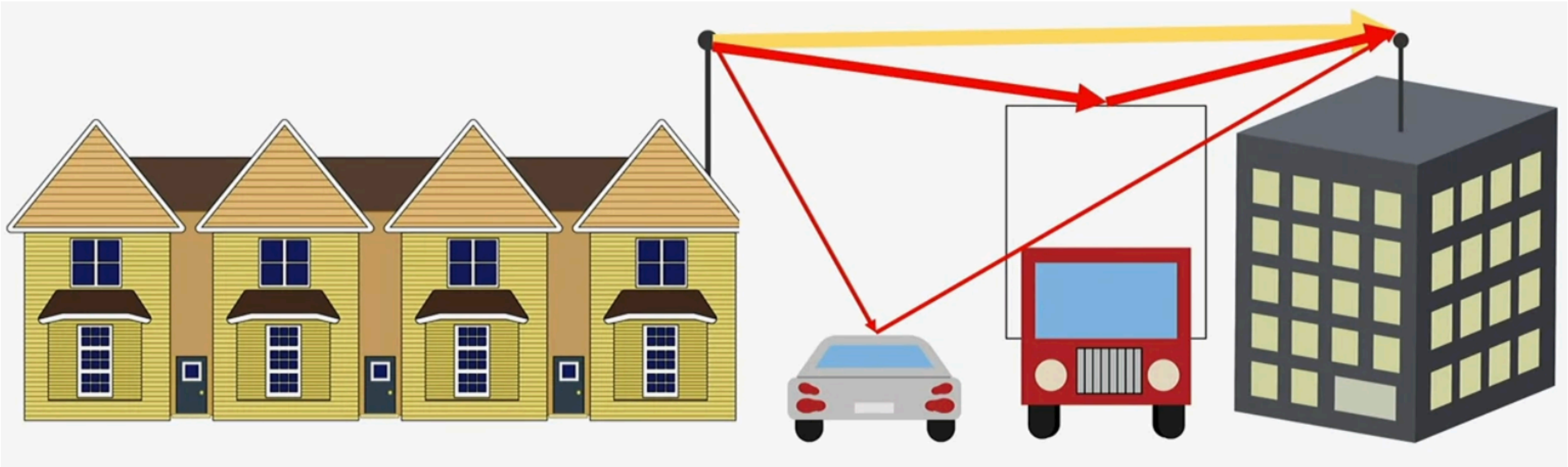


- Peaks and valleys of the attenuation are due to wave vibrations matching or mismatching the vibrational frequencies of the atmosphere.



- The composition of the air can have a big effect on signal propagation. For example, heavy rain tends to attenuate over a larger range of frequencies. If you have a lighter rain like a drizzle, it looks like the lower line, attenuation is weaker. Fog has a strong attenuation impact, also visible light is highly attenuated.

Interference changes with distance from line of sight path



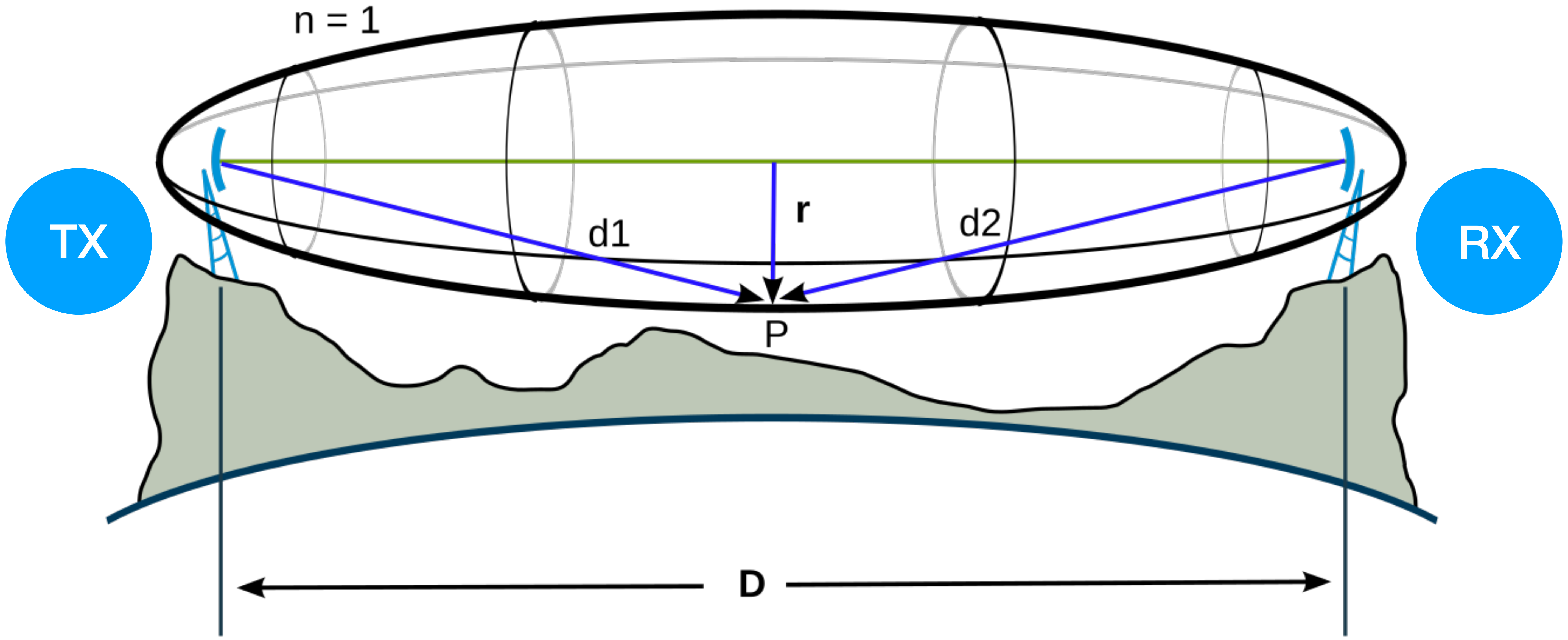
- As we have seen, even directional antennas actually propagate in all direction.
- The signal is transmitted to the receiver and it is stronger along the **line of sight (LOS)** and weaker as it gets farther from it.
- Signal gets reflected, diffracted, refracted etc on objects standing between the transmitter and receiver antennas, creating a **multi path effect**

Fresnel Zones (1)

- Fresnel zones are ellipsoids between transmitter TX and receiver RX.

- The first zone is the smallest.
  The $n-$th zone has equation:

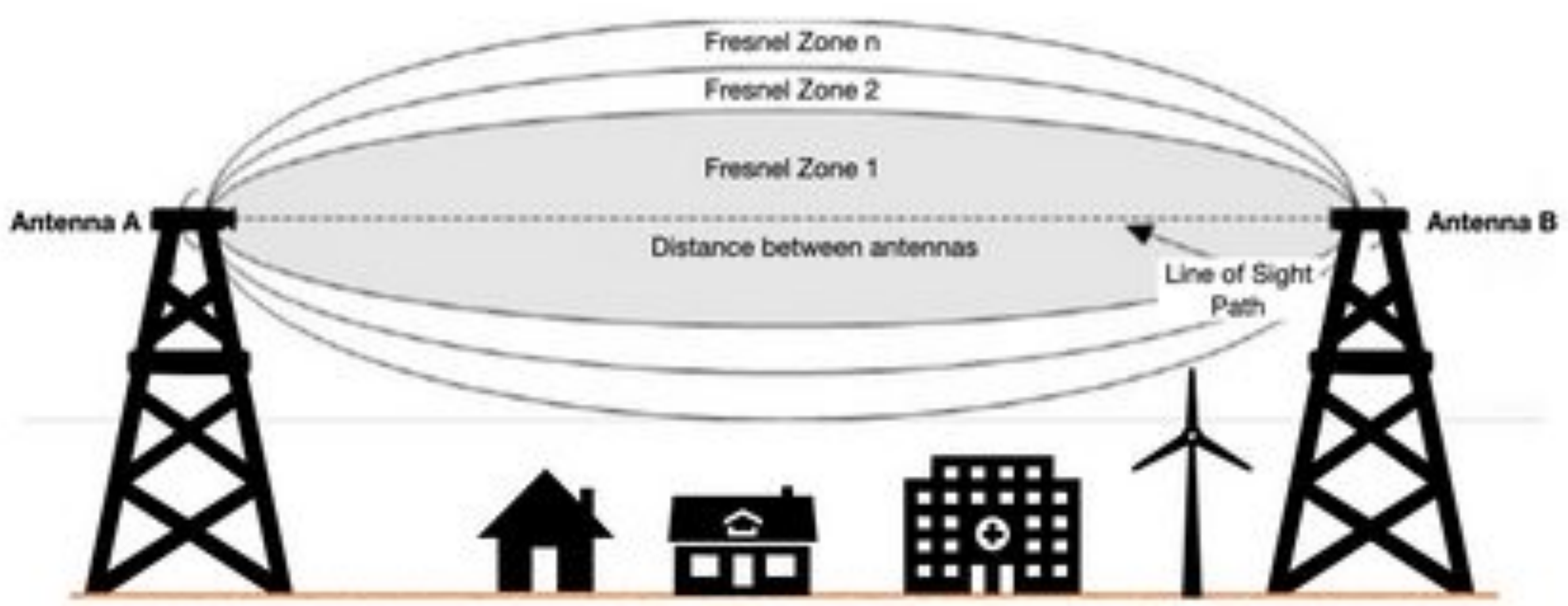
$$r_n = \sqrt{\frac{n\lambda d_1 d_2}{d_1 + d_2}}$$

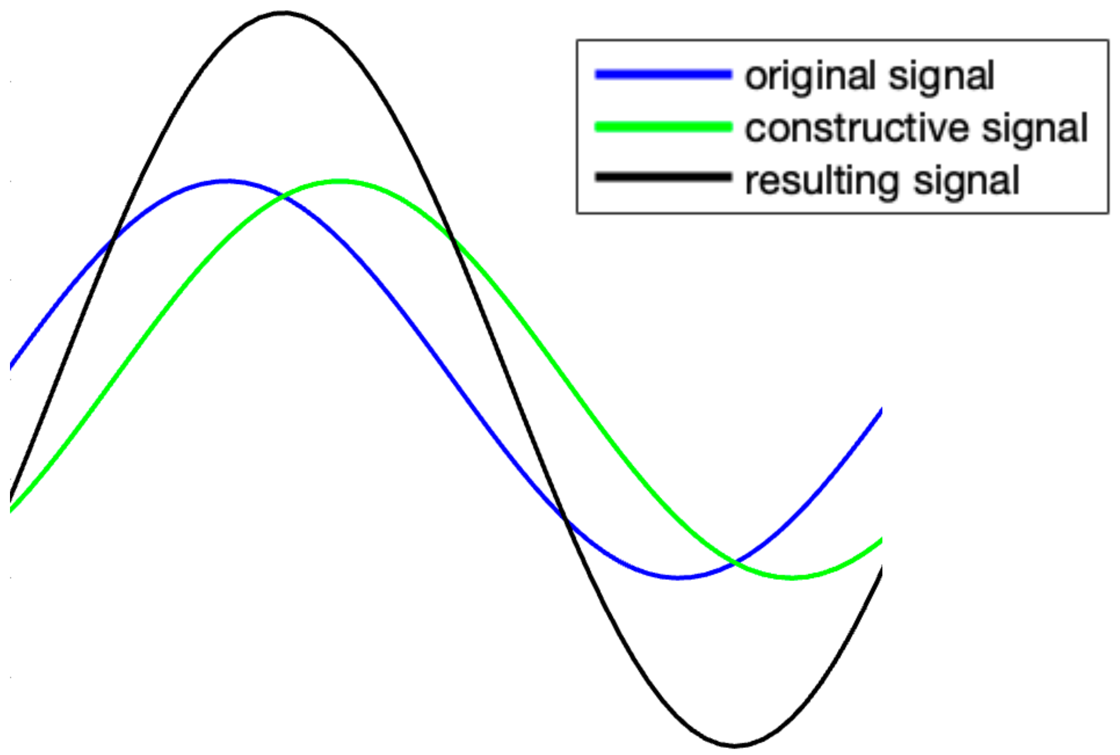


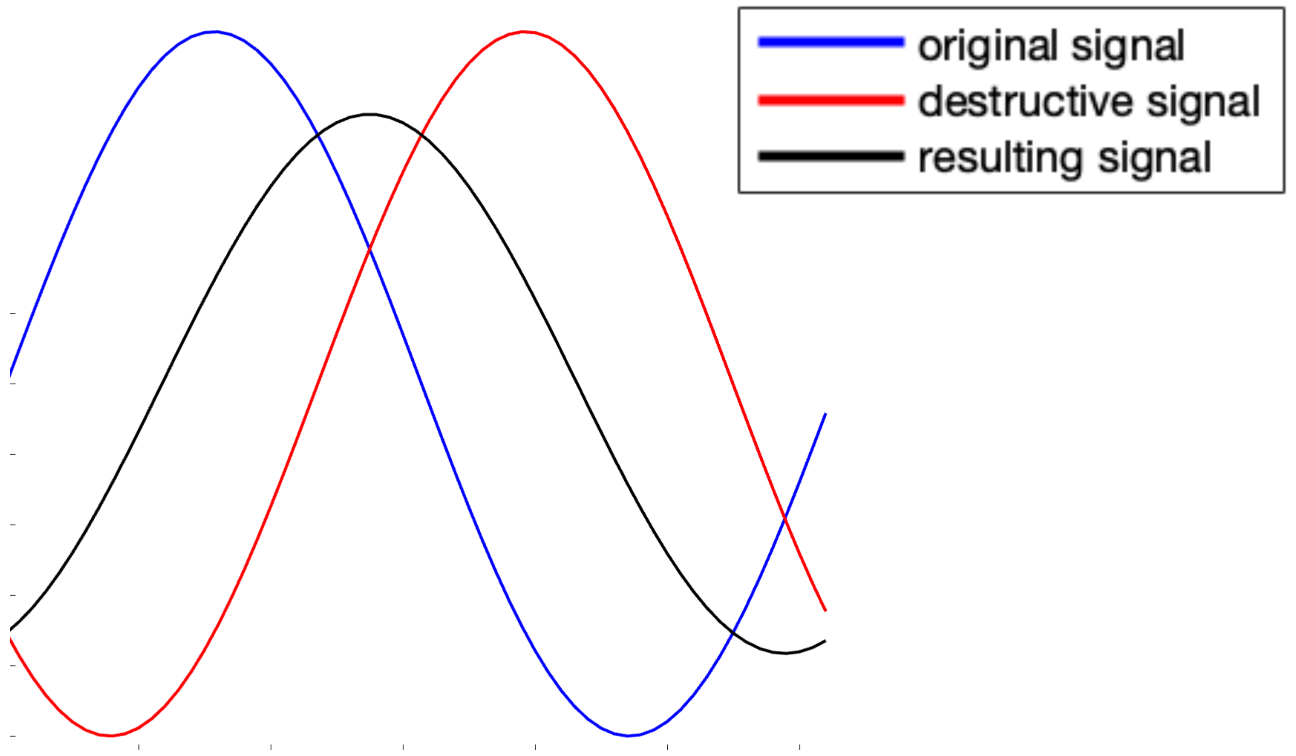
- r_n = ray of the n-th Fresnel zone at point P, λ = wavelength, d_1, d_2 distance of the generic point P to the TX and RX.

Fresnel Zones (2)

- Objects reflecting the signal in the odd zones have a disruptive effect, while objects reflecting the signal in the even zones have a constructive effect







- It can be proved that interferences with objects in the first zone create waves out of phase by $\theta \in [0,\pi]$ objects in the second zone create waves out of phase by $\theta \in [\pi,2\pi]$, etc.

- Usually we are interested only in the first three zones, since the signal is very weakened outside the third Fresnel zone.

Is attenuation good or bad? (1)

- In some environments, there are many nodes that need to transmit. For instance, this building or any office.

- Consider the example of an office with two access points (APs)

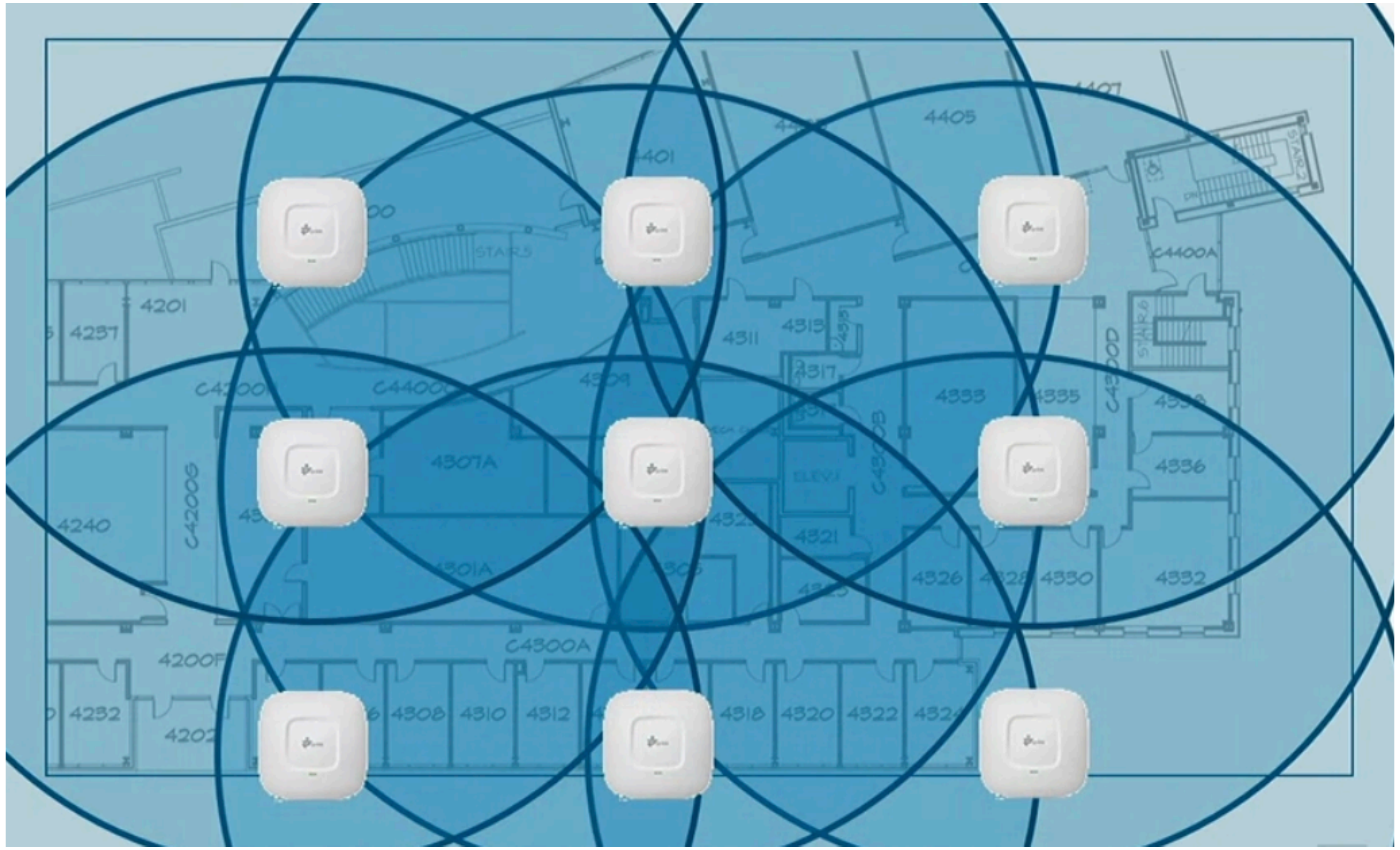
- What if there is not enough capacity? i.e., what if the two access points cannot keep up with all the transmissions in the office?

Is attenuation good or bad? (2)

- One possible solution: deploy more access points!

- This is not a good solution, as they are going to interfere with one another, causing many collisions.

- Counter intuitive issue: increased transmitter density can lower available bandwidth.

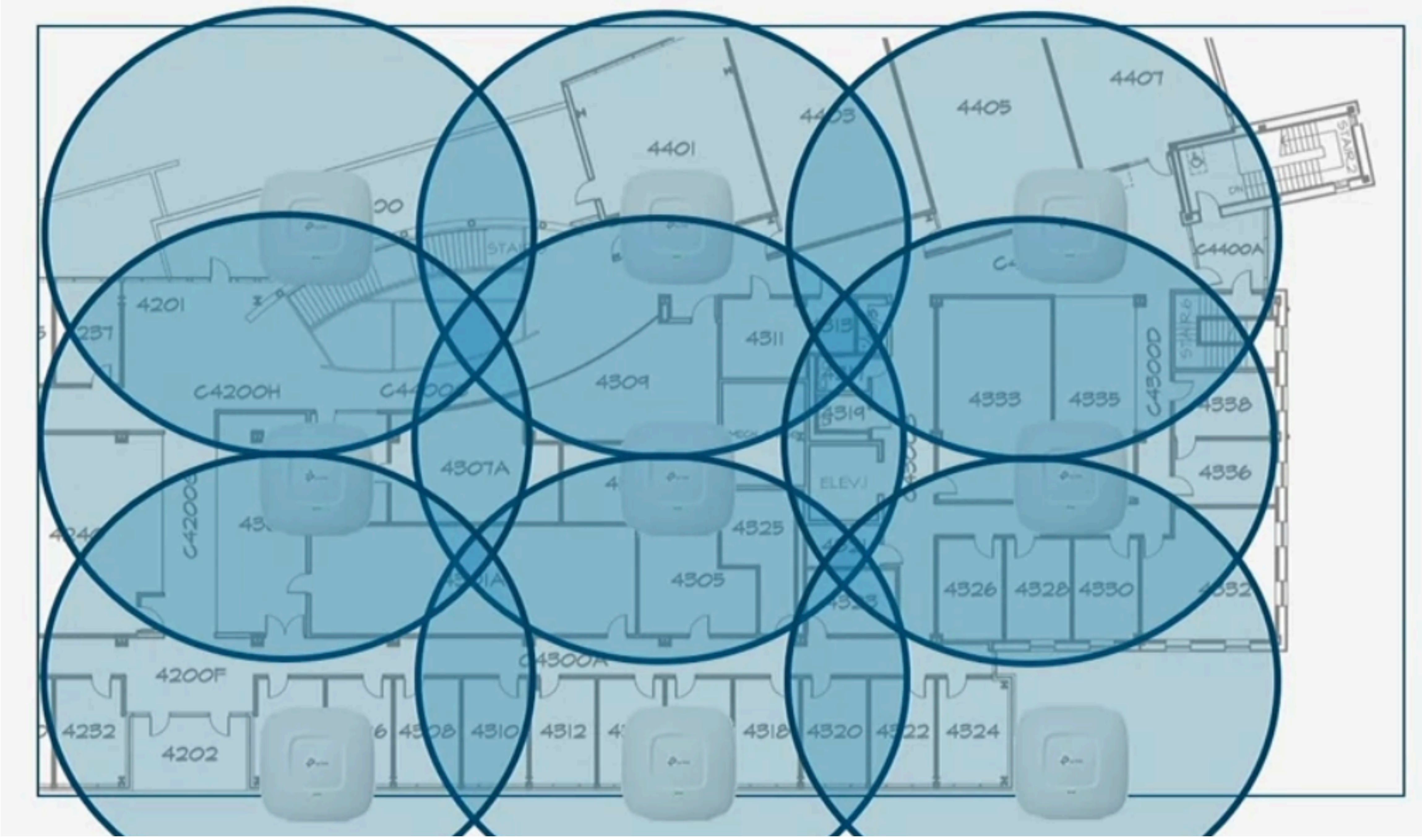


Is attenuation good or bad? (3)

- In dense environments, more attenuation is good!

- In this example, more attenuation is caused by walls and objects in the office.

- More attenuation means that the APs' transmission range gets smaller, leading to less interference between each other.

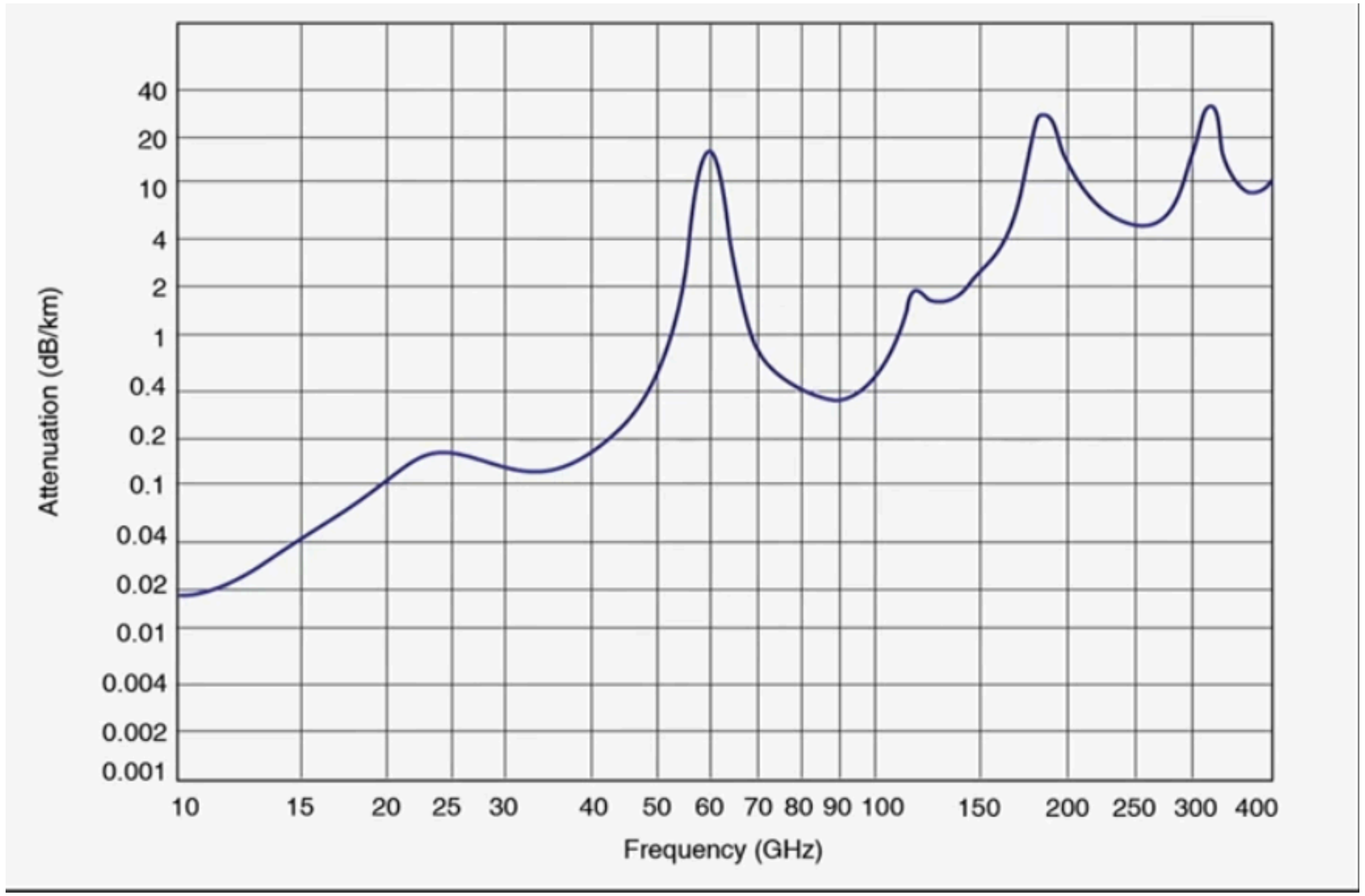


Atmospheric attenuation in wifi spectrum

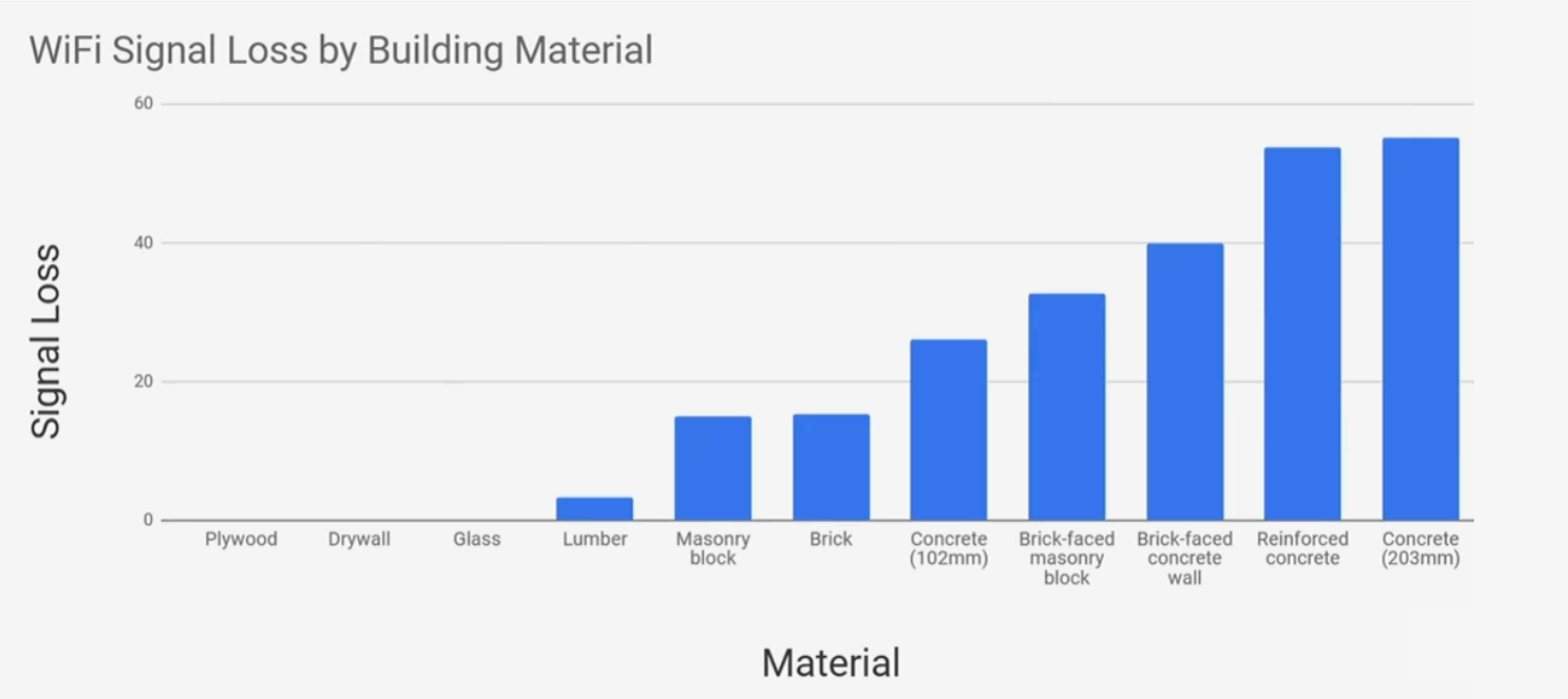
- The figure shows a zoom of the figure shown some slides ago.

- There are some peaks in the attenuation when varying frequencies, e.g., there is a peak at 60 GHz.

- 60 GHz is often chosen by engineers for transmitting in dense environments.



How physical objects affect wireless signals (2.4GHz)



Water attenuation

- 2.4GHz attenuation in water to 1/3 signal strength:
  - pure distilled water: 8.01 km (low attenuation)
  - drinking water (contains calciums and salts): 44.54 m
  - sea water (has a lot of salt in it) 8.91 mm

- Humans are more similar to sea water.

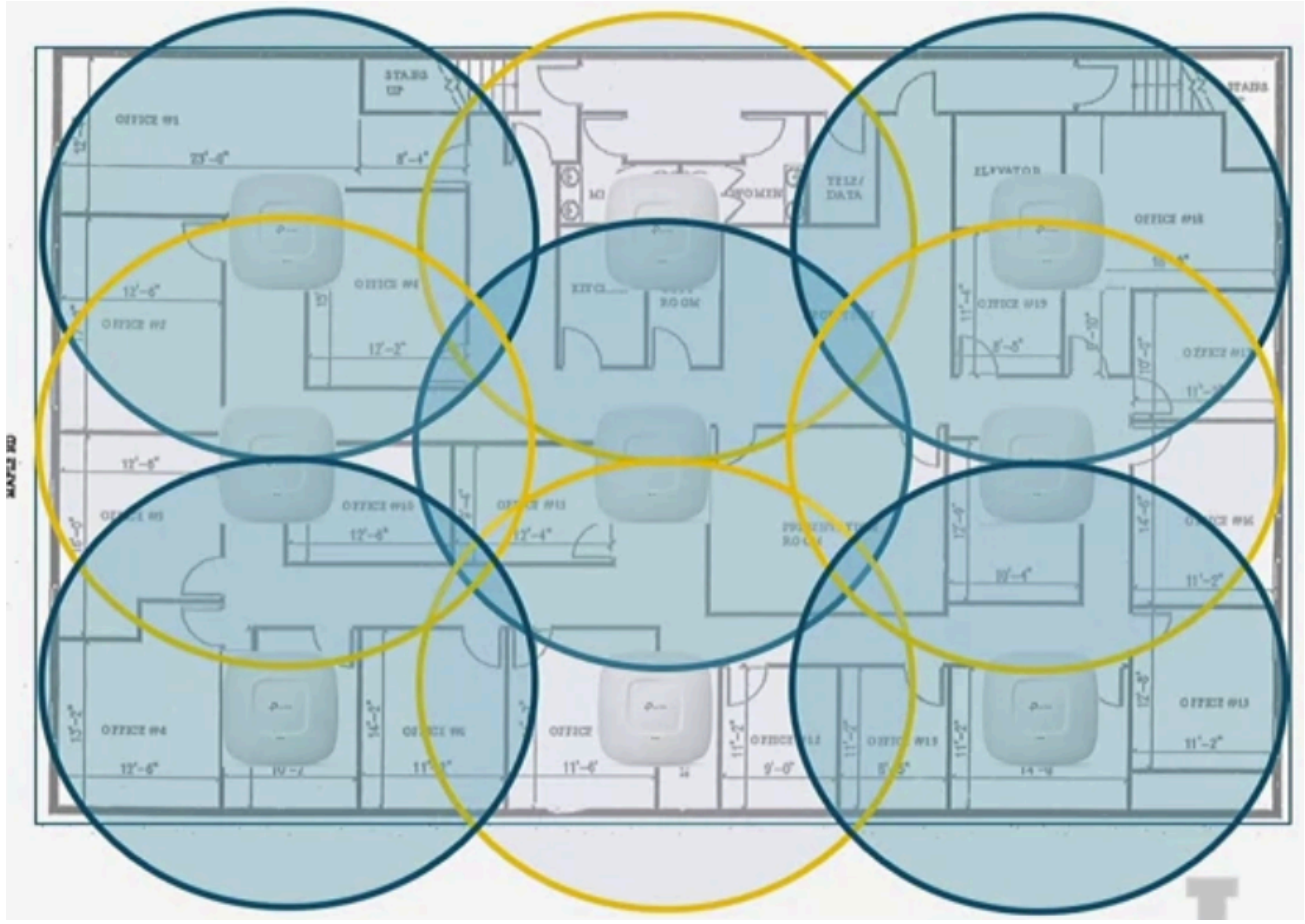
- Elements can store water, like trees and soil. This is the reason why some outdoor environments can work perfectly fine but the signal strength will drop drastically if it starts raining.

Spectrum Division

- Back to the overcrowded office scenario. A possible solution for guaranteeing enough bandwidth to all nodes while avoiding collision is **spectrum division.**

- Many standards (e.g., 802.11) support different channels, i.e., different wavelengths, to avoid collision.

- How to deploy APs and to set up wavelengths for transmitting/ receiving in an efficient way?



Vertex Colouring Problem (background)

- Decisional problem: Given a graph G and M different colours, is it possible to colour each vertex of G with one of the M colours such that no two neighbour nodes have the same colour? (NP-complete)

- Optimisation problem: Given a graph G, what is the minimum number of colours that we must use so that each vertex is coloured and its colour is different from the colours of its neighbours? (NP-hard)

- Minimum number of colours for the vertex colouring problem is called the **chromatic number of $G, \chi(G)$.**

Vertex Colouring Problem

- Create a graph whose vertices are APs. An edge between two nodes exists if the two corresponding APs are within radio range with each other. Colours represent different channel frequencies.

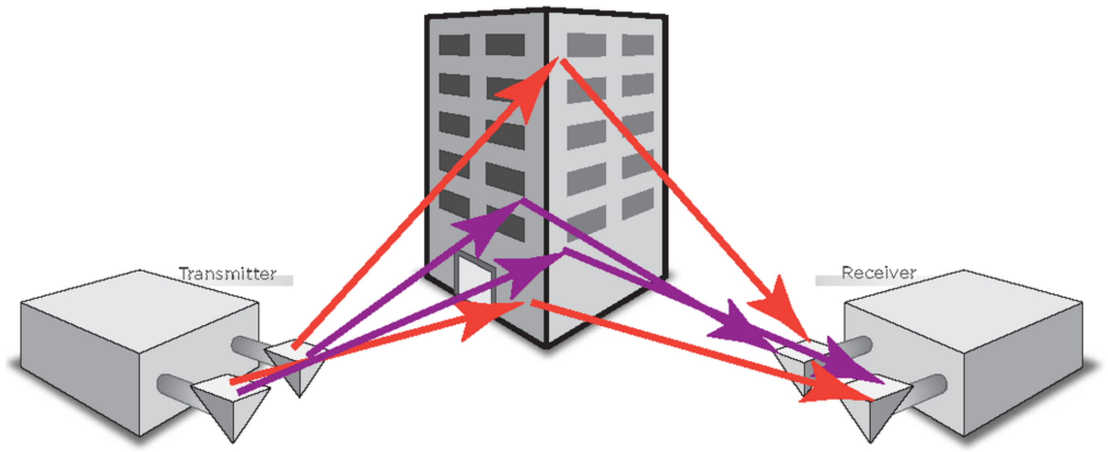
- We can also create a weighted graph where edge weights are the percentage of overlap of two APs.

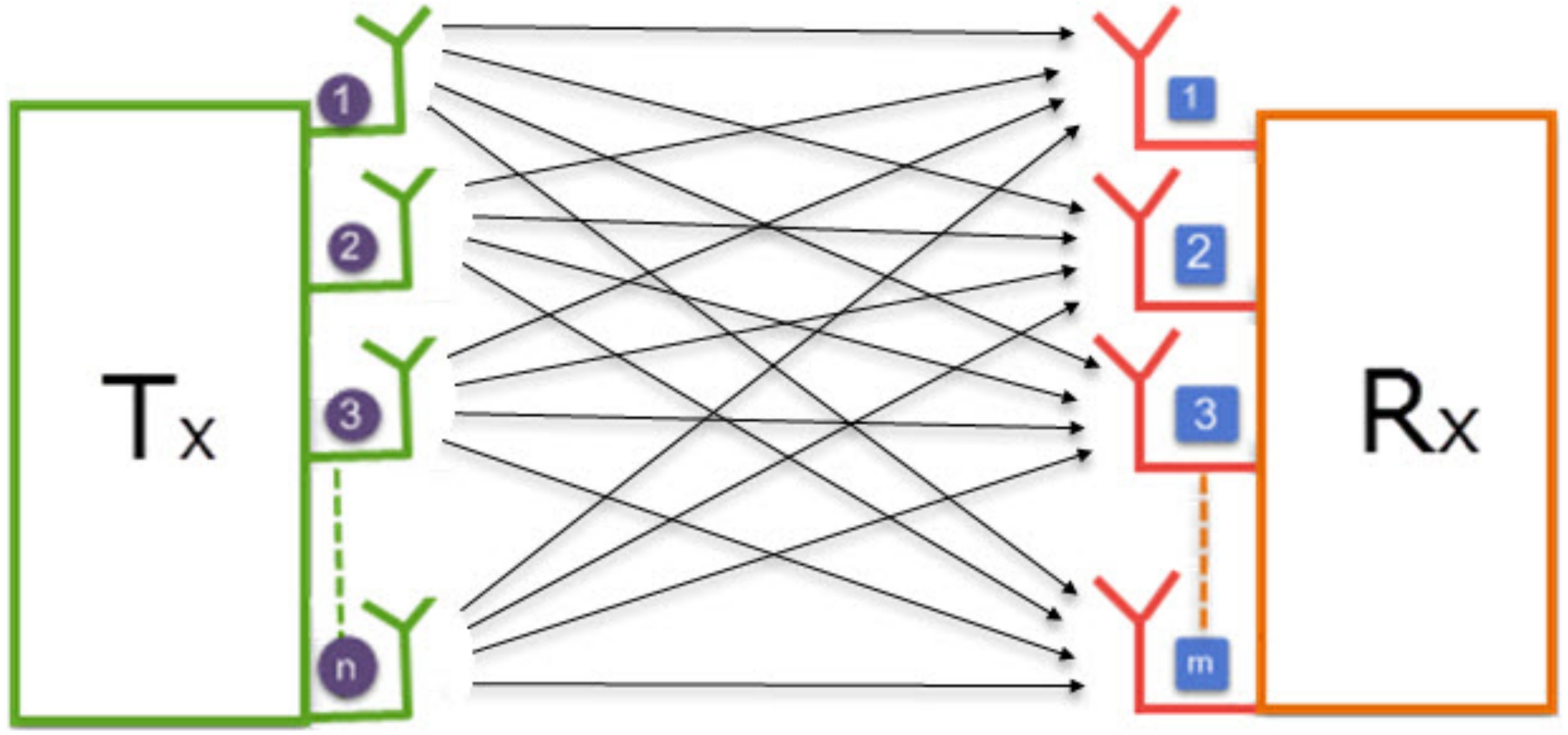
- Reading: “Spectrum graph coloring to improve Wi-Fi channel assignment in a real-world scenario via edge contraction” https://www.sciencedirect.com/science/article/pii/S0166218X18306528

5.1.5 Multiple Input Multiple Output (MIMO)

MIMO

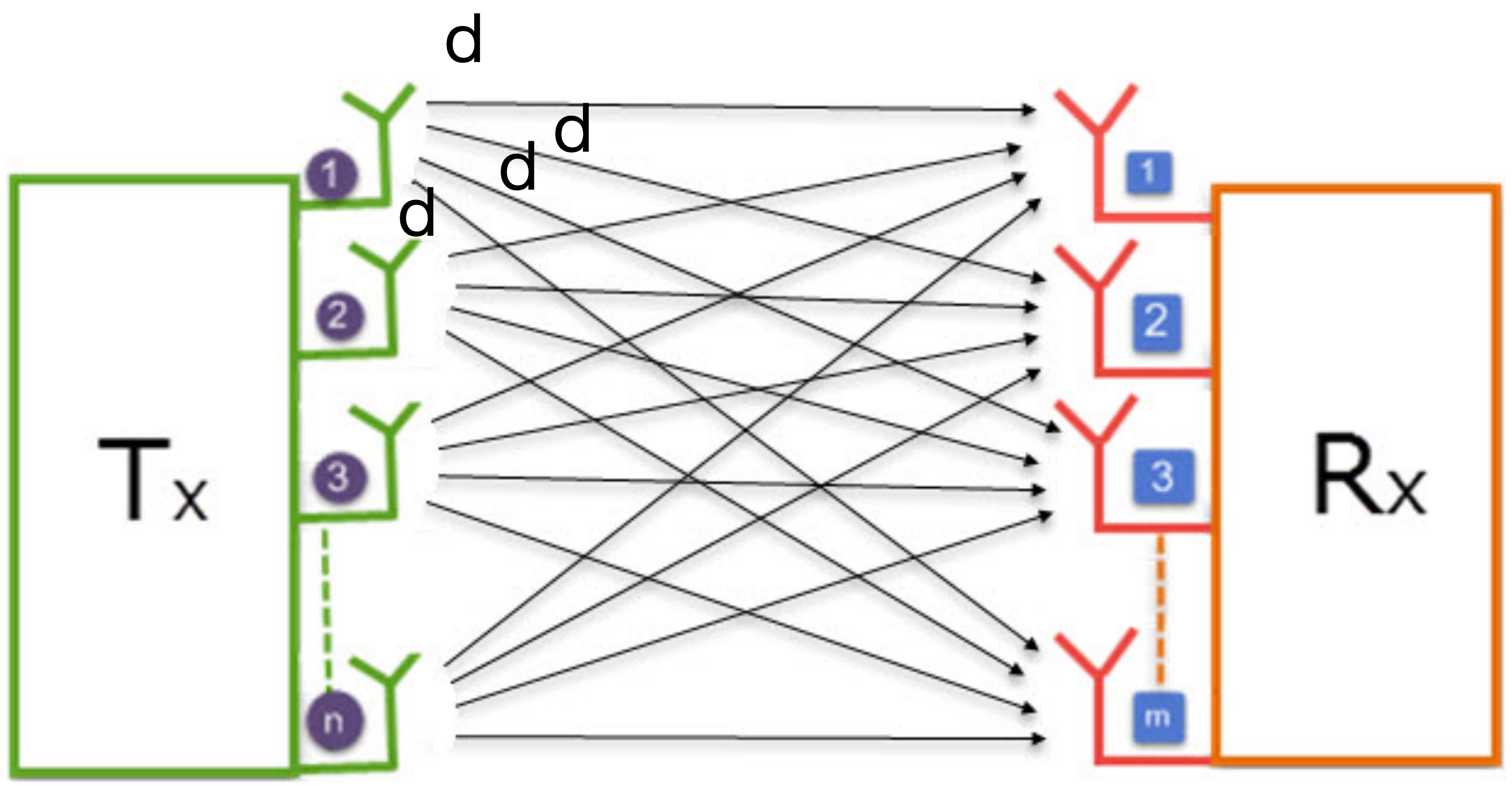
- MIMO (Multiple Input Multiple Output) is a wireless technology that uses **multiple antennas** at both the transmitter and receiver to improve communication performance.
- It significantly increases data rates, reliability, and coverage by contrasting multi path fading effects.
- There are several techniques used in a MIMO wireless network. We will briefly talk about:
    - Spatial diversity
    - Spatial multiplexing
    - Beamforming





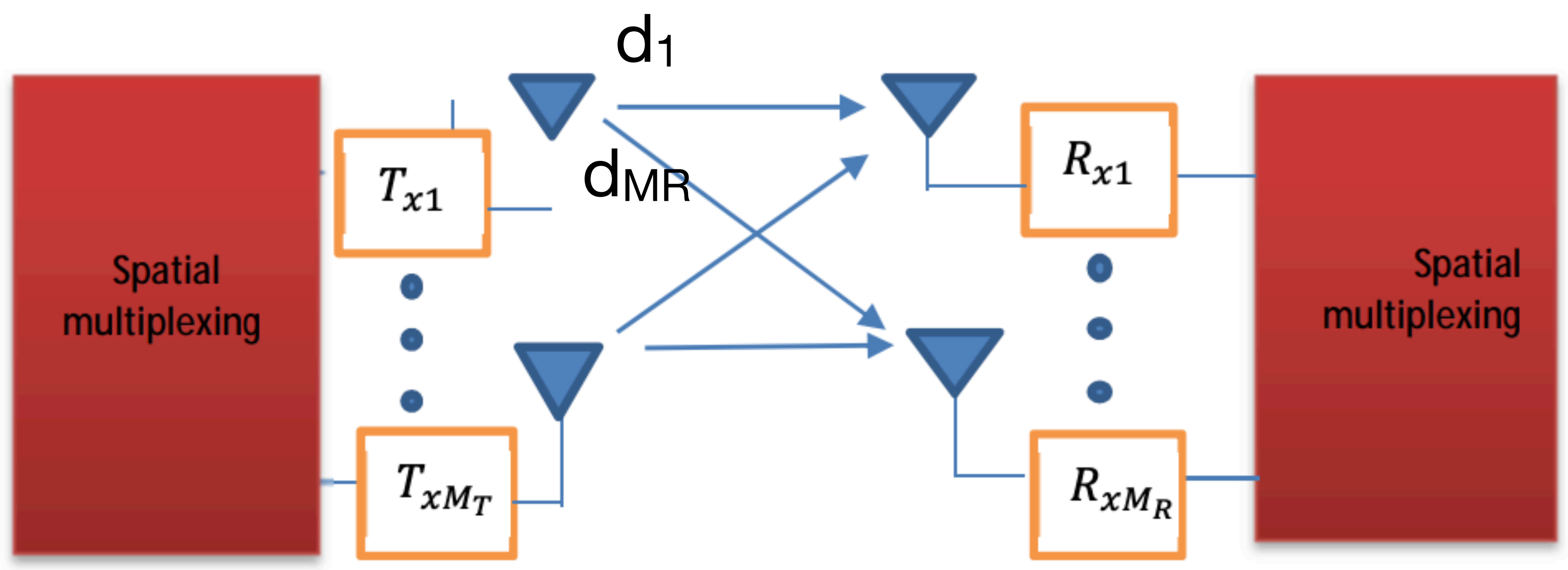
Spatial Diversity

- Signals are subject to multipath fading, which can cause signal loss.
- **Spatial diversity** reduces fading by sending multiple, redundant versions of the signal through different paths in parallel, in this way:
  - The probability that at least one copy of the signal arrives with good quality increases.
  - Antennas at the receiver can capture different bits and pieces of the signal (which might scattered or diffracted) and recompose it with signal processing techniques.
- Improved reliability and robustness.
- Supported by 4G LTE antennas.



Spatial Multiplexing

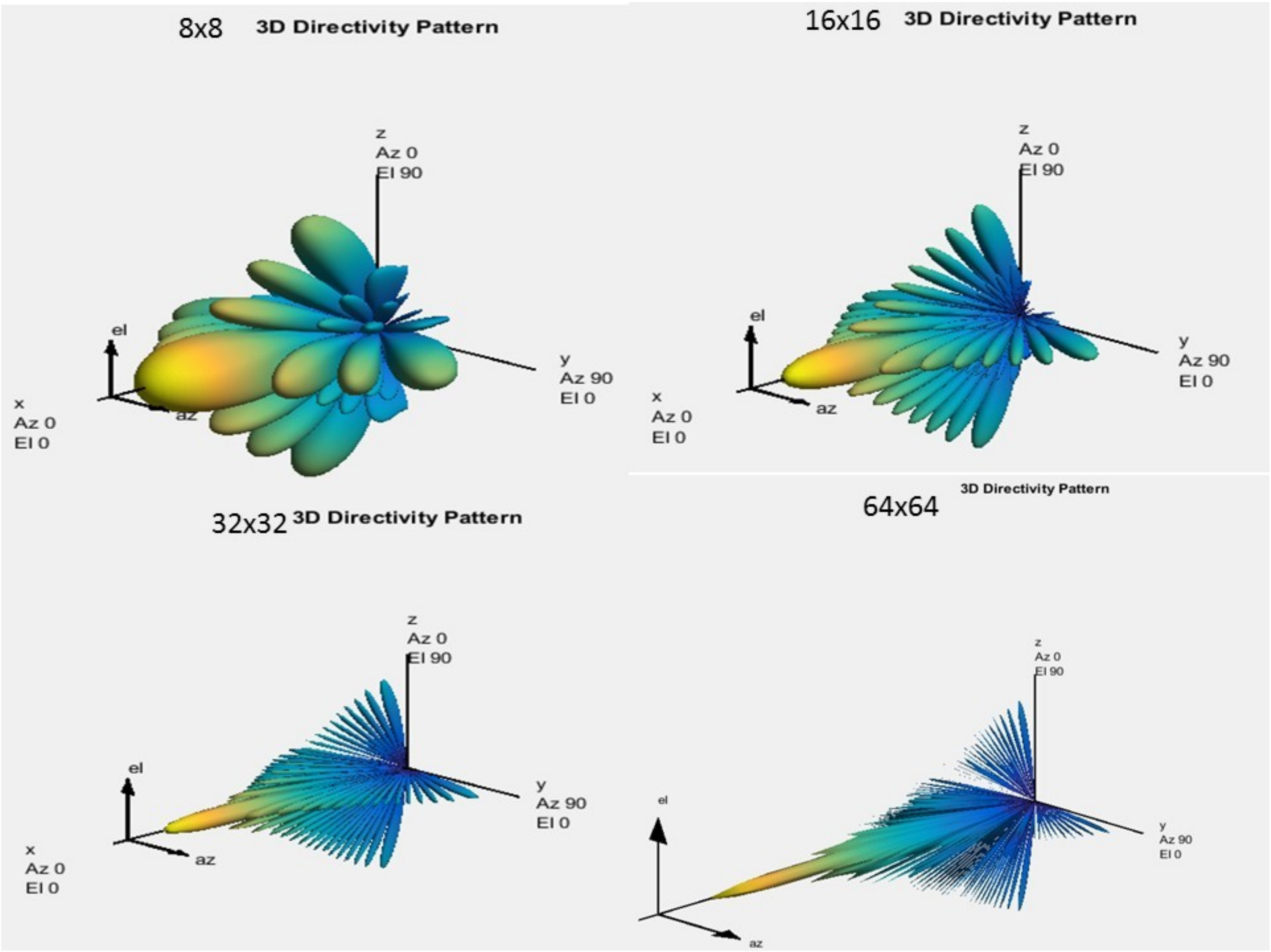
- Instead of contrasting multi path fading, spatial multiplexing increases throughput by leveraging this phenomenon.
- The transmitter **splits the data stream into multiple independent sub-streams** and sends them simultaneously through different antennas.
- Due to multi path and attenuation, these signals can arrive at different times or with different power due to constructive or destructing interferences.
- The receiver sees a linear combination of the transmitted signals at each of its antenna. Signal processing techniques are exploited to reconstruct the original signal.



- Used in Wi-Fi 6, 4G LTE, 5G networks.

Beamforming

- With beamforming, the transmitter directs signals toward specific devices instead of broadcasting them in all directions.
- Each antenna transmits the signal.
    - Each antenna adjusts the phase and the amplitude of each signal so that the produced signals add constructively in the desired direction and destructively in other directions.
    - This way, that they concentrate the energy into a narrow beam (lobe).
- It enhances signal strength, range, and efficiency by focusing energy towards a desired point.
- Used in Wi-Fi, 4G LTE, 5G, and radar systems.



Beamforming with different numbers of transmitting and receiving antennas

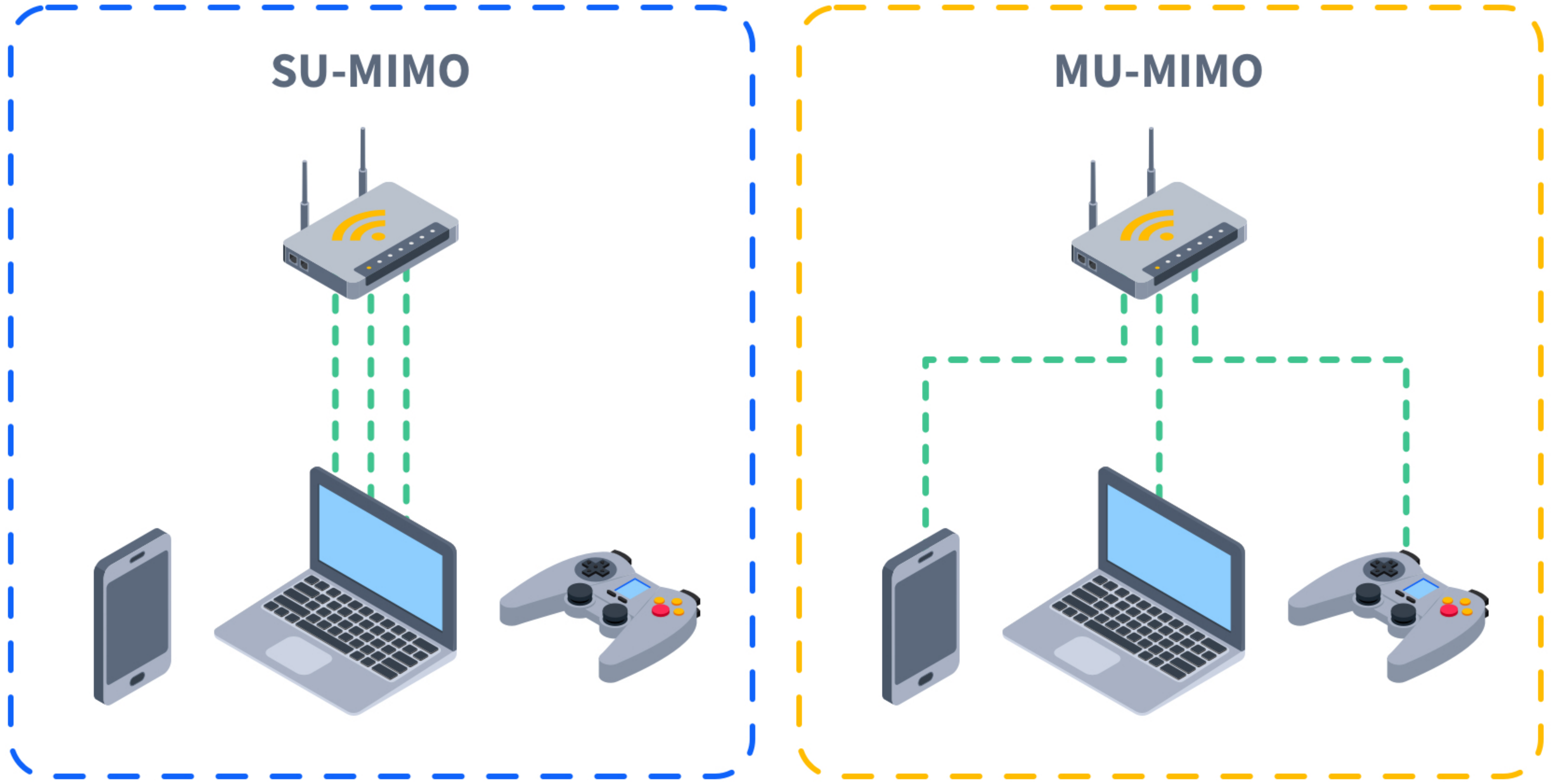
Multi-User (MU)-MIMO

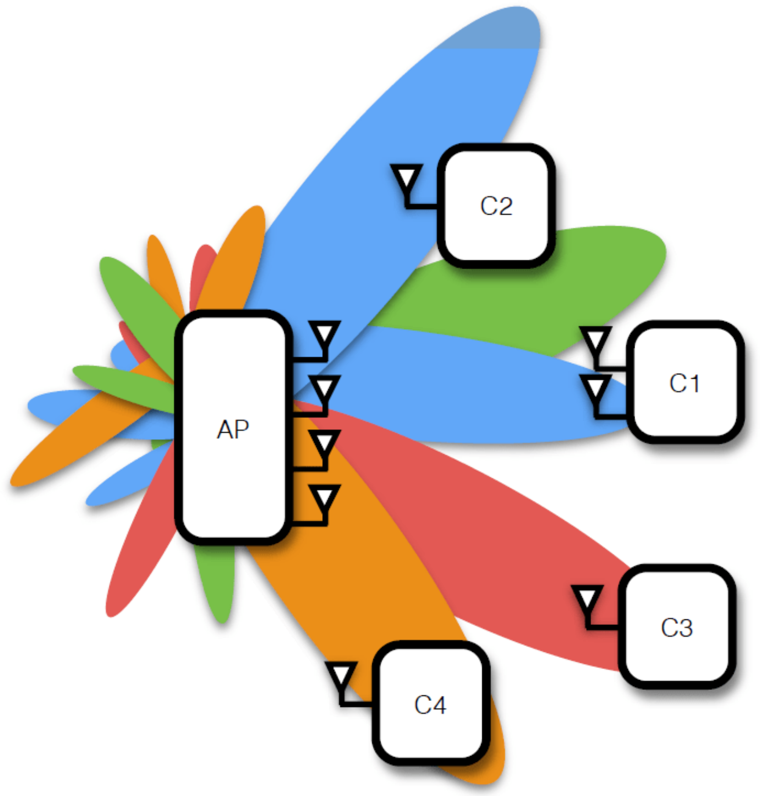
- Beamforming is particularly effective to implement a multiple-user MIMO.

- The transmitter can use its multiple antennas to send data to multiple users simultaneously.

- Used in WiFi to send data to multiple devices.

- In 5G networks, we talk about Massive MU-MIMO: the transmitter forms multiple beams to serve many users at once.



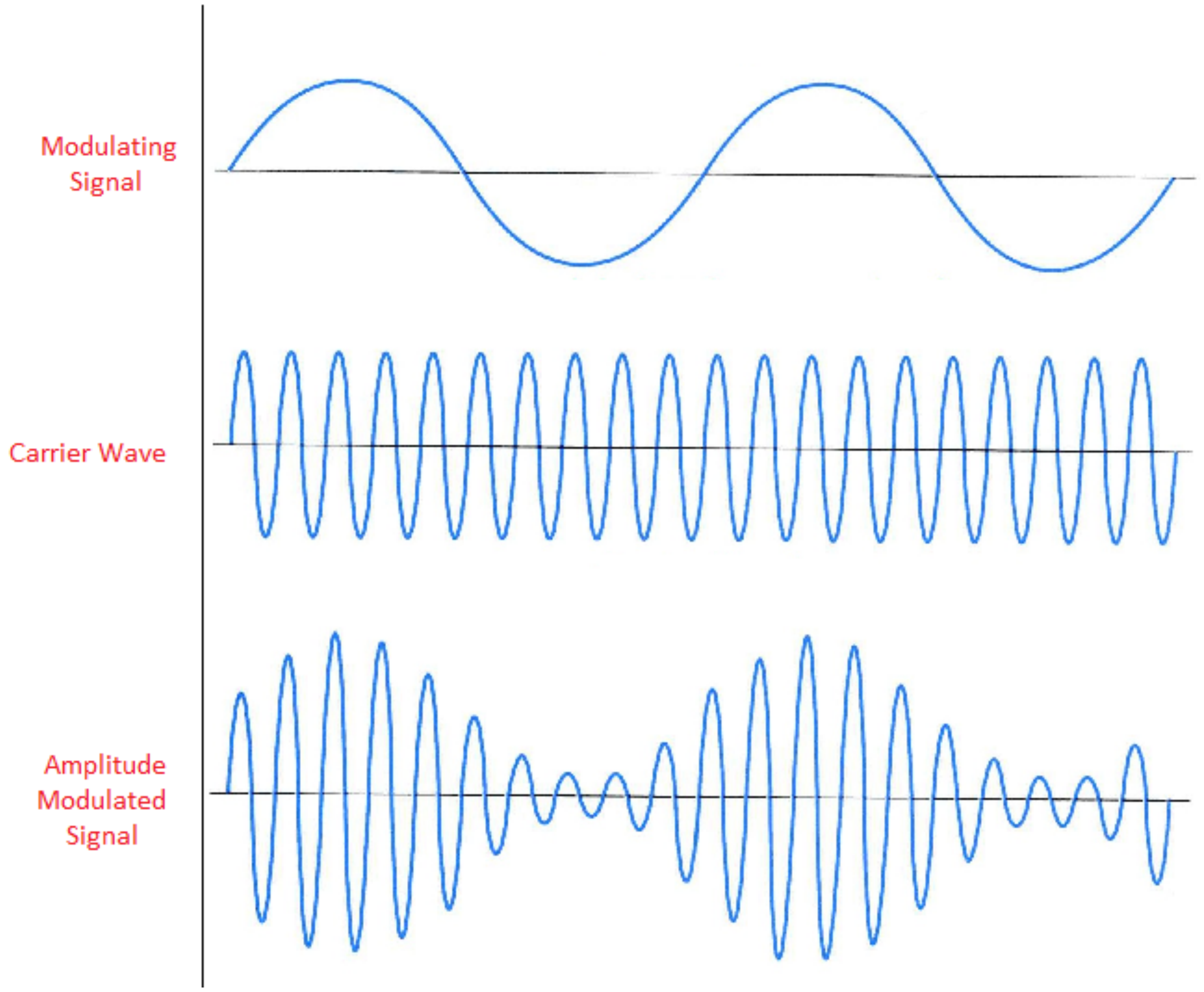


Multi-user MIMO

5.1.6. Modulation

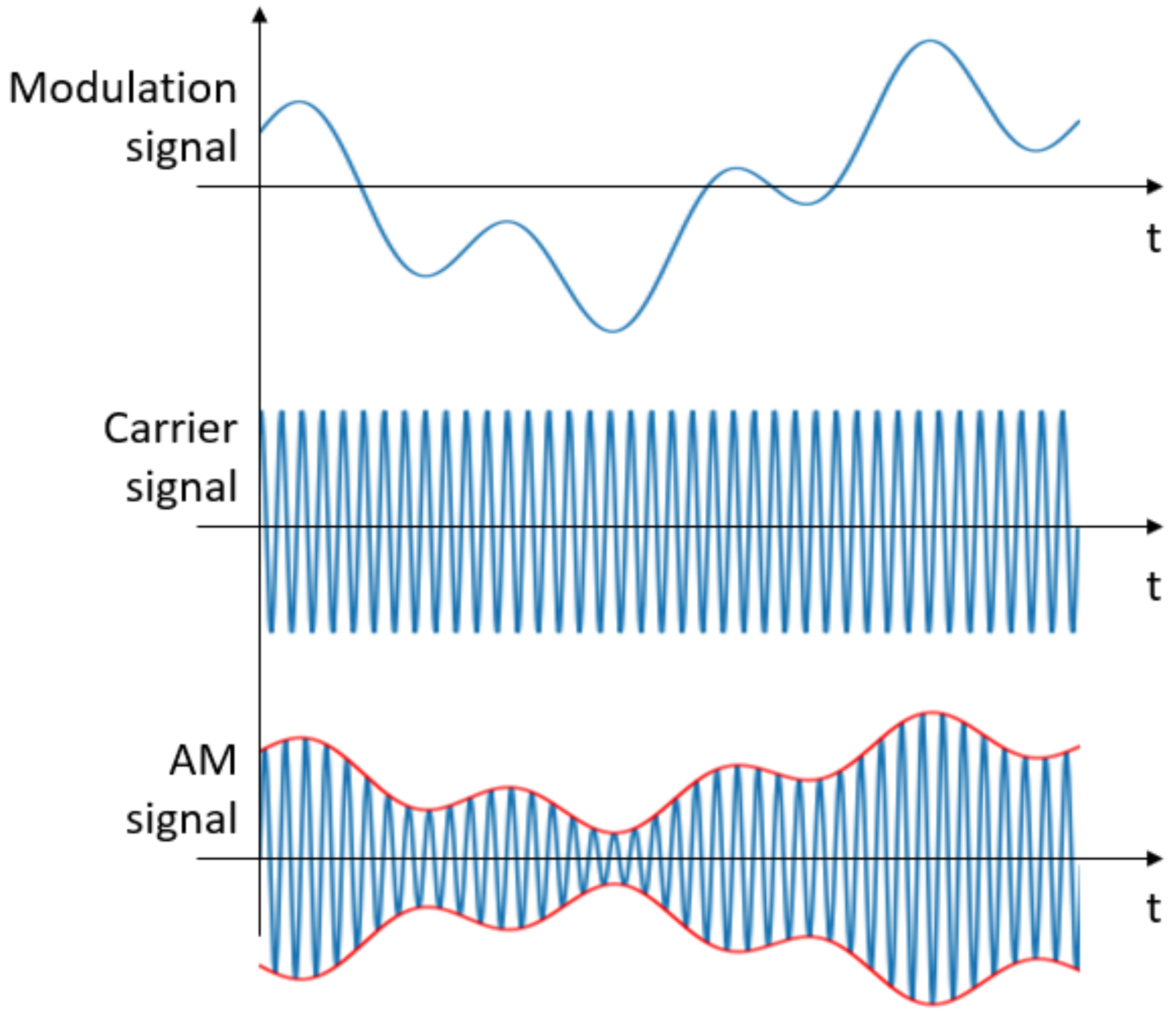
Carrier and Message signals

- Modulation is the process of varying one or more parameters (frequency f, amplitude A, phase ϕ) of a waveform (a sine) to encode data. $g(t) = A_t \sin(2\pi f_t + \phi_t)$.
- Information is encoded into a signal (**modulating** or **message signal**).
  If the source of information produces digital data, this signal is obtained by a DAC.
- The **carrier signal** is a high frequency signal that is combined with the modulating signal and altered in some of its features.
  The frequency of the signal represents the central frequency of the resulting signal.

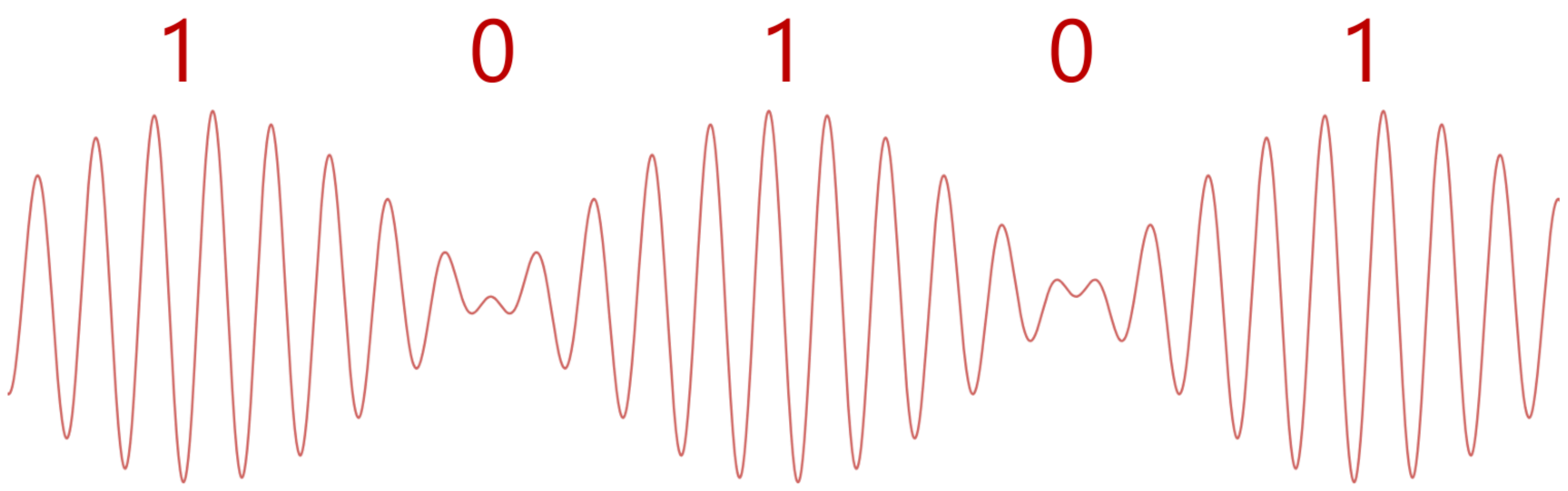


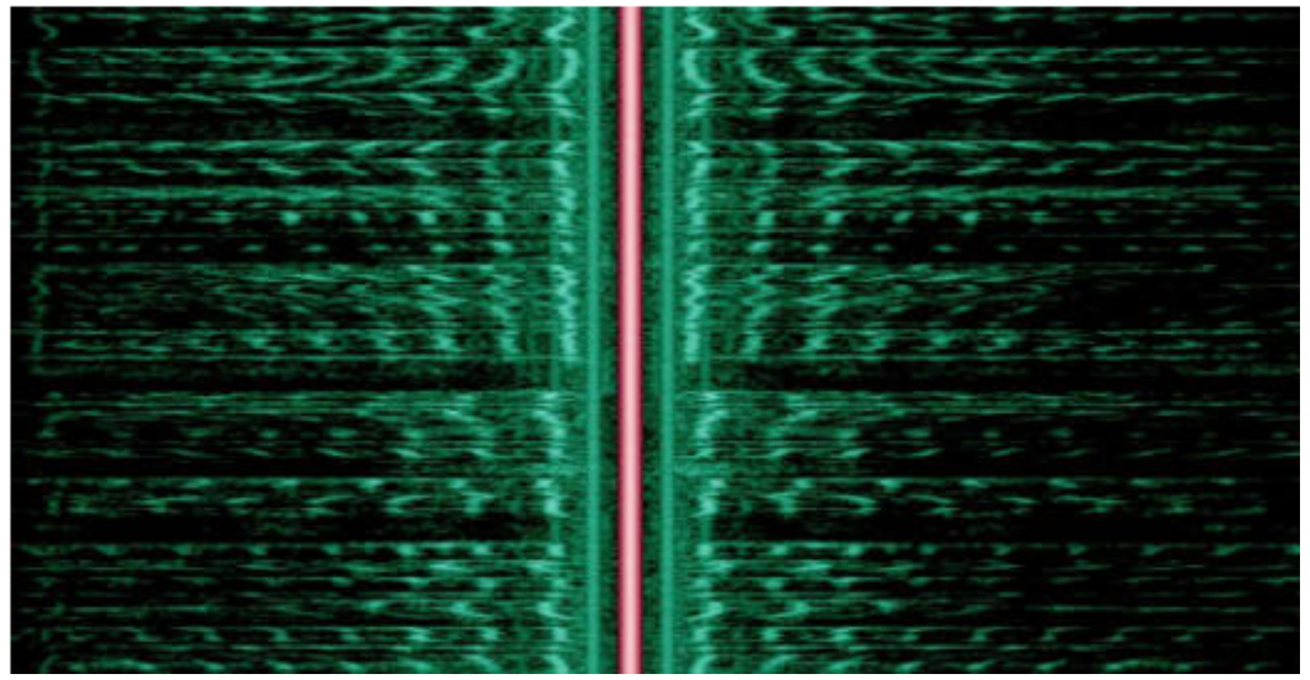
Amplitude Modulation

- Amplitude Modulation (AM) is widely used for analog data transmission, e.g., voice or video, as the AM radio.
- The amplitude of the carrier signal is modified to encode the modulation signal.
- For binary signals, we use the term "Amplitude shift keying".



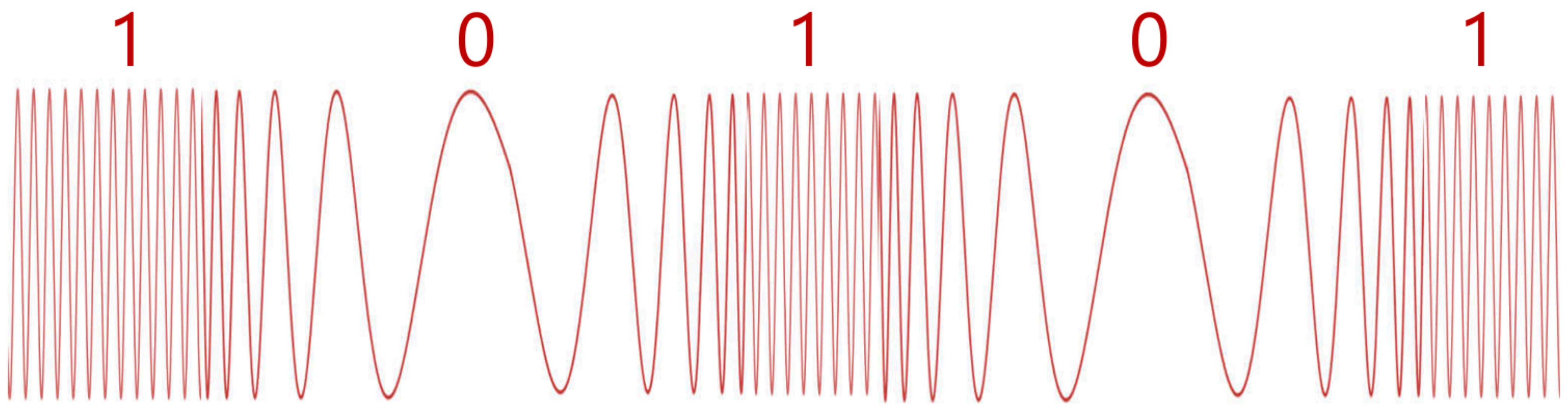
Amplitude Shift Keying (ASK)

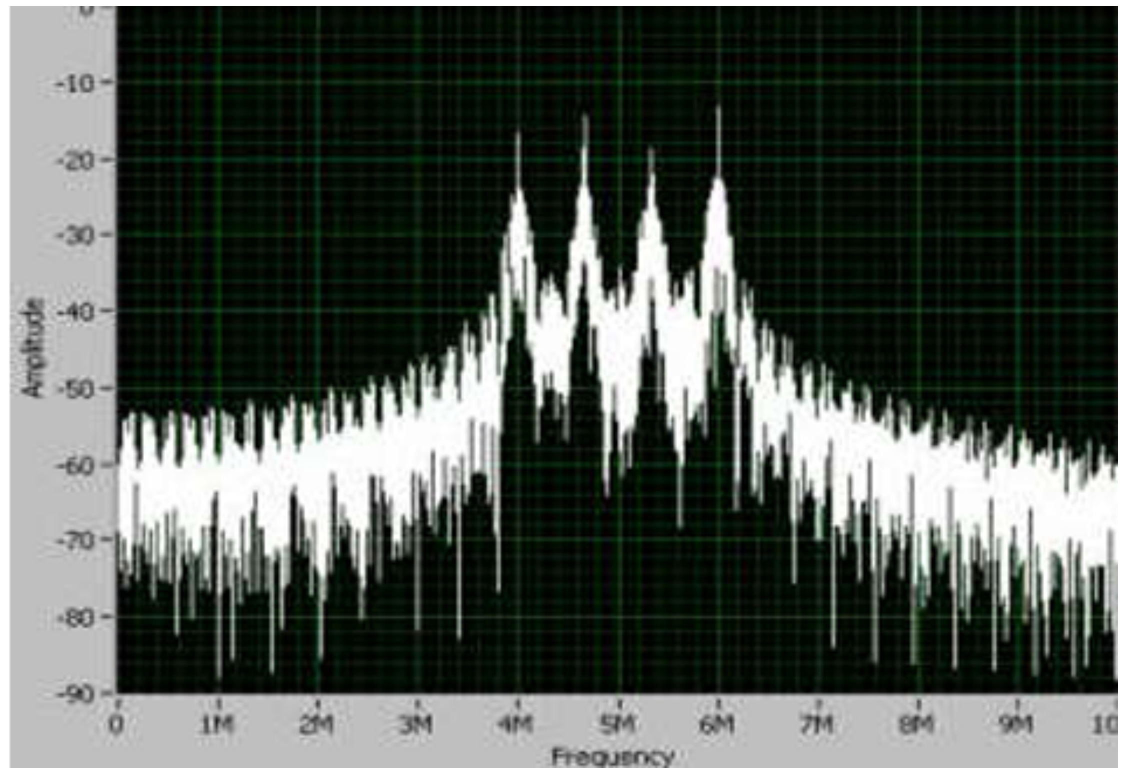




- "Amplitude shift keying" is used for binary signals.
  - 1s correspond to high amplitude intervals, 0s to low amplitude intervals.
- Pros: Cheap hardware
- Cons: more susceptible to noise.

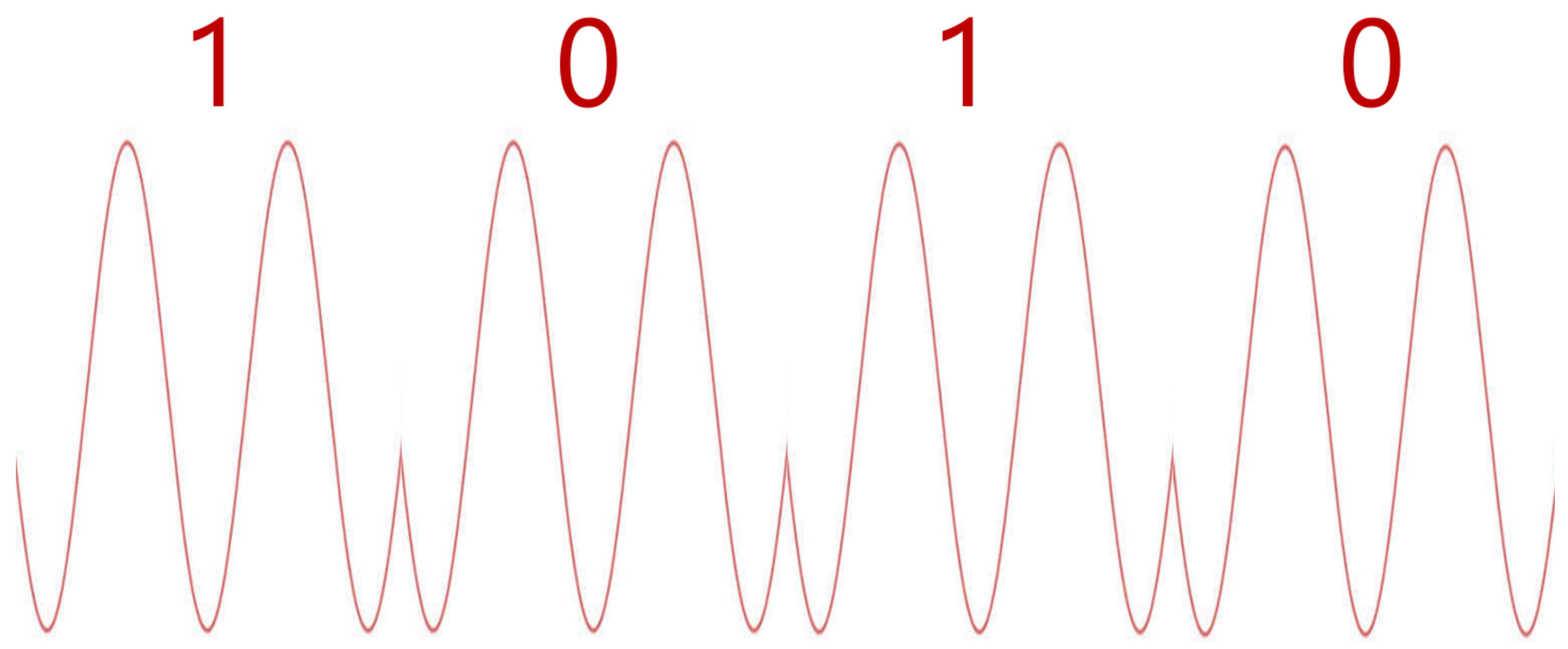
Frequency Shift Keying (FSK)

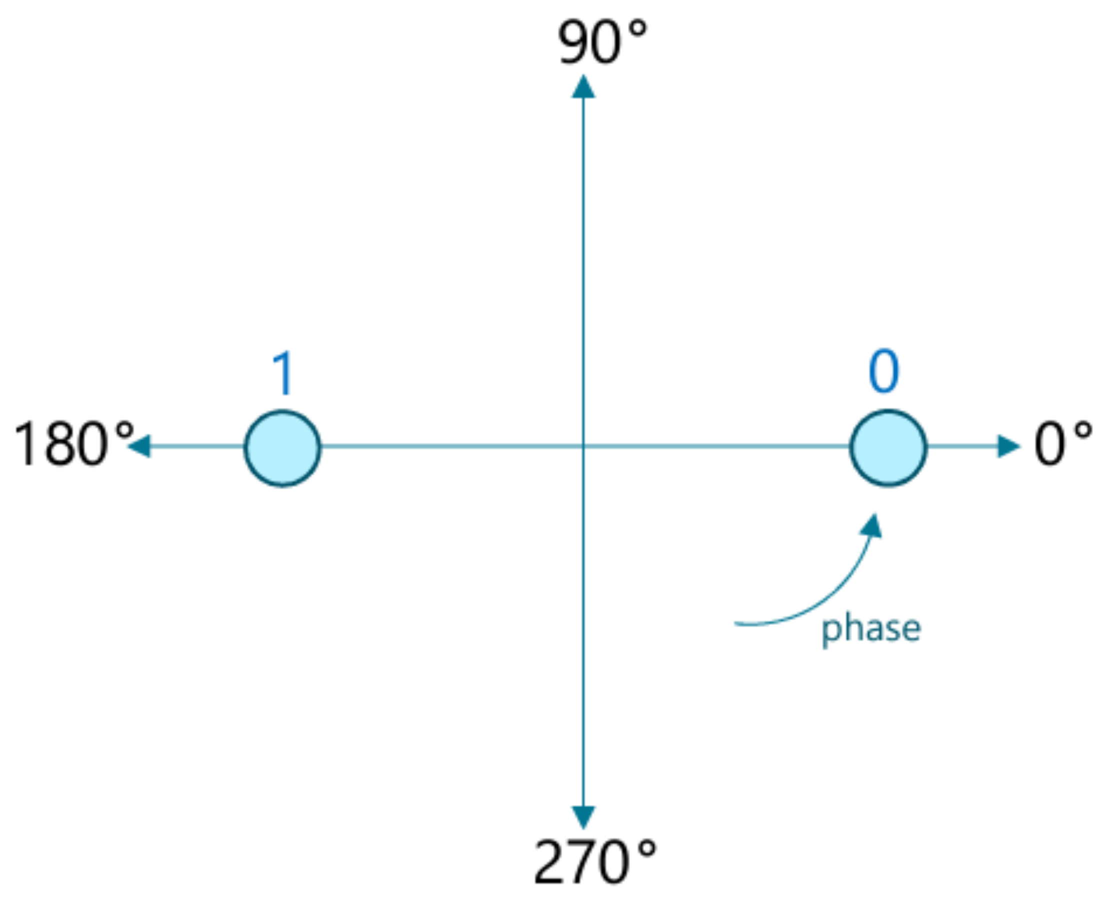




- Vary the frequency of signals and uses two different frequencies to represent 0s and 1s.
- Pros: less sensitive to noise, robust to variations in attenuation
- Cons: requires larger bandwidth.

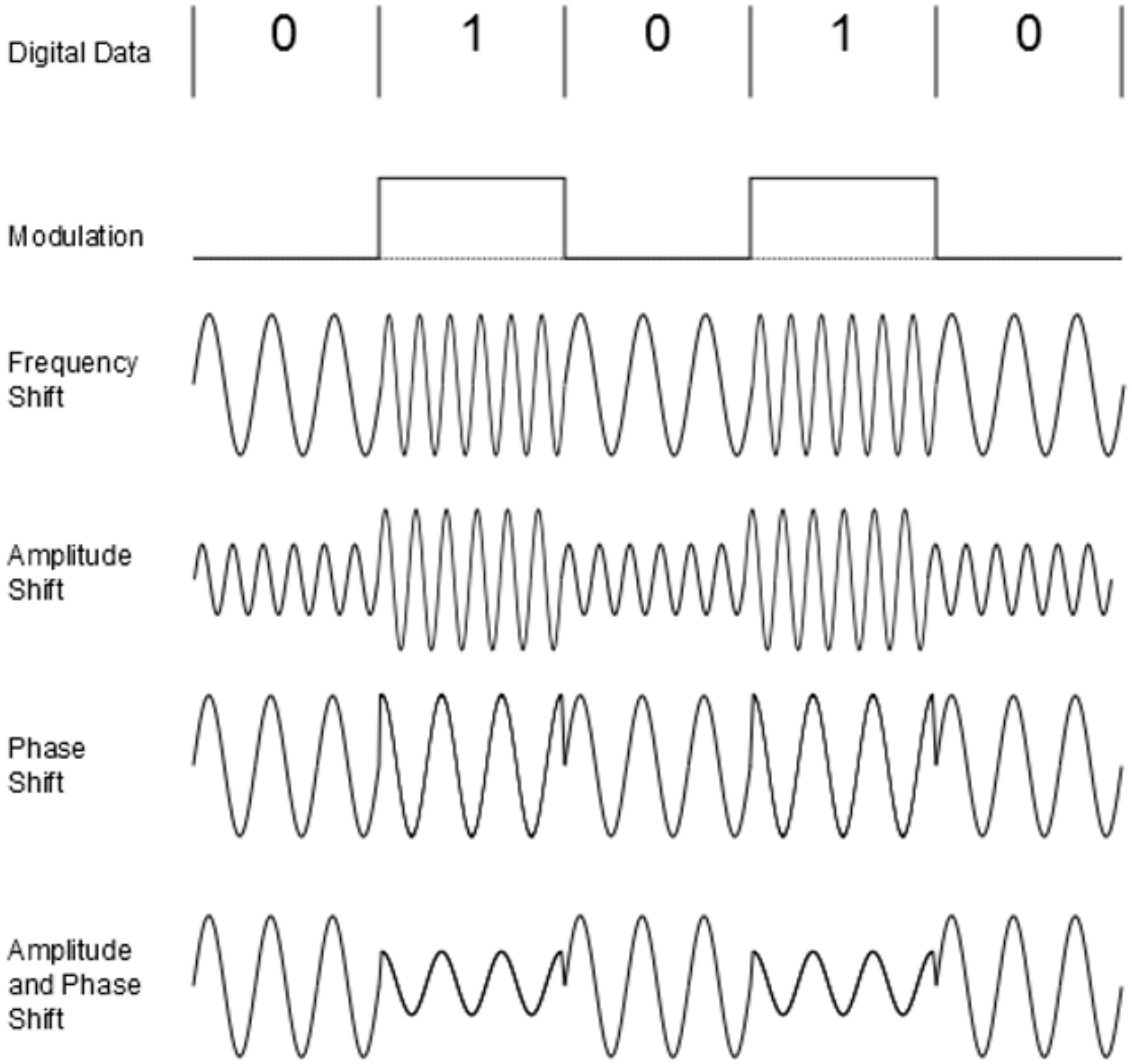
Phase Shift Keying (PSK)





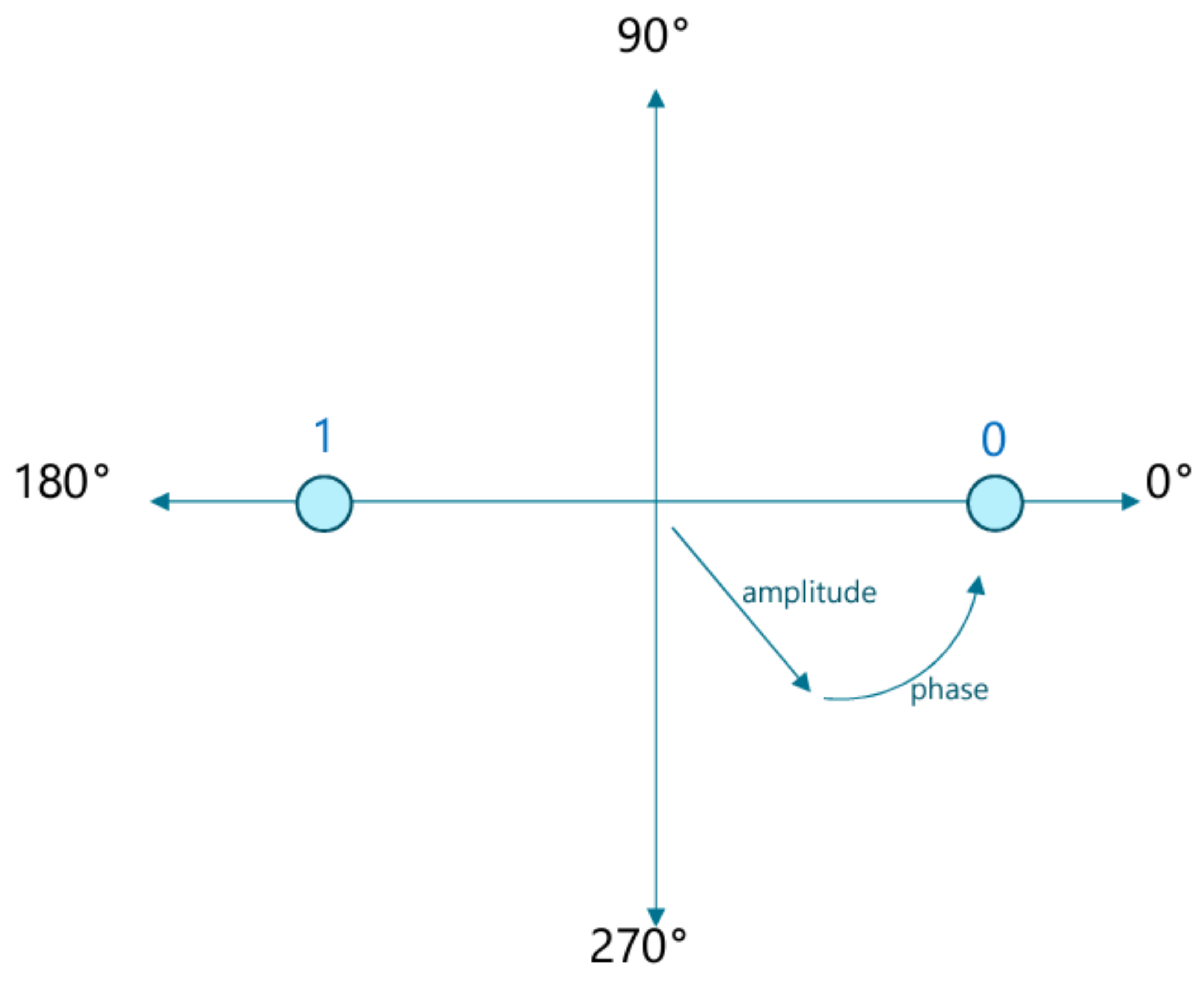
- Vary the phase of the signal by moving the signal 180° out of phase for 0s and 1s.
- Pros: Power efficient, less susceptible to errors than ASK
- Cons: Can be difficult to implement, worse noise rejection than FSK.

DAC


Digital Data
0
1
0
1
0
Modulation
Frequency Shift
Amplitude Shift
Phase Shift
Amplitude and Phase Shift

Constellation Diagram (1)

- Each point of a constellation diagram represents a signal “symbol” or “word” encoded in polar coordinates, where the **angle** is the **phase** and the **distance from the origin** represents the **amplitude of the signal,** i.e., its **strength**.

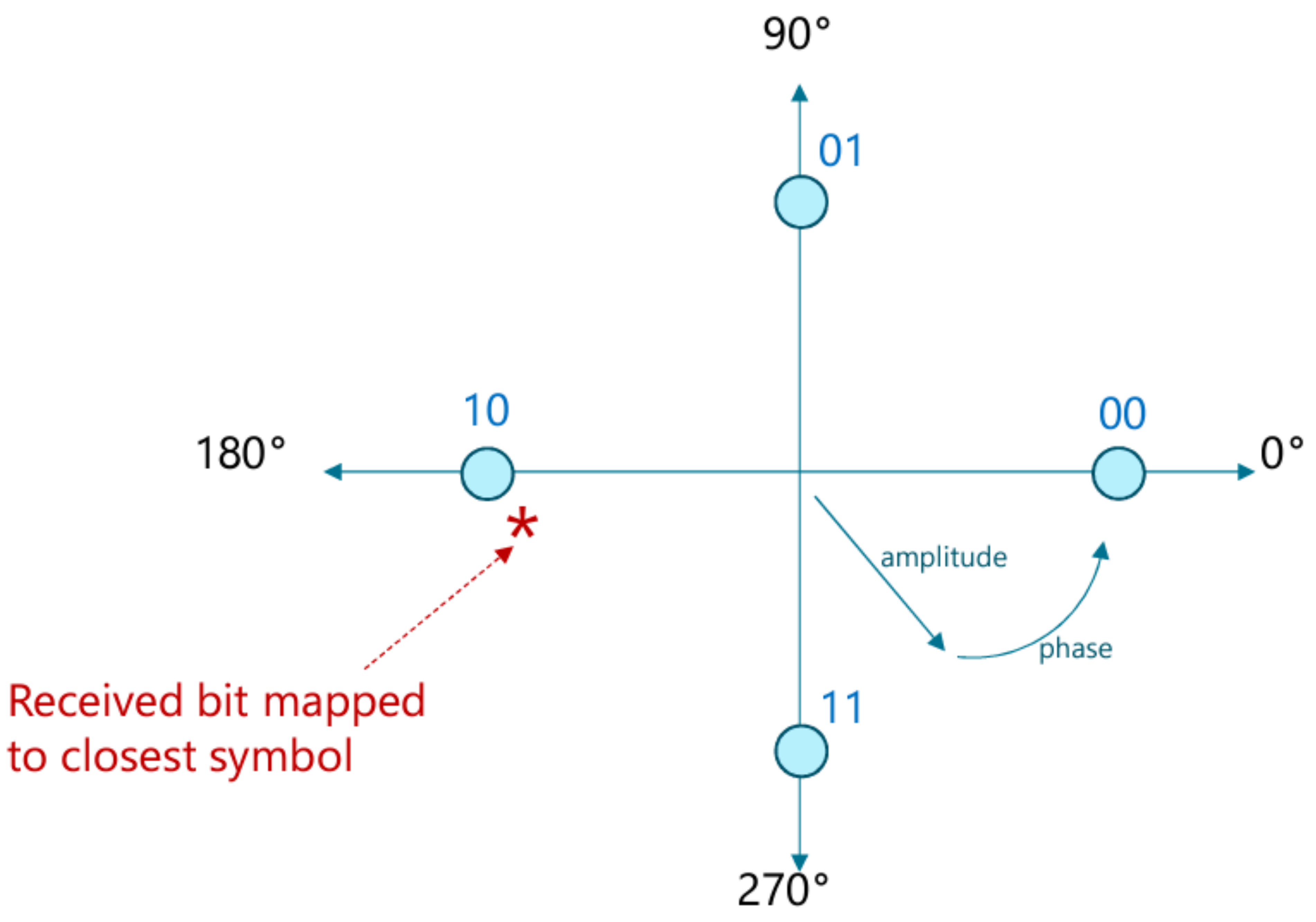


- Constellation diagrams can be used to describe a particular modulation scheme.

- In this example, we encode 0s with an in phase signal and 1s with signals having a 180° shift. Both have fixed amplitude.
  This scheme is called 2-PSK.
  We can do more…

Constellation Diagram (2)

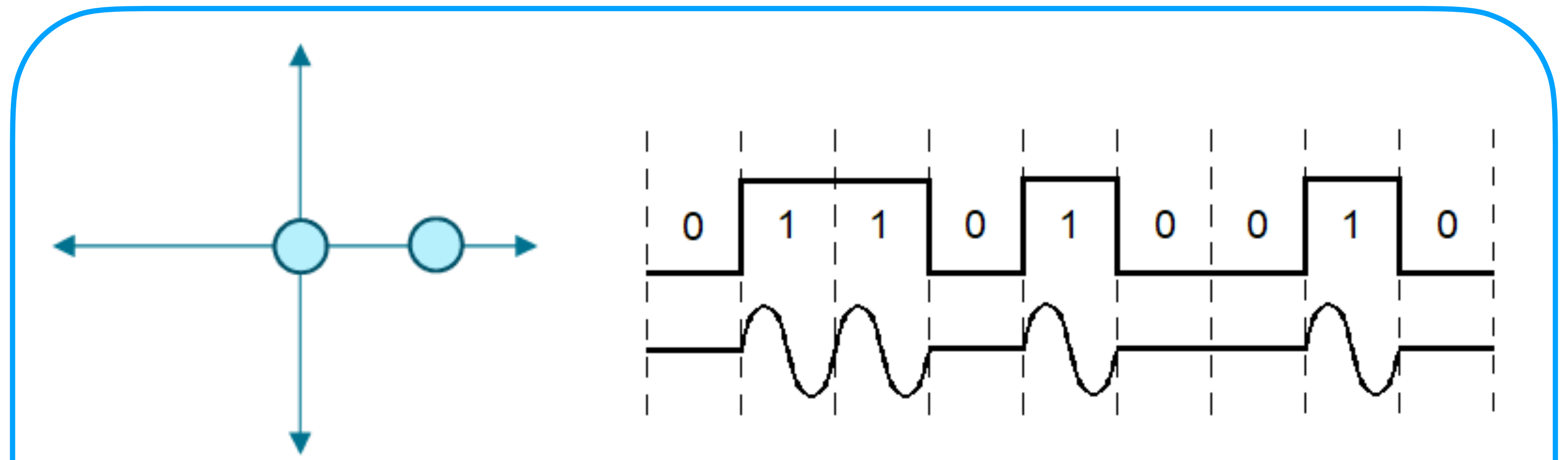
- If we allow 90° phase shifts, then we get a 4-PSK. This allows us to send bits twice as fast as encoding two bits instead of one.



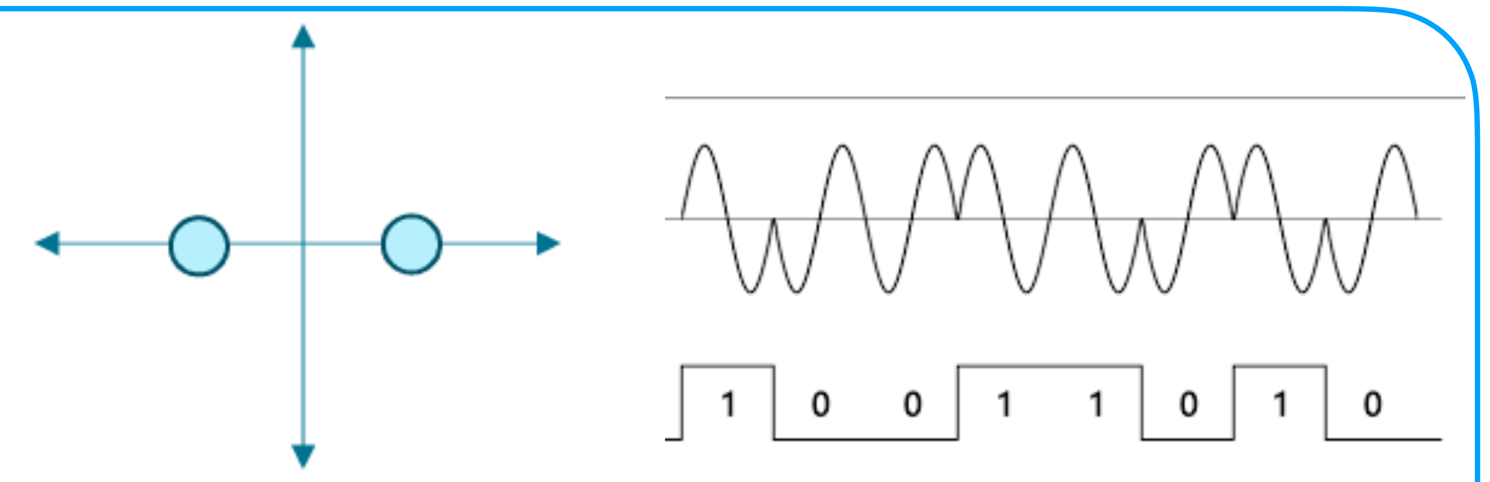
- Constellation Diagram can also be used for visualising received information. Assume the receiver receives a signal that is mapped in point *. There is a chance that the transmitter was trying to send 11, but most likely it sent 10.

- Receivers implement estimation algorithms (e.g., maximum likelihood) to recover the received signal.

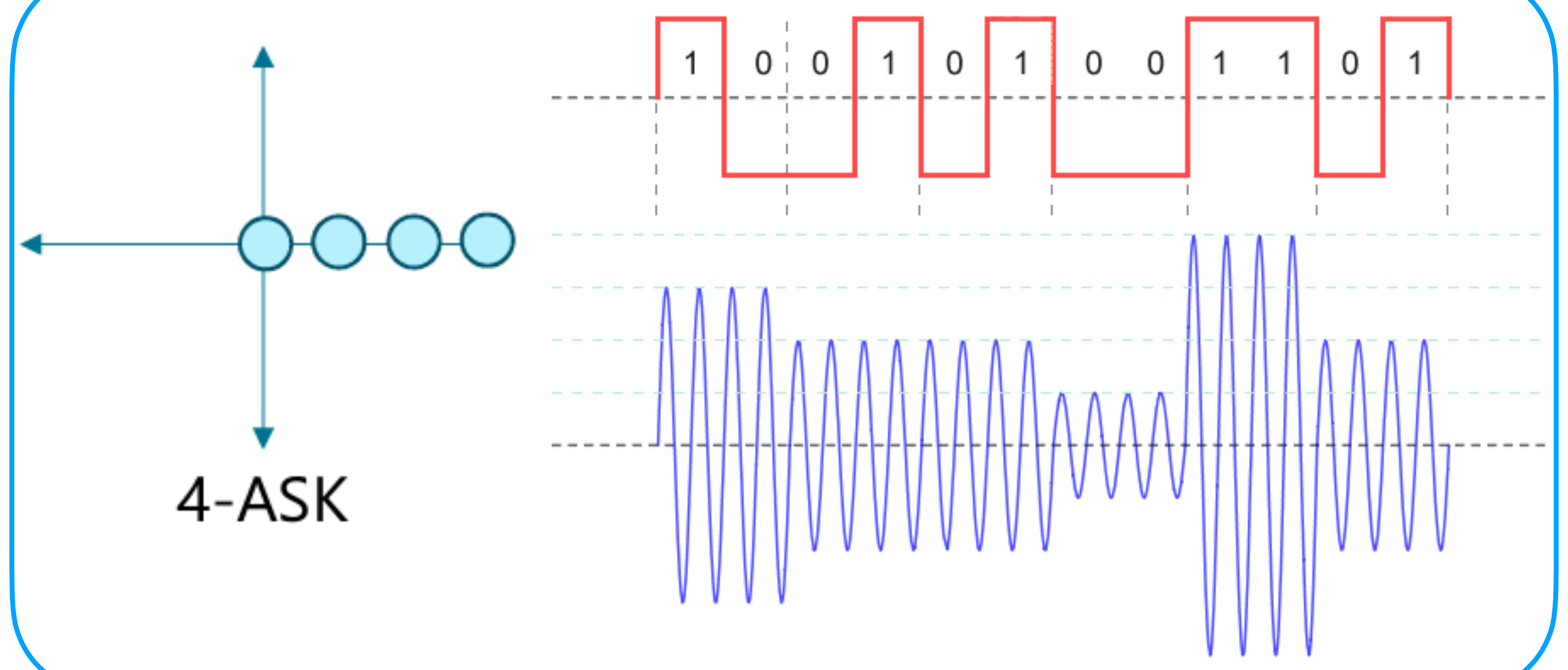
Common Phase-Amplitude Modulations



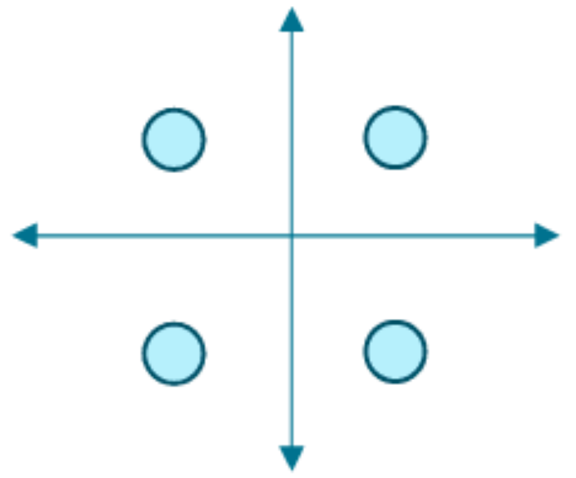
On-Off Keying
(OOK)



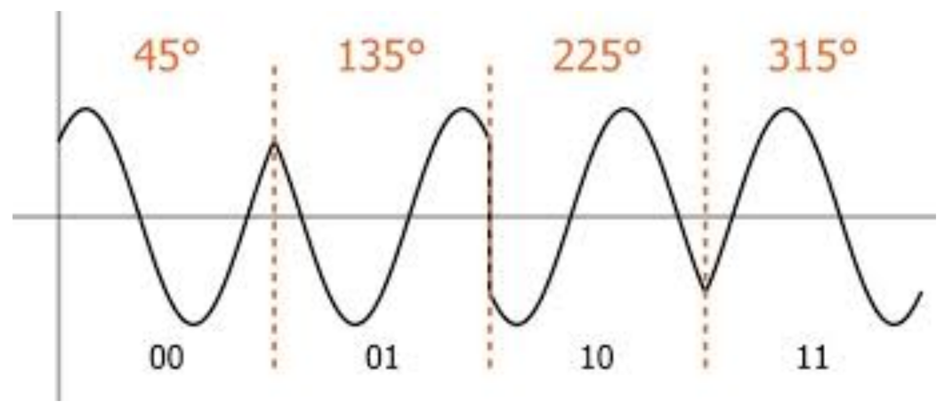
Binary Phase-Shift Keying
(BPSK)

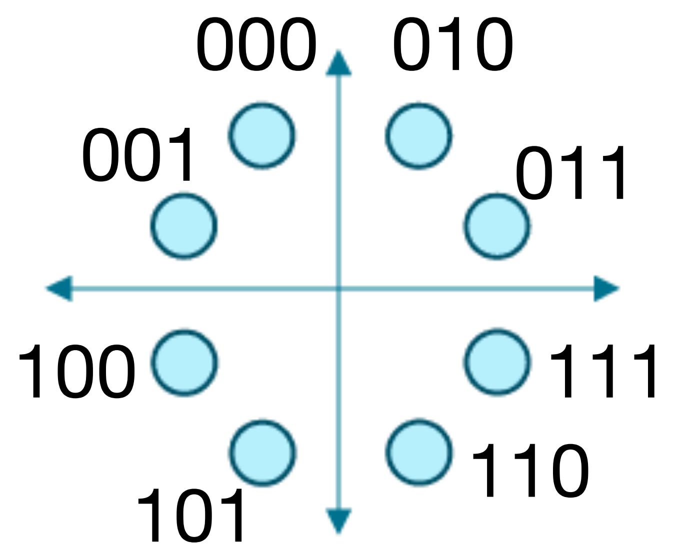


4-ASK

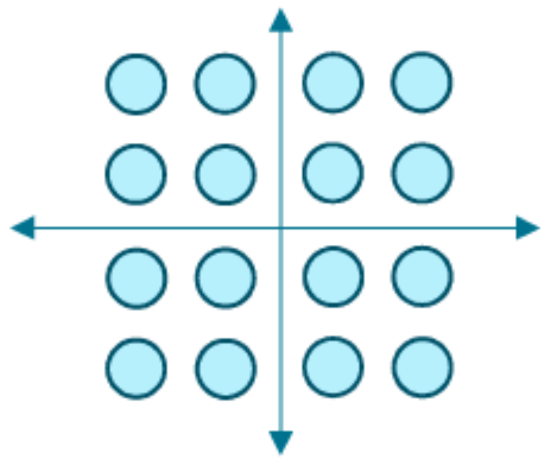


Quadrature Phase
Shift Keying
(QPSK)

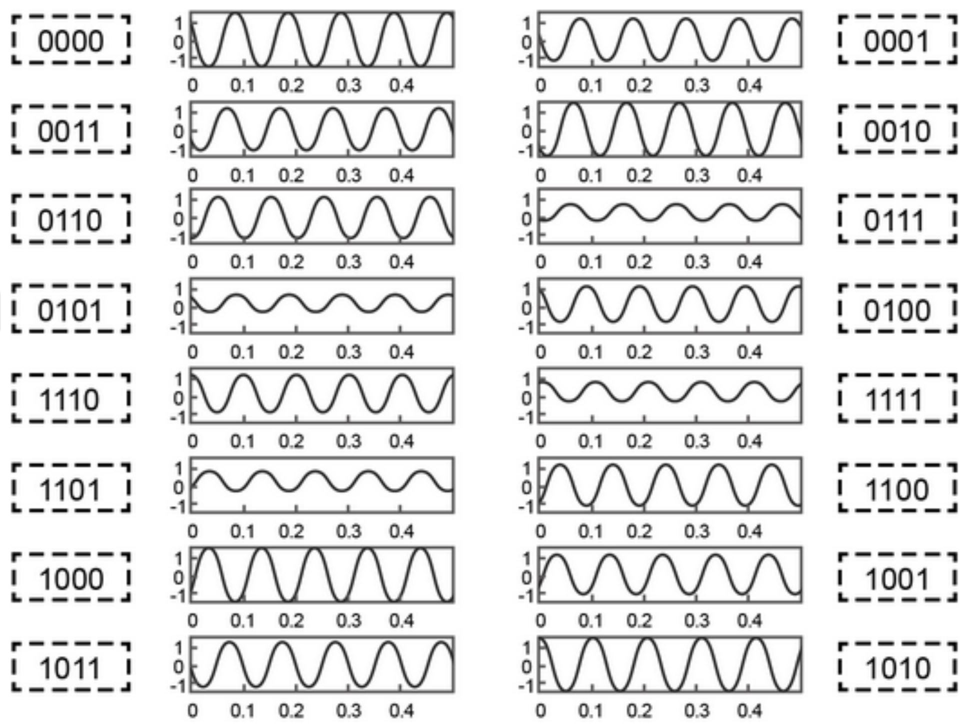




8-PSK



16-Quadrature
Amplitude Modulation
(16-QAM)



And so on… up to 16384-QAM

Constellation Diagram and Beyond

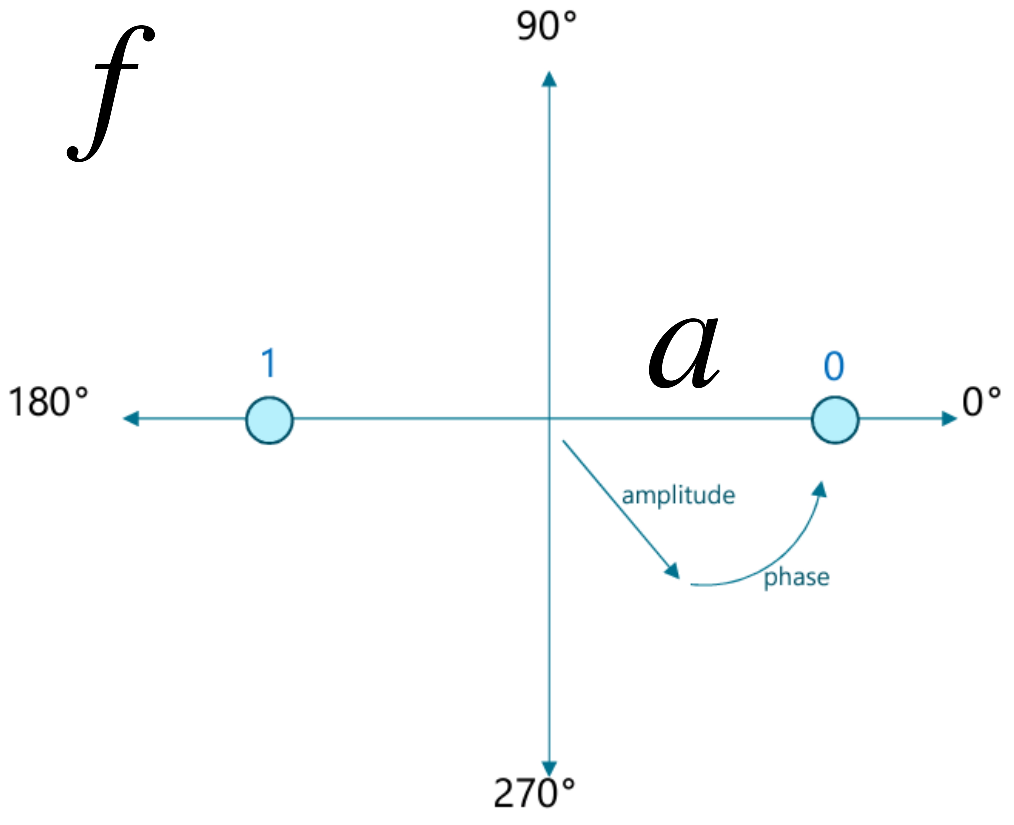
- Each point in a constellation diagram is called **symbol** or **word.**
- Nevertheless, not all "symbols" fit a constellation diagram
  - Constellation diagrams encode signal amplitude and phase for a fixed frequency, but we can modulate signals also using different frequencies.
- **A symbol is a sine wave with a certain frequency, an amplitude and a phase.**
- By changing these parameters, we can create an encoding to words. For example:

  - In 2-PSK:

$$(f, a, 0°) \xrightarrow{\text{2-PSK}} 0$$

$$(f, a, 180°) \xrightarrow{\text{2-PSK}} 1$$

  - A word is a bit



Constellation Diagram and Beyond

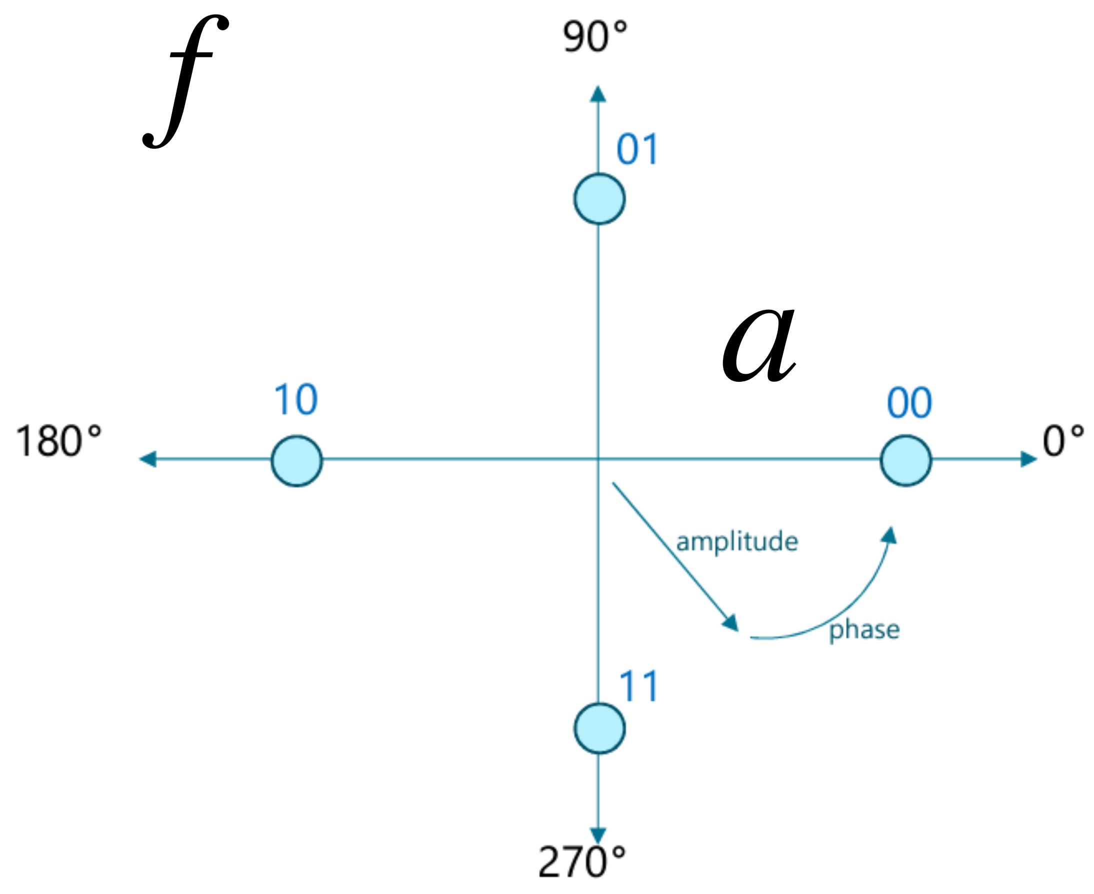
- Each point in a constellation diagram is called **symbol** or **word.**
- Nevertheless, not all "symbols" fit a constellation diagram
  - Constellation diagrams encode signal amplitude and phase for a fixed frequency, but we can modulate signals also using different frequencies.
- **A symbol is a sine wave with a certain frequency, an amplitude and a phase.**
- By changing these parameters, we can create an encoding to words. For example:

- In 4-PSK:

$$(f, a, 0°) \xrightarrow{\text{4-PSK}} 00,\ (f, a, 90°) \xrightarrow{\text{4-PSK}} 01$$

$$(f, a, 180°) \xrightarrow{\text{4-PSK}} 10,\ (f, a, 270°) \xrightarrow{\text{4-PSK}} 11$$

- A word is 2 bits



Constellation Diagram and Beyond

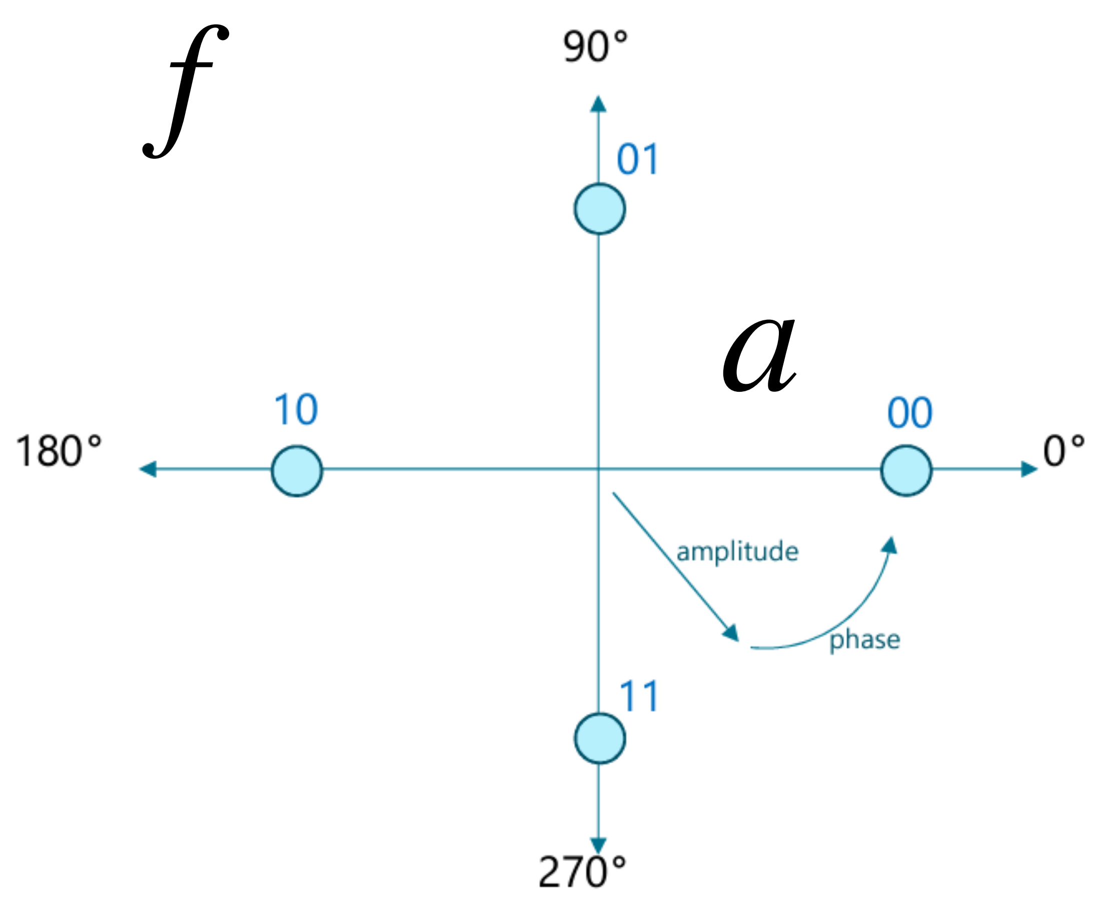
- Each point in a constellation diagram is called **symbol** or **word.**
- Nevertheless, not all “symbols” fit a constellation diagram
  - Constellation diagrams encode signal amplitude and phase for a fixed frequency, but we can modulate signals also using different frequencies.
- **A symbol is a sine wave with a certain frequency, an amplitude and a phase.**
- By changing these parameters, we can create an encoding to words. For example:

- In 4-PSK:

$$(f, a, 0°) \xrightarrow{\text{4-PSK}} 00,\ (f, a, 90°) \xrightarrow{\text{4-PSK}} 01$$

$$(f, a, 180°) \xrightarrow{\text{4-PSK}} 10,\ (f, a, 270°) \xrightarrow{\text{4-PSK}} 11$$

- A word is 2 bits



Here just the phase is changing (a and f are fixed) but other parameters can change too

Constellation Diagram and Beyond

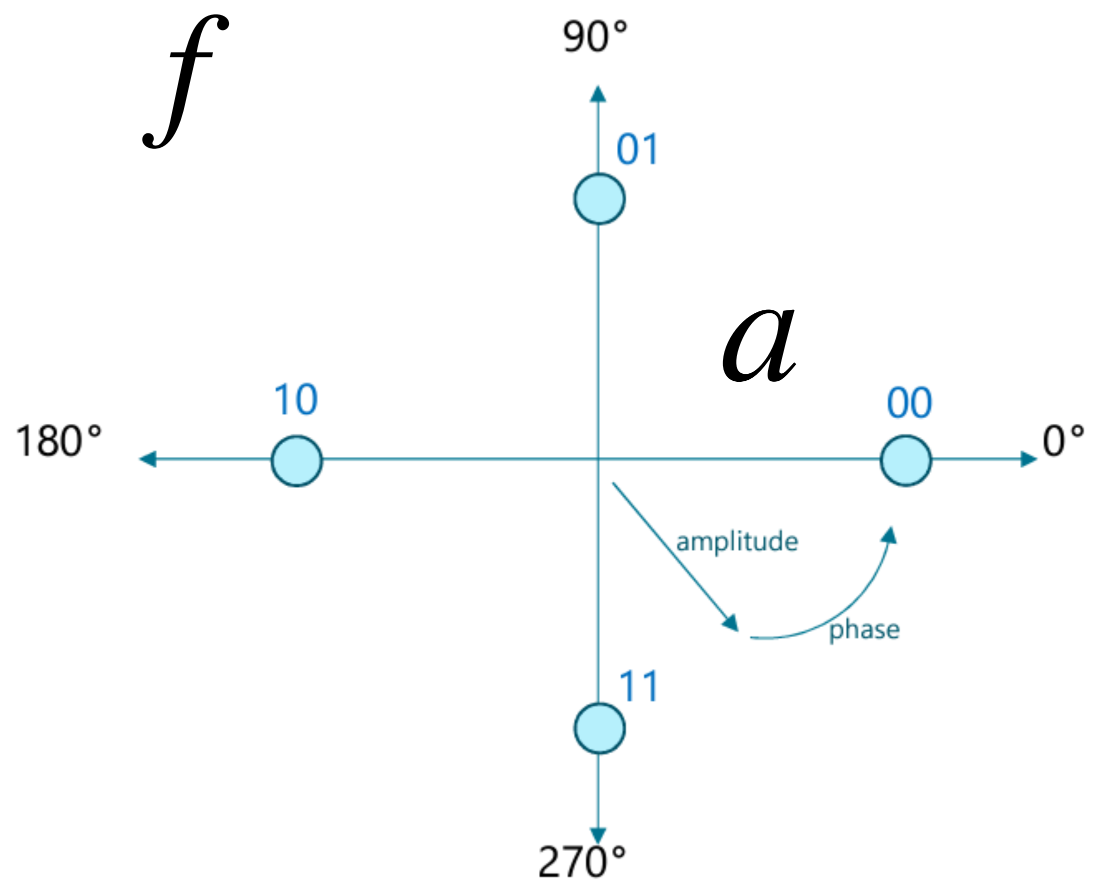
- Each point in a constellation diagram is called **symbol** or **word.**
- Nevertheless, not all "symbols" fit a constellation diagram
  - Constellation diagrams encode signal amplitude and phase for a fixed frequency, but we can modulate signals also using different frequencies.
- **A symbol is a sine wave with a certain frequency, an amplitude and a phase.**
- By changing these parameters, we can create an encoding to words. For example:

- In 4-PSK:

$$(f, a, 0°) \xrightarrow{\text{4-PSK}} 00,\ (f, a, 90°) \xrightarrow{\text{4-PSK}} 01$$

$$(f, a, 180°) \xrightarrow{\text{4-PSK}} 10,\ (f, a, 270°) \xrightarrow{\text{4-PSK}} 11$$

- A word is 2 bits



TRANSMISSION RATE=log(D*N)
D = number of diagrams
N = number of points in each diagram

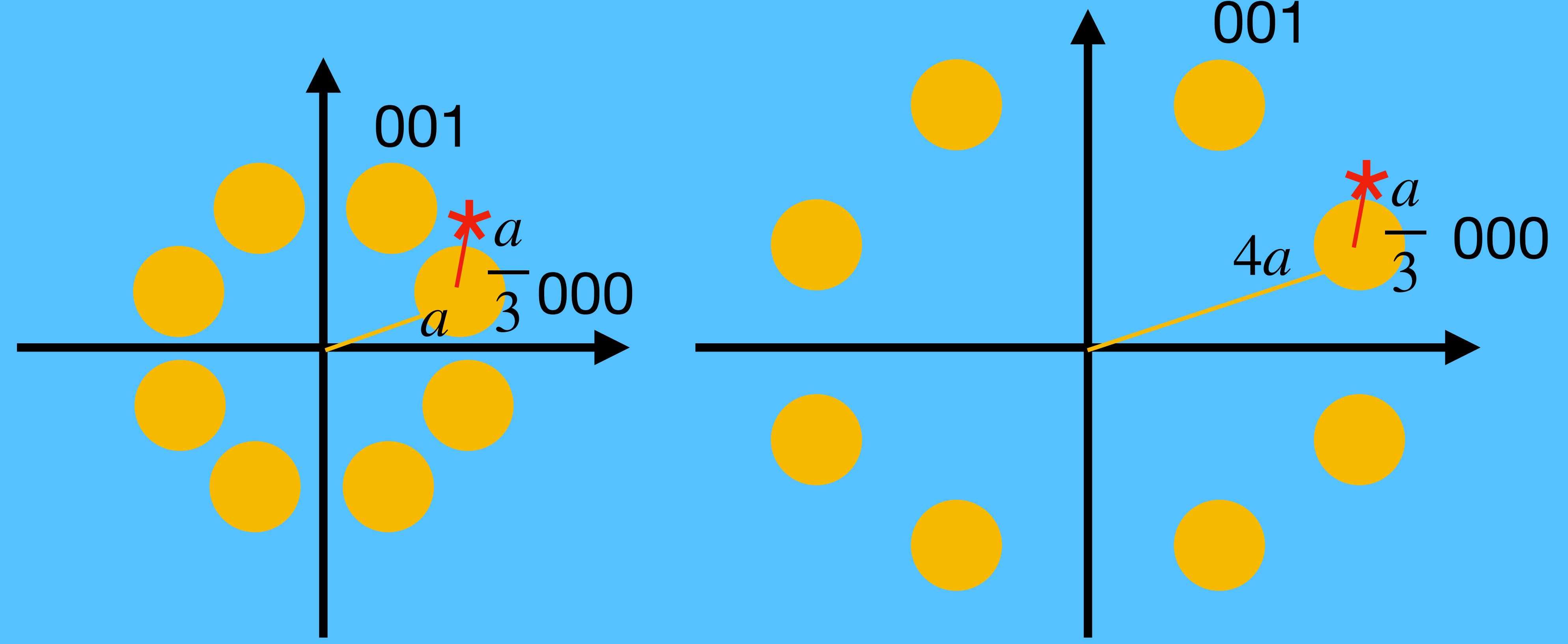
Symbol rate and bit rate

- **Symbol rate, SR**, is the number of symbols that can be transmitted per second.
- **Bit rate, BR**, is the number of bits that can be transmitted per second.
- Easy relationship: BR = SR x Transmission rate (number of bits encoded by a symbol, recall the Information Theory lecture).
  - Longer sequences encoded per symbol means higher bit rate

- **High order modulation allows more bits per symbol (potentially higher maximum channel capacity), but if symbols are close together in the diagram (i.e., they have similar amplitude/phase), small noise fluctuations can cause the receiver to decode the wrong symbol, increasing the Bit Error Rate (BER).**
  - $\text{BER} = \dfrac{\text{\# mistakenly received bits}}{\text{total \# transmitted bits}}$

    `10011011` ➡

    ➡ `00011111  BER=0.25`



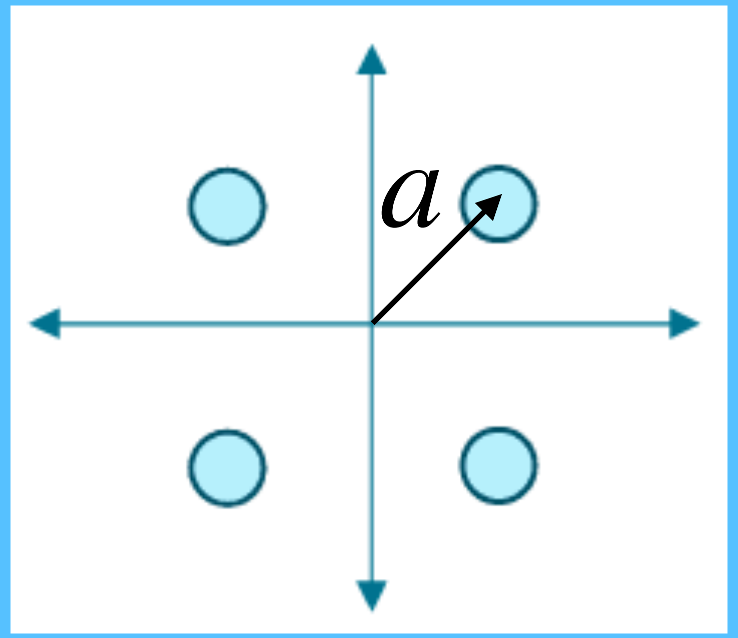
  - Note that the data rate of a channel is the actual speed at which data is transmitted over the channel, whereas its capacity is the maximum possible rate that can be achieved without errors.

- How can we decrease bit error rate?
  - **By increasing the amplitude of the symbols**, which means, **increase the points' power** (of course this has to correspond to larger distance in the diagram).

- The power of a symbol, or a point, i in the constellation diagram (P_i) is such that $P_i \sim |a_i|^2$.

- The power of the signal is the average of the power of all the points in the constellation diagram,

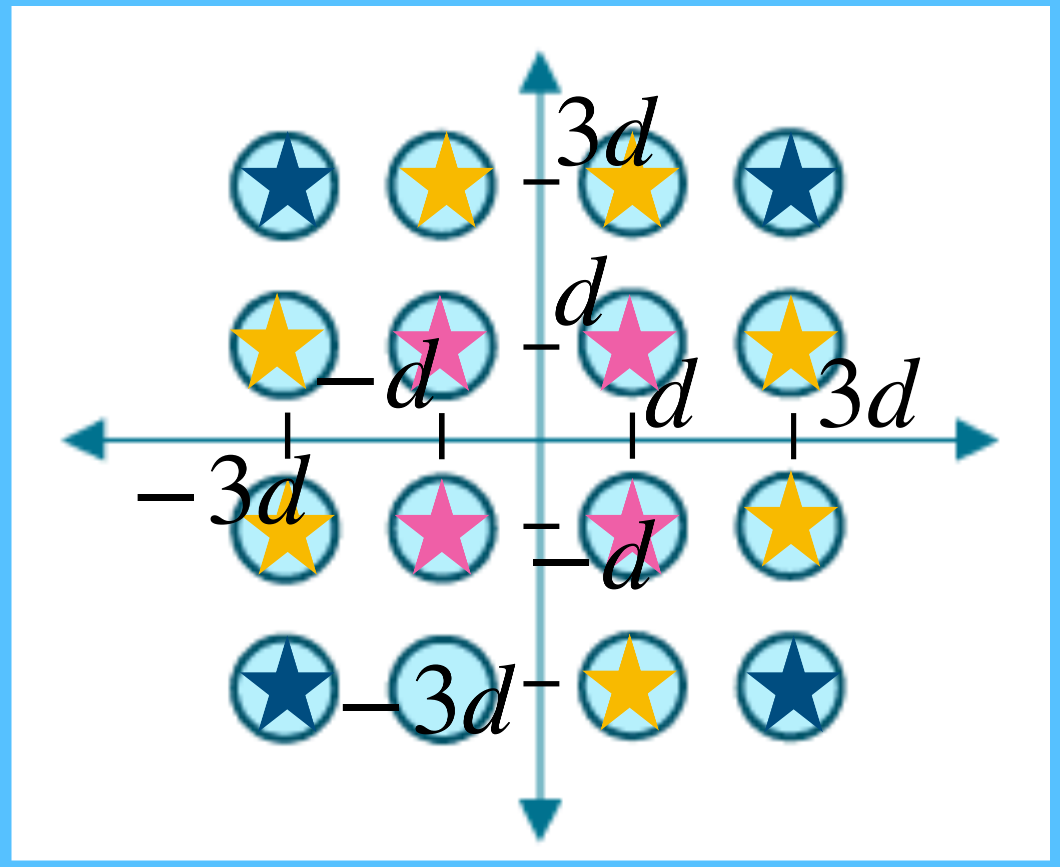
$$P_{signal} = \frac{1}{N}\sum_{i=1}^{N} P_i$$



QPSK

$P_i \sim a_i^2$

$P_{signal} \sim a_i^2$



16-QAM

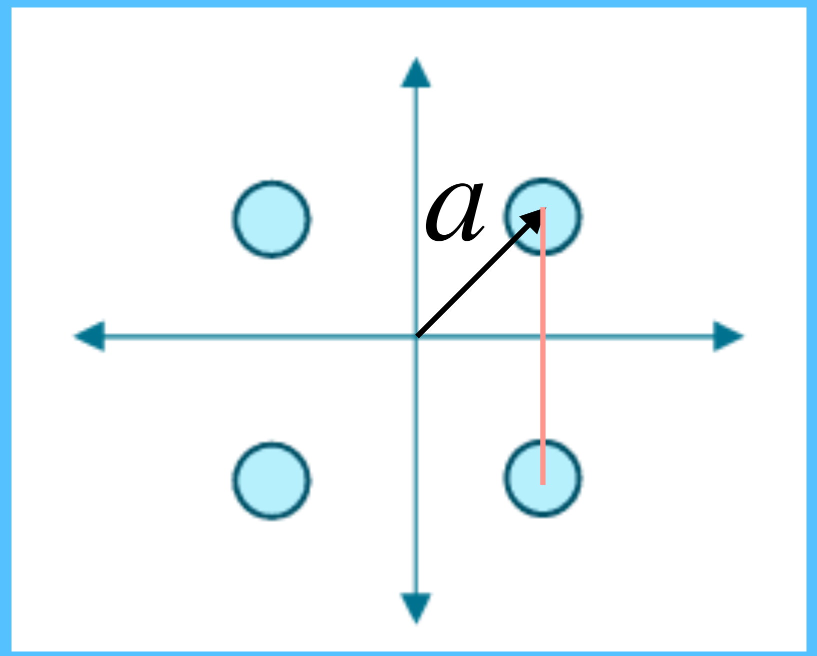
★ $P_i \sim 2d^2$

★ $P_i \sim 18d^2$

★ $P_i \sim 9d^2 + d^2$

$$P_{signal} \sim \frac{1}{16}(4 \cdot 2d^2 + 4 \cdot 18d^2 + 8 \cdot 10d^2)$$
$$= 5d^2$$

- Minimum distance in constellation diagram is the minimum Euclidean distance between two points (symbols).



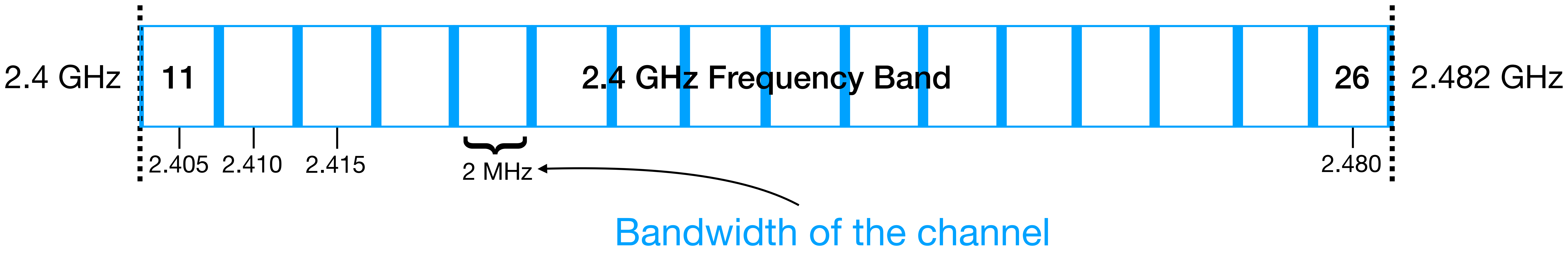
- So the Bit Error Rate decreases when SNR grows, since SNR $= \dfrac{P_{signal}}{P_{noise}}$.
    - Explicit formulas formalise the relation between BER, SNR and minimum distance (we will not see them).

- Is this surprising?
    - NO. **Shannon-Hartley theorem** says that $C = B \log_2(1 + SNR)$ so, to increase capacity with **modulation**, we still must increase bandwidth or SNR.
    - **Why more bandwidth means higher data rate?**

More about bandwidth and frequencies

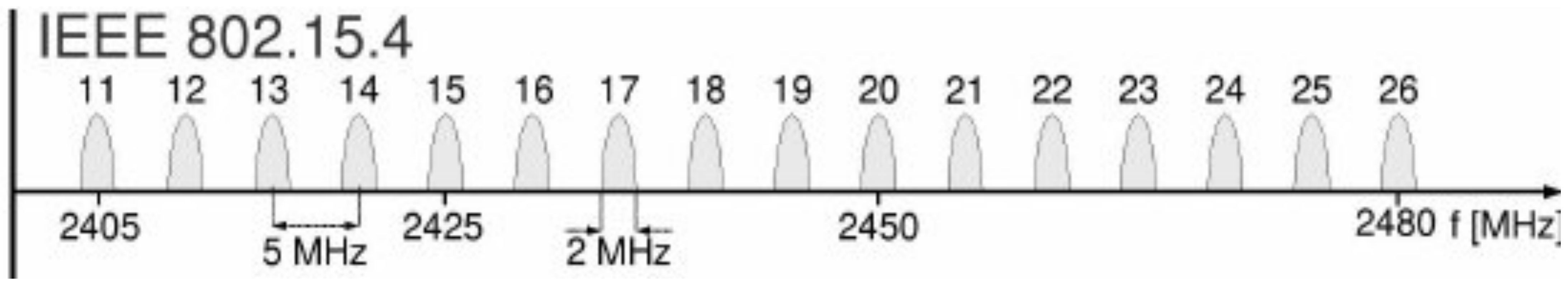
- We will see that protocols and standards operate at certain frequencies.
- For instance, the IEEE 802.15.4 standard can operate at 2.4GHz with 16 channels.
- 2.4GHz is the **frequency band**, which is not a single frequency, but a range of frequencies.
- In 802.15.4, the band is divided into 16 channels, each 2MHz wide, i.e., with 2MHz bandwidth.
- The center frequencies of these 16 channels range from 2.405 GHz to 2.480 GHz and are centered 5 MHz apart.

- In the 802.15.4 standard, channels are numbered from 11 to 26.



- Although numbers and parameters change, we can build similar schemes for other standards.

- Actually, 802.15.4 looks more like this



Frequency Division multiplexing

Back to our question

- Why more bandwidth means higher data rate?
  - Because more bandwidth means supporting more frequencies, hence more parallel channels!
- But we have seen many modulation schemes where the frequency is fixed and only the phase and/or the amplitude change!
  - We have seen that we need the signal to be powerful enough to increase the SNR, but still, in those cases, we can transmit a signal with 0 bandwidth!
  - …

Fourier series

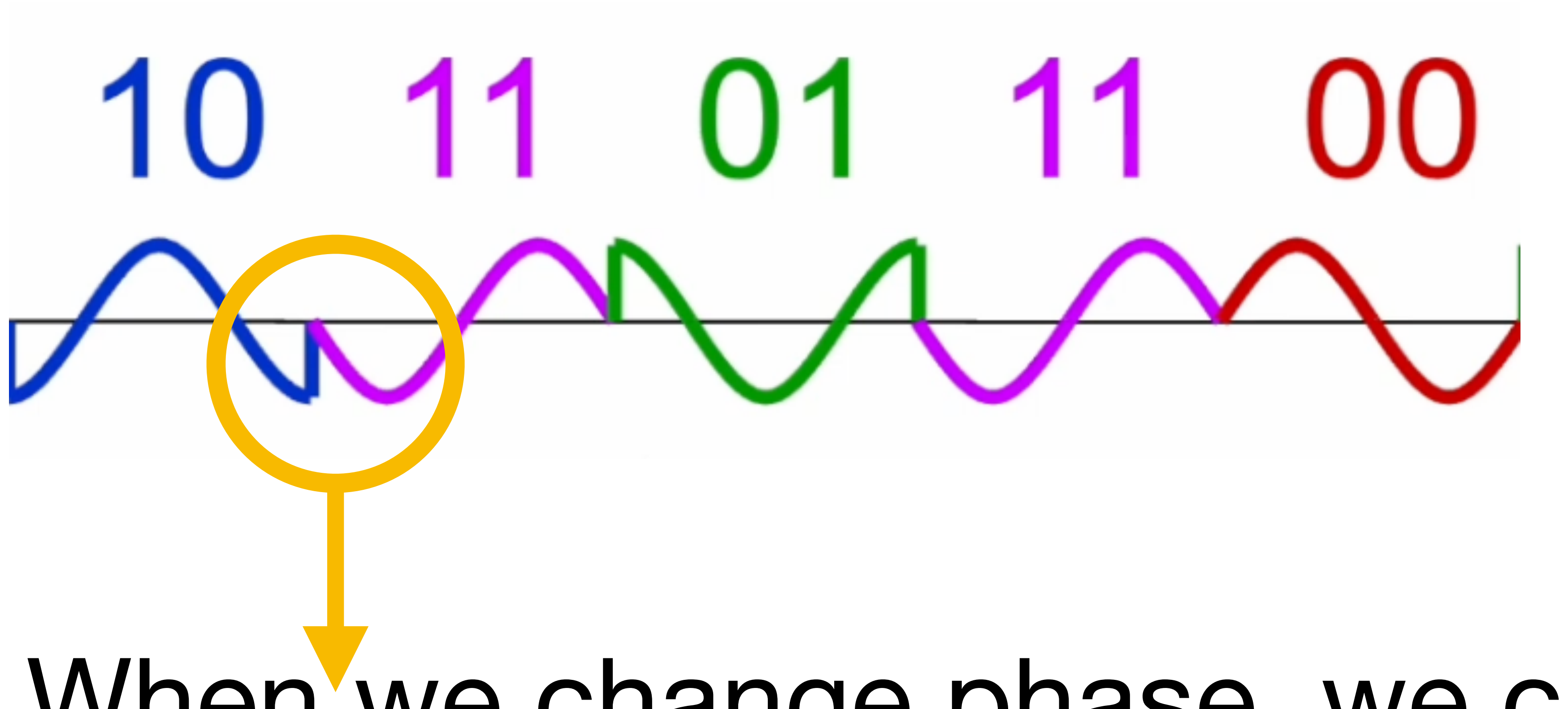
- Any periodic function $f(x)$ of period T and integrable in $[0,T]$ can be written as the following infinite sum of sines and cosines:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos(k\omega x) + b_k \sin(k\omega x) \right)$$

where $\omega = 2\pi/T$ and

$$a_k = \frac{2}{T}\int_0^T f(x)\cos(k\omega x)dx,\ b_k = \frac{2}{T}\int_0^T f(x)\sin(k\omega x)dx$$

Quadrature phase shift keying



- When we change phase, we can have discontinuity.
  To modulate discontinuity, we need high frequency waves!

- The Fourier Transform of a square wave (which approximates these phase flips) shows a wide range of frequency components. Fourier Series Animation Video

- This means that we need different frequencies (i.e., positive bandwidth) to enforce phase shifts, even for modulation schemes with fixed frequencies (like QPSK).

An experiment

- https://onlinetonegenerator.com/

Summary

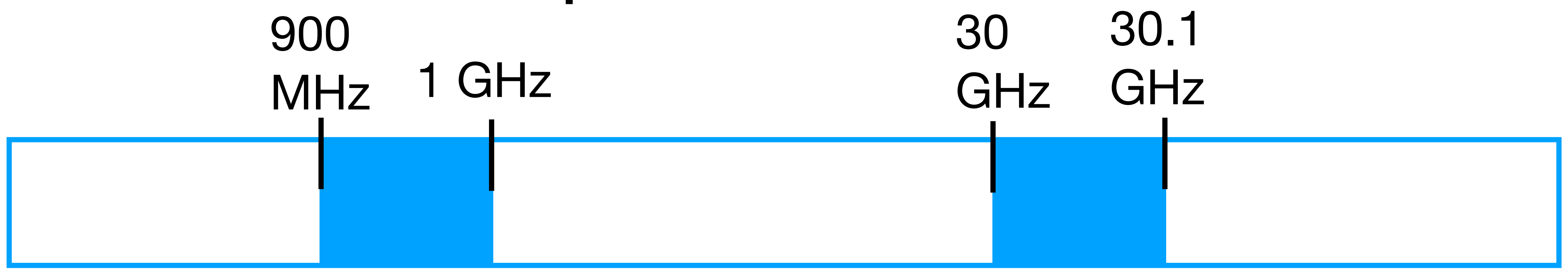
- More signal power implies higher SNR, hence less errors.
- More bandwidth means **more bits per symbol** and more parallel signals (both increase bit rate)
  - Both increase the maximum channel capacity (bandwidth does it more)!
- Devices supporting large bandwidth signals require more complex circuits that consume more electrical power.

Channel multiplexing

- We can decide to split the frequency band into **many channels with narrow bands**.
  - More parallelism, i.e., more users can share the band
  - Narrowband signals require less power
  - Lower data rate in each channel
- … or into **fewer wideband channels**.
  - Higher bandwidth per channel (more data rate in each channel), efficient for high-speed applications
  - Less users can use the channel simultaneously and require more power

Last note on channel bandwidth

- We have seen that the bandwidth of a channel is the difference between the highest supported frequency f_H and the lowest supported frequency f_l.
- Consider two channels with the same bandwidth B and SNR, but operating at different frequencies.



B = 100 MHz, does not depend on absolute values of f_H and f_l

- The Shannon Hartley theorem says that the maximum bps that they can achieve is the same.
- Nevertheless, in practice, channels **operating at higher frequencies can deliver more bps** (more efficient antenna design, modulation, wider available bandwidth - low frequency bands are crowded).

- Additional modulation schemes exist. We will talk about some of them as we encounter them.

END OF FIRST MIDTERM PROGRAM

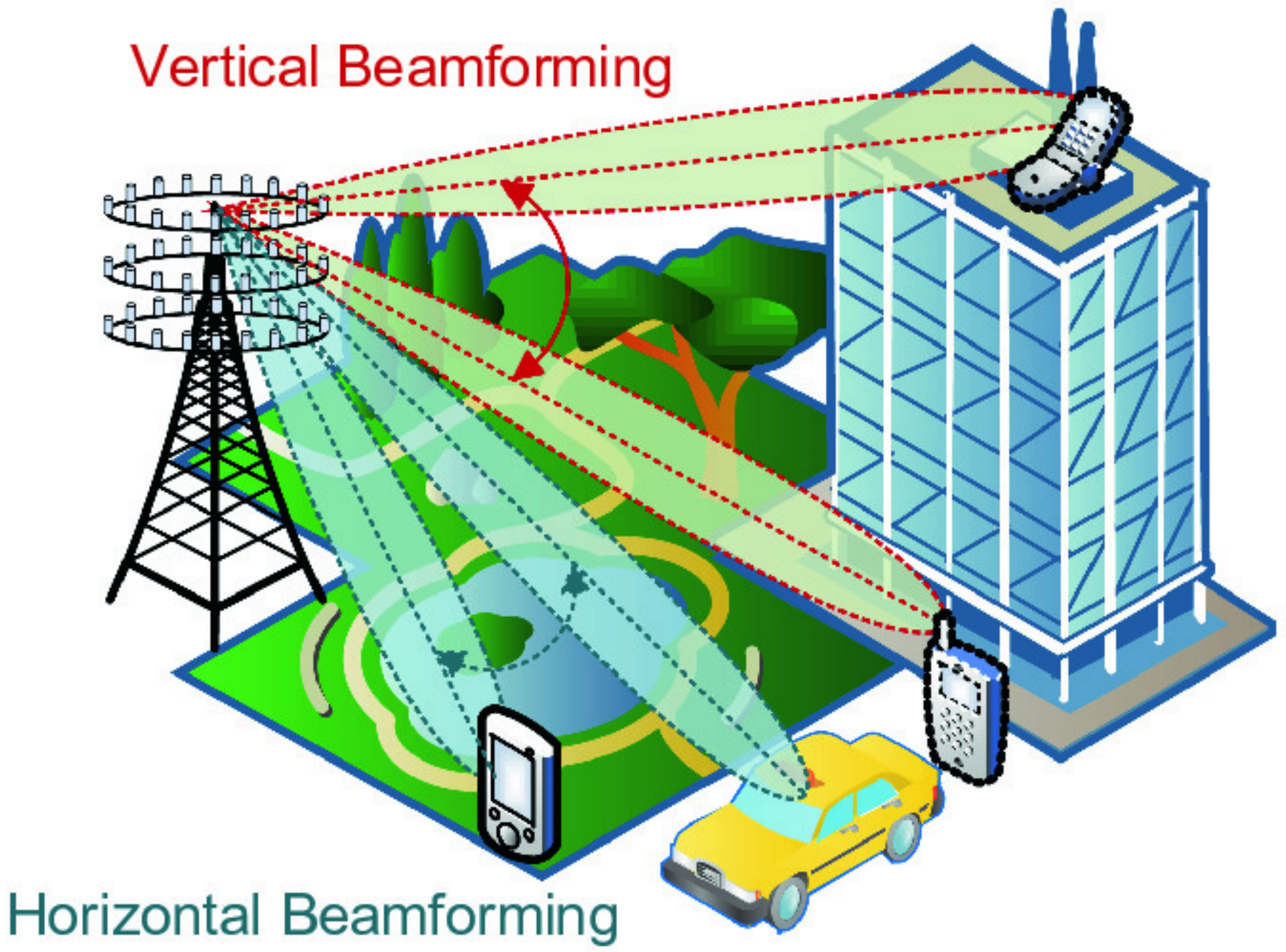
5.1.7 Multiplexing

Multiplexing in wireless communication

- Multiplexing is a transmission technique that allows to send multiple different signals over the same channel.
- Allows parallel communication with multiple users, even without MIMO.
- Leveraged by MAC layer protocols for managing the multiple access.
    - Space division multiplexing (SDM)
    - Frequency division multiplexing (FDM)
    - Time division multiplexing (TDM)
    - Code division multiplexing (CDM)

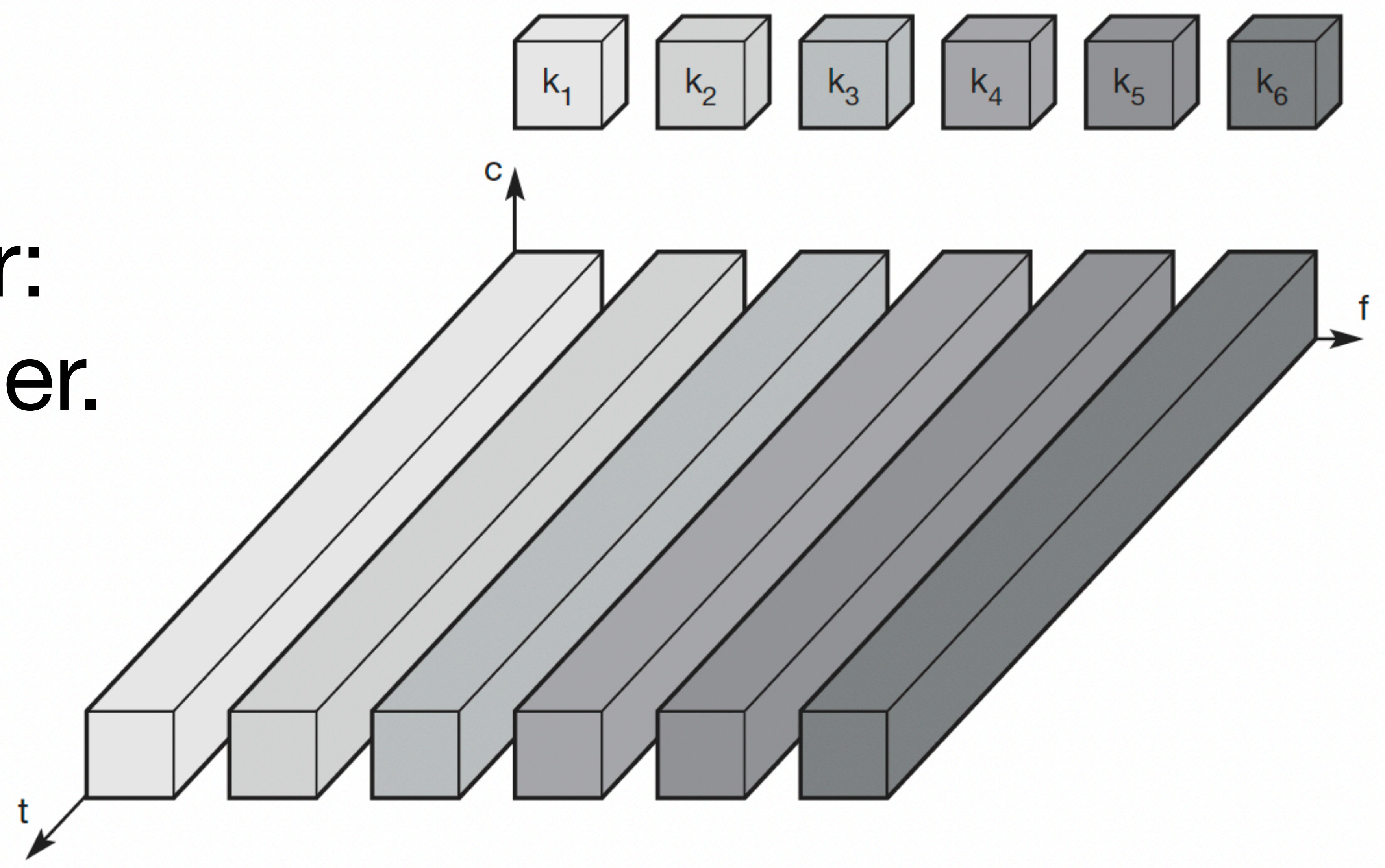
Space Division Multiplexing (SDM)

- SPD is achieved with MIMO technology.
- The channels that are created between a receiver and a transmitter are physically separated, as antennas at the receiver and the transmitter are at a certain distance from each other and can be oriented towards different directions.
  - Smart antennas, i.e., arrays of antennas connected with each other at the same node, can coordinate and adjust themselves and perform beamforming, to send and receive multiple signals in parallel.
- Not to be confused with spatial multiplexing



Frequency Division Multiplexing (FDM)

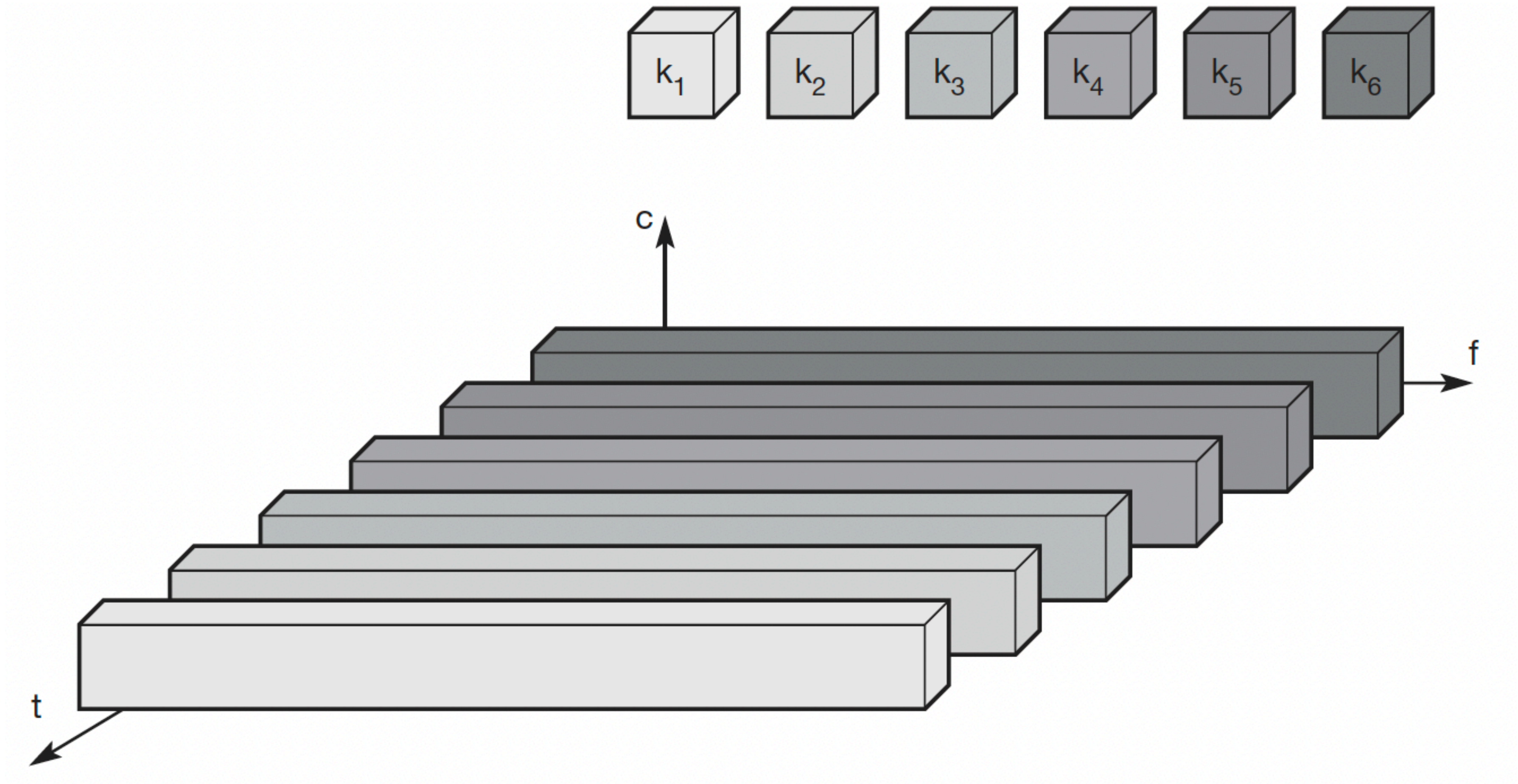
- Frequency division multiplexing (FDM) subdivides the frequency range into several non-overlapping frequency bands (mentioned it before).
- Different sub channels have their own frequency band.
- Senders using a certain frequency band can use this band continuously.
- Guard bands are needed to avoid frequency band overlapping (“adjacent channel interference”).
- Simple multiplexing scheme, does not need complex coordination between sender and receiver: the receiver only has to tune in to the specific sender.
  - works good for FM radios, but not for mobile communication (imagine having one channel for each mobile device).
- Does not require MIMO.

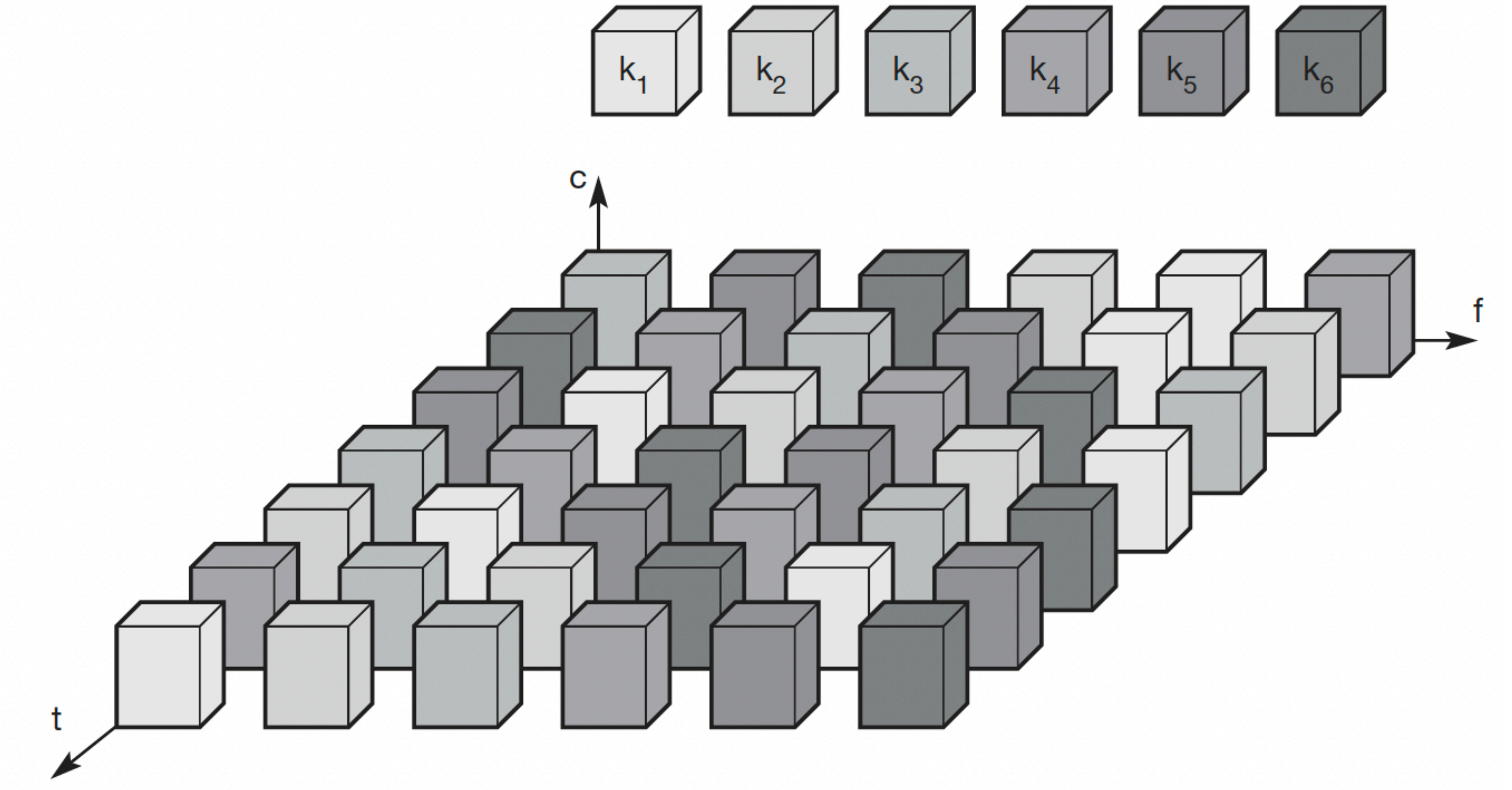


Time Division Multiplexing (TDM)

- Each channel is given the whole bandwidth for a certain amount of time.
- All senders use the same frequency but at different points in time.
- Guard spaces (i.e., time gaps) have to separate the time slots when the senders use the medium.
- Does not require MIMO.

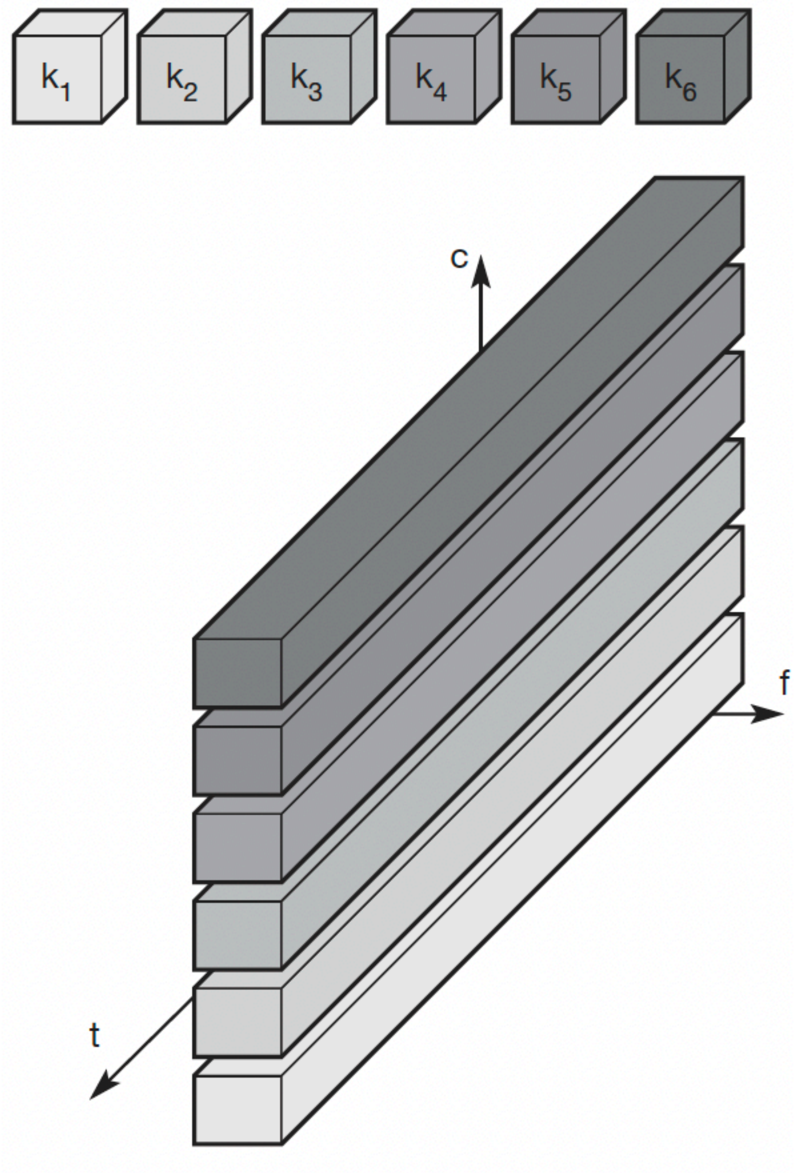
- Can be combined with FDM.





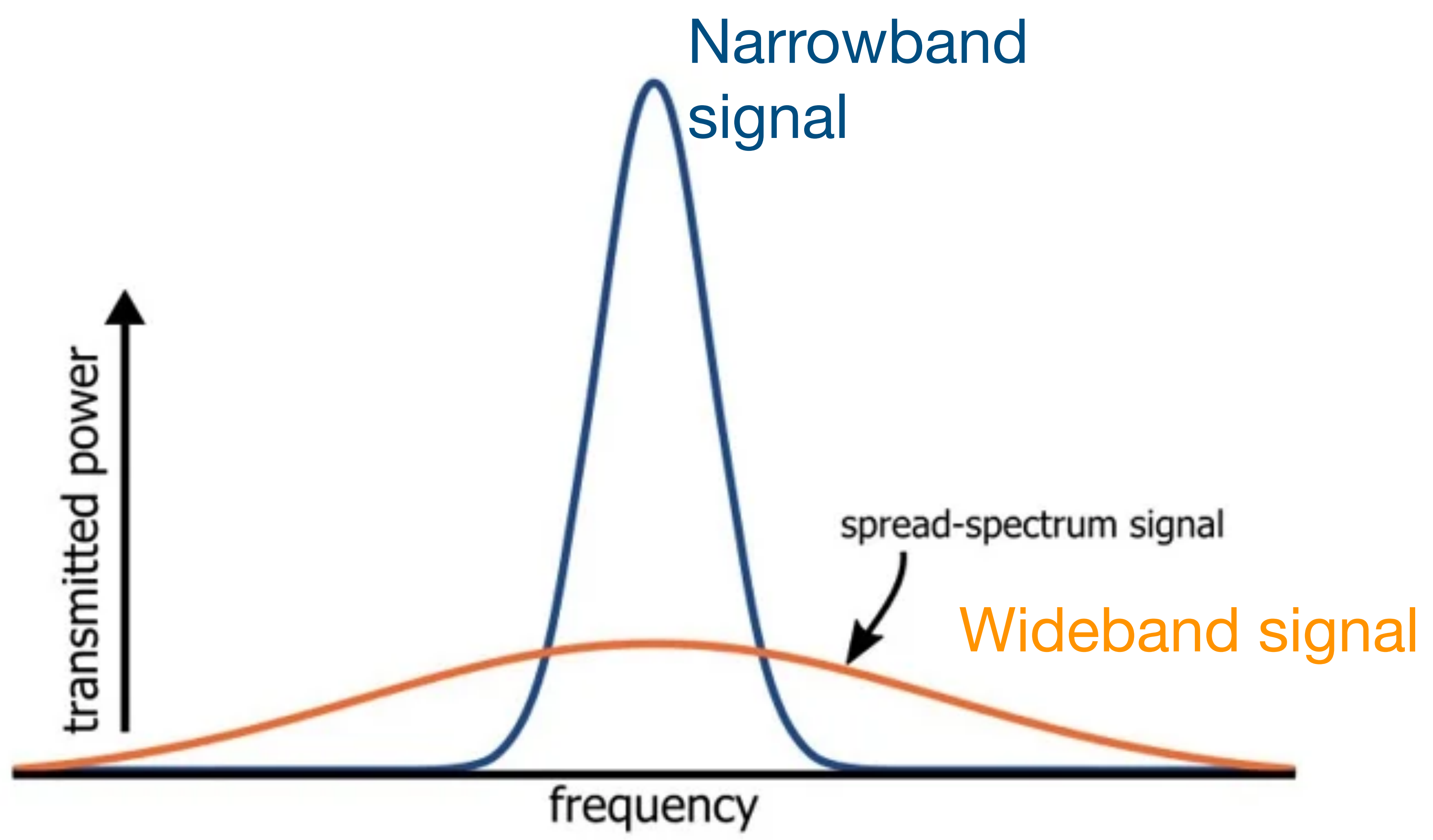
Code Division Multiplexing (1)

- Code Division Multiplexing allows different channels to use the same frequency and the same time for transmission.
- Separation is now achieved by assigning each channel its own 'code'.
- Every day life example: people in the same room talking simultaneously at the same voice level in different languages.
- To achieve this, CDM uses **spread spectrum techniques**.



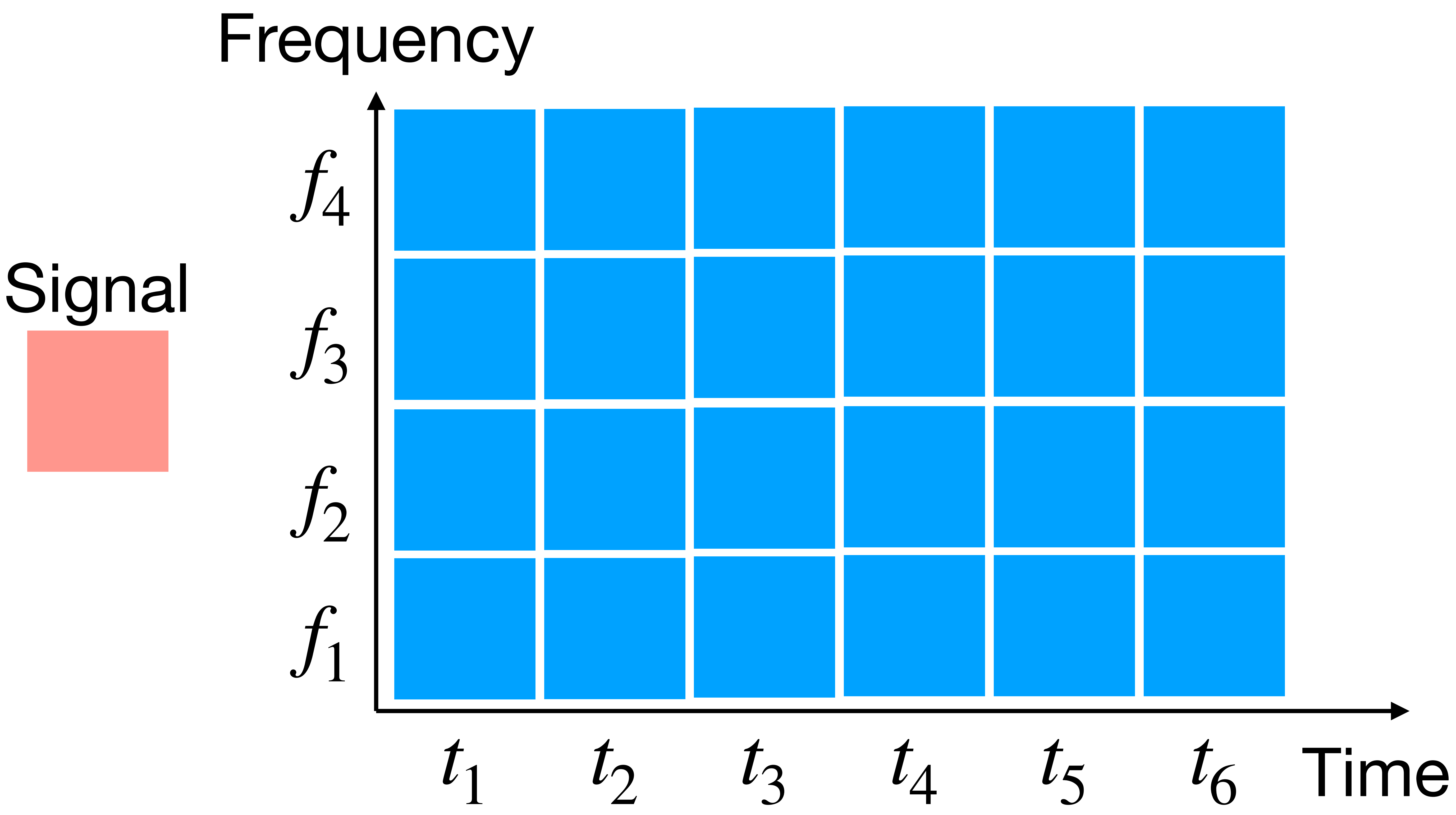
Spread Spectrum

- Spread spectrum techniques involve spreading the bandwidth to transmit data.
- Narrowband signals are transformed into spread-spectrum signals.
  - The power of the signal remains the same
  - The resulting signal is barely distinguishable from noise (higher security).
  - The receiver has a way to distinguish the original signal.
- Two main techniques: Frequency Hopping SS (FHSS) and Direct Sequence SS (DSSS).



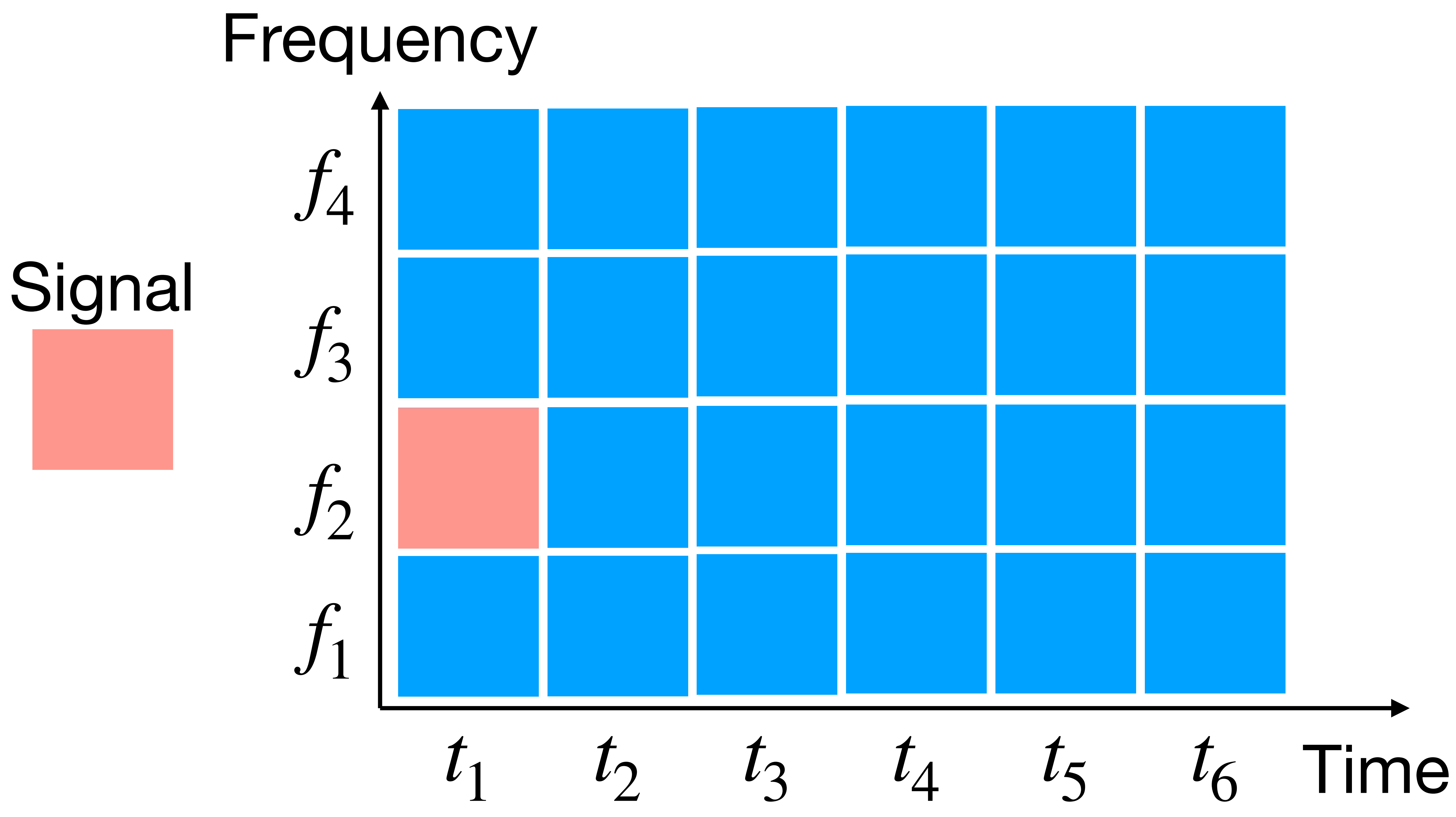
Frequency Hopping Spread Spectrum (FHSS) (1)

- Uses a combination of Frequency Division Multiplexing and Time Division Multiplexing, with guard bands.



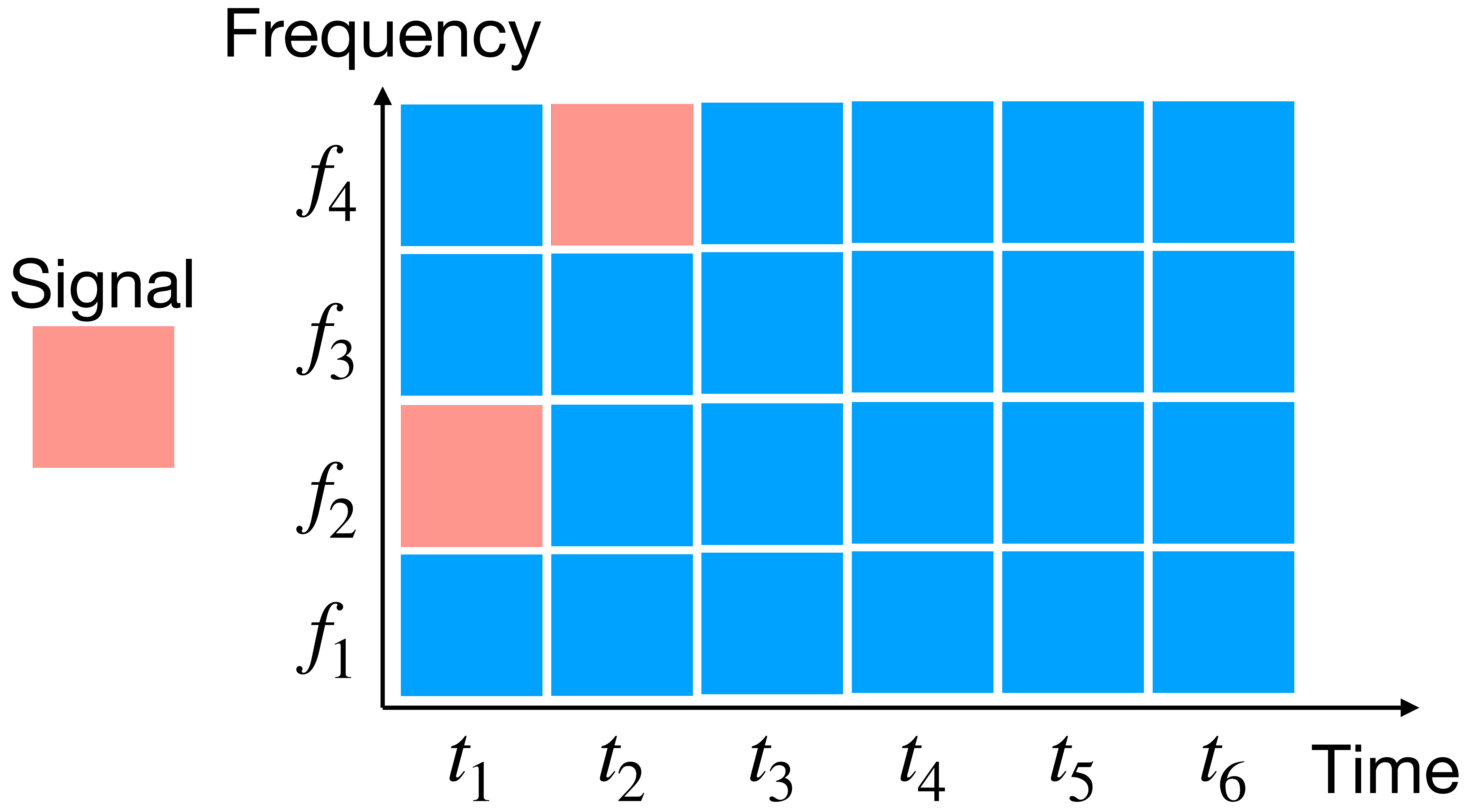
Frequency Hopping Spread Spectrum (FHSS) (1)

- Uses a combination of Frequency Division Multiplexing and Time Division Multiplexing, with guard bands.



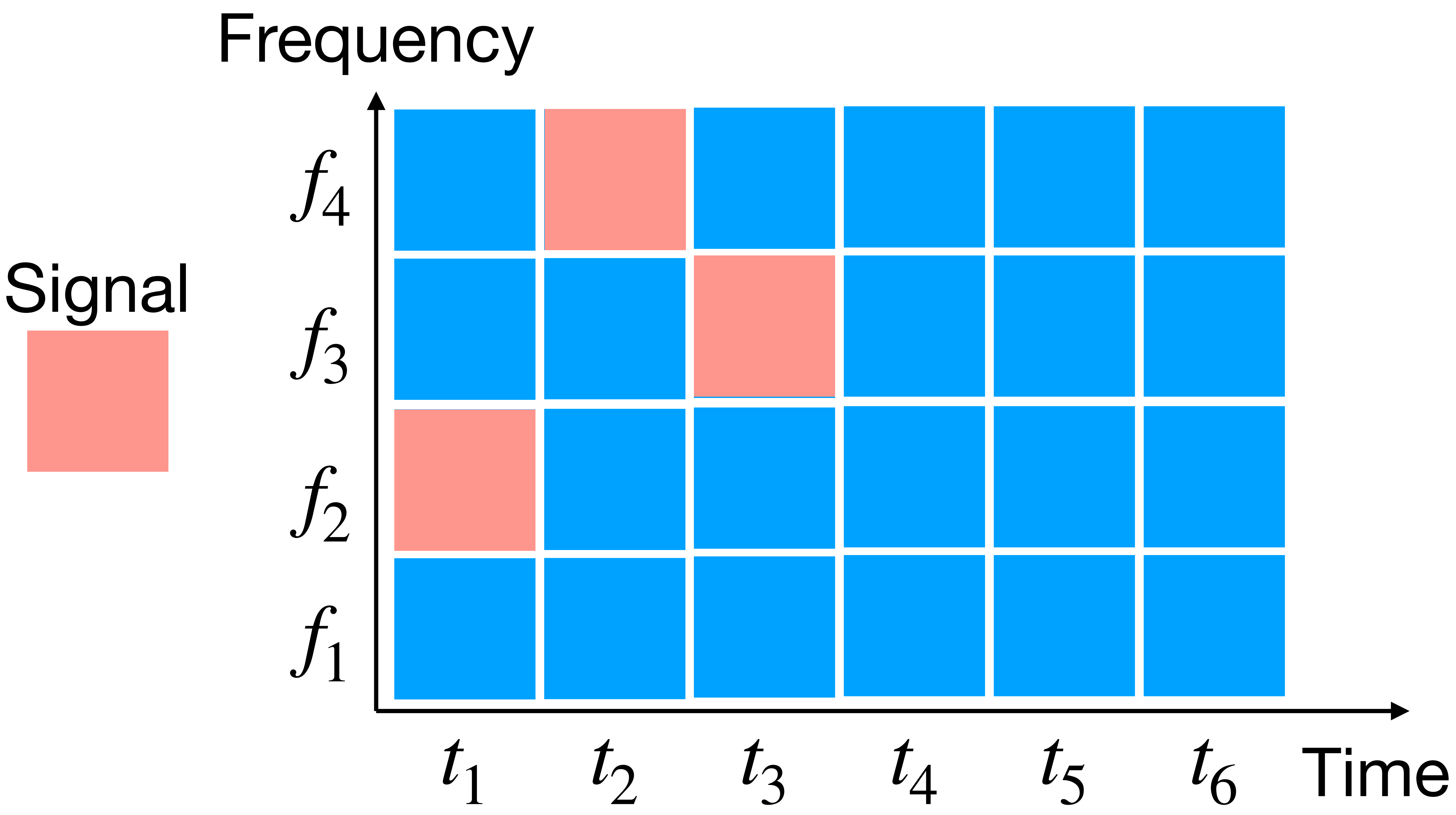
Frequency Hopping Spread Spectrum (FHSS) (1)

- Uses a combination of Frequency Division Multiplexing and Time Division Multiplexing, with guard bands.



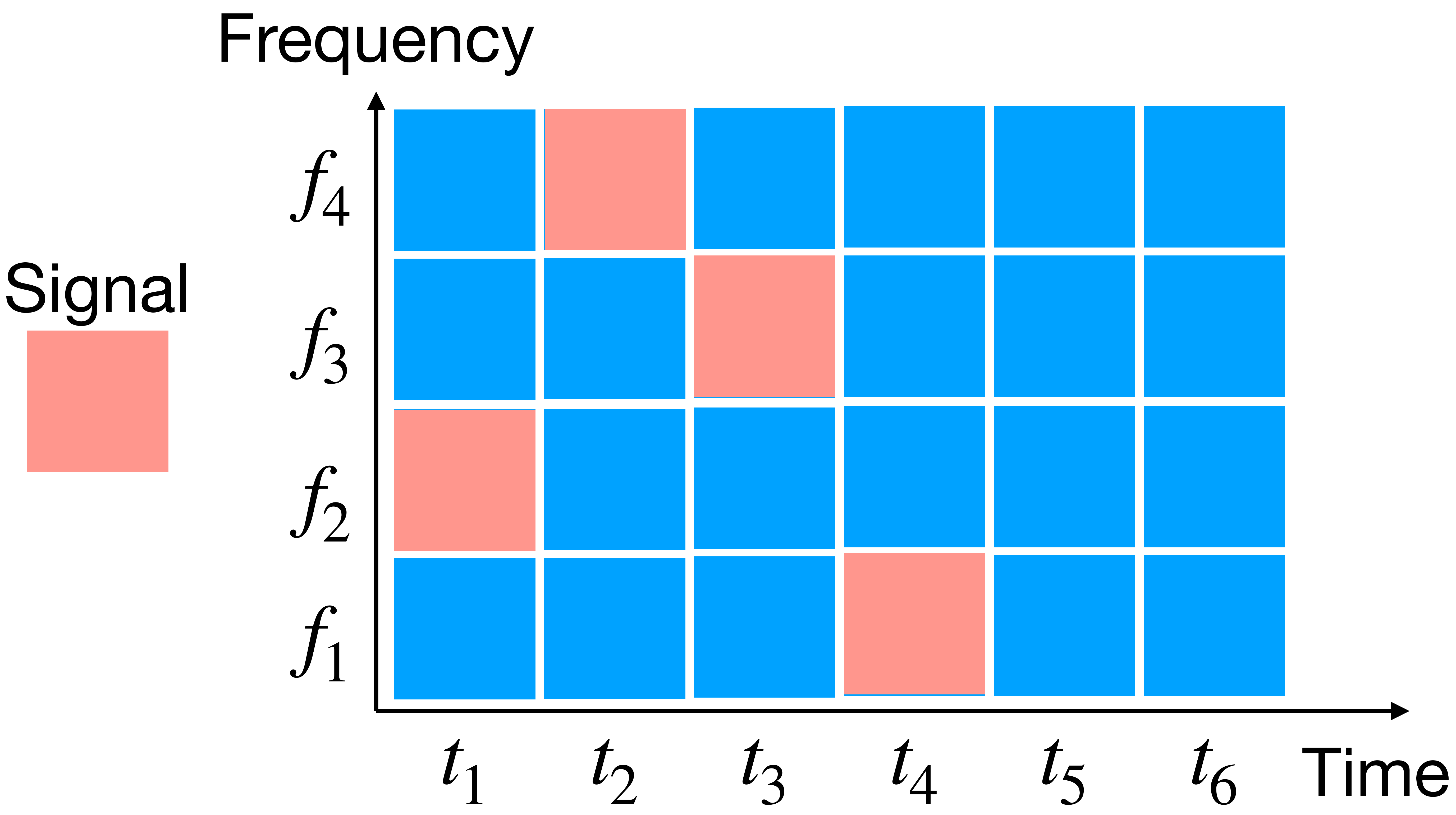
Frequency Hopping Spread Spectrum (FHSS) (1)

- Uses a combination of Frequency Division Multiplexing and Time Division Multiplexing, with guard bands.



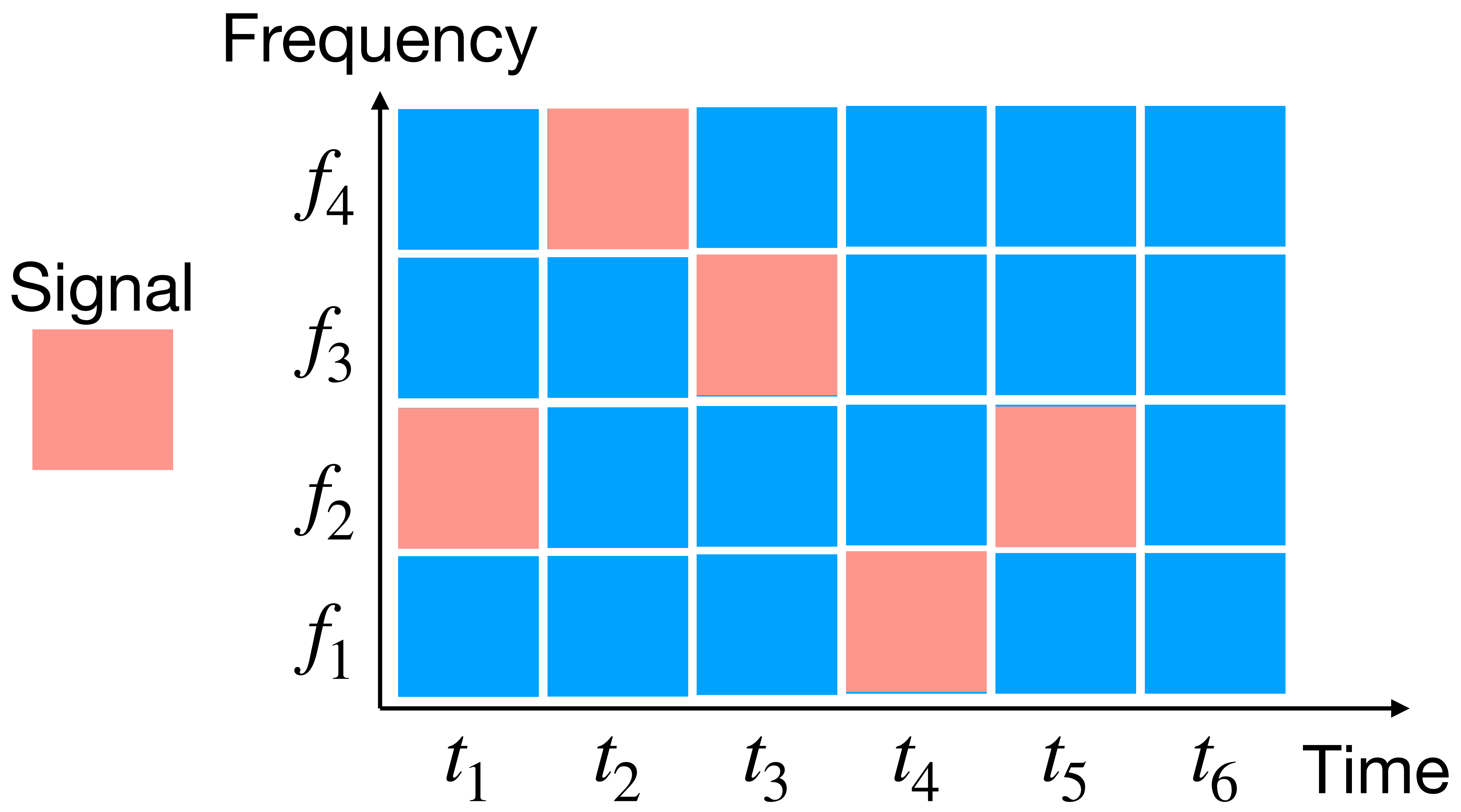
Frequency Hopping Spread Spectrum (FHSS) (1)

- Uses a combination of Frequency Division Multiplexing and Time Division Multiplexing, with guard bands.



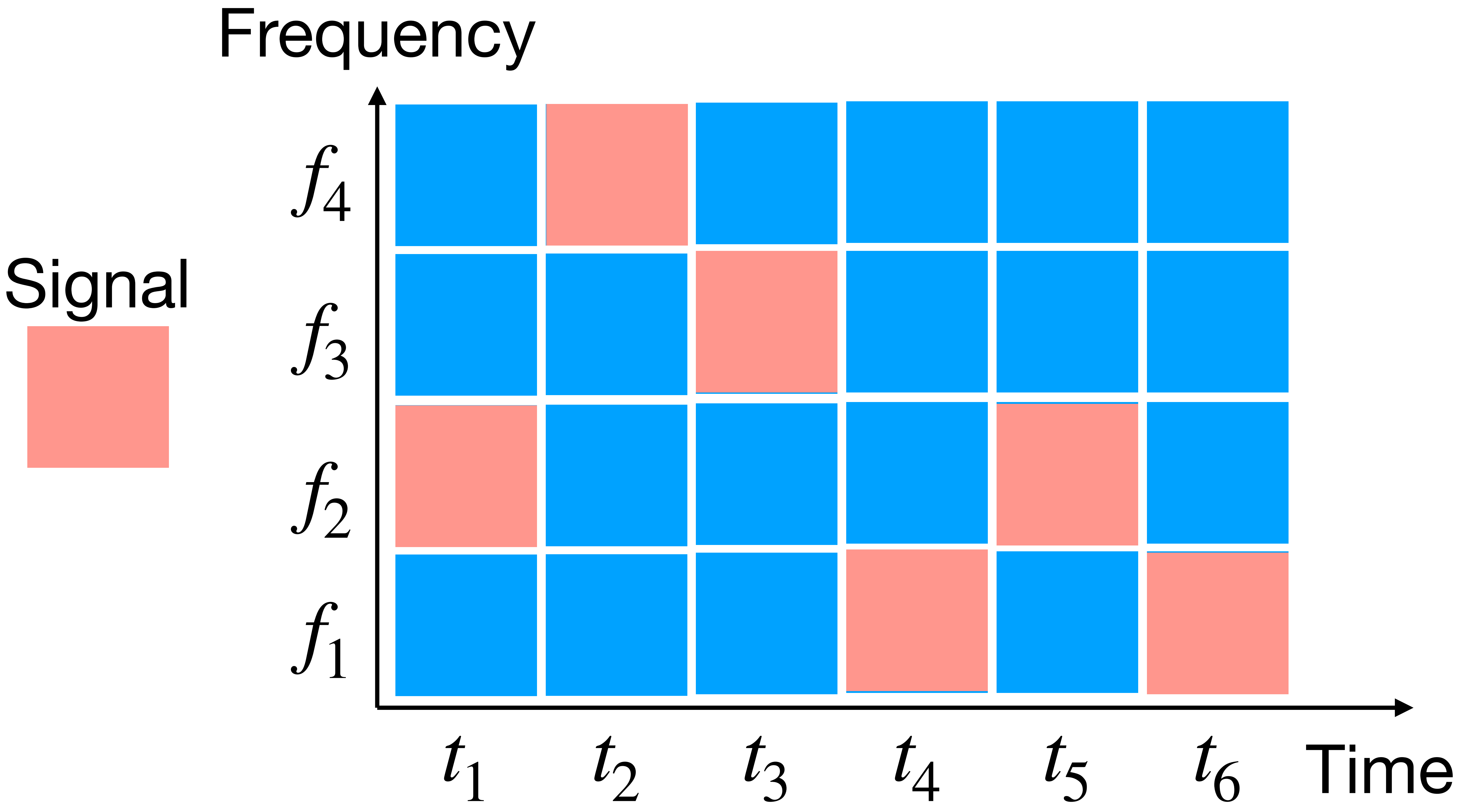
Frequency Hopping Spread Spectrum (FHSS) (1)

- Uses a combination of Frequency Division Multiplexing and Time Division Multiplexing, with guard bands.



Frequency Hopping Spread Spectrum (FHSS) (1)

- Uses a combination of Frequency Division Multiplexing and Time Division Multiplexing, with guard bands.



The pattern of channel usage is called **hopping sequence.**

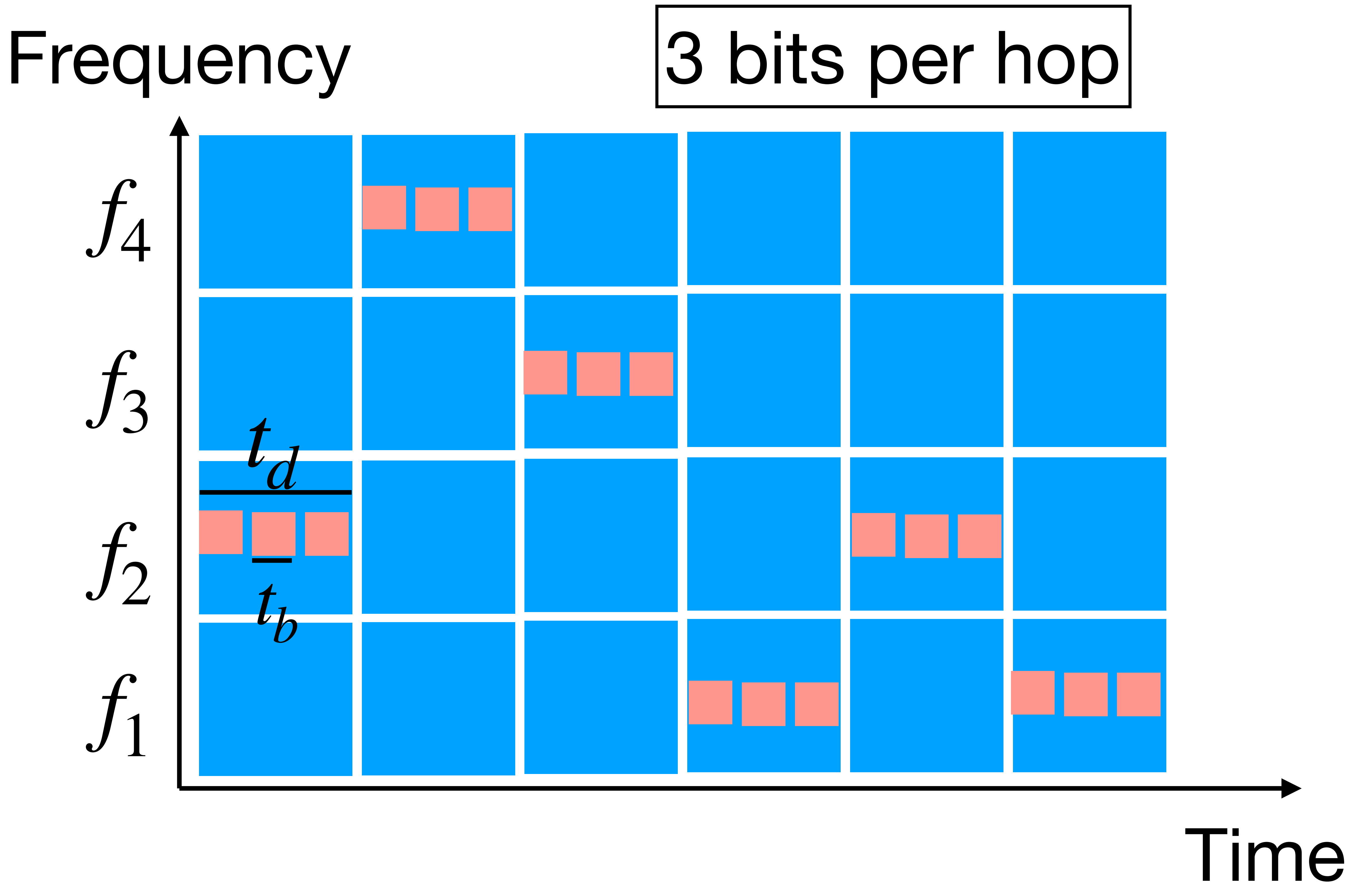
$$f_2 \; f_4 \; f_3 \; f_1 \; f_2 \; f_1$$

The time spent on a channel with a certain frequency is called **dwell time,** t_d.

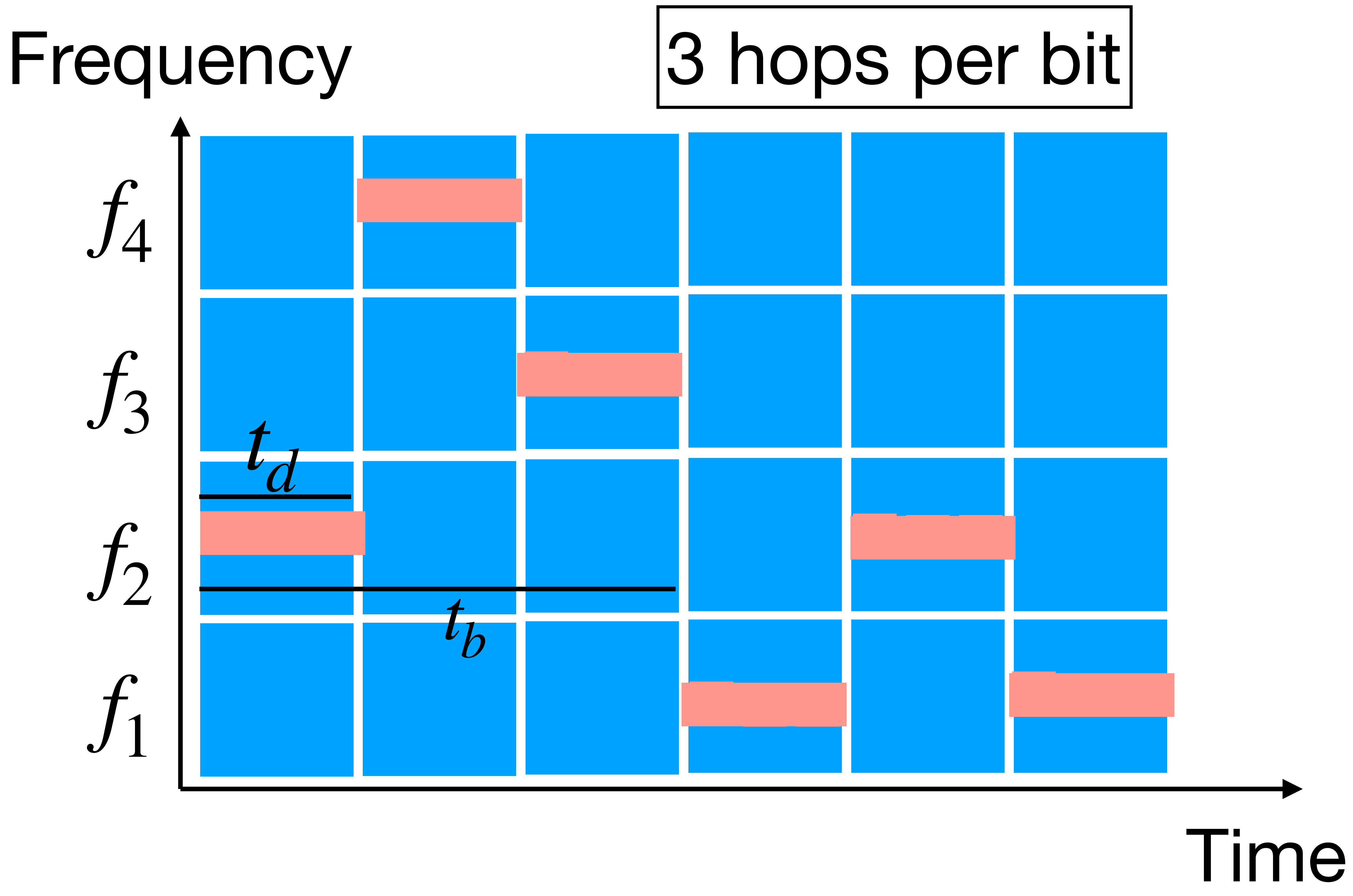
The time required for sending one bit is t_b.

Frequency Hopping Spread Spectrum (FHSS) (2)

- If t_b **<** t_d -> SLOW HOPPING



- If t_b **>** t_d -> FAST HOPPING

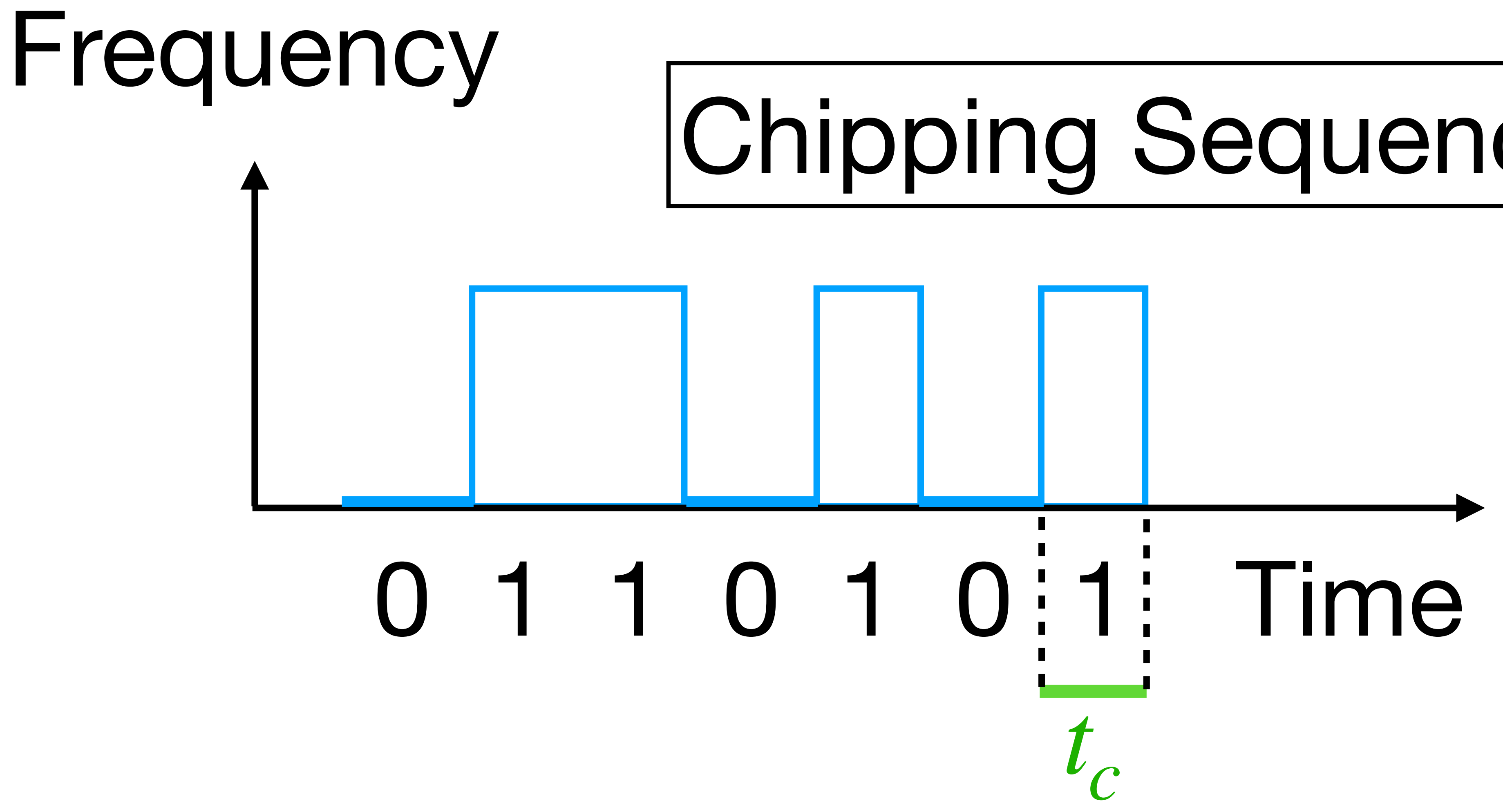


The receiver of an FHSS system has to know the hopping sequence and must stay synchronized.

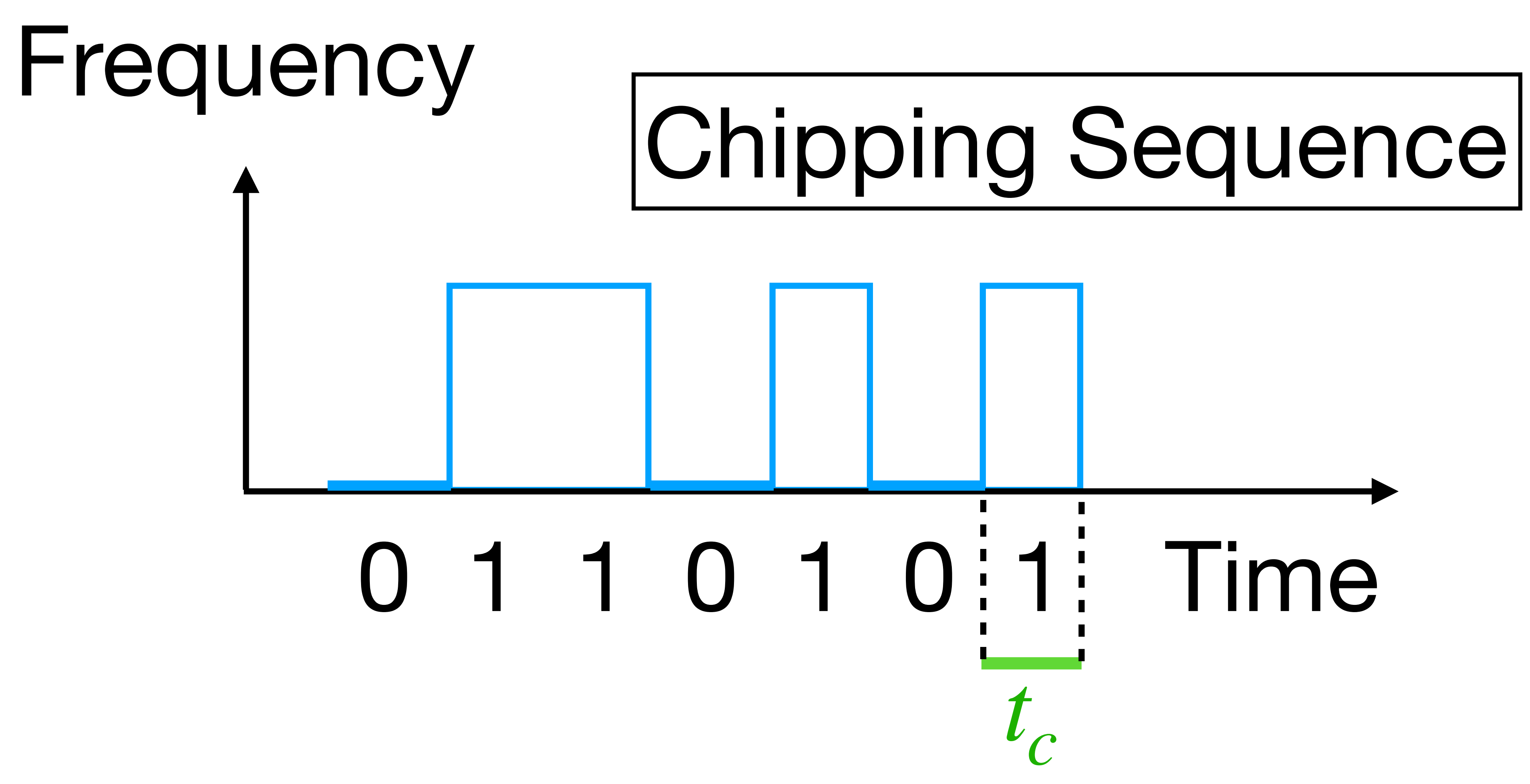
Direct Sequence Spread Spectrum (DSSS) (1)

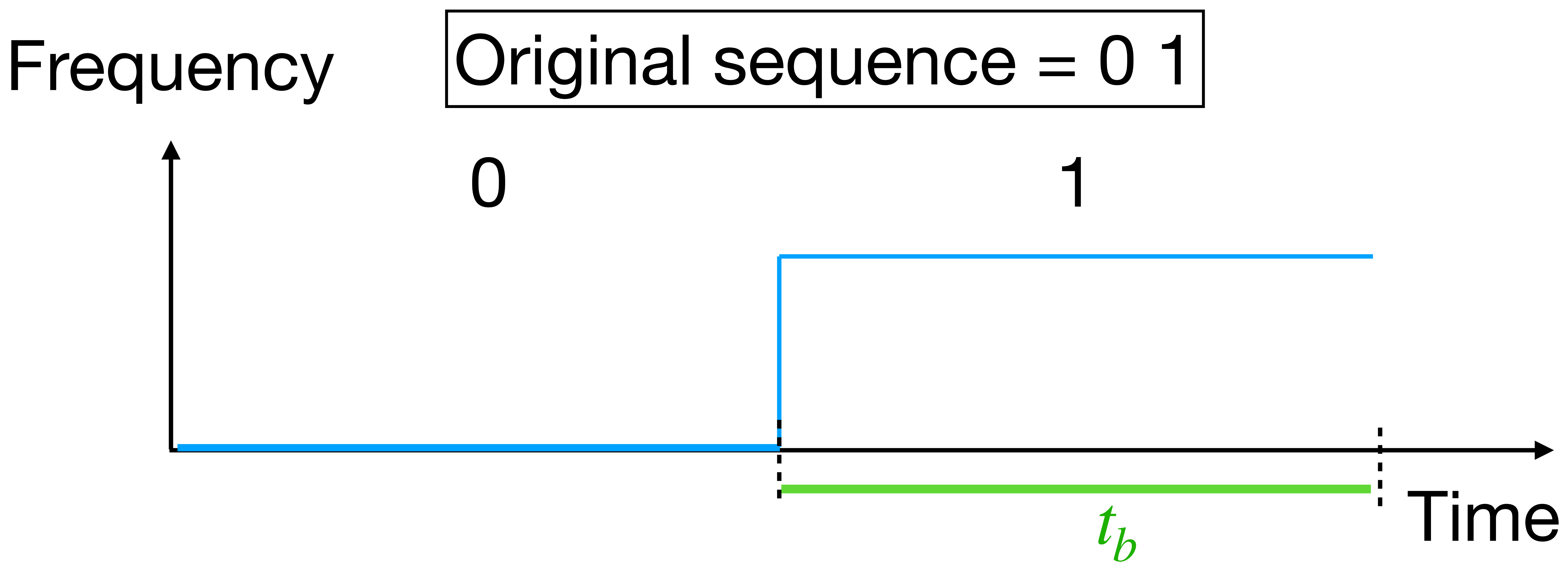
- Direct sequence spread spectrum (DSSS) systems take a user bit stream and perform an (XOR) with a so-called **chipping sequence.**
- It consists of a sequence of smaller pulses, called **chips**, with a duration t_c.
- t_b: duration of a bit (inverse of the bit rate). t_b **>** t_c**.**
- If the chipping sequence is generated properly it appears as random noise.

Direct Sequence Spread Spectrum (DSSS) (2)

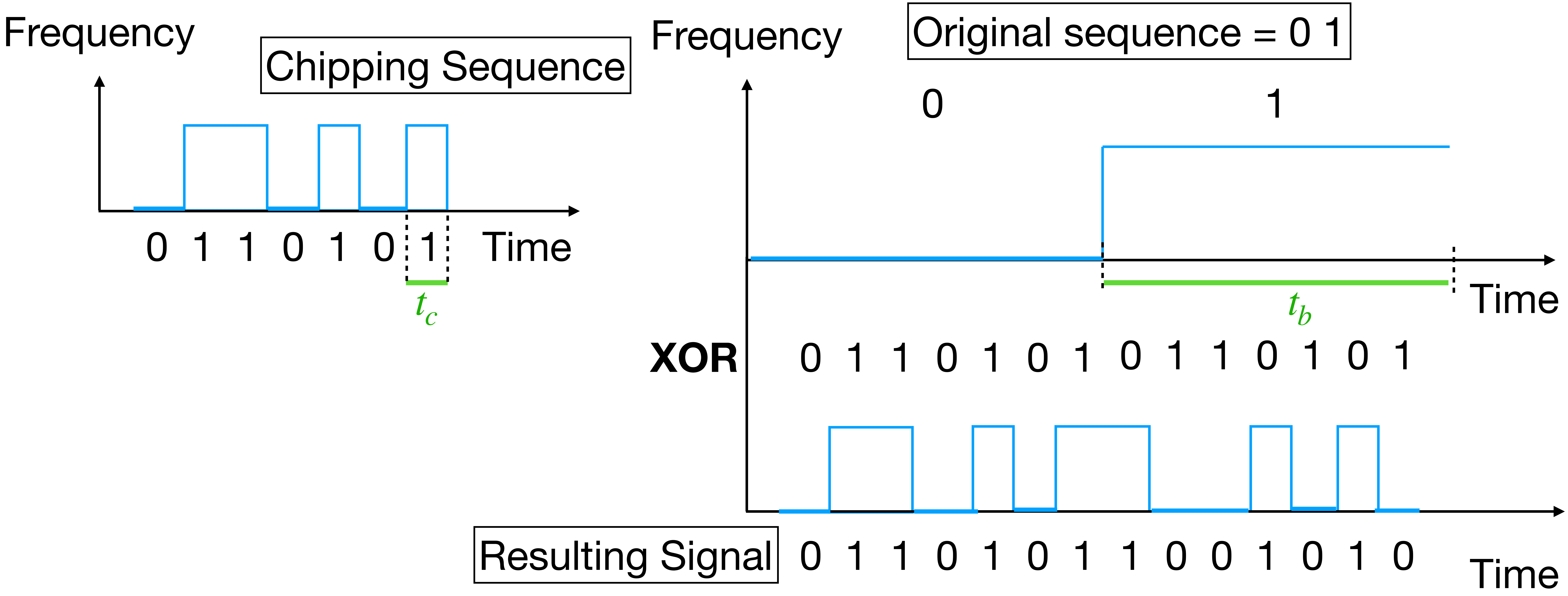


Direct Sequence Spread Spectrum (DSSS) (2)





Direct Sequence Spread Spectrum (DSSS) (2)



Direct Sequence Spread Spectrum (DSSS) (3)

- The receiver generates the same pseudo random chipping sequence as the transmitter.
- Sequences at the sender and receiver have to be precisely synchronized: the receiver calculates the product of the chipping sequence with the incoming signal.

Received signal	0 1 1 0 1 0 1 1 0 0 1 0 1 0
	XOR
Chipping sequence (repeated)	0 1 1 0 1 0 1 0 1 1 0 1 0 1
	0 0 0 0 0 0 0 1 1 1 1 1 1 1
	0 1

Direct Sequence Spread Spectrum (DSSS) (3)

- The receiver generates the same pseudo random chipping sequence as the transmitter.
- Sequences at the sender and receiver have to be precisely synchronized: the receiver calculates the product of the chipping sequence with the incoming signal.

Received signal	0	1	1	0	1	0	1	1	0	0	1	0	1	0
							XOR							
Chipping sequence (repeated)	0	1	1	0	1	0	1	0	1	1	0	1	0	1
	0	0	0	0	0	0	0	1	1	1	1	1	1	1
				0							1			

During propagation, errors might occur here

Direct Sequence Spread Spectrum (DSSS) (3)

- The receiver generates the same pseudo random chipping sequence as the transmitter.
- Sequences at the sender and receiver have to be precisely synchronized: the receiver calculates the product of the chipping sequence with the incoming signal.

Received signal 0 1 1 0 1 0 1 1 0 0 1 0 1 0

During propagation, errors might occur here

XOR

Chipping sequence (repeated) 0 1 1 0 1 0 1 0 1 1 0 1 0 1

---

0 0 0 0 0 0 0 1 1 1 1 1 1 1

0 1

Resulting in ambiguous information here

Direct Sequence Spread Spectrum (DSSS) (3)

- The receiver generates the same pseudo random chipping sequence as the transmitter.

- Sequences at the sender and receiver have to be precisely synchronized: the receiver calculates the product of the chipping sequence with the incoming signal.

Received signal 0 1 1 0 1 0 1 1 0 0 1 0 1 0

During propagation, errors might occur here

XOR

Chipping sequence (repeated) 0 1 1 0 1 0 1 0 1 1 0 1 0 1

The receiver has correlator and decision modules

---

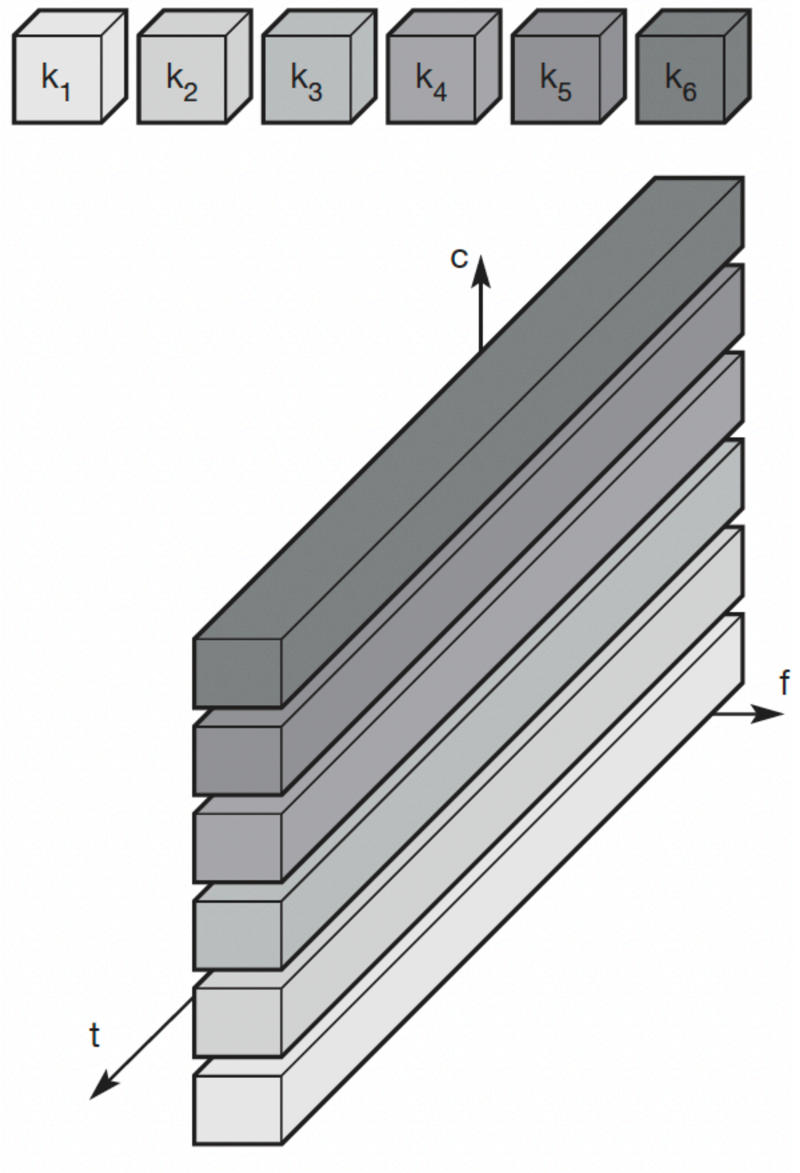
0 0 0 0 0 0 0 1 1 1 1 1 1 1

0 1

Resulting in ambiguous information here

Code Division Multiplexing (2)

- Each channel is assigned with a different code, which can be either a different hopping sequence or a different chipping sequence, depending on the spread spectrum technique used.
- In DSSS, to minimize overlaps between channels, chip codes must distant enough, in particular, they should be orthogonal (their dot product should be equal to 0).



Bibliography

- More about antennas components: https://www.industrialnetworking.com/pdf/Antenna-Patterns.pdf (Cisco)
- Jochen H. Schiller. “Mobile Communications” ADDISON-WESLEY
- IoT Communication, Illinois Urbana-Champaign