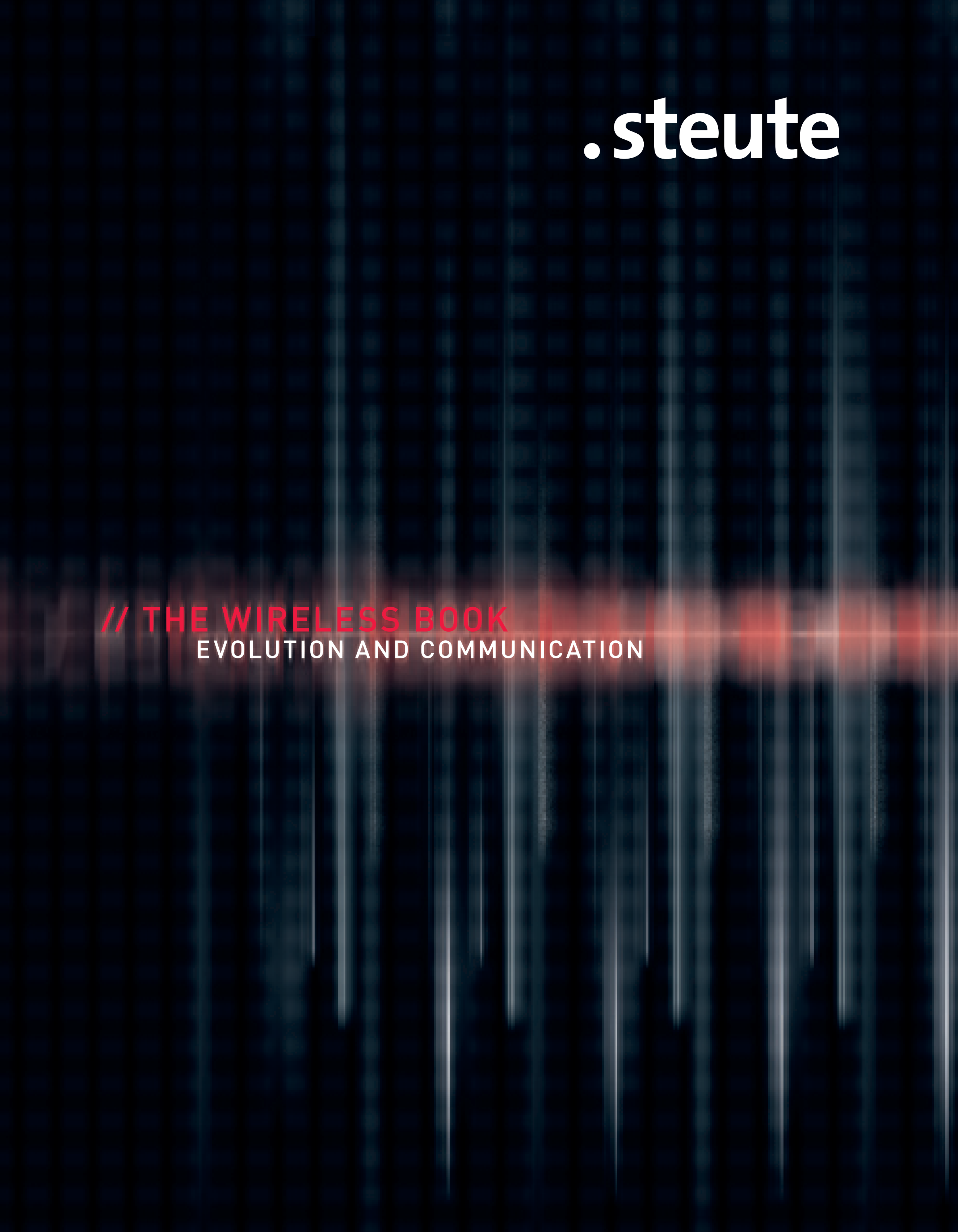


.steute

// THE WIRELESS BOOK

EVOLUTION AND COMMUNICATION

// THE WIRELESS BOOK

EVOLUTION AND COMMUNICATION

// Contents

// Foreword by the management

Communication via electromagnetic waves – referred to in short as radio or wireless communication – has a compelling story behind it. From the first attempts to generate electromagnetic waves with an electric arc to the most modern low-power ASIC transmitters, much has been discovered and researched and ultimately converted into usable solutions. This is especially true for the field of consumer electronics. From radios through mobile phones up to wireless thermometers, we are now surrounded by wireless devices.

Wireless is also not a new topic in the industry. But, for many engineers the construction of wireless technologies or the integration of wireless devices in the control of industrial plants is not familiar territory. There is a lack of technical information on the theory and practice of wireless communications in an industrial environment. And there is often a lack of confidence that wireless solutions can achieve the desired availability under the sometimes unfavourable industrial environments.

These fears are unfounded, as shown by steute's extensive experience with wireless automation even in hazardous areas and with wireless operating systems in demanding applications field such as medical technology.

Nevertheless, there is still a need for information about wireless in industry. This book hopes to make a contribution to fill this information gap in a very basic, always practical and entertaining way.

With this book we are also fulfilling a desire of ours. For the curiosity at steute, always delving deeper into the fundamental areas of our daily development work, knows no bounds. Our main task is to transform this knowledge and the resulting ideas into smart, sensible and practical products. That is precisely what we strive for with our range of wireless solutions for machinery and process equipment as well as for medical technology. The basis which we build on is explained with this wireless book.

We are pleased that we could win two experienced wireless professionals for this project in Dipl.-Ing. Wolfram Gebhardt and Prof. Dr.-Ing. Jörg Wollert, who were responsible for most of the content, we owe them a wholehearted thank you. The book owes its existence to their impressive expertise and the equally important rare gift of sharing and conveying this knowledge.

We wish you an informative and enjoyable reading!

Marc Stanesby
Managing Director, steute Schaltgeräte GmbH & Co. KG

// Introduction

Wireless – a trend technology

The cultural development of humanity is essentially determined by communication. A major problem of linguistic communication, however, lies in the limited range of sound waves. A virtually unlimited range and a variety of information transmission with sound, image and data and also a high coverage became possible only with the development of wireless communication: the radio broadcast.

Albert Einstein said in his speech in Berlin at the opening of the Funkausstellung (consumer electronics fair) in 1930: »Hearing the radio, one wonders how people got hold of this wonderful tool of communication. Think of Maxwell, who showed the existence of electric waves in a mathematical way, and Hertz who produced and demonstrated them first with the aid of the spark«.

Today, wireless is a technology with wide-spread use – ranging from mobile telephony through wireless Internet access to the car keyless access. With the success of many small wireless devices in the consumer market and in telecommunications, it is easy to understand that ever new application areas have been developed in industrial communication in the last ten years.

Hardly a provider of fieldbus technology or intelligent components could escape the general trend in terms of freedom from wires and cables. This is not surprising, since the need for wireless standards in the IT world is acknowledged and industrial automation always places more emphasis on mobility and flexibility.

Wireless in industrial automation – Opportunities and risks

Particularly in automation, the freedom from cables and connectors brings not only advantages but also risks that are often unknown or ignored. Whether the new technology nonetheless achieves an enthusiastic reception depends on the significant increase in comfort, flexibility and mobility. This increase enables a wave of innovation for the man-machine interface: Completely new approaches in the operation of appliances, machinery and equipment become possible.

But the use of wireless technology needs to be planned prudently. The applications and the possible implementations are too different. There is no single universal wireless technology for all applications and the simple replacement of a cable by a radio link will in many cases lead a project to failure – out of ignorance and exaggerated expectations.

Goal: The right wireless technology for each and every application.

Consequently, uncertainties are the norm in practice. Which of the many technologies, Bluetooth through EnOcean, Nanonet, WIMAX and WLAN up to Z-Wave and ZigBee, will eventually find favour with the customer depends not only on qualified consulting, but also on the extensive technical knowledge of the user himself. Not every technique will meet every requirement.

The potential for success also lies hidden here – finding the right technology for the right application. For end-users, to begin with, it is irrelevant what this technology is and how it works in detail – he just wants to have a reliably functioning device that satisfies his requirements of availability and performance. The consumer market has already recognised this and has responded to it.

Great creativity of the user

But for all the creativity of the manufacturers, often the users are still more creative. Systems are used in applications in ways the manufacturer would not have even dared to dream about – but nonetheless, the user expects reliable functionality.

To summarise, from a technical point of view, one can assert that a wireless solution can be found for almost every problem. However, a responsible use of the wireless technology is necessary for a successful application. Because, even the greatest creativity cannot abolish the laws of physics. The principle still applies also when you are enthusiastic about wireless technology in automation and consumer technology: where a cable is possible, it should also be used. To use wireless just because there is wireless does not make sense. In fact, the benefit should be immediately recognisable. A brief overview of the technological opportunities and limitations are provided in this work.

// The evolution of communication

01

Evolution and communication

// Evolution and communication

Nature has developed a variety of methods for the transmission or communication of information. The highly differentiated syllable language of the people has created a fundamental prerequisite for the development of civilisations.

In both cases, exciting interactions happened between the communication partners and their methods of communication. While the natural systems of information transmission use mechanical, chemical, acoustic and optical techniques, humans, with the invention of the electric processes and particularly the wireless method for communication, have accomplished their own grandeur in the development of humanity and its culture. This becomes clear when one envisions the various communication methods.

Mechanical communication

Presumably, the oldest sense is the sense of touch. It is used by sea anemones, corals, spiders and other simple organisms to capture prey. Information signals are transmitted by mechanical contact or the sense of vibrations. Many animals sense a natural disaster, such as the finest earthquake waves, and take safety precautions. Humans use the sense of touch as well, e.g. in Braille or as physical contact between mother and child. Mechanical communication can bridge only very short distances or none at all.

Chemical communication

Signals can also be transmitted as odours or taste. Plants scream for help with scents if they are attacked by predators and summon their natural enemies. Or they change to less appetising tastes so that the predators abandon them. Bees and hummingbirds orient themselves by aromas while foraging. Animals and people use odour and taste of food and foodstuffs in order to find food or to check for salubrity. Well known are pheromones as scented attractants for mate detection over some distance and in complex environments. It is not only animal mothers who recognise their children by smell – this phenomenon has been demonstrated even in humans.

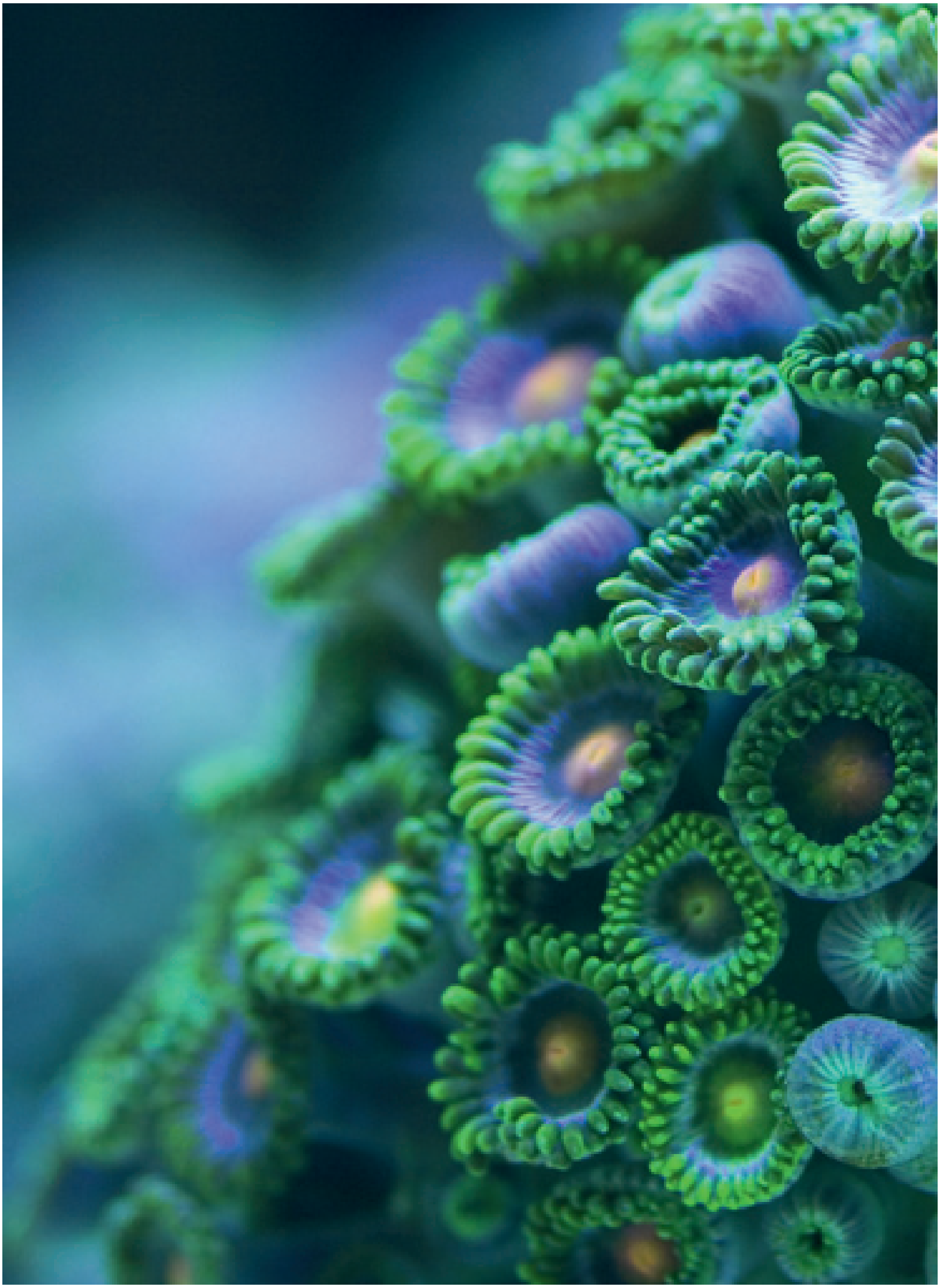
The important thing is that chemical communication in contrast to mechanical communication is capable of covering distances – from close distances up to several hundred metres.

Acoustic communication

Acoustic communication has differentiated itself in a typal manner and adapted flexibly in its range of information (its intelligence). Animals transmit quite complex content through conspecific sounds. Even by the choice of sound frequencies, some species such as elephants and whales bridge long distances of up to 40 km by infrasound in order to maintain social cohesion.

The invention of language has brought the biggest surge in development for humans on the path to Homo sapiens. In addition, humans have developed other acoustic techniques for transmission of information – for example, the yodelling, which is useful in an environment disturbed by the echo effect over distances of up to about 1000 metres, albeit with limited information density.

A significant surge in the development of acoustic communication was the use of drums, megaphones, trumpets, horns and other devices. They allow the transmission of information over long distances or in noisy environments (such as stadiums or in the din of battle) to communicate with sufficient reliability.



In addition, human beings make use of acoustic communication for their art (music) and for medical (ultrasound) or therapeutic (psychosomatic) treatment.

Acoustic communication has three main benefits:

- Good distance coverage and range
- Good spatial coverage (local accessibility in air and water) and a moderate assertiveness (vol.)
- Flexibly adaptable to the users and their variously sophisticated languages and applications.

Therefore, acoustic communication represents a substantial innovation on the path of humanity and for the cultural development of humans.

Optical communication

The evolution of life took place under constant sunlight. Therefore almost all plants and animals and even humans respond to optical information. Most living things are even dependent on light for their existence.

By their shape and colour, plants call on active volunteers to maintain the species (pollination). Animals recognise their environment and their counterparts by the sense of sight. Some species change their colour for the purpose of camouflage or communication with potential partners, some like fireflies can even glow actively.

Animals and humans use light for social communication through their body language. If the dog wags its tail, it is pleased. If the cat does the same, it is excited or displeased. Therefore, misunderstandings often occur between the two. Only humans – if they are familiar with the visual language of dog and cat – understand both correctly. All social animals maintain a fairly sophisticated body language. The facial expression is particularly well developed among the primates; monkeys express more through facial expressions than through sounds.

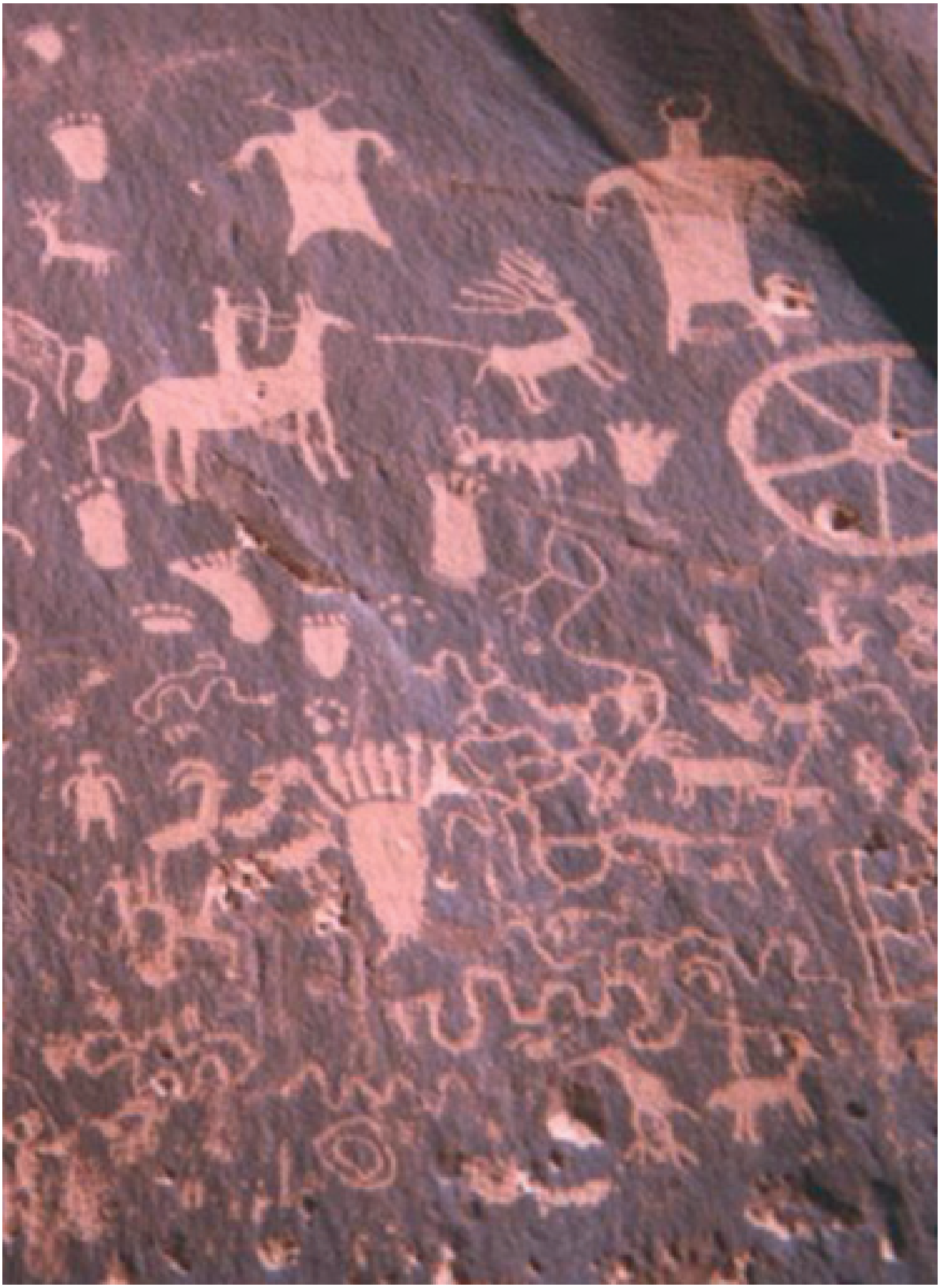
Optical communication was initially critical for survival for humans, because the upright posture provided a better view and they could communicate through sign language, for example, when hunting over medium distances, depending on visibility.

They invented light signals which bridged long distances, through fire alarms especially at night and smoke signals by day, in so far as the meaning of the signals was previously agreed upon and then codified. In seafaring too, optical communication was the instrument of choice for distances up to 40 nautical miles (75 km). Beacons on towers and flag signals were used. The Romans also used a distance limit for optical communication from watchtower to watchtower.

A highlight of this development was the unfolding of the performing arts for the visual communication of abstract issues from hunting to religion.

The most important evolutionary leap was the invention of writing – first as a pure pictorial script, later as a syllabic script (hieroglyphs) and then alphabetic script. Thus documentation was available and the communication could be shifted from real-time (online) to deferred time (offline).

The written communication as a pattern of optical communication with the advantages of documentation and storage, the overcoming of any distances (possibly with loss of time) and the representation of concrete and abstract issues are now at the heart of human communication.



The advantages of optical communication are thus:

- Long range (visual contact)
- High spatial coverage, but limited to unobstructed view
- Assertiveness and confidentiality (directivity option)
- Can be documented and preserved over long periods in written form
- New perspectives for art, culture and science (e.g. astronomy)

Summarising the importance of communication for humanity

- The communication processes developed in the course of evolution ran parallel to the developmental stages of life.
- Older communication methods were not replaced but expanded by more sophisticated methods.
- All communication methods together formed a system that was adapted to the respective life form of the species.
- Acoustic and optical communications set the stage for the development of humanity in terms of art, culture and science.
- Humans continued to use all developed evolutionary communication methods.
- An expansion of communication was possible, but only with respect to certain features such as range, spatial coverage, confidentiality and assertiveness against environmental influences.
- Technical communication methods, especially wireless, liberated the natural process of existing bonds, thus bringing a new major wave of development.

// The history of wireless communication

02

The discovery of electromagnetic waves

The first wireless applications

Milestones in the development of radio

Maxwell‘s equations in the Gaussian cgs system

Ampere´s law

$$\nabla \cdot H = 4\pi \frac{j}{c} + \frac{1}{c}\frac{\partial D}{\partial t}$$

Law of induction

$$\nabla \cdot E = -\frac{1}{c}\frac{\partial B}{\partial t}$$

Coulomb´s law

$$\nabla \cdot D = 4\pi p$$

Gauss´ law of magnetism

$$\nabla \cdot B = 0$$

Maxwell‘s equations in modern notation

// The discovery of radio waves

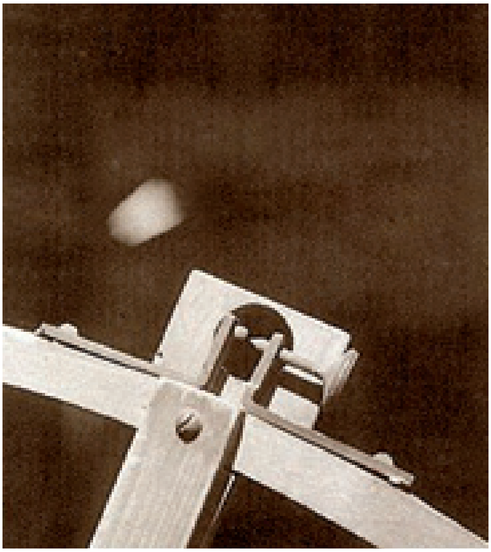
Humans discovered electricity only relatively recently – which is due to the fact that humans did not sense it, i.e. lacked a sensory organ for the perception of electricity. In ancient Greece, it was observed for the first time that amber rubbed with organic materials (towels, cat fur) drew out non-metallic particles such as dust or papyrus pieces. Today we know that the amber was electrically charged. In ancient Greek »amber« means »electron«, the name for the electrically charged elementary particles.

The discovery of magnetism goes back probably to the time of transition from antiquity to the Middle Ages. Thus a freely suspended piece of magnetic iron ore was used as a compass pointing north for centuries by sailors before the Viking Age in the 7th and 8th centuries. Such pieces of ore were rare, some were found on the Babilonie, a Germanic castle on the northern edge of the Wiehengebirge (near the steute headquarters), where iron smelting was already underway in the La Tène period. The systematic study of electromagnetism began with the Englishman Michael Faraday. He discovered electromagnetic induction in 1814 and demonstrated a way to generate electricity. In 1836/37 he developed from it the theory of electric and magnetic fields which could penetrate non-metallic material and free space. Since he could not prove his theory mathematically, science did not believe him at first.

Maxwell‘s equations

It was 30 years later (1864/1865), that the Scottish mathematician and physicist James Clark Maxwell developed his major differential equations. Maxwell‘s equations showed that the interaction of electric and magnetic fields could be established. In 1866 it was purely mathematically derived from his equations that there must be alternating electromagnetic fields and waves, which propagate at the speed of light.

Heinrich Hertz



Test set-up for experimental verification of electromagnetic waves

Heinrich Hertz

The actual proof of the predicted waves was acquired 22 years later, in 1888, through Heinrich Hertz.

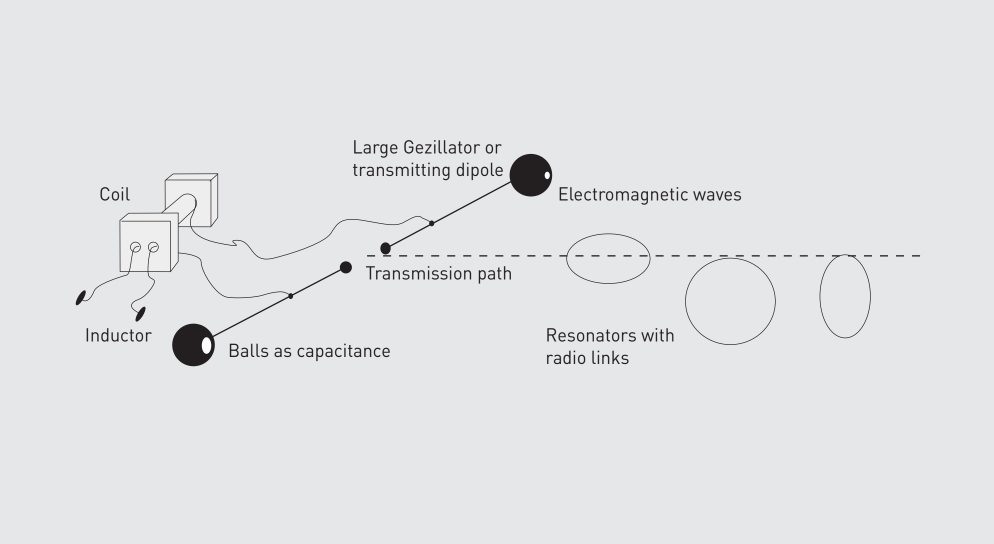
Hertz, who graduated at the age of 23 years, studied in Berlin under famous scholars such as Hermann von Helmholtz and Robert Kirchhoff. Helmholtz, who recognised the talent of his doctoral students, was an enthusiastic supporter of Maxwell's theory and his field equations. He gained Heinrich Hertz for the project to experimentally demonstrate the electromagnetic waves.

Hertz, who was appointed a full professor of experimental physics at the Karlsruhe Institute of Technology at the age of 26 years, began his research work in 1887 at the age of 30 years. At that time it was already known that electrical discharges execute electrical oscillations through a coil, which, when interrupted by a spark gap, produce electric sparks. When he investigated the sparks at his first attempt, he accidentally discovered that a second coil in the vicinity, which was not connected electrically with the first coil, also emitted sparks.

Starting point: A chance discovery

From this accidental discovery Hertz had to conclude that an energy in the form of vibrations or radiations had been transmitted from the spark gap and the coil through space to the second coil and which then led to the tiny spark. Today we know that this was the first radio link, whereas it is uncertain whether Hertz knew that he had just discovered electromagnetic waves.

In any case, Hertz now continued to research more specifically. To simplify the experiment, instead of the second coil he now modified a simple wire ring, which was also interrupted by a spark gap. Hertz called it a resonator, because he knew that a resonance had to exist between the sending spark coil and the receiving ring.



Heinrich Hertz developed the first dipole

In order to detect the faint sparks at the resonator at all, he had to install an adjustable micrometer and darken the room.

In the next step, he built a more powerful transmitter. It consisted of a stronger spark coil, on which were connected two elongated wires with a spark gap in the middle and metal balls at the ends. That was the first dipole.

Now the energy was sufficient to carry out further investigations. He noted that there were areas within the laboratory where the sparks were stronger at the resonator and places where no spark was produced. He concluded logically that they were stationary waves, whereby the waves received directly from the transmitting spark coil overlapped with the waves reflected on the zinc plates on the opposite wall.

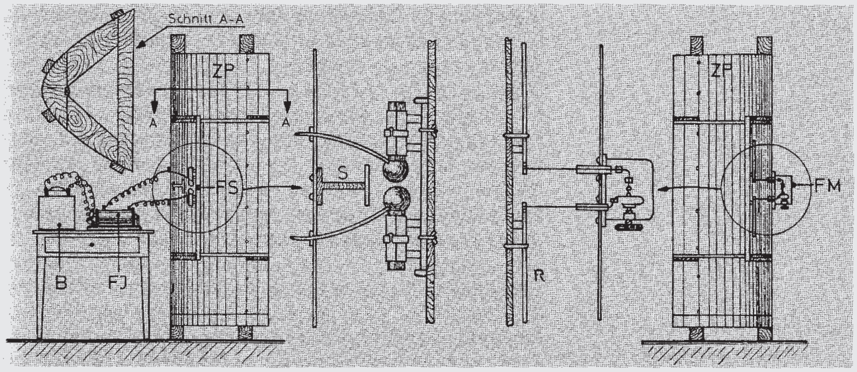
Indisputable scientific evidence

That was the irrefutable scientific proof for the wave characteristic of energy transmission. Hertz was even able to detect the wavelength of about nine metres. By comparing the waves on a taut wire in free space, he showed that the free waves propagate approximately at the speed of light.

Consequently, he then studied the properties of these waves. First, he found that the sparking at the resonator is dependent on the direction in which this wire ring is held with respect to the transmitting wires. The wire ring was held in the three axes of the room, parallel or perpendicular to the wires and also transverse to the wires, which then showed different sparks.

Initial studies of the properties of radio waves

In another experiment, Heinrich Hertz used a 2 x 2 m wide rectangular wooden frame, on which he stretched parallel wires. He held these wires parallel to the transmitting wires (antenna), so there was no spark. But when he turned the frame by 90 degrees, the sparking was back. With it he realised that they were transverse waves,



The first directional radio link: Test set-up of Heinrich Hertz

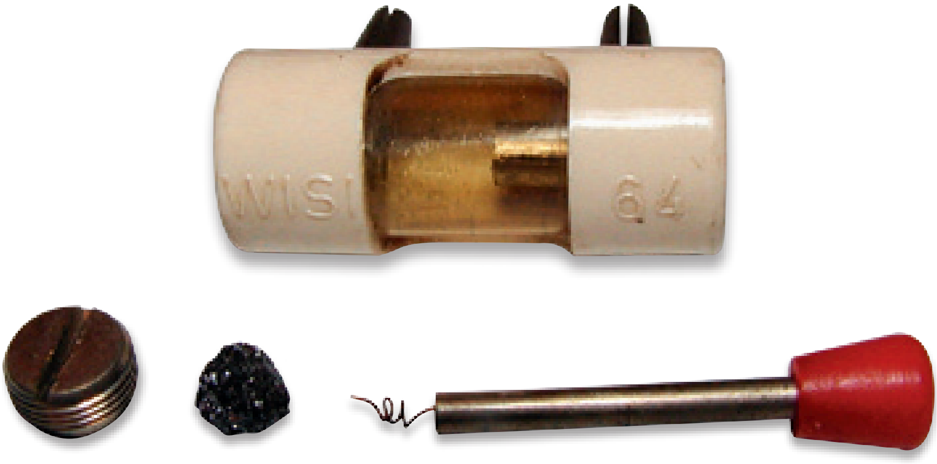
which were additionally more polarised like light waves.

Then he examined whether these waves could bend like light. For this purpose he placed in the propagation path a large prism made of hard pitch with the base of a regular triangle with a side length of 1.3 metres and a height of 65 cm. This part alone weighed about 600 kg. It was found that the radio waves were diffracted by the glass prism just like light waves and had roughly the anticipated diffraction angle.

Finally, he brought the transmitter dipole in the focal line of a cylindrical parabolic mirror made of zinc and a receiving dipole in an identical reflector with 1.5 m aperture width. Here, he had mathematically adapted the dimensions of the reflectors with the length of the dipoles. This system then operated at a wavelength of 66 cm.

The measured radio values were high within the focused radio path, but no sparks were detected outside. Thus the focusing analogous to light was demonstrated and the first directional radio link invented, albeit with a range of about 16 metres.

Through his focussed research Hertz not only demonstrated electromagnetic waves, but also studied their properties extensively. In the process he found and identified the same properties as in light waves: »Light is an electrical phenomenon …« This was revolutionary. A practical application of his radio waves, however, could not be imagined by Heinrich Hertz. The unit of frequency is named Hertz (Hz) in his honour.



Plug-in crystal detector with pyrite crystal and spike

// The first wireless applications

Hertz's experiments triggered a wave of experiments in Europe, which had among other things objectives to increase the range of radio waves. The previous radio link in the reception ring with magnifier was too insensitive and required a very high transmission power.

In Paris in 1890, Edouard Branley utilised a detector to increase the sensitivity of the receiver, which was called a Radioconducteur or coherer. Branley also recognised the importance of the antenna. He called it »air wire«. The Englishman Oliver Joseph Lodge increased the sensitivity further by introducing a resonance circuit. In 1894 in Petersburg, the Russian Alexander Stepanovich Popov raised the sensitivity of the receiving circuit by inserting a sensitive relay which switched on an electric bell or controlled a Morse writer known from the former wired telegraphic method. The first words that he transmitted on the occasion of a demonstration on 24/03/1896 at the University of St. Petersburg with this Morse code telegraph were »Heinrich Hertz«.

End of the 19th century: Wireless transmission covers several kilometres

With Popov's circuit, the Italian Guglielmo Marconi achieved a radio transmission of several kilometres. Through the mediation of his English mother, he received a British patent on this circuit in 1897 and support for continued successful wireless experiments. The increasing range aroused the interest of shipping companies. Soon he equipped ships and shore stations, founded his own company in 1900 and had a de facto monopoly for marine radio links. Marconi became famous in 1901 for the first successful transmission of a radio telegram from England across the Atlantic to Newfoundland in America. Many unsuccessful attempts had preceded it.

On the German side, Professor Adolf Slaby and his assistant Georg Graf von Arco were concerned

with radio experiments in 1897. Kaiser Wilhelm II was very interested in the technology and attended a demonstration by Slaby on 27 August 1897. The emperor wanted to remain independent of Marconi‘s monopoly and encouraged Slaby‘s trials. Thus, two months later, on 7 October 1897, Slaby set up a »world record« of 21 km, whereby antennas of the size between 300 and 500 m long were carried by captive balloons. In the spring of 1898, the radio link covered a distance of 60 km from Berlin to Jüterbog.

Industrialisation of wireless technology

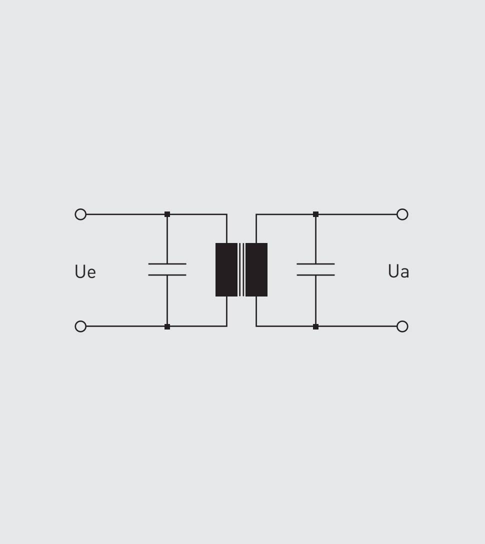
In the same year German companies were established, giving rise to the »Radio Telegraphic Department« of AEG led by Graf von Arco and the »Funkentelegraphie GmbH« in Cologne. Professor Ferdinand Braun of the Technical University in Strasbourg founded the »Gesellschaft für Telegraphie System Professor Braun und Siemens & Halske m.b.H« (Telebraun) in Hamburg, and wireless technology was introduced in military applications – first in the Navy, a little later in the air force and the army.

// Milestones in the development of radio

The effort to increase coverage and reliability of radio communication has led to many small improvements as well as taking some wrong turns.

Increased sensitivity of the receiver

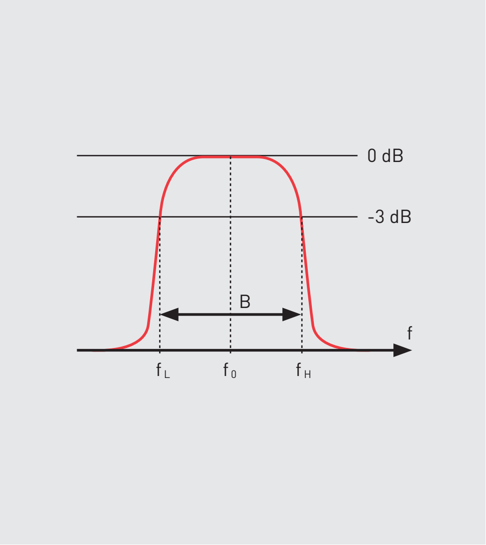
The first task was to increase the sensitivity of the receiver. This was achieved with better detectors. In 1903, Wilhelm Schloemilch invented the »electrolytic cell« with two platinum electrodes in dilute sulphuric acid. This cell worked very well as a rectifier with a small bias voltage and quickly replaced the coherers. From 1906, Professor Braun‘s crystal detector replaced the electrolytic cell as the receiving rectifier. Of the four materials: pyrite; silicon; molybdenum and carborundum, pyrite and carborundum proved to be the most suitable. They were on the one hand sensitive enough to receive weak signals, but on the other hand, robust enough to withstand their own strong transmitter signals and they prevailed as crystal detectors. By the way, the receiving rectifier was the indispensable element for later sound reception with a headset.



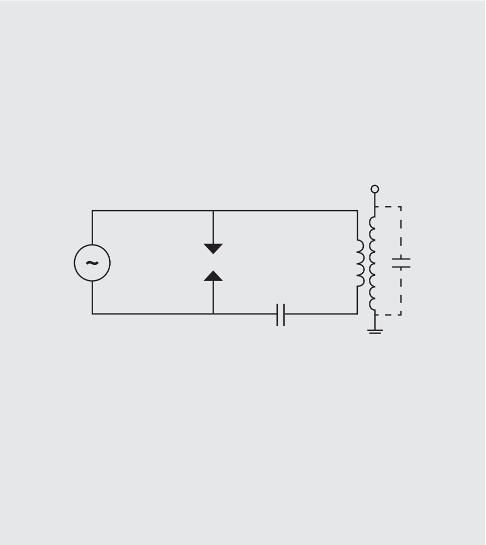
Circuit diagram of a bandpass filter

With the invention of the wave meter in 1902, one could selectively measure the wavelengths/frequencies and tune to optimise the radios. The wave meter consisted of a resonant circuit with variable capacitor and plug-in voice coil, which could swing a galvanometer with resonance via a detector. Even today absorption frequency instruments such as the grid dip meter or the satellite finder work in this way.

In 1904 the selectivity was improved by incorporating a second resonant circuit into the receiver. The tuned antenna circuit and the good filtering through the dual-circuit band filter reduced interference from neighbouring channels and



Frequency response of a bandpass filter

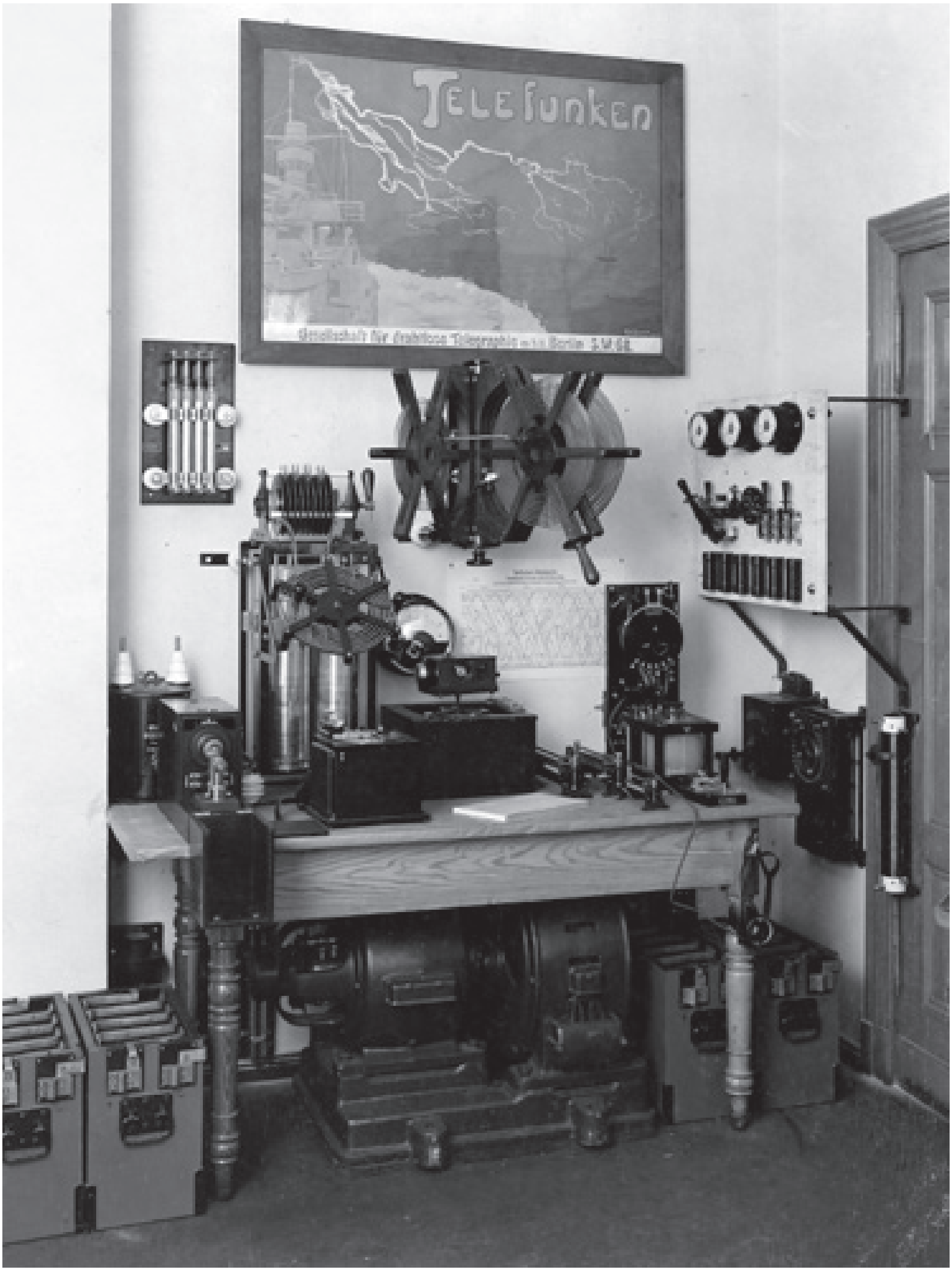


Simplified circuit diagram of a quenched spark transmitter

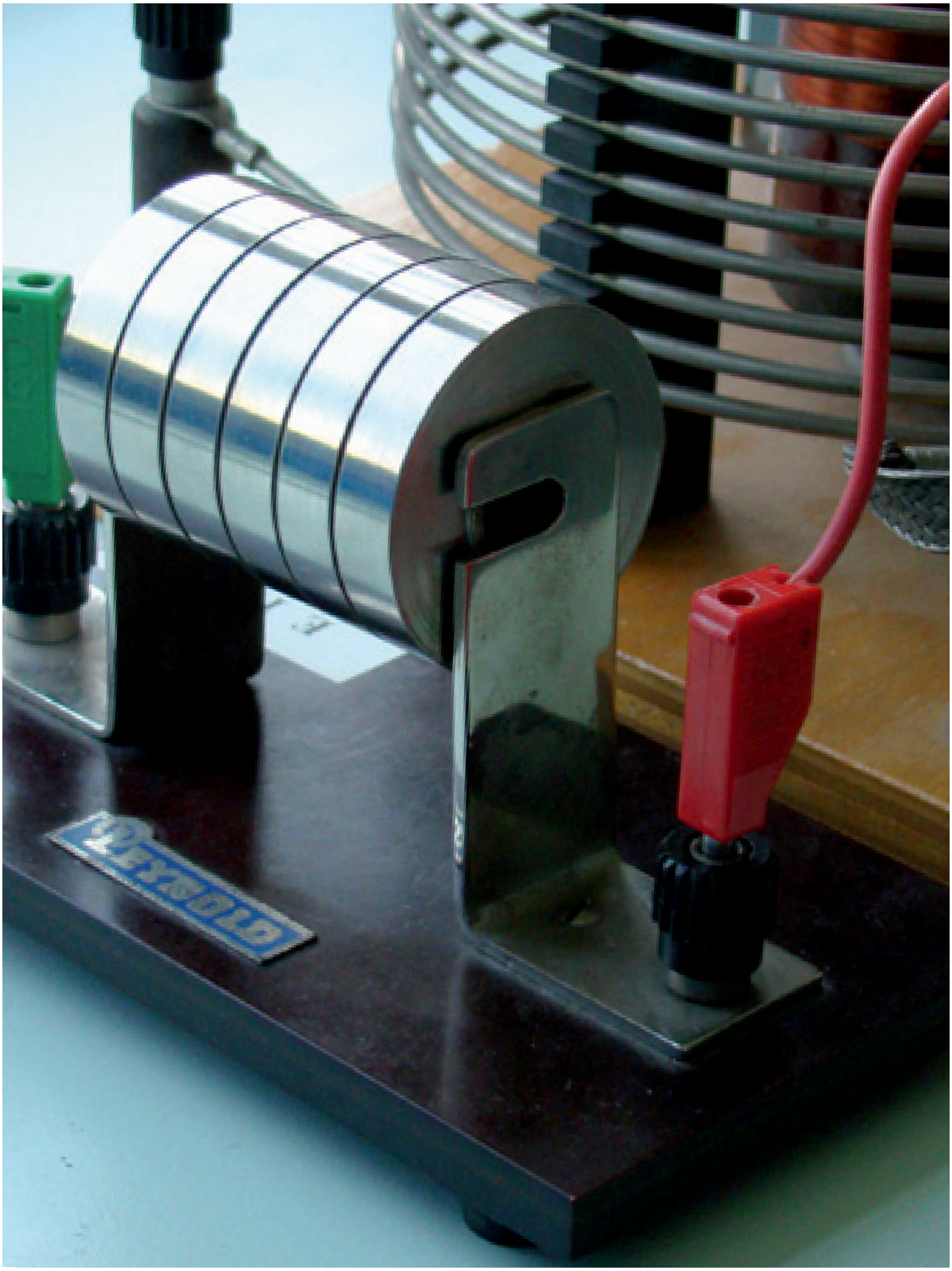
enabled the parallel radio operation of multiple wavelengths.

Improvement of efficiency and performance

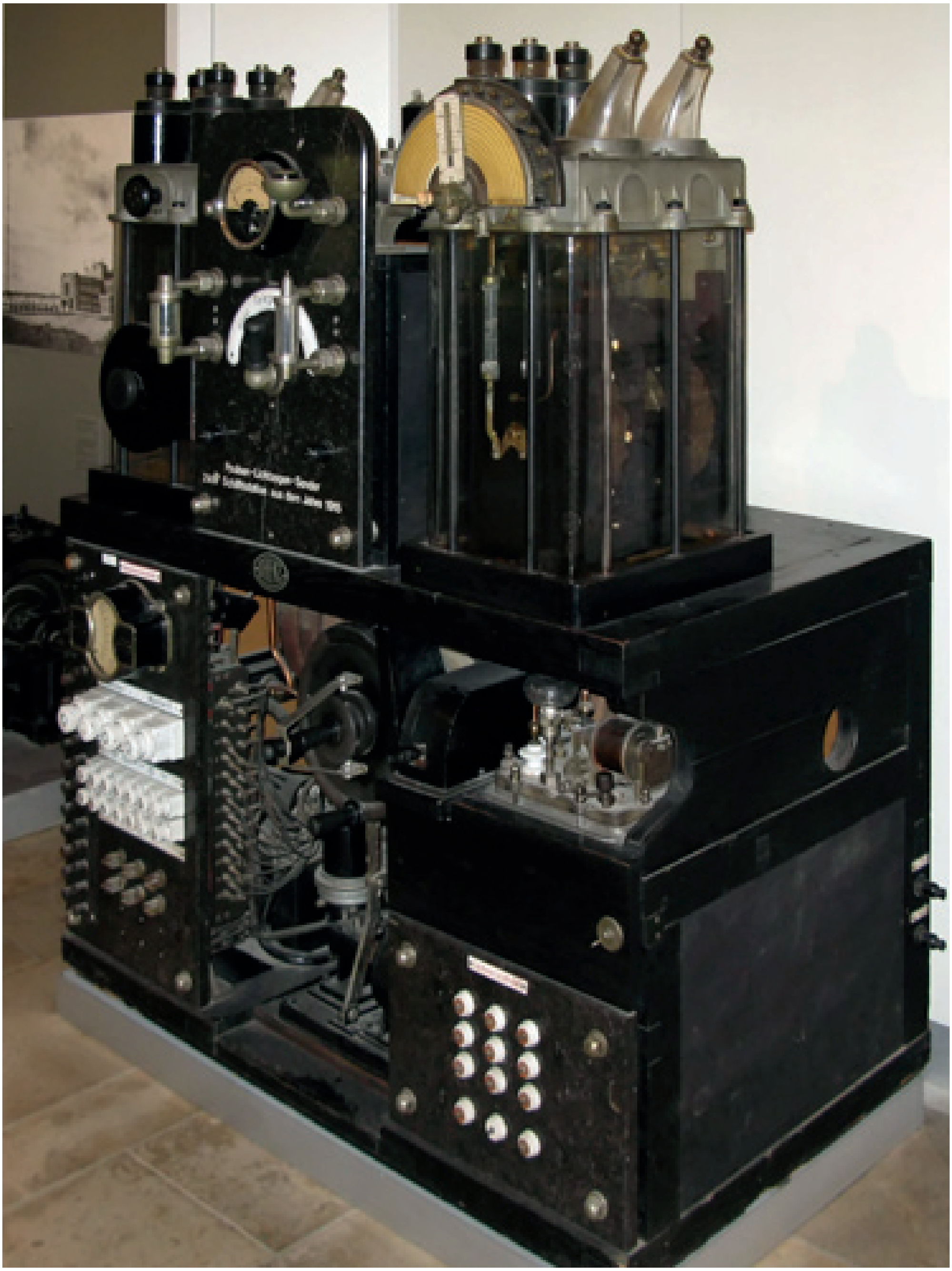
Attempts were made to improve the efficiency and performance of the transmission. The waves triggered by the spark coil subsided very quickly. Only when the spark duration of the so-called spark gap transmitters was shortened by blowing on the spark gap, they oscillated longer until the next spark. By the series connection of many spark gaps with small electrode spacings, the spark went out without blowing. Thus the quenched spark transmitter was built by Prof. Max Wien with an almost continuous emission, but it was still modulated like a sawtooth voltage. The careful tuning between the antenna resonant circuit and the transmitter resonant circuit resulted in a reduction of the occupied bandwidth and thus an added efficiency of 50 to 70%. In 1909, a range of 4600 km was achieved with a quenched spark transmitter of 25 kW power. But all these transmitters emitted attenuated waves, which were not suitable for a sound or even music transmission.



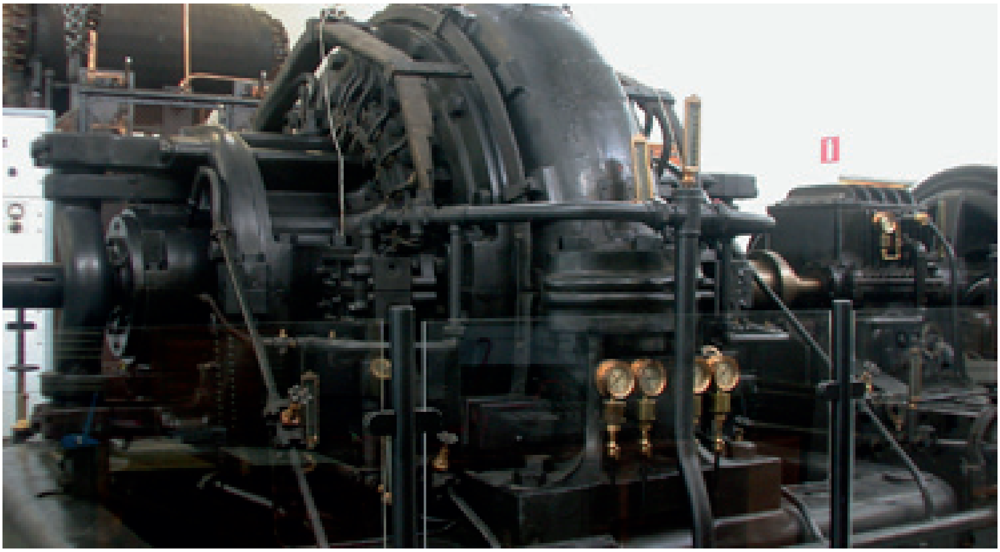
Telefunken quenched spark transmitter



View of a quenched spark gap with Tesla transformer



View of a Poulsen arc transmitter



Alexanderson transmitter, left generator, next right multiplier

Invention of the arc transmitter

Attempts were now made to generate unattenuated waves. It was known since 1900 that an arc powered by direct current also generated high-frequencies (singing arc). The negative resistance of the arc leads to vibrations which are dependent on the surrounding circuitry.

The Danish physicist Valdemar Poulsen mounted an arc in a tuned resonant circuit of an electric arc and thus developed a powerful high-frequency generator, which was coupled to an antenna circuit. In 1903 he invented the arc transmitter, which has a narrower band and fewer auxiliary and harmonic waves. This improved the transmission range and the simultaneous operation of multiple transmitters in the same frequency band. The antenna circuit could now be interrupted by a direct current switch for telegraphy or modulated via microphones for speach.

1904: The first voice communication

Poulsen succeeded in making a voice communication by radio for the first time in 1904. In 1906, the advanced technology was published, which constituted the basis for today‘s radio and wireless technology.

The typical transmission power of German army equipment was 1.5 kW in mobile units and 4 kW in fixed facilities, and in the Imperial Navy, 6 kW on large ships. Radio stations with 100 kW were used in the civilian sector from 1910 onwards for continental radio communications, particularly in the United States. Probably the strongest arc transmitter was set up in 1922/23 in Malabar/Java (now Indonesia). With a primary power consumption of 2400 kW, the transmitter covered a distance of 11,500 km to link up with the main station in Kootwijk/Netherlands.

Various receiver and power amp tubes

Tube type RE 144 of Telefunken (approx. 1920)

»Transmission equipment« for overseas radio

The development of transmission equipment took place in parallel, in which high frequency output was generated directly by rapidly rotating alternators. The upper limit of the frequency was determined by the mechanics. A distance of 310 km was achieved at 81 kHz and 1 kW of power. Until 1916, circuit improvements for frequency multiplication led to frequencies up to 150 kHz with outputs up to 250 kW and later up to 1000 kW. These transmitters were used for overseas wireless connections and the emerging news services.

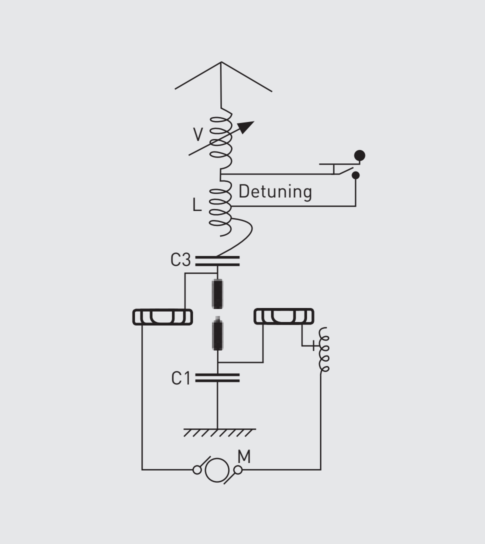
Unattenuated transmission stations broadcast a very clean high frequency virtually without intrinsic noise. They were identifiable in the previous receivers only by the absence of noise. Radio call signs were heard as faint cracking noise.

Therefore, the silent receiving signal had to be made audible artificially by a buzzer or ticker. So if »the news came over the ticker,« then either a Poulsen arc transmitter or a transmission device was behind the radio news service.

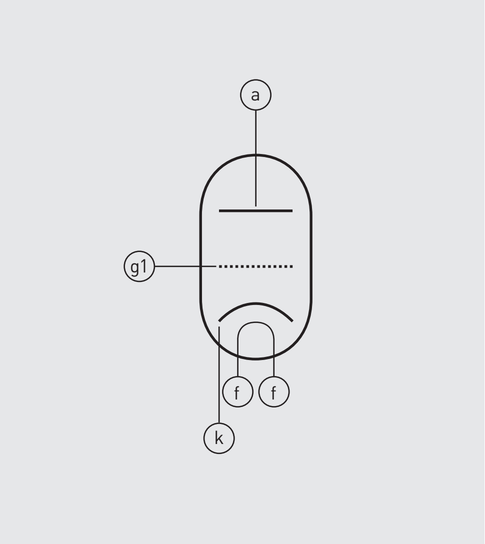
Development thrust through electronic amplifier tubes

The next significant increase in power for both the transmitter and the receiver was achieved by electronic amplifier tubes. Professor Arthur Wehnelt in 1904 worked in Erlangen on a gas-filled »valve tube« in which a hot cathode emitted electrons, which were collected by an anode and which were suitable for the rectification of alternating currents. Lee de Forest announced his Audion tube in a patent in the U.S. in 1907, a grid-controlled tripole tube with thermionic cathode in an evacuated glass bulb. The patent was not known in Europe until 1912. Moreover, his Audion could not afford any amplification.

In Germany, Robert von Lieben developed a gas-filled tripole tube as an amplifier. Only after



Simplified circuit diagram of a Poulsen arc transmitter



Simplified circuit diagram of a triode

the development of the highly efficient molecular air pump by W. Gaede in 1911, did it became possible to build a high-vacuum tripole tube with useful amplification factors. Key patents for the feedback audion and the tube transmitter were issued in 1913 in three countries independently and almost simultaneously.

The audion tube improved the sensitivity of the radio receiver by several orders of magnitude. This enabled the transmission power to be equally reduced so long as the distance remained constant. The transmission tubes initially had only low outputs of several watts. By 1918, transmission tubes were manufactured with 1.5 kW, a water-cooled transmission tube of 5 kW was being tested, and a Telefunken transmitter already reached an RF power of 10 kW. Telefunken at that time already manufactured about 1000 receiver tubes and 100 transmitter tubes per day.

After the First World War, the age of the »dinosaur transmitter« was over and the time of tube radio had begun. Technically elegant electronics replaced the power escalation. The move to ever higher frequencies could begin. Public radio broadcasting began in 1923.

// Properties of electromagnetic waves

03

Basic properties

Propagation

Signal noise

// Basic properties

The study of radio waves and their utilisation aimed early on in the direction of the widest possible range. In this feature the radio is far superior to sound waves. Whereas a megaphone of 25 watts at best achieves a range of 250 m, decimetre waves range from 25 mW up to 2500 m. Thus, their range is higher by a factor of 1:10,000.

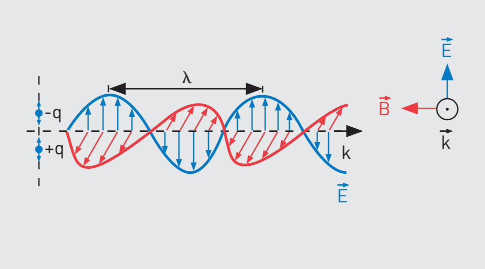
Another feature of universal significance is the propagation of radio waves in all directions, thus covering large areas. This is essential, for example, for public broadcasting. The radio coverage is particularly evident in the supply of, for example, all of Europe by a single television satellite.

Wireless connections can also be operated while in movement, thus with transportation and people travelling with it. And, radio goes to places that are not otherwise reached. It ranges from measuring points in inaccessible areas through industrial processes at short distances up to the most distant objects in space for research.

Another important feature of radio is its freedom from cables, which may pose obstacles. Wireless can make a contribution to workplace safety and in medicine, for example, to support the doctor effectively during surgical operations.

Frequency and wavelength

An electromagnetic wave consists of coupled electric and magnetic fields, which are ideally homogeneously spread in all directions in space. The electric field E and the magnetic field B are perpendicular to each other and move in the direction of the wave vector k. In a vacuum, the wave propagates at the speed of light c_o. For technical applications, the frequency f or the wavelength l is of importance. In a vacuum the correlation between wavelength and frequency results in:



λ - Wave length

c_o - Propagation velocity in vacuum

c_o= 299792,5 km/s; f - Frequency in Hz (1/s)

$$\lambda = \frac{c_0}{f}$$

Propagation of an electromagnetic wave

$$\lambda = c_o / f$$

// Propagation property of electric waves

The propagation of electric waves is influenced by various factors. Important here is the free field attenuation and the interference overlaying it.

Given an ideal monopole in a transmitter, the radiated power P_S is distributed with increasing distance over an ever-growing spherical surface. The radiation density S_r diminishes with it quadratically.

Radiation density

If a receiver would like to receive the signal, an antenna is used for capturing the radiated power. Logically, only a fraction of the radiated power can be captured. This is known as the notional antenna aperture A_W.
The free field attenuation L_{FS} is then calculated from the ratio of the radiated to the received transmission power (see diagram above right).

$$S_r = \frac{P_S}{4\pi r^2}$$

$$P_E = S_r \cdot A_W$$

Top: Radiation density in the range r
Below: Received energy in the range r

$$L_{FS} = \frac{P_S}{P_E} = \left(\frac{4\pi r}{\lambda}\right)^2 \cdot \frac{1}{G}$$

Calculation of free field attenuation L_{FS}

1/G describes here the product of the antenna gains, which will not be discussed further at this point.

Three issues are essentially crucial in this derivation:

1. The free field attenuation and thus the received power depends quadratically on the distance.
2. The free field attenuation depends quadratically on the frequency.
3. The free field attenuation depends on the effective antenna aperture.

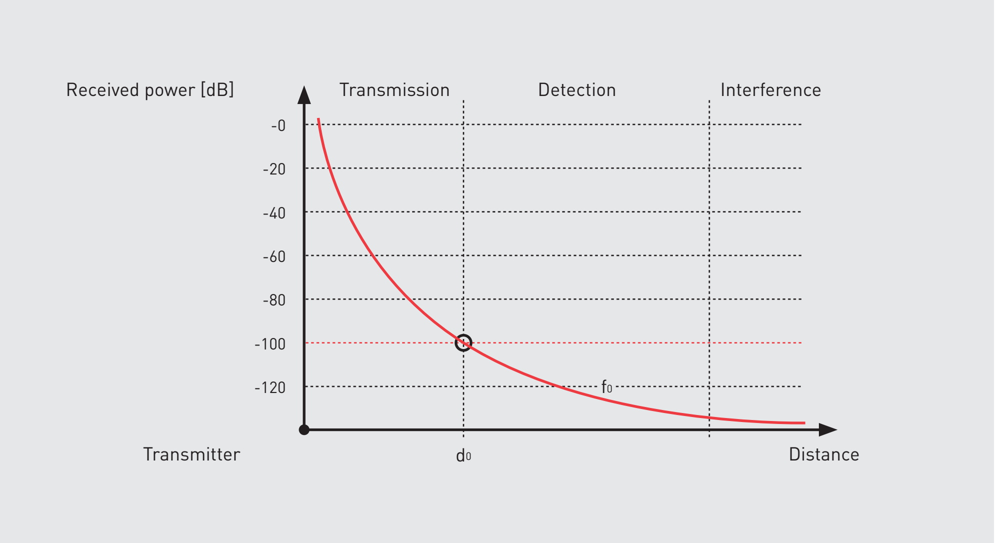
The illustration on the following page shows an example of the received power of a receiver at a given transmitter So. The sensitivity limit of the receiver typically amounts to about 100 dBm according to the data sheet.

If the received power is better than the sensitivity, one says that the receiver is in the transmission range. If the received power decreases further, a data exchange can only sporadically take place – the system is in the detection area. None of the signals can be used for a communication from a certain distance. Here the transmission signal only contributes to interference and a general noise level.

Range of radio signals

In practice, the radio range depends on many factors, so the manufacturer usually cannot lay down precisely reliable specifications for all possible applications. But when you know all the factors, you can precisely calculate the range. Unfortunately, not all factors are generally known and hence one has to introduce estimates or limits.

A series of pictures will explain how you can still come to reliable conclusions. For this purpose, the coverage is entered in kilometres in the logarithmic scale in a diagram on the x-axis and the respective power level to be considered is also shown in the logarithmic scale in dB on the y-axis.



The received power depends quadratically on the distance.

The logarithmic scales have the advantage that the quadratic decrease of the transmission level can be represented by a straight line.

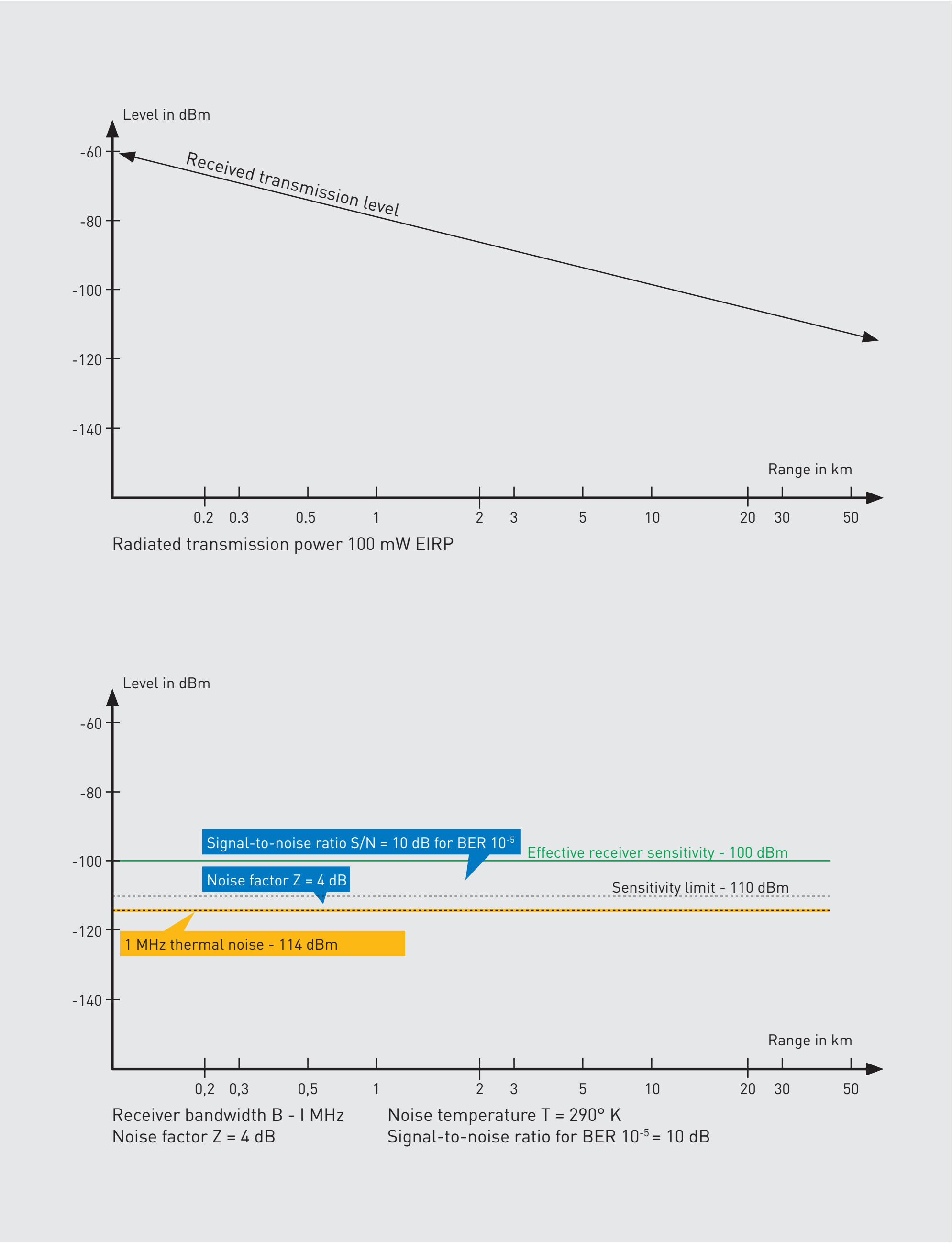
Basic conditions are:

- Antenna gains are not taken into account in all diagrams, but as spherical radiation. It is based on a standardised transmission power of 100 mW EIRP (effective isotropic radiated power)
- Special propagation conditions such as overreach have been omitted for clarity.
- The receiver bandwidth is set at 1 MHz, the receiver temperature at 290°K (17°C).
- Natural noise from the environment is ignored, because it is below the noise level of the receiver.
- The noise factor Z of the receiver is assumed favourable with 4 dB.
- 10 dB is selected as minimum signal-to-noise ratio for a bit error rate (BER) of less than 10^{-5}.

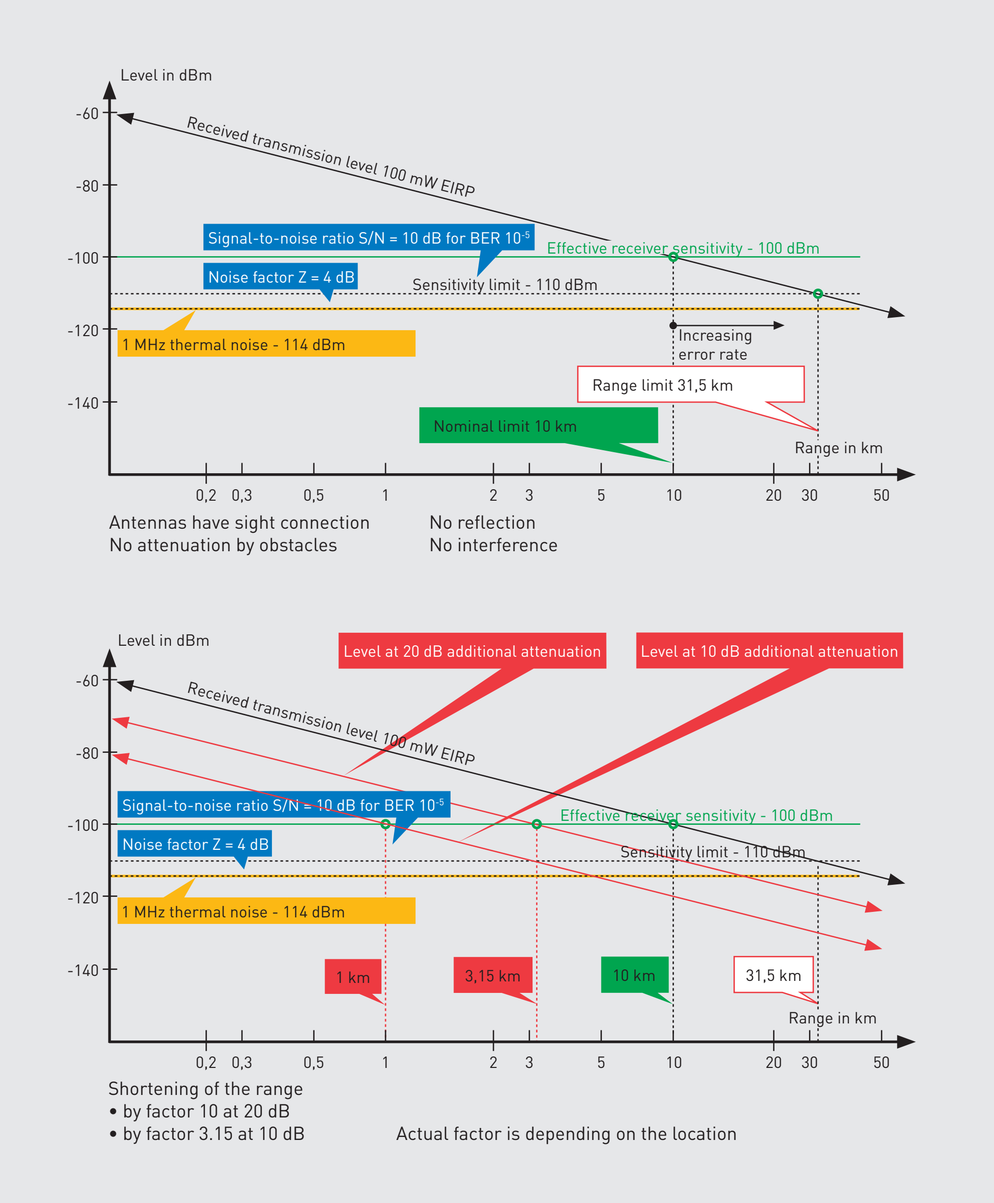
// Signal noise

With the use of sensitive superheterodyne receivers from 1918, an unexpected threshold was reached, where no increase in sensitivity was possible. The problem was the receiver noise. Researchers examined the relationship and found out that this was mainly thermal noise.

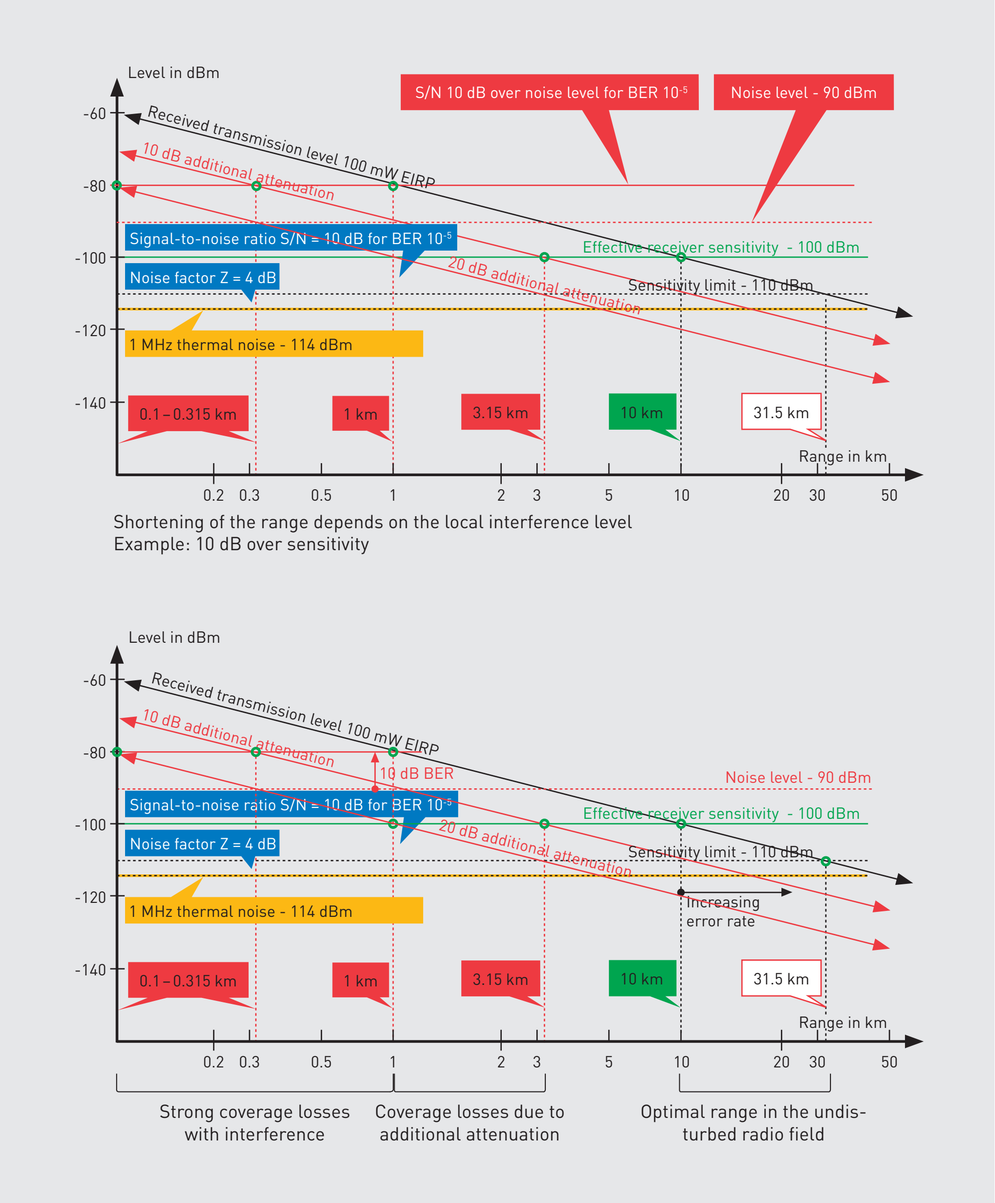
Atoms and molecules oscillate in any physical state, which is known as Brownian motion. Their oscillation amplitude varies with the fourth power of the absolute temperature. The electrons oscillating with the atoms or molecules constitute a high frequency alternating current, which leads to an AC voltage at the input resistor or the impedance of the input stage of the receiver and is perceived as noise. The noise power depends on the frequency interval B of the receiver input circuit.



Top: Diagram 1 shows how the received transmission level decreases with distance.
Bottom: Diagram 2 shows the noise level of the receivers, which are independent of the distance.



Top: Diagram 3 shows the theoretically achievable range as intersection of transmission level and receiver sensitivity.
Bottom: Diagram 4 considers the additional attenuation of the transmission level with 10 dB and 20 dB. A shortening of the range occurs by factor 3.15 at 10 dB and by factor 10 at 20 dB.



Top: Diagram 5 now considers additional interference level that exceeds the threshold of sensitivity of the receiver and therefore becomes effective. The result is a further shortening of the range.
Bottom: Diagram 6 is a precisely calculated example of an SRD application (short range devices; model remote control) in the ISM frequency band of 2.4 GHz.

Thermal noise power
at the receiver input

$$N = k \cdot T \cdot B$$

Noise after the first stage

$$N = z \cdot k \cdot T \cdot B$$

Noise levels at the receiver

The so-called spectral noise power density N_o, that is the noise power per frequency interval, is independent of the frequency over a wide frequency range; it is, in this context, called the white noise or Johnson-Nyquist noise. This leads to the formula for the thermal noise power: Here the Boltzmann constant

$$k = 1{,}38 \cdot 10^{-23}\ \mathrm{Ws} / {}^\circ\mathrm{K}$$

The second effect, the so-called shot noise, arises in the first amplifier stage through the electrons of the amplifier current. These must be individually raised over a potential barrier and then drop into a potential sink, wherein a tiny pulse is generated. The sum of these pulses sounds like that of shot pellets falling on a hard surface – hence the name. This noise component is called receiver noise figure z, which typically lies in the range between 4 and 40. The result is a so-called noise factor z between 6 dB and 16 dB. The noise factor z is thus a measure of quality of an input amplifier. Low noise amplifiers are below 10 dB.

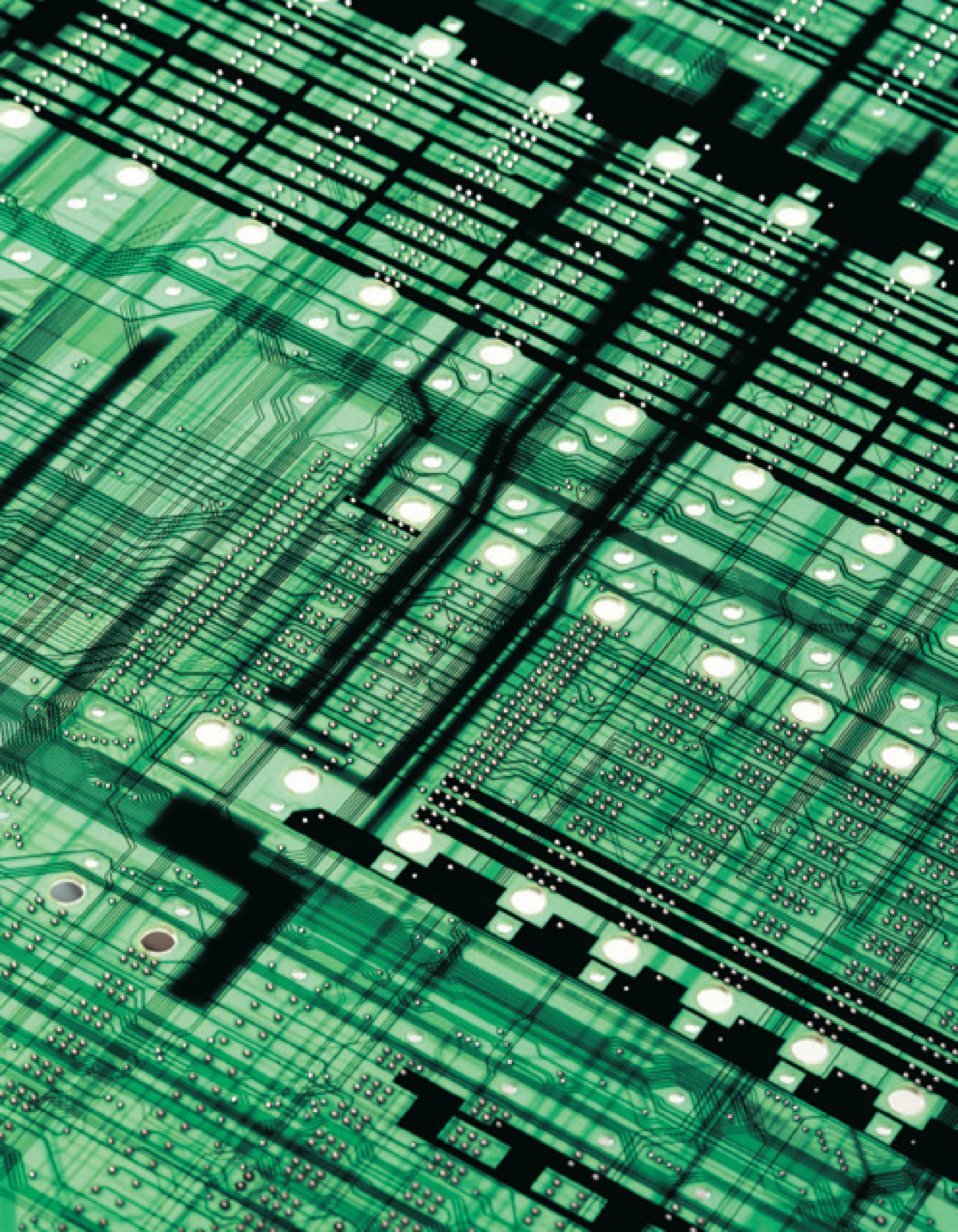
From both noise components we finally arrive at the formula for the receiver noise as shown opposite.

You can reduce the noise only in a limited manner:

- The easiest way is to limit the bandwidth B in the input of the receiver to the minimum level.
- The second possibility is the cooling of the input stage to a substantially lower temperature T.
- The third way is parametric amplifiers, in which the energy for the first amplifier stage is available not as current, but through the noise-free quantums of a pump frequency higher by factor 10 ($E = h \cdot v$) .

Such expensive amplifiers are used only in exceptional cases, such as in space research. Otherwise, for the required frequency range selected low-noise preamplifiers with semiconductors such as gallium arsenide field-effect transistors are used.

Other noise components such as flicker noise of electron tubes and the 1/f-noise (pink noise) are not important in this context.



// Basics of wireless technology: Frequencies

04

Historical development

Co-existence of wireless services

// Historical development of the use of frequencies

The invention of radio led to a huge demand internationally for frequencies to support the new services being established. The initial restriction to long wave and medium wave – short wave was still considered unsuitable for long-distance communications – quickly led to mutual interferences.

The International Wireless Telegraph Convention signed in Berlin in 1906 was the first step towards better organisation of frequency usage. When amateur radio operators discovered the large coverage of short waves, the demand also rose for short wave frequencies. The number of radio services grew more quickly than the range of frequencies. The Geneva Frequency Plan of 1923 was the first attempt at regulation.

The growing »scramble for frequencies«

The gap between availability and demand continued to rise. Frequency allocations were inadequate for most nations and they stopped observing the regulations laid down in the Frequency Plan. The general free-for-all caused real chaos. Hardly any radio frequencies remained undisrupted by radio stations in other countries, particularly in the evenings. The introduction of frequency grids, performance limits and the wide spatial separation of transmitters using the same frequency generally only brought partial relief in subsequent Frequency Plans.

The use of heterodyne and superheterodyne technology to reduce transmitting power where necessary increased the sensitivity and selectivity of receivers, leading to marked improvements. The exploitation of higher frequencies for radio and television in the VHF range caused a shift away from RF/FS, considerably relieving the old range of wavelengths.

Current use of frequencies – the role of the ITU

Using frequencies is a sovereign regulatory activity and is subject to government control. This means that not everybody has the right to emit signals in a self-defined frequency spectrum. The use of all frequency bands is thus subject to approval of national and international authorities. Non-compliance of the legislation in force is a criminal offence.

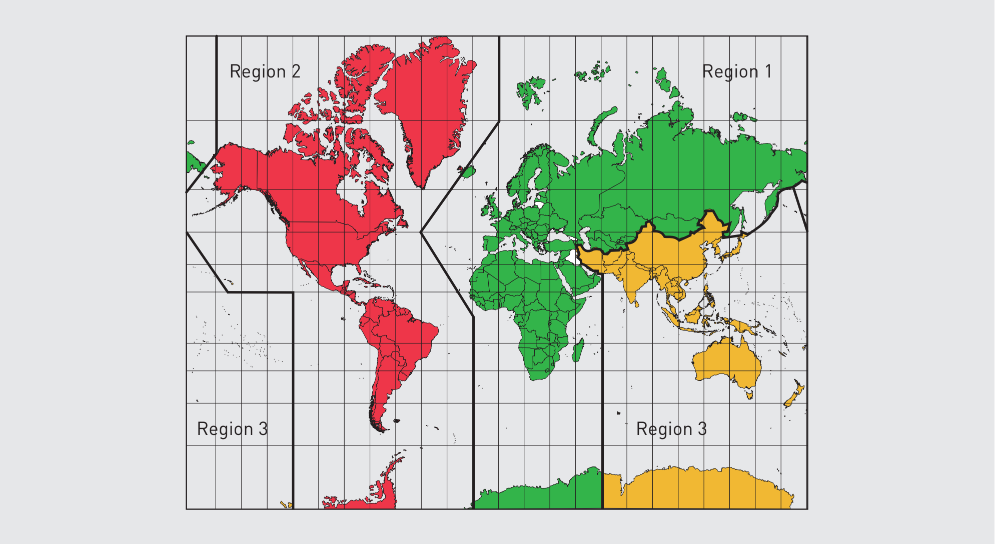
The ITU (International Telecommunication Union) is responsible for international frequency co-ordination and was established to keep abreast of highly dynamic developments in the use of wireless technology across the world, and to aim towards widespread harmonisation. Its objective is to encourage efficient use of resources and prevent interferences between radio services in different countries. Frequency management is based on the ITU's regulations for radio services (Radio Regulations). The Regulations define three international frequency regions:

1. Europe and Africa
2. North, Central and South America
3. Asia and Australia

Radio range is allocated according to an international Frequency Allocation Plan, which allocates both the radio categories and frequency ranges. There are also individual international agreements through which the ITU is aiming to harmonise some frequency ranges that are rated as excellent and used for special services.

European harmonisation

At the European level, the CEPT (Conférence Européenne des Administrations des Postes et Télécommunications) is implementing the relevant standards. The aim here is also to achieve full harmonisation, at least at the European level. A European frequency allocation and usage plan has been in place since 2002. It is to be implemented and become legally binding in



ITU radio regions

accordance with European law and the relevant statutory provisions of each member country.

European Community member states are encouraged to follow the guidelines and rulings of the European Parliament and Council with respect to the regulatory framework on frequency policy, as well as the R&TTE directive and service-specific guidelines and rulings, such as the GSM guideline or the UMTS ruling.

The EU directive »Radio and telecommunications terminal equipment« (1999/5/EC) is a legally binding policy concerning the marketing of relevant devices in Europe and has since become a part of European and member state legislation. The directive allows faster access to the market, as equipment is not subject to extensive national test procedures and regulatory approval. Instead, harmonised conformity assessment procedures are in place to enable the majority of products to be brought into circulation with self-declarations from manufacturers.

National frequency range assignment

As the use of frequencies is a sovereign activity, individual countries implement the guidelines as appropriate for them. In Germany, the guidelines have become part of the German frequency range assignment plan (FreqBZP), which allocates radio ranges to radio services and other applications, as appropriate. The specifications are subject to statutory regulation under the auspices of Germany's Federal Network Agency. In Section 46 of the TKG (German Telecommunications Act), the Agency specifies its frequency usage plan (FreqNP) for individual frequency ranges. Details of all the usage of frequency ranges from 9 kHz up to 275 GHz are contained in the 660 pages of the plan (frequency allocation charts 1 to 486). To delve further into the subject at this point would far exceed the scope of this book.

It is however worth mentioning that the FreqNP allocates special frequencies for equipment used in medical, scientific and industrial (MSI)

Frequency range in MHz	Maximum channel bandwidth/ Channel pattern in kHz	Maximum equivalent transmission power (ERP)/Maximum magnetic field strength	Relative frequency usage duration/ Listen before Talk (LBT)
a) 6.765–6.795	No restriction	42 dBµA/m at a distance of 10 m	No restriction
b) 13.553–13.567	No restriction	42 dBµA/m at a distance of 10 m	No restriction
c) 26.957–27.283	No restriction	42 dBµA/m at a distance of 10 m	No restriction
d) 40.660–40.700	No restriction	10 mW	No restriction
e) 433.050–434.790	No restriction	10 mW	No restriction
f) 868.000–868.600	No restriction/ Narrowband and broadband modulation	25 mW	≤1.0 % or LBT
g) 868.700–869.200	No restriction/ Narrowband and broadband modulation	25 mW	≤1.0 % or LBT
h) 869.300–869.400	25	10 mW	No restriction
i) 869.400–869.650	25/Narrowband and broadband modulation	500 mW	≤10 % or LBT
j) 869.700–870.000	No restriction/ Narrowband and broadband modulation	5 mW	No restriction
k) 2.400–2.4835	No restriction	10 mW	No restriction
l) 5.725–5.875	No restriction	25 mW	No restriction
m) 24.000–24.250	No restriction	100 mW	No restriction
n) 61.000–61.500	No restriction	100 mW	No restriction
o) 122.000–123.000	No restriction	100 mW	No restriction
p) 244.000–246.000	No restriction	100 mW	No restriction

Frequency usage parameters for short range devices (SRD)

From	To	Type	Remarks
6.765 MHz	6.795 MHz	A	SRD
13.553 MHz	13.567 MHz	B	SRD
26.957 MHz	27.283 MHz	B	SRD
40.66 MHz	40,70 MHz	B	SRD
433.05 MHz	434.79 MHz	A	SRD, only region 1 (Europe, Africa, successor states of the USSR and Mongolia)
902 MHz	928 MHz	B	only region 2 (North- and Southamerica)
2.400 GHz	2.500 GHz	B	Region 1, 2 and 3
5.725 GHz	5.875 GHz	B	/
24 GHz	24.25 GHz	B	/
61 GHz	61.5 GHz	A	/
122 GHz	123 GHz	A	/
244 GHz	246 GHz	A	/

Different frequency ranges and their prevalence

applications. These are subject to simplified provisions due to their low effective transmission power.

ISM band

The so-called ISM bands have special significance and may be used licence-free for industrial, scientific, medical and domestic applications. They have a low effective transmission power. Type approval means that the only requirements that must be observed are those relating to transmitting power and interference for adjacent frequency ranges. The devices used for these bands are also known as short range devices (SRD).

There are two different ISM band types:
Type A:
Applications must be approved by the relevant national authorities. Blanket approval may also be granted.

Type B:
Applications do not require approval but must be able to handle any faults that arise. ISM bands in different frequencies are defined by the International Telecommunication Union (ITU). Unfortunately, harmonisation is not yet at the stage where all ISM frequencies are available everywhere. Furthermore, countries may approve their own frequency ranges.

To sum up, this has two effects.

1. The 2.4 GHz band is available internationally as an exclusive band and thus meets the requirements for standard wireless products worldwide. This is very positive on the one hand,

but on the other the different technologies have to compete for the narrow frequency band. WLAN (IEEE 802.11b and g), Bluetooth and ZigBee (IEEE 802.15.4) are the main representatives of the 2.4 GHz category.

2. The favourable propagation characteristics of the sub-GHz range make it particularly attractive. ISM bands of 433 MHz and 900 MHz belong to this range, although the two bands are not available as standard. In Europe, the range from 863 ... 870 MHz is still available as an SRD band for special applications. This frequently results in the development of products that can be used optionally in bands 433 MHz (Region 1), 868 MHz (Europe) and 915 MHz (Region 2).

// Co-existence of radio services

There are very few services that can claim exclusive right of use for frequency bands. Most services are required to share the frequency bands with others. This can lead to mutual interferences, which is particularly problematic in the case of ISM frequencies.

The regulations set out in the FreqNP and the technical specifications contained in the EN, ENELEC and ETSI European standards are in place to prevent such interferences. The aim of these specifications is to afford all competing services established in the same frequency band the highest possible usage by means of »traffic rules«. Or, put another way, the relevant frequency band should be fully but not excessively utilised.

FreqNP regulations

The most well-known provision of the FreqNP is the cap on transmission power, whether at a constant value or at different power values depending on the time of day. So, for example, MW transmitters may transmit using higher powers through the day than at night, when the coverage gets wider as a result of propagation conditions. The second most important provision is the time limit on broadcasting rights, either as a percentage or an hourly calculation.

The next provision is a geographical limitation (territorial coordination), for example the stipulations in VHF frontier agreements to prevent interference in neighbouring territories. France, however, still operates a very powerful broadcast station in Strasbourg, which transmits all the way to southern Germany.

Finally, there are regulations on the shared use of frequencies, which apply in particular to situations where there is an overlap in the activities carried out by radio services. For example, in search and rescue operations, military aeronautical radio-navigation services can have shared use of frequencies 156.3 MHz, 156.375 MHz, 156.5 MHz and 156.675 MHz, which belong to the VHF maritime mobile service. This regulation enables SAR services to work together at sea. Only the German Federal Armed Forces have SAR aeronautical radio-navigation services (helicopters and search planes). These must communicate with ships at sea during deployment. There are similar provisions for public safety organisations and authorities that share certain contact frequencies or even use overlaps on shared frequency band borders.

As a general rule, allocation is normally in terms of primary and secondary users. Generally speaking, secondary services must not disrupt primary services and are not entitled to protection from subordinate services.

European standards regulations

The European Telecommunications Standards Institute (ETSI) is one of the three largest standards organisations in Europe. They define the regulations that should allow the highest

usage possible by a large number of users or services in the same frequency band. A good example is the ETSI EN 300 328 V1.7.1 (2006-10) standard for the 2.4 GH band.

Regulations on shared frequency usage:

- Power spectral density (e. g. 100 mW/MHz)
- Maximum dwell time at a particular frequency (e. g. 0.4 sec.)
- Minimum and maximum number of individual frequencies in the allocated frequency band (e. g. 14 frequencies for non-adaptive and 20 for adaptive frequency hoppers)
- Adaptive frequency agility (aims for a uniformly loaded spectrum, low-traffic frequencies/ channels or unused gaps in the band are preferred)

Rules of access, Spectrum Access Techniques:

- Detect and avoid (DAA)
- Listen before talk (LBT)
- Media access protocol

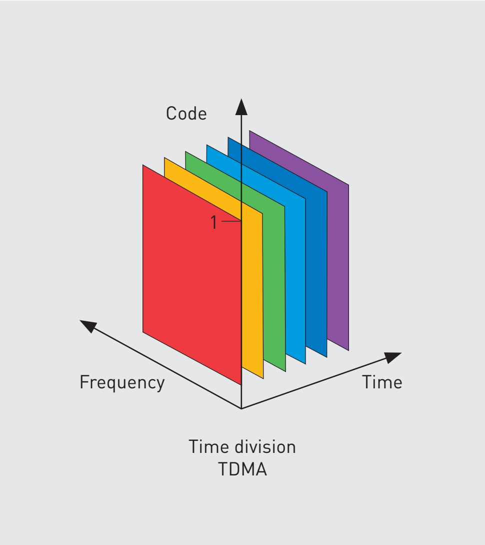
LAN regulations

Alongside the special Radio Regulations, there are also the same administered broadcasting rights as those for wired closed networks (local area networks, LAN). These either have fixed time slots in msec range or »handover« tokens that can now also be used in wireless local area networks (WLAN). This means that all users in the network, according to their data usage, are either allocated a time slot or they can send their data in the fixed user sequence.

In summary, these operational regulations optimise traffic intensity for competing users through fixed or adaptive »traffic rules«, making a valuable contribution to improvements in radio communications.

Multiplex methods for multiple use of resources

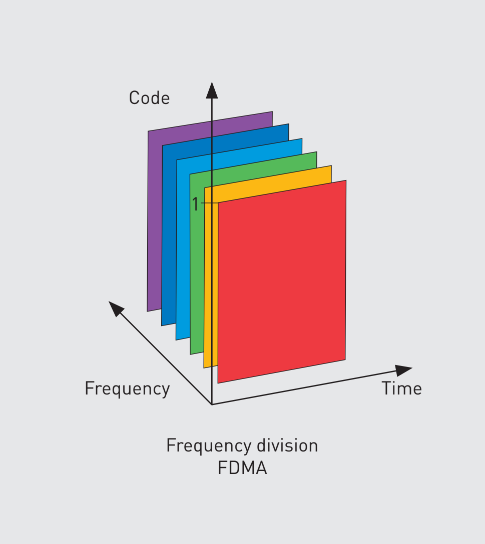
In radio communications, multiplex methods are used for multiple use of the time bandwidth resource. When access is given to this resource using different frequencies within the allocated spectrum, this is known as Frequency Division Multiple Access (FDMA). If access is provided over time, this is known as Time Division Multiple Access (TDMA). Digital processes based on spread spectrum techniques that use the whole spectrum as well as the whole time slot assign unique codes and the process is thus known as Code Division Multiple Access (CDMA).



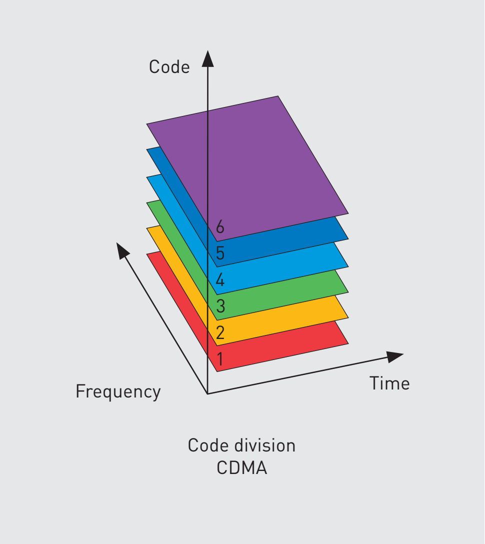
Different multiplex methods allow optimum use of frequencies.

FDMA

In FDMA, part of the spectrum or a channel is available to the user during the entire time slot and the other channels are available to other users. The most well-known example is carrier frequency technology, which allocates lots of telephone channels in a radio-relay or cable system. Two sets of message content can be »multiplexed« in a single channel, for example during stereo transmission in VHF radio broadcasting or sound and image transmission in television broadcasting.



Different multiplex methods allow optimum use of frequencies.



Different multiplex methods allow optimum use of frequencies.

FDMA (except in the case of radio communication services) does not usually involve fixed allocation of a channel to a user for all time slots (synchronous FDMA), just for the duration of the transmission of information. This is why regulations are required - so-called protocols - to ensure that during connection, set-up channels that are already being used remain protected and only free ones are allocated (asynchronous FDMA). This makes usage more economical.

Multiple use of the same frequency at the same time is only possible in FDMA through spatial separation - either through spacing or using beam antennas. Information is modulated directly onto the radio frequency, and not initially on a pulse sequence.

TDMA

In TDMA, the whole spectrum is available for some of the time slots; in the other time slots other users occupy the whole spectrum. Fixed allocation of time slots to users (synchronous TDMA) is not required. This allocation is only needed for the duration of the transmission of information. A user with substantial information requirements can therefore occupy several time slots, provided the other users have correspondingly fewer requirements. This is why a protocol is also necessary here, to protect occupied time slots and only allow free ones to be used (asynchronous TDMA).

Multiple use of the same time slots and the whole spectrum again means spatial separation is required, either through spacing or using directive radio links. Here too, information is modulated directly on the radio frequency and not initially on a pulse sequence.

CDMA

In CDMA, useful information is not modulated directly on the radio frequencies. First, this is done on a code pulse train with a significantly higher bit rate. During this process, bandwidth and time slot demands can become so high that the entire time slot/bandwidth resource is required.

Multiple usage is still possible, but only if the unique characteristics of the code are used. Other users simply use another code. As each receiver knows its code, it can select from multiple voices, decode its message again and prepare it for further processing.

The technical skill lies in optimum coding, which should enable the best possible separation of different users. This is known as cross-correlation. However, the right receiver should be able to quickly and securely synchronise itself to its transmitter. This is known as autocorrelation.

Hybrid approaches

In modern radio systems, combinations of the previously mentioned technologies are frequently used. This means additional advantages, such as increased protection against interception and interference.

Common examples of these hybrid systems are FHSS frequency hopping processes such as Bluetooth, which use both time-slot and frequency-division multiplex technology.

To determine an environment's suitability for co-existence, the relevant channel definitions must be carefully analysed and their susceptibility to interference assessed accordingly.

Conclusion and outlook: the demand for frequencies is still growing

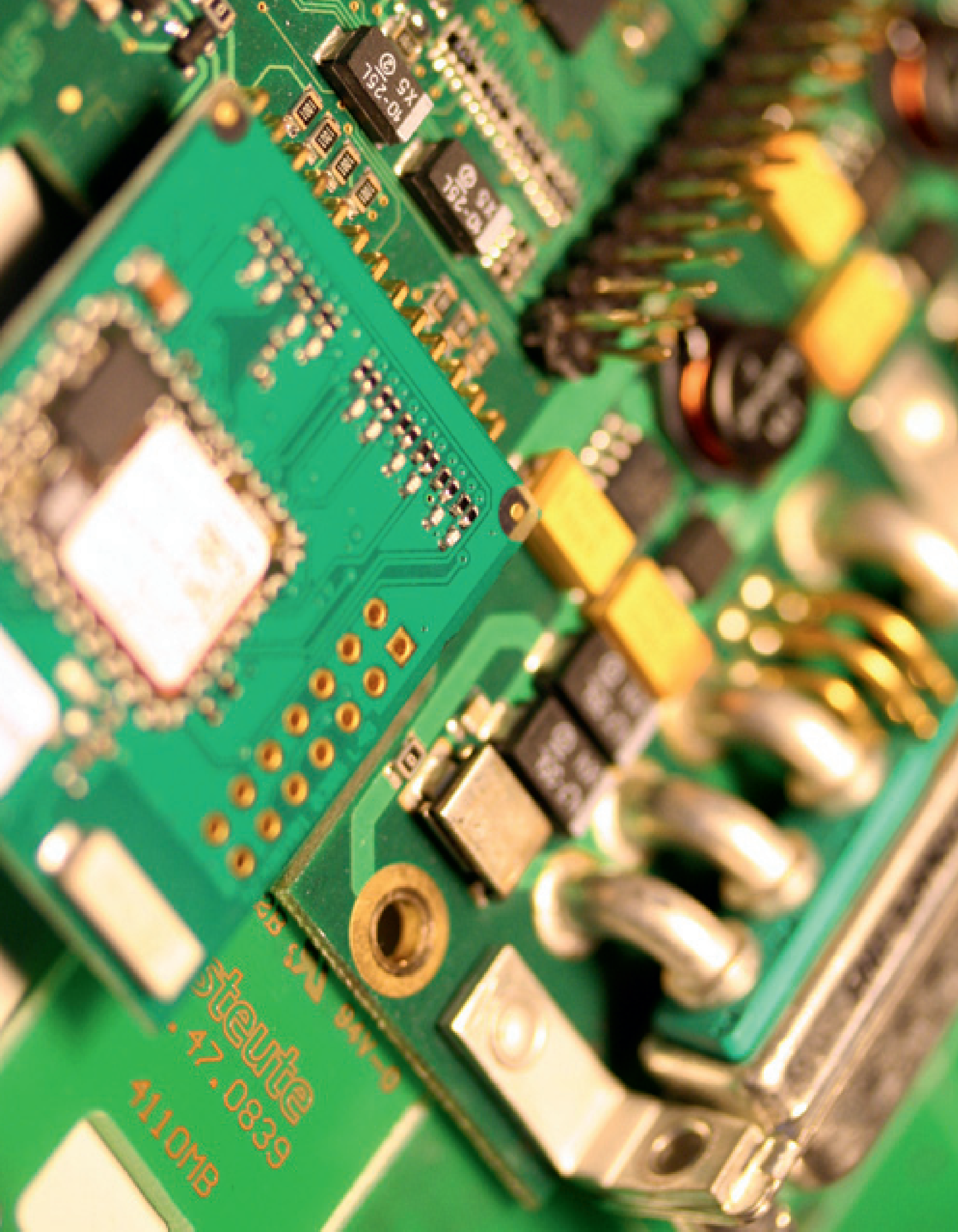
Ultimately, the multiplex processes described here are only necessary because radio network resources are limited. Even today, the dual problem of increasing demand for frequencies and a spectrum that is not freely extendable is still prevalent. Around 40 different radio systems and 2 million radio services require coordination. Furthermore, radio is now being used in outer space. Wireless communication saw globalisation coming and the division of the world into only three wireless communication zones is to some extent outdated. Rules must now be the same and global standardisation implemented. In its Radio Regulations, the ITU has pro-actively established allocations and developed rules for the 9 kHz to 400 GHz spectrum, even though the current technology is still not capable of making the range fully accessible. But that is only a question of time.

In conclusion, the following can be acknowledged:

- As a limited resource, the frequency spectrum must be handled economically and with care in the future
- Services must be assigned to the specified spectrum propagation conditions
- Radio should only be used if there is no other option, such as a broadband cable (in this case, requirements take precedence over convenience)
- All technology options should be fully exploited and measures towards improvement must continue to be pursued

Goals of optimisation:

- Noise, transmitting power, range
- Channel usage, modulation
- Data reduction, redundancy, security
- Operating procedures, access and multiple use
- Protection measures


steute
.47.0839
4110MB

// Wireless coverage

05

Jamming and interferences

Short description of interference protection procedures

Propagation models

Attenuation

Special propagation characteristics

Radio topology

// Jamming and interferences

In addition to pure free field attenuation there are also other influencing quantities that occur in a real environment, although these are difficult to describe. Radio waves basically propagate in a straight line, in the same way as light waves. The fact that the wave characteristics of radio and light are similar means they also have similar effects.

Shadowing effect

This is an effect whereby free propagation of the signal between transmitter and receiver is interrupted by natural obstacles (mountains, ravines) or artificial obstacles (all types of construction, temporary structures, etc.). Devices located in dead spots have a markedly reduced signal level.

Diffraction

This is an effect relating to the behaviour of waves at an obstacle, whereby a wave propagates in a geometric shadow area of the obstacle. Diffraction creates new waves that can constructively or destructively overlap each other, leading to interferences.

Fading

This is an effect whereby refraction or reflection at objects means a signal reaches the receiver via several paths out of phase. Fading losses caused by multipath propagation are of major significance in the case of moving objects.

Reflection

If waves encounter smooth surfaces in relation to their wavelength, they are reflected. Generally speaking, only part of the wave is reflected. The remaining wave portion propagates in the reflection medium. The law of reflection applies to smooth surfaces; diffuse scattering occurs on rough surfaces. The reflected waves overlap each other constructively or destructively.

Scattering

If electromagnetic waves encounter objects that diffract or reflect the wave field, the wave field is scattered. Forward and back scattering occurs on the objects, depending on the surface type of the scatter object.

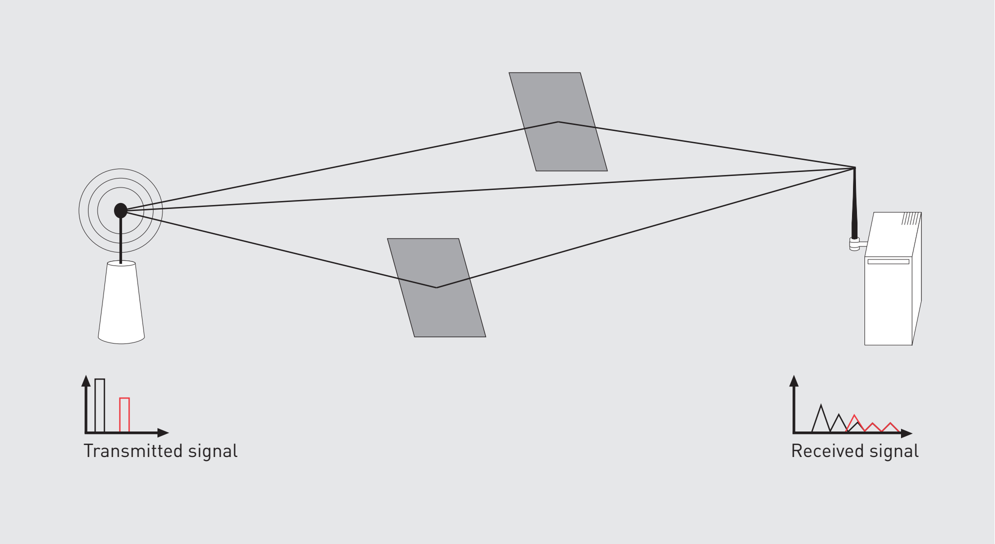
All these interference factors, which manifest themselves to varying degrees depending on frequency, modulation, interference object and environment, can impair radio reception. This can be avoided if the antennae have line of sight. It is for this reason that antennae should be positioned high enough so that they are within each other’s field of vision, but not so high that other antennae are in their line of sight and thus disrupted.

Individual transmission power competes with:

- Broadcast stations from other services that have equal or preferential rights
- Technical equipment such as electric motors, where their commutator sparks
- Electronic devices with a smaller range, such as electronic codes and keys
- Natural sources of interference such as the cosmos, sun, moon and atmosphere.

These all bring energy into the radio field and must be drowned out by individual transmitters.

So that interferences of this type do not cause an interruption in the transfer of information, or cause falsified information to be transferred, various interference protection measures have been developed. These are described briefly in the next chapter.



Multipath propagation creates more difficult receiving conditions

// Short description of interference protection measures

1. Short-term measure (Burst Communcations)

The data is transmitted via the radio field, with a higher bit rate in the shortest possible time to prevent the interferer from detecting or from taking advantage of the response time to cause the interference.

2. Disguised radio transmission

In their bandwidths, transmitted signals are spread wide apart in such a way that the power spectral density breaks down under the natural contact noise, preventing radio communications from being detected by the interferer. Only the authorised receiver can despread the signal correctly again and distinguish it from the noise and interference level.

3. Frequency Hopping

The »frequency hopping spread spectrum« (FHSS) transmits short bursts of data on continuously changing frequency channels. The interferer cannot track them because it has insufficient response time. But if the interference is preventative and over a wide band, the interferer must spread its power so widely that the power spectral density is no longer sufficient to effectively disrupt the high performance of the hopper. FHSS is also a way of allowing several users to simultaneously occupy the same frequency band, provided their collision rate remains under 40 %.

4. Direct sequence spread spectrum

»Direct Sequence Spread Spectrum« (DSSS) modulates a higher pulse rate over the full spectrum of a binary random sequence (direct sequence) with the information signal and transmits it to the radio signal. This minimises the power spectral density, as is the case with disguised radio transmission, but it must not drop below the natural noise level. This process is very effective against all potential scrambling methods. The noise signals for authorised receivers are attenuated by the despreading process to the same level as contributed by the spread ratio of direct sequence bit rate (chip rate) to information bit rate. This draws the interferer into the complexity of a power escalation.

5. Chirp

A chirp in the context of signal processing is a signal whose frequency changes in time. Band-spreading modulation techniques such as Chirp Spread Spectrum (CSS) are suitable for retrieving information signals from noise or interference through pulse compression at the receiver end using SAW (Surface Acoustic Wave) filters. Technical applications are available mainly for SAR (Synthetic Aperture Radar) systems in aircraft and satellites.

6. OFDM

OFDM (Orthogonal Frequency Division Multiplex) is a modulation technique that uses several orthogonal carrier signals within the radio channel for digital data transfer. If narrowband interference occurs within the channel bandwidth of the OFDM signal spectrum, the carrier signals affected by the interference can be excluded from the data transmission. The total data transmission rate then only drops by a small proportion. OFDM is therefore more suitable for preventing random interferences and frequency swept interferers (chirps) than for preventing broadband interferences.

// Propagation models

If we consider the previously described characteristics, an exact description of radio propagation in real conditions would be difficult. Various channel models can be used to assess the comparability of ranges for different devices.

- LOS (Line of Sight) – describes the pure line of sight between transmitter and receiver. An interference-free line of sight is usually required here.
- NLOS (Non Line of Sight) – describes a communication relationship in which transmitter and receiver are out of each other's field of vision. This means it is practically impossible to compare systems, as basic conditions are not standardised.
- AWGN (Additive White Gaussian Noise). An AWGN channel is an idealised channel with pure LOS connection, where the actual interference is reproduced by an additive white noise. AWGN channels are frequently used in mathematical models, but do not reflect reality.

In a real environment, radio signal interference through overlapping always occurs. This means the transmitted radio waves reach a receiver via different paths, a process known as multipath propagation. This is an unavoidable occurrence in real environments. Different modulation techniques are sensitive to reflection and multipath scenarios in different ways.

// Attenuation

To achieve reliable propagation situations, it is not sufficient to use the typically specified LOS ranges for radio systems as a yardstick. The idealised conditions are corrupted in the real environment by walls and roofs, meaning that as a general rule only much smaller ranges are usually reached.

Examination of current building materials reveals very different types of attenuation behaviour, which is also dependent on frequency. The following illustrations show several diagrams taken from the BSI guideline »Electromagnetic shielding of buildings«, which provides a qualified description of the attenuation behaviour of different building materials and constructions.

It is clear from the diagrams that no qualified statements can intuitively be made about the attenuation behaviour of building materials.

Generally speaking, low frequencies transmit far better through walls than high frequencies. See the table on the right to get an initial idea.

Material	Thickness	Attenuation
Wood	< 30 cm	1 ... 10 %
Plaster (board)	< 10 cm	1 ... 10 %
Glass (uncoated)	< 5 cm	1 ... 10 %
Particle board	< 30 cm	30 %
Pumice	< 30 cm	10 %
Aerated concrete blocks	< 30 cm	20 %
Brick	< 30 cm	35 %
Reinforced concrete	< 30 cm	30 ... 90 %
Metal grille	< 1 cm	90 ... 100 %
Metall, Alu lamination	< 1 cm	100 %
Rain, fog, snow	/	60 ... 90 %

Attenuation behaviour of individual materials

// Special propagation characteristics

If several radio users are active in the same environment, this can produce other negative effects that may add to existing ones. For example, the 'hidden station problem' effect. This refers to a communication relationship in which a user cannot detect an active communication, due to being outside the transmission and recognition range.

An example:
Station A communicates with station B. Station C would also like to contact station B, but is outside A's recognition range. C cannot detect that A is in communication with B. If C now starts communicating with B, a conflict can arise in station B as C's communication with B is sending uncoordinated signals to A. This usually results in an interruption in communication at both ends.

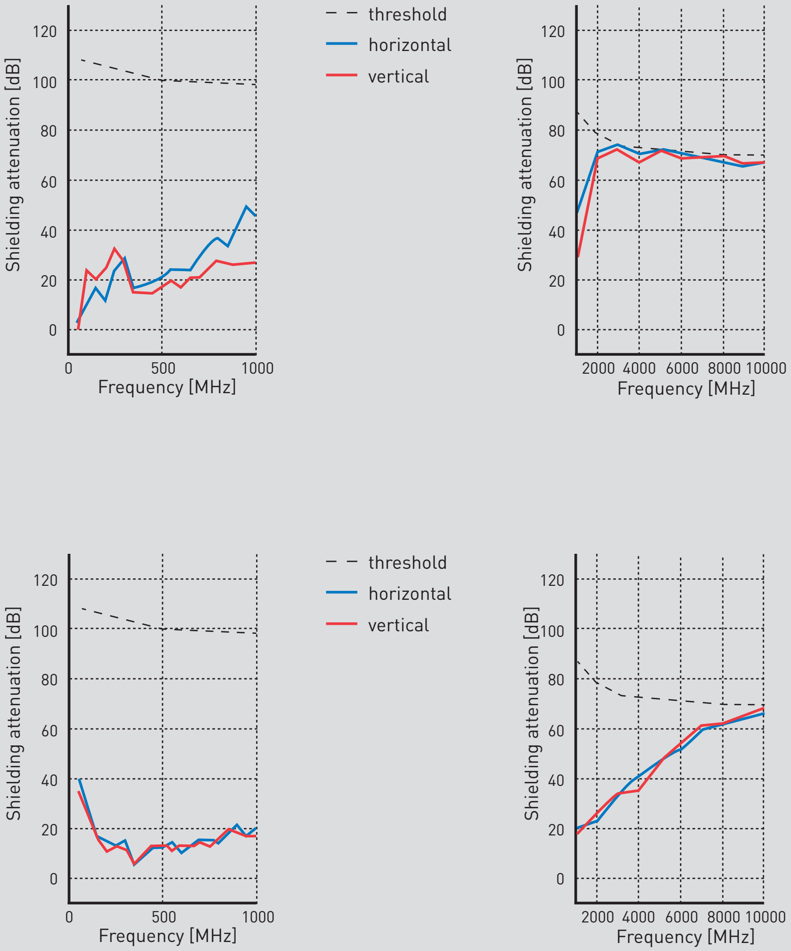
Another communication scenario is the 'exposed station' problem. This occurs when a station is affected by communications between other devices. In the case illustrated, station B communicates actively with station A. Station C would like to establish communication with D. C notices that the channel is occupied by B, even though communication was established mainly with D. Communication cannot be established.

// Radio topology

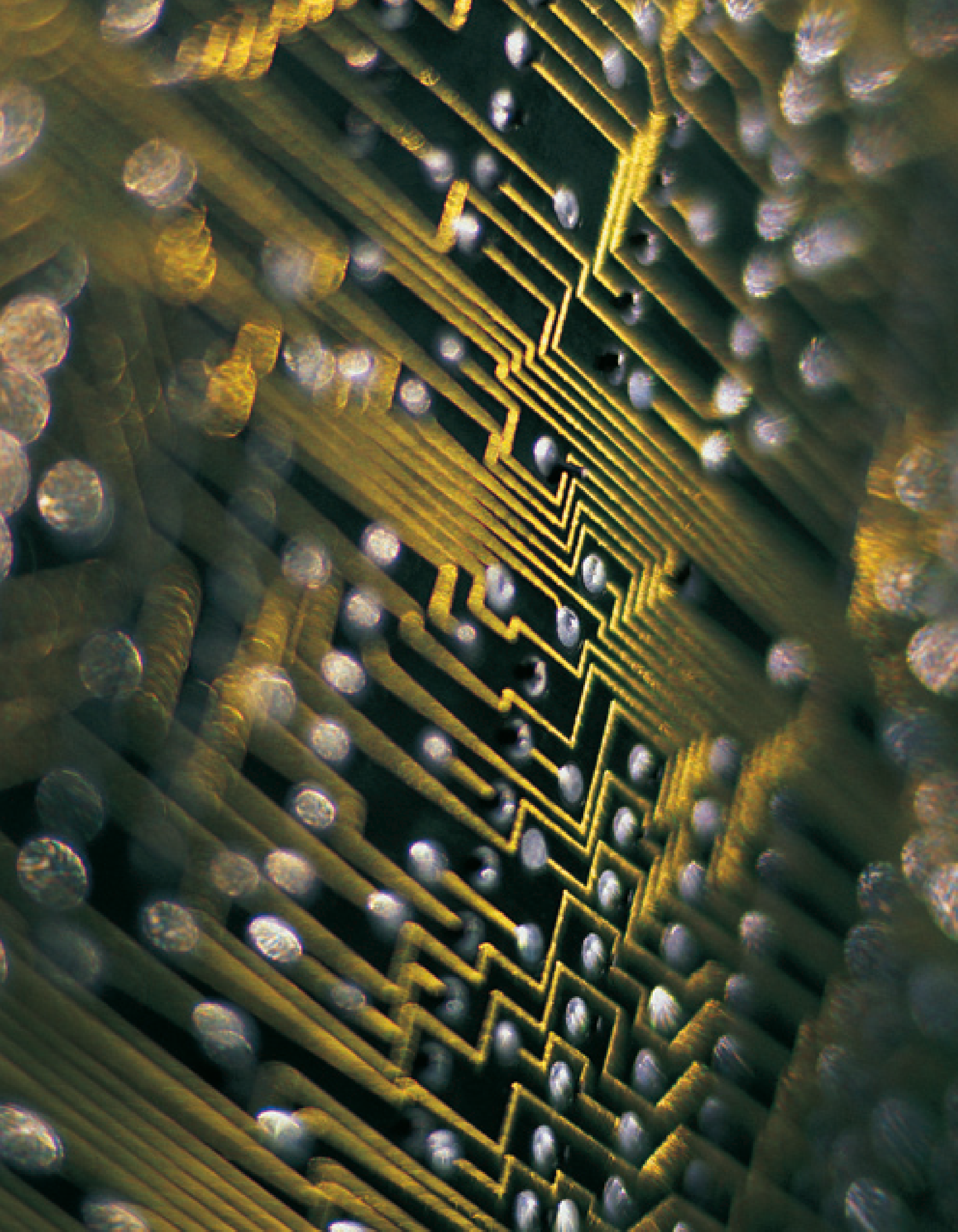
A central aspect of using wireless technologies is configuring all the possible network topologies. The ideal network would be infrastructure-free and enable mobile units to communicate easily with one another, without any problems. But are these types of networks really optimal? A good example is mobile telephony. The illusion of telephoning from mobile phone to mobile phone occurs in reality via a complex managed network. A radio link only exists between base stations and mobile devices. The remaining communication is processed via conventional infrastructure networks. The optimum structure thus lies in the right combination of direct and infrastructure-free cabling.

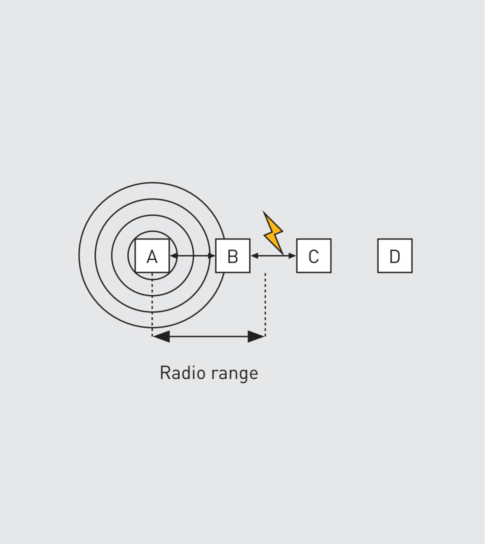
Fixed infrastructure

Compared with wireless systems, conventional fixed wired networks have important advantages. The structure is fixed and only dependent on the connected systems, meaning that behaviour can be predicted if the topology and technology are familiar. Field bus systems and Ethernet networks demonstrate that fixed infrastructures are relatively easy to manage.

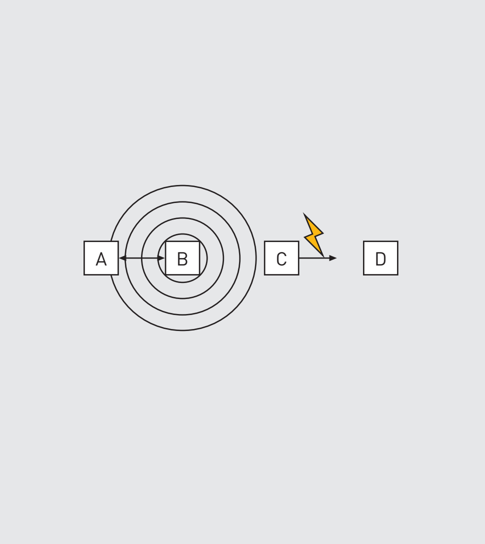


Top: frequency-dependent attenuation behaviour of a sand-lime brick wall (source: BSI).
Bottom: frequency-dependent attenuation behaviour of a 20 cm steel-reinforced concrete surface (source: BSI)





The ‘hidden station’ problem affects stations outside the radio range



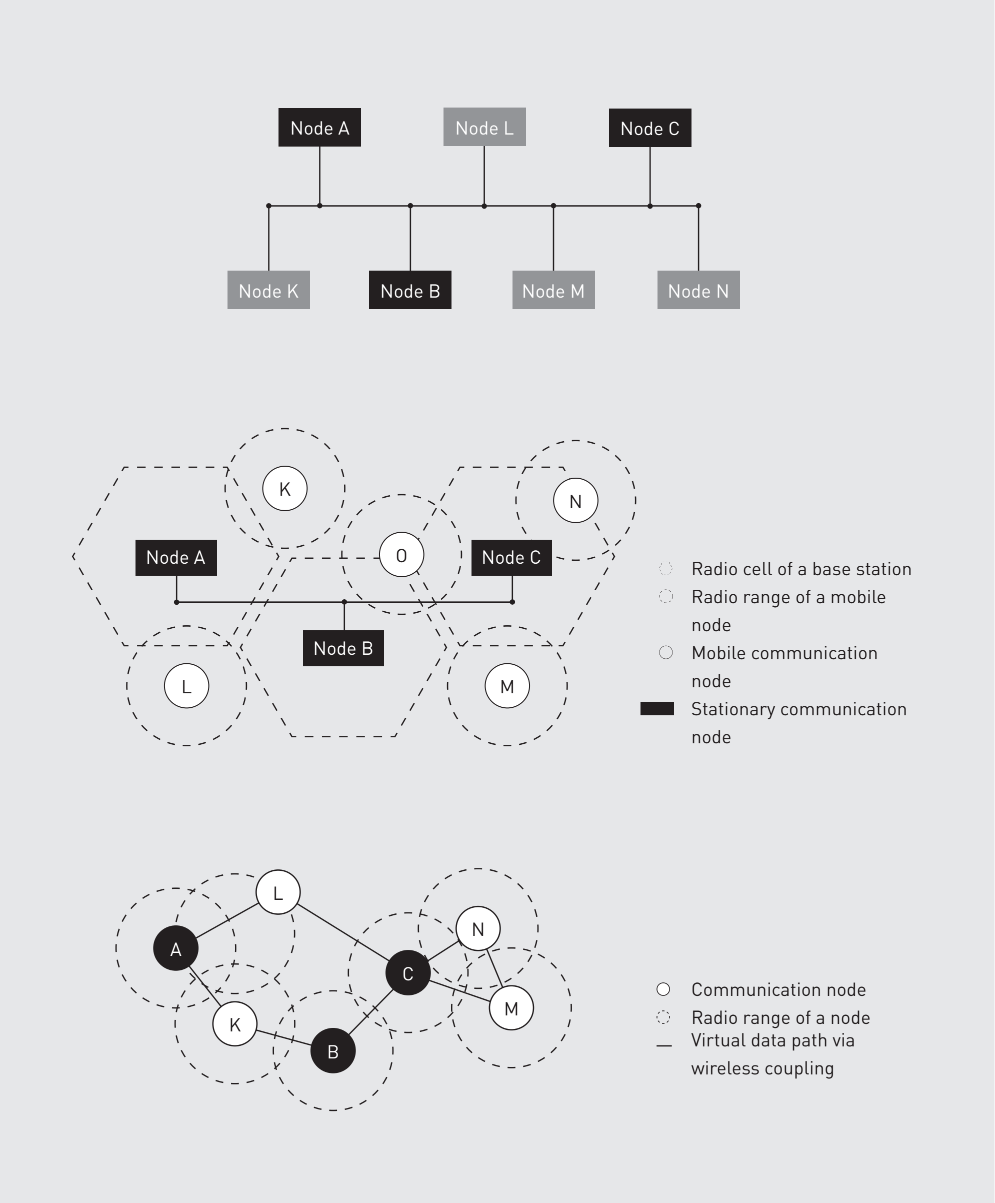
The ‘exposed station’ problem limits the number of possible connections

Mobile infrastructure networks

Making a fixed structure more flexible means infrastructure networks can be used for mobile applications. The network is converted to an infrastructure network and the different access points are connected via the infrastructure network. Depending on the structural features, access points can be managed centrally or decentrally and coverage of the area is either exhaustive or selective. In the case of cellular GSM networks, complete coverage of the area with adjoining radio cells is essential. In the case of individual applications, incomplete cellular structures are also possible because the radio cells are already statically connected.

An important advantage of this type of structure is how relatively easy it is to predict behaviour of the systems. Communication behaviour can be predicted by the single hop from the mobile elements to the base station, using the route knowledge from the fixed infrastructure. Data security and accounting procedures in an infrastructure network are comparatively easy to manage.

Special requirements are applicable for mobile devices. There are no problems here in terms of linking mobile devices to an individual base station, but if a mobile device moves across the entire network over radio cell limits, this will present a challenge. On the one hand, it must be possible to detect mobile devices at any given point of the network and they must be connected to the network - a process known as roaming. On the other hand, communication of a moved device should not be interrupted, even when leaving a network cell range. The device should therefore be passed from one stationary cell to another, completely transparently. This is known as handover. Both behaviour patterns are usually desirable for mobile devices. In practice, allowances are made for this behaviour technically and it occurs, both in GSM networks and with WLAN802.11, as described in the WIFI specifications.


Node A
Node L
Node C
Node K
Node B
Node M
Node N
K
N
O
Node A
Node C
Node B
L
M
Radio cell of a base station
Radio range of a mobile node
Mobile communication node
Stationary communication node
L
A
N
C
M
K
B
Communication node
Radio range of a node
Virtual data path via wireless coupling

Top: Fixed linear cabling is easy to manage.
Centre: Infrastructure networks are easy to manage and are suitable for use in a variety of wireless systems.
Bottom: Ad-hoc networks will in future enable infrastructure-free communication across pico-cellular structures.

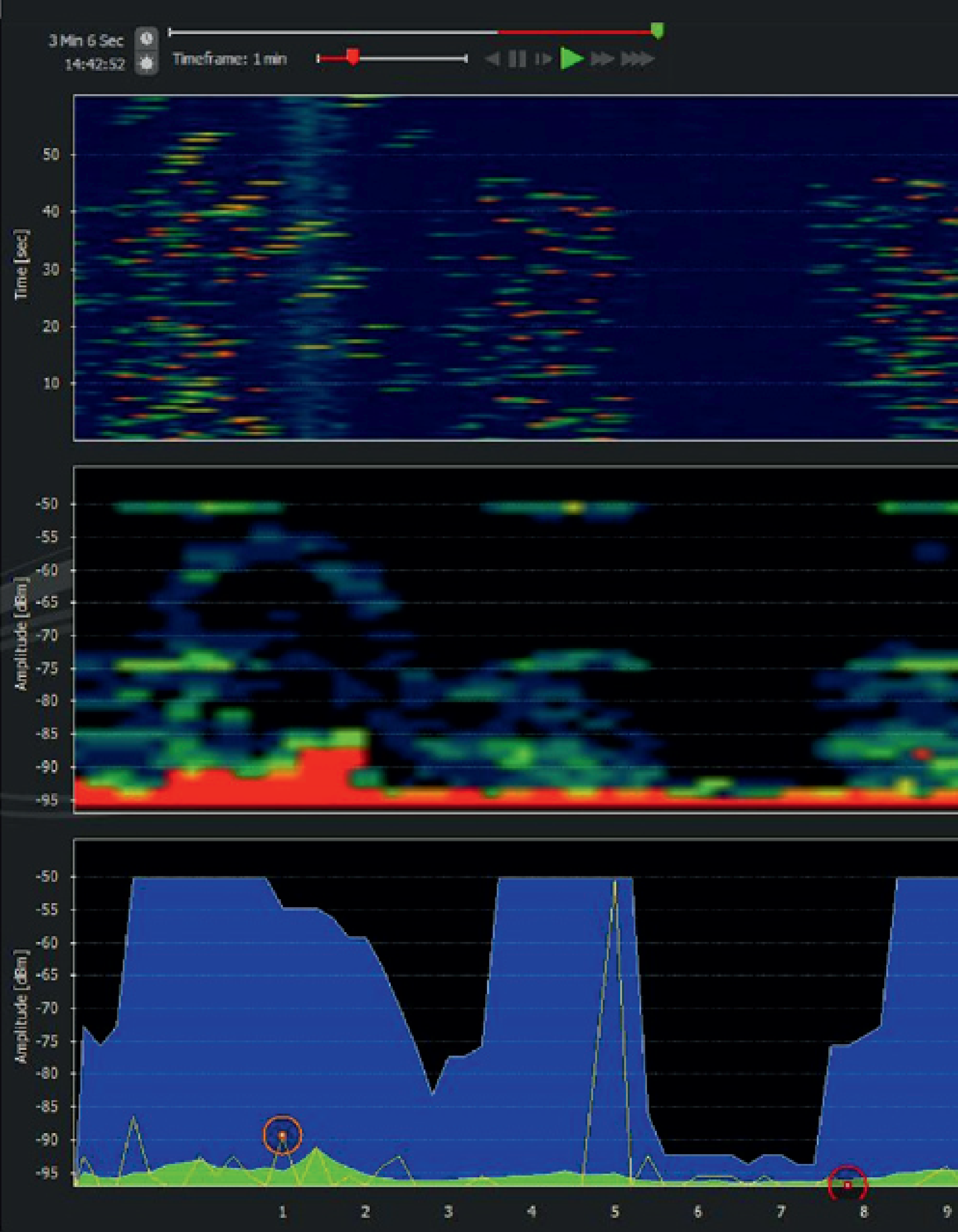
Peer-to-peer and ad-hoc networks

Networks where no infrastructure components are used are known as peer-to-peer networks. The mobile devices communicate directly with one another, without having to rely on infrastructure components. This behaviour is now commonly applied in Bluetooth and WLAN 802.11 networks. Not using any infrastructure is certainly advantageous in terms of cost-effectiveness. The problem is security of access to the network, as it is not usually known which transmitter or computer is integrating into the network.

A special form of these so-called peer-to-peer networks is the ad-hoc network, whereby the stations organise themselves independently and may forward messages to other stations within the network, without using the full range themselves. This particular form of self-organising networks is being researched and shows promise of exciting new application scenarios. In the case of packet switching in particular, the data packet travel paths can reach all over the world, depending on the network topology, although this can result in loss of network control if there is an overflow of foreign data packets in an individual network.

The main difference between this and peer-to-peer networks is that with ad-hoc networks the mobile devices have their own sub-functions, such as routers and gateways, and forward messages completely transparently. Message switching is fully delegated to the terminal equipment, which not only processes the individual data traffic but also communicates it from other stations in a tandem-type operation. Furthermore, ad-hoc networks build themselves dynamically. This means that a network can only be considered static for incremental periods.

To date, not all ad-hoc network problems have been resolved. In terms of research, approaches are being developed through autonomous mobile platforms and multi-hop networks. RFC 2501 defines the standard for these networks under MANET (Mobile Ad Hoc Networking).


3 Min 6 Sec
14:42:52
Timeframe: 1 min
Time [sec]
50
40
30
20
10
Amplitude [dBm]
-50
-55
-60
-65
-70
-75
-80
-85
-90
-95
Amplitude [dBm]
-50
-55
-60
-65
-70
-75
-80
-85
-90
-95
1
2
3
4
5
6
7
8
9

// The frequency channel

06

Channel utilisation by modulation

Digitisation, data reduction, redundancy and error protection

// Channel utilisation by modulation

Previously, the external optimisation of wireless systems was regarded with frequency, power and site coordination. But now, it is about the internal optimisation, the best possible use of the allocated bandwidth. In order to do this, we have to select the proper modulation process.

Amplitude and Single Side Band modulation (AM/SSB)

In times of analogue transmission, the first thing was to modulate a transmitter or carrier frequency in its amplitude. This is called amplitude modulation (AM). Two side bands are formed in the process having the same information, namely HF carrier plus tone frequency (or low frequency, LF) and HF carrier minus LF. We can omit one of side bands without losing information. This is called Single Side Band (SSB) modulation. There are also variants with carrier frequency, reduced carrier frequency and suppressed carrier frequency.

As a result, it is possible to cope with half of the bandwidth and to utilise the other half otherwise. We can for example double the low-frequency range from 3.4 kHz (voice channel) to 7 kHz (music channel), or we can transfer two different sound channels into one radio wireless channel.

Single side band transmission is common with modern shortwave radio. In individual cases, we find medium and long wave modulation. Data transfer via short wave takes place by modulation of 2n sound frequencies with so-called parallel modems (modulator-demodulator, for n = whole numbers).

It is advantageous if the selective fading in the transmission channel does not affect all sound channels at the same time. With a customised error protection process, this results in good transmission safety. The other process uses a fast serial bit transmission with delay equalisation (serial modem). Both processes have specific advantages and disadvantages.

Broadband multi-channel systems for telephone transmission via directional radio and cable are still using today the single side band technology with the carrier frequency technology. This results in simultaneous transmission of 12, 24, 48, 60, 120, 240, 480 or 960 phone channels. Analogue TV also uses single side band technology. This technology is not fully suited for high-quality video and audio transmissions.

Frequency modulation (FM)

Introduction of the ultra-short wave or Very High Frequency (VHF) resulted in a lot more bandwidth - instead of 1 MHz with medium wave, first MW 12 MHz and later 20 MHz. The opportunity was used to introduce the frequency modulation FM. In the process, the transmission frequency resonates at constant amplitude in the LF pulse around the centre frequency. Instead of amplitude, maximum deviation takes place from the centre frequency, the frequency hub.

FM is immune to many errors, particularly against lightning, because the amplitude interferences are clipped. Besides, by widening the channel bandwidth to 300 kHz, compared to the AF bandwidth of 20 kHz for music, 1:15 or 12 dB is gained. This frequency spread leads to a considerable increase in transmission quality which opened the door for Hifi. However, this was not an economic usage of frequency but rather the opposite. A bonus effect of this large bandwidth was the later development of stereo transmissions. Due to the quasi-optical spread, FM channels can be allocated numerously. FM channels established themselves worldwide as the band for radio transmissions and has relieved the load on other frequency bands.

Digital modulation processes

Modulation processes use information in the form of bits. Bits are transmitted here as »In« state and no bits as »Out« state. The simplest form is the Morse code. In this case, the transmission frequency is switched on or off, or to be precise, the amplitude is shifted – as the expression amplitude shift keying (ASK) suggests. Shifting of the transmission frequency between two discrete values with continuing amplitude is called Amplitude Shift Keying (ASK).

A common form is Soft Frequency Shift Keying, in which the bits are formed in advance by a Gaussian filter, the so-called Gaussian Frequency Shift Keying (GFSK). The GFSK process is used for example by DECT, Bluetooth and WLAN. The modulation of the transmission frequency phase (around 180°) at constant amplitude and frequency is the third and most important digital modulation process. It is called phase modulation (Phase Shift Keying, PSK; see below).

Hybrid forms of digital modulation

In addition to the basic forms of digital modulation, there are also mixed forms with analogue processes. Discrete Multitone Transmission (DMT) is a multi-carrier process, where many discrete analogue audio frequencies within the transmission channel are parallel keyed and transmitted. This procedure is used in ADSL technology for data transmission in the ISDN channel.

The quadrature amplitude modulation (QAM) is a type of modulation that combines amplitude modulation and phase modulation. In this process, two independent signals are imposed on the same carrier oscillator. In principle, each signal is modulated via amplitude modulation on a carrier of the same frequency, however, with a phase that is shifted by 90°. Subsequently, the two carrier vibrations, which were modulated in such a way, are added. OFDM (Orthogonal Frequency Division Multiplex) is a specialised implementation of the multi carrier process. Process information that needs to be transmitted with a high data rate is first allocated to multiple partial data streams with low data rate. These multiple partial streams are modulated, each with a conventional modulation process, such as the quadrature amplitude modulation with low bandwidth. Subsequently, the individual carrier signals are added. This minimises the influence of sub-carrier signals on each other.

Phase modulation PM

Phase modulation is closely related to frequency and angle modulation. Phase and frequency modulation are forms of angle modulation. It is known as phase shift keying with the abbreviation PSK. In doing so, the phase of a sine wave oscillation (carrier) is modulated by phase shift. It is called binary phase modulation (BPSK) if there is switching (shift keying) between two phase layers. Typically, phase layers 0° and 180° are equivalent to the binary states »0« and »1«.

In a multi-level phase modulation, a-shift modulation represents a sequence of multiple bits. With 4-PSK (QPSK), switching takes place between 4 phase stages, e.g. 45°, 135°, 225° and 315°. Each phase shift equals 2 bits. With 8-PSK there are 3 bits, and so on. 4-PSK is used for example when transmitting a facsimile via the telephone network. With GSM expansion EDGE (Enhanced Data Rates for GSM Evolution), 8-PSK is used for instance, and this triples the gross transmission rate compared to a binary modulation in the otherwise identical channel.

QPSK is used with signal transmission in digital satellite channels, terrestrial transmission of digital signals and also with wired transmission procedures. This type of modulation is now also used for HSDPA (High Speed Downlink Packet Access) technology in UMTS networks. The data rate increased from 384 Kbit/s to approx. 2 mbps and in mobile communication and allows data

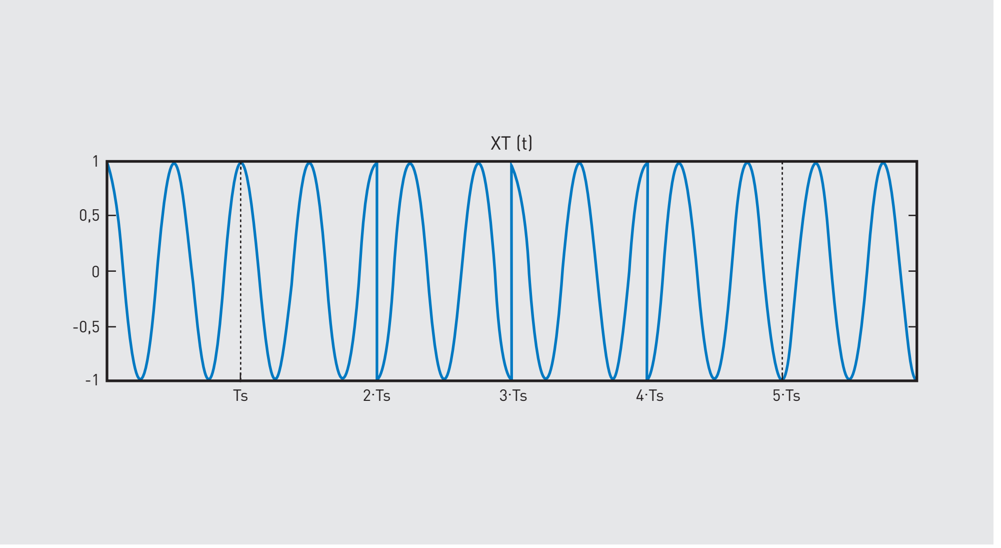


Image shows the binary signal ooLoLL in form of phase shifts.

rates comparable to DSL fixed networks. Physically, QPSK provides the same results as 4-QAM, which uses amplitude rather than phase modulation, and therefore should not be confused with it.

Phase modulation can be adapted quite successfully to the transmission characteristics of a given channel and has prevailed globally for the digital transmission of voice, radio, television and data. The receiver must no longer detect the exact value of an analogue amplitude or frequency. It is enough to decide whether a bit is present or not. This makes the phase modulation stable against natural disturbances in the transmission channel. The signal-to-noise ratio can be reduced, thus leading to larger ranges with the same transmission power, or to a reduction of transmission power within a preset range. The key advantage of all digital modulation processes is the transfer of a maximum possible data rate through the optimal adaptation to the transmission channel's physical properties.

Pulse shaping

Steep pulse flanks of the digital information present cause the formation of harmonics, with each modulation process, which can extend beyond the allocated bandwidth and interfere with adjacent channels.

A pulse is an amplitude modification in the time dimension. If converted into the frequency dimension, it results in a basic frequency and a range of harmonic frequencies. This can be mathematically proved with the Fourier transformation. The steeper the flanks of the pulse and the smaller the pulse, the higher is the power percentage of the harmonics.

The receiver of digitally modulated information requires no super steep flanks. It is sufficient to know whether a potential difference exists for a pulse in the time window that exceeds a certain threshold or not. The receiver can subsequently recreate the steep flanks.

The problem has two solutions. The first solution consists in lowering the data rate so much that the harmonics remain in the allowed range. The other and better solution consists in making pulses »softer« by allowing the flanks to slightly rise and fall. To achieve this, the »hard« pulses are sent through a pulse-forming filter.

In practice, three pulse forms prevail. The cosine pulse is the positive half wave of a cosine oscillation. The cosine square pulse is a bell curve as displayed between the two minimum values of a cosine base oscillation, however, with double basic frequency. The Gaussian pulse is a broad bell curve, like the Gaussian normal distribution; but the pulse on the flanks is clipped. This pulse is regarded as that with the lowest percentage of harmonics.

If we strip away more and more harmonics from a rectangular signal, which alternately consists of an equally long bit, only the sinusoidal fundamental oscillation will remain at the end. By doing so, a complete oscillation carries 2 bit information. In short: 1 Hertz = 2 bit. If the bits do not occur in alternating sequence, more bandwidth is necessary. There is a limit, where we can still transmit as many bits as possible through the narrow channel. Depending on the limiting conditions, it lies between 1.3 and 1.6 bits per Hertz. The resulting harmonics reach into the adjacent channel, but they are weak interferers and cannot damage the robust data transmission.

Many digital services make use of this, including digital satellite television. For example, digital aircraft radio transmits 16 Kbit/sec in the 25 kHz grid, which are 1.5 bits per Hertz.

Polarisation

An often used option for optimisation consists of using a radio channel twice, made possible by separating both transmission channels only through the horizontal or vertical polarisation of radio waves in the same frequency band. This is a standard method in satellite technology.

// Digitisation, data reduction, redundancy and error protection

Analogue signals

For the transmission of acoustic or optical information via a radio channel, the particular signal must first be put in electrical form. This can be accomplished by using a microphone or an older, still analogue video camera. It converts sound vibrations into electrical oscillations. But many measured values still result in analogue form, even if they exist directly as electrical values.

These signals can be used directly for the analogue modulation of radio channels and then occupy a typical bandwidth, as shown in the previous chapter. Analogue transmission is fortunately narrow-banded, but unfortunately very sensitive to errors in the transmission channel, which cannot be eliminated later.

If several radio fields are cascaded in a row to bridge long distances, the interfering components are added, and the connection degenerates until it cannot be used any longer. Therefore the CCIR (now ITU-R) defined the dummy reference circle of 2500 km for directional radio and cable and determined interference components in such a way that it is still possible to obtain understandable remote transmissions with 3 Pico watt per km (3 pW/km) as well as to manage technology demands to make sure that devices are still affordable. However, the transmission limit of 2500 km range remains.

Digital signals

When transferring with digital modulation, interferences in the radio channel can be hidden on the receiving side. This means that the received signal is retrieved with few errors. Thus, there is no longer any range limit. This is one of the major reasons why digital transmissions are increasingly popular. But in order to do so, the resulting analogous information must be digitised first. To achieve this purpose, several processes were tried and tested. Delta Modulation and Pulse Code Modulation are two of them and they are briefly explained here due to their importance.

Delta modulation samples and compares analogue electrical sound signals, for example 16,000 times per second. With increasing amplitude, binary 1 is sent. If it drops at equal value, binary 0 (0101010 sequence) is sent. In other words, the bit rate is 16 kbps. That is enough for voice transmission but not for music transmission. For music transmission, we need higher position indicating rates, preparation and post processing of the signal, e.g. pre-emphasis and de-emphasis.

With the Pulse Code Modulation (PCM), the analogue value is also sampled at constant time intervals. The value is measured with a scale and issued as a digital number. In the ISDN phone channel, language is sampled 8000 times per second and output is a byte of 8-bit length. Accordingly, the benchmark is 28 = 256 subdivisions. Consequently, the bit rate in the phone channel is 8 bits x 8000/sec = 64k bps. Pre- and de-emphasis are here the applied tools and it results in a very natural language, like a sound »from the next room«, though the participant might be speaking from the other side of the globe.

PCM was already invented by Reeves and patented in 1939. Back then, there were no technical possibilities to put it into practice. But today, in the age of the integrated digital building blocks, modulation is economical and useful. Digital technology is now cheaper than analogue technology. This is another important reason for the success of the digital data transmission.

It can be mathematically proven that both digital modulations can reproduce the original analogue signal clearly and free of error, apart from the quantification noise which is a product of the amplitude's quantification levels.

Based on a bandwidth of 3.4 kHz per analogue voice channel, the bandwidth demands of the digital voice channels are far greater. Calculating 1.5 bits per hertz, the Delta modulation is around 25 kHz and the PCM is around 100 kHz bandwidth. This means that a data reduction is necessary.

Data reduction

Data reduction is the application of processes which reduce the amount of data according to purpose without noticeably worsening the information content.

Here is an example: A DIN A4 page 210 x 297 mm is sampled with 600 pixels per inch (pixels). That makes 35,925,120 pixels. Each point has 24-bit colour depth. This results in 862.2 mbps (in bitmap format .bmp). The page is not printed and blank, so we could instead send just the »empty page« or »all white points«. That would be less than 1 Kbit, and the data reduction factor would be around 1:1 million.

In reality, the factor is smaller but the reduction is still worth it. It can be either used to reduce the bandwidth or to reduce the transmission time. International research institutes and international expert groups, such as the »Joint Picture Expert Group JPEG« are entrusted with this task, and they have already developed good compression processes, such as the JPEG-2 or JPEG-4.

```
    if (x > y) {
        temp=y/x;
        ans=x*sqrt(1.0+temp*temp);
} else {
        temp=x/y;
        ans=y*sqrt(1.0+temp*temp);
}
return ans;

fcomplex Csqrt(fcomplex z)
{
    fcomplex c;
    float w;
    if ((z.r == 0.0) && (z.i = 0.0)) {
            c.r=0.0;
            c.i=0.0;
    } else {
            w = sqrt((sqrt(z.r*z.r + z.i
            if (z.r >= 0.0){
                    c.r=w;
                    c.i=z.i/(2.0*w);
            } else {
                    c.i=(z.i >= 0)
                    c.r=z.i/(2.0*c.i);
```

The Moving Picture Expert Group (MPEG) processes data reduction of images (also moving). Their standards record not only video data but also the associated audio data. Their standards are the basis for digital television (Digital Video Broadcast DVB) and DVD.

PC and the Internet use these compressed data formats widely. Such processes enable, for example, videoconferencing over the ISDN telephone line or Internet surfing with the mobile phone. By the way, »facsimile transmission« (fax) via phone line was the historical predecessor of this data reduction.

Redundancy and error protection

Redundancy literally means abundance. Virtually any information contains redundancy. This can be seen by covering up the upper or lower third of a printed word; if we still can read the word, the covered part is redundant.

In a broader sense, redundancy refers to those measures that remove unnecessary parts from information, until only the essential data remains for the transmission. Freed up data can be filled out with wisely selected control data space and incoming bit errors can be detected and corrected by the receiver.

There are basically two processes to realise this. With the first and older process, the receiver detects the errors based on the attached parity bits and is able to automatically request a repetition of the erroneous data word, which is then immediately submitted. This procedure is called »Automated Request« (ARQ) and is the most commonly used procedure as far as two-sided connections (duplex) are concerned.

The second error protection process is the preliminary Forward Error Correction (FEC). This procedure provides so many additional bits to the useful information that the receiver is able to correct most common errors. A residual error rate remains. This process is often used in unilaterally set up wireless connections (Simplex), if the remote station does not respond.

For both processes, there are a large number of variants which are adapted to each typical transmission characteristic and error marker frequency of the wireless channel, and can generate significant amounts of additional data. Error protection is »paid« either with additional bandwidth or with additional transmission time.

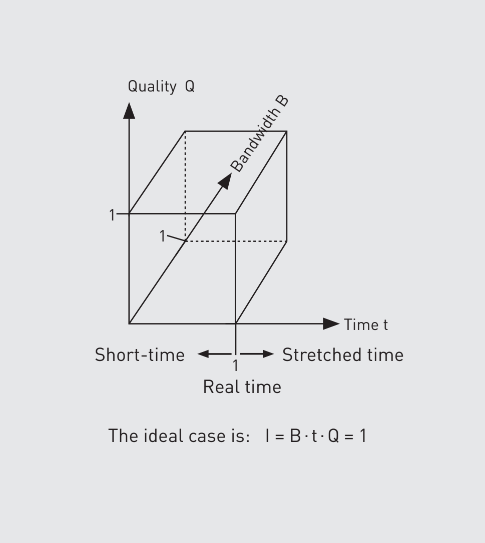
The transmission theorem

With real time or online transmission the data is transmitted without delay so that the data rate of the information source determines the transmission bandwidth simultaneously.

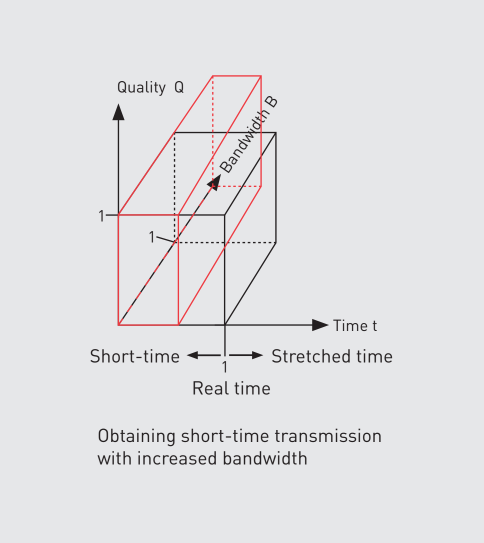
If the radio channel is only available for a short time (only a few msec. at Meteorscatter), the data must be temporarily stored in order to shortly transmit at a higher data rate and bandwidth (offline transmission, see image 2).

Besides this short-time transmission there are also stretched time transmissions where the time is increased but the required bandwith is smaller (also offline transmission). This occurs, for example, with radio links of research satellites to earth that need to operate with small bandwidths because of the long distance and the low transmission power. This ensures that the power spectral density on earth is sufficient to receive the data. The transmission time and the transmission bandwidth are the two degrees of freedom that can be configured to solve transmission problems.

A given data volume I should be transmitted via a radio channel with bandwidth B within time t error-free, i.e. with quality Q = 1.
The time bandwidth relation is the time law of electrical communications engineering according



Transmission theorem, image 1



Transmission theorem, image 2

to K. Küpfmüller. It means that the product of transmission time t and the required bandwidth B is a constant K for every transmission channel.

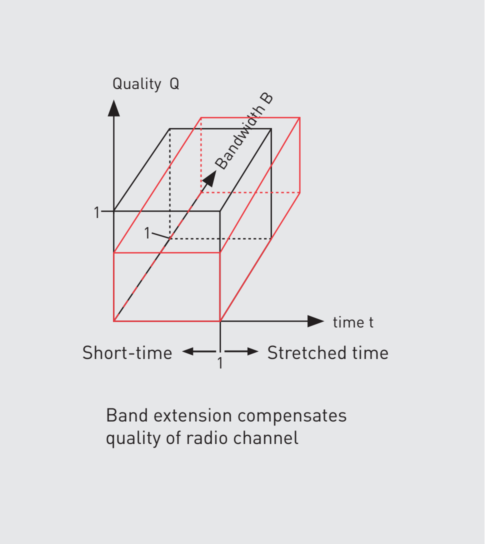
$$K = B \cdot t$$

This leads to the fact that both can be replaced by each other as long as K is constant. In the ideal case the constant is equal to the information volume I, i.e. K = B · t = I. If the specified data amount I, which protects against the natural error rate (in the radio channel), must be enlarged to avoid arbitrary interference (ECM), a second rule arises. Thereafter, bandwidth and signal-to-noise ratio of the specified transmission channel are interchangeable. The product from bandwidth (and/or time) and signal-to-noise ratio (quality factor Q) is therefore a constant, which also depends on the extent of the disturbance. In other words, interference suppression costs time or bandwidth.
This results in a summary rule, whereby the product of bandwidth B, time t and best Q result in a constant Kq, and all three dimensions of this cube are interchangeable with each other.

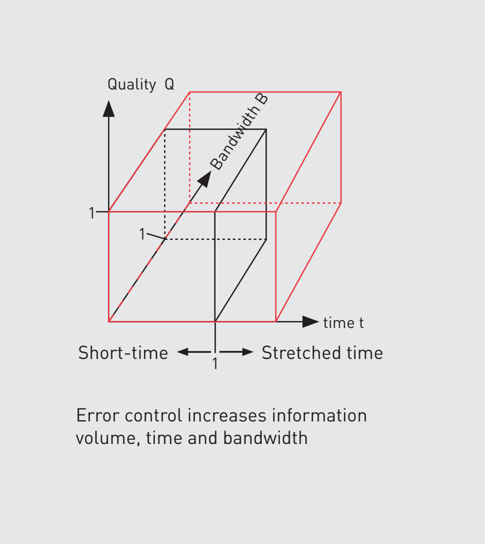
$$Kq = Q \cdot B \cdot t$$

This rule is called the transmission theorem. Graphically, it can be presented as cuboids with the dimensions Q, B and t.

In the best and undisturbed case, the cuboids with quality dimension Q = 1 has the information content I, namely Kq = Q B t = I (image 1). Volume I increases when there is disturbance, e.g. by additional bits. It has the result that more time or bandwidth is needed for the transmission. Protective measure against ECM by bandwidth spreading is discussed below.



Transmission theorem, image 3


Quality Q
Bandwidth B
1
1
time t
Short-time
Stretched time
1
Error control increases information volume, time and bandwidth

Transmission theorem, image 4

Channel gain (link budget)

The preceding sections explained that radiated transmission power, gain and directive efficiency of the transmitting antenna, free field attenuation, gain of the receiving antenna and receiver sensitivity have a major impact on the transmission power. It is helpful to consider the entire chain of action, including preceding properties and circuitry of antenna technology.

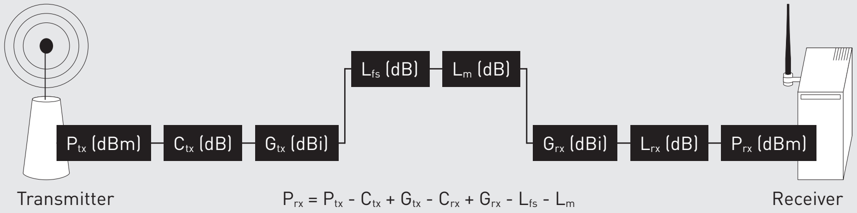
Consideration of all components and influencing variables results in the so-called channel gain (link budget).

In this regard, free field attenuation was already described. However, considering them alone will not go far enough. With regard to the transmitter, ERP (Emitted Radio Power) is used as the definition for transmission power. It consists of transmission power P_{Tx}, cable and/or coupling attenuation C_{Tx} and antenna gain G_{Tx}. Thus, transmission power is only one of several cause variables. Other important parameters that together form C_{Tx} are hardware design and antenna connection. The antenna also has significant influence on transmission characteristics.

The same applies to the receiver. Its hardware has significant influence on the antenna, its connection and of course on the received power.

Commercially available consumer devices, like WLAN or Bluetooth USB sticks demonstrate a spread of almost 20 dB, despite identical wireless chipsets. These differences stem from a more or less good hardware design and the selection of a more or less high quality antenna. Assuming an attenuation of 6 dB leads to halving the transmission range, in the case of a 20 dB attenuation, the range is reduced by 90 %. A range is for example down to 10 instead of the desired 100 metres. In practice, such spreads and limitations can be found.

To sum up, we find that link budget matters are important for wireless systems – and this in-



The channel gain (link budget) includes the entire chain of action.

cludes the entire chain of action as described. In industrial wireless systems, much importance is placed on an optimal adjustment of the hardware design in order to take advantage of every opportunity that is hidden in a technical system.

This is one reason why industrial wireless systems are often significantly superior to consumer market products, although identical receiver building blocks are used.

Bit error rate and sensitivity

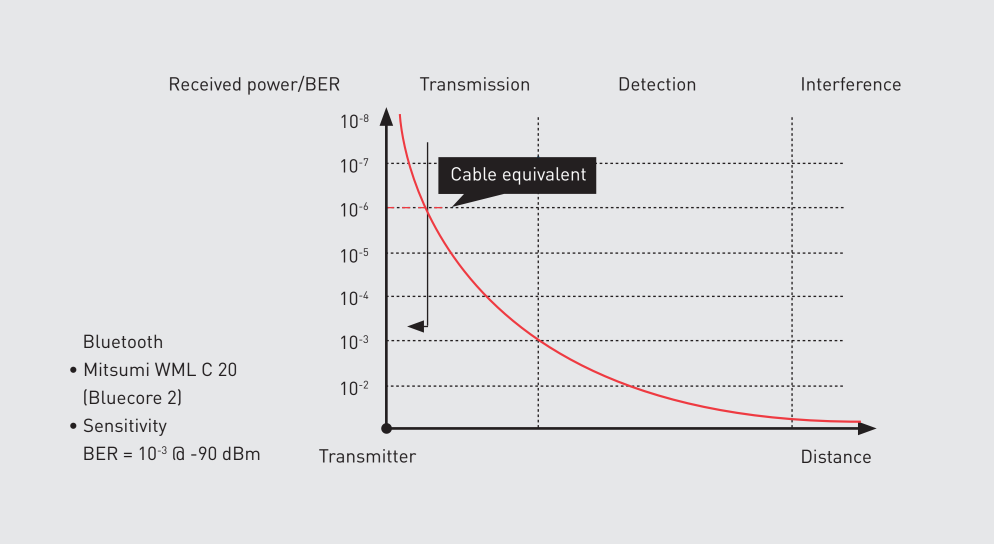
In general, the signal/noise ratio is specified after the Nyquist criterion as a limiting parameter for the transmission of systems. However, more practical is the consideration of the bit error rate (BER - Bit Error Rate) in wireless systems.

It is proven that BER is inversely proportional to the received power under the same coding conditions. In the process, the sensitivities of wireless receivers are also specified in a BER and/or PLR (Packet Loss Rate). The example of a Bluetooth module (Mitsumi WML C20) shown below demonstrates a sensitivity of 90 dBm at a bit error rate of 10^{-3}.

A bit error rate of 10^{-3} means that every thousandth bit does not arrive in the desired form. Threshold BER 10^{-3} also describes the limits of the transmission and detection area. With communication that is no longer measurable, a wireless system only contributes to ground noise, in other words, contributes to general interference.

By comparing these values with a »normal« cable, this BER has a magnitude of approx. 10^{-5}. It is two orders of magnitude better than a wireless system in the threshold area. When transmitter and receiver are close enough together, cable (in comparison) is just as reliable as wireless.

Considering the derivation of the free field attenuation in detail, we will discover (beside the quadratic dependence of the distance) also a

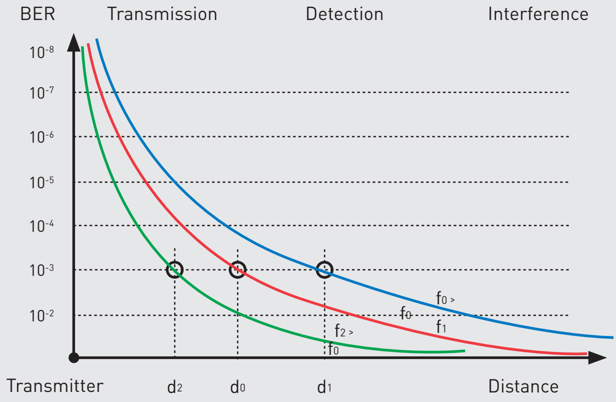


Correlation between received power and distance

quadratic dependence of the frequency. This means that at low frequencies, free field attenuation is much smaller than at high frequencies.

In case of identical transmission power and receiver sensitivity, we can assume that systems using a smaller frequency have a significantly longer range. Or at low frequencies, we can obtain identical ranges with much lower transmission power, which have major advantages for low-power devices.

On the other hand, high frequencies allow the construction of antennae with high antenna gain at reasonable dimensions. This is the requirement for radar systems and directional radio systems with large ranges, including satellite communications.



Correlation between receiving power and distance with different frequencies

// Wireless in industrial automation

07

Introduction

Motivation for the implementation of wireless technologies

Selection parameters

Range and mobility

// Wireless technologies for industrial automation

In addition to general challenges involved in wireless technology, the selection of the ideal type of wireless technology for industrial automation needs to be made. Selection cannot be decided on generalisation, making this far from simple. Numerous differentiated wireless technologies each have their own specific advantages. Ultimately, what is relevant for the solution of individual problems simply shows the various perspectives on requirements.

The implementation of radio or wireless technologies is, without doubt, only significant if costs can be reduced by increased mobility - decreased time in business operation, faster service, reduced installation time or facilitation of use.

Under common discussion in this regard is the simple expansion of existing Ethernet networks via WLAN according to IEEE 802.11 a, b, g standards or future innovative technologies. This is a possible expansion of classical Ethernet technology, though only covering one aspect of wireless privileges relating to MES (Manufacturing Execution Systems) and ERP (Enterprise Resource Planning).

The sensor-actuator level, chiefly responsible for real-time processing, plays an important role in automation technology. Here, the use of wireless technology has to be planned and implemented very carefully and responsibly.

It is immediately evident, without going into technical detail, that the use of wireless technology is complex. In summarising the general demands of this technology, one can identify various differentiated fields of application for wireless devices.

Motivation for the implementation of wireless technology in automation

Faster installation

Installation involves the connection of sensors and actuators, or field operating devices. Portable or mobile equipment must generally be installed via flexible cables and with plug-in connectors. In the case of high speed or locally flexible devices, this is problematic. Moreover, the location of displays or operating terminals frequently changes, making the importance of wireless all the more relevant.

Simplified operation

Technical devices are often provided with service interfaces that allow for comfortable operation. Unfortunately, the electrical and logical specifications are as varied as the types of devices available. By using mobile communication with a standard interface, time spent on installation is significantly reduced. Furthermore, machine diagnosis and monitoring devices can be run independently.

Prompt service

In the event of mechanical failure, immediate intervention is possible as a result of wireless communication with service or emergency personnel. With wireless service interfaces, service personnel are able to analyse a malfunction there and then, and if need be, connect to a service desktop or parts supplier.

Mobile data capture

An important motivation for implementing wireless terminal devices is the possibility of remote or mobile data capturing. In all aspects of logistics, as well as in production and service companies, the use of mobile, wireless data exchange is of great importance. A wireless system is the only solution in achieving comprehensive data processing, saving time, effort and paper, while ensuring accuracy in central databases. With RFID (Radio Frequency Identification), the location of goods in a logistic chain can easily be identified.

Integrated control of manufacturing and production

Today, even flexible production and manufacturing facilities are hard-wired. Adjustment to new production requirements generally involves expensive retooling. The softening of rigid structures using strong modularisation and data transfer via wireless connections simplify communication structures and allows for flexible plant design.

// Selection parameters for the choice of wireless technologies

What is generally desirable in wireless technology is a comprehensive and readily available system. This is not always achievable, with varying requirements on technology and transmission power. If a large coverage with a high rate of transmission is needed, a greater system-related transmission capacity is required. Low frequencies lack the required bandwidth, and while high frequencies may provide more bandwidth, the free field attenuation requires more transmission power. Low power is not possible.

A low-power wireless system with a long life battery is usually operated in a sub-GHz field due to the ideal propagation conditions. Similarly, the communication may only occur sporadically, if the battery is not to be changed frequently. As a result, high volume data exchange is not possible.

The choice of one system over another should therefore be carefully considered. It has become clear that a differential view of some peripheral requirements simplifies the prioritising of requirements:

Frequency band

The question of frequency band already points one in an important direction. Is worldwide accessibility a priority? Are there legal limitations for certain frequencies in the areas of use? Is there a need to use a specific frequency band due to the nature of propagation?

Modulation

The various modulation processes differ in resilience with regard to jamming and interference. Generally, the choice of modulation is dependent on the technological demands and the range of freedom these demands allow for.

Regulatory conditions

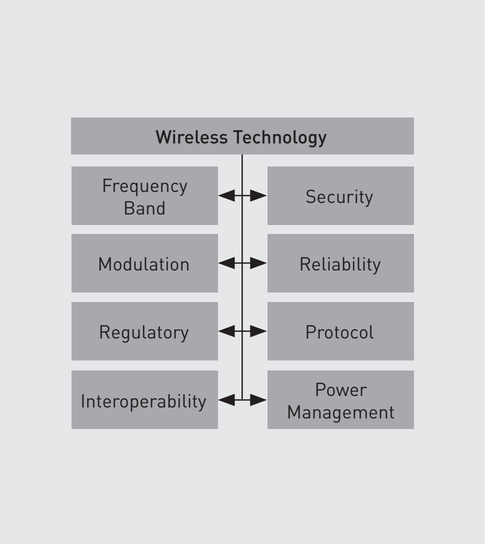
The regulatory conditions are a major deciding factor. As stated earlier, radio is a matter of national regulation. This involves a differential view of which peripheral conditions are appropriate for the field of application in question. Furthermore, it is necessary to explain the measures that need to be taken regarding required approval.

Interoperability

If a coupling is necessary or required for example with consumer devices, one should consider how interoperability can be achieved. Extensive interoperability testing is often necessary to obtain sufficient quality.

Data security

Security is becoming an increasingly important issue in wireless technology. Encoding or authentication of communication is a major challenge. Sometimes a powerful cryptography is already offered at chip level, in other cases, appropriate software is required. One must be aware of the required security options and what they entail.



Criteria for deciding on the ideal wireless system

Reliability

A reliable system requires that data communication is carried out successfully under all circumstances. Frequently time constraints are a prominent factor and communication must be successfully completed in a specified time. The shorter the time-span, the more complex and elaborate it becomes.

Protocols

If a wireless system is to be integrated into an existing system, it is usually necessary to adapt standardised protocols. In this case the peripheral conditions and the communication processes have to be known, not least because many standard protocols such as TCP call for a processor with high computing power.

Power Management

Low power is becoming increasingly important. It is of paramount importance to consider which devices can be energy-saving in a wireless network.

Overview of relevant performance parameters

The choice of wireless system should definitely include considerations of the field of application and the level of performance to be achieved. Five parameters regarding these aspects have turned out to be particularly relevant in industrial communication:

Real-time cycle

It defines the cycle length during which a system can be sampled. The time varys from a few milliseconds for production automation to any number of seconds in the process industry.

Duty cycle, pulse duty factor

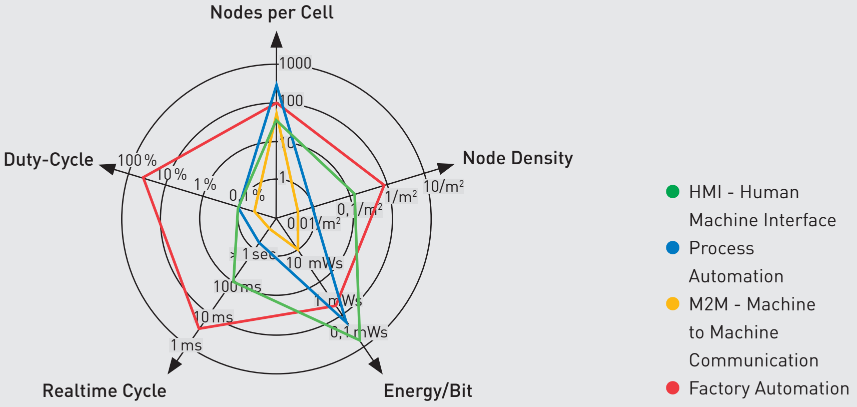
This refers to the operating life of a period for, as an example, a communication. The duty factor is dimensionless and is stated in a range of 0 % - 100 %. A duty cycle of 20 % means that 20 % of the interval period is being used. In a duty cycle of 100 %, the entire interval is being used. Especially for low-bandwidth systems such as the 868 MHz SRD band, there is a restriction of, for example, 0.1 % duty factor.

Nodes per Cell

It describes the number of communication devices within a communication cell. In production automation, some 10 nodes are typical within a network segment. Sometimes they are limited, but so too is the number of participating devices in the system. Bluetooth, for example, has a limit of 7 active devices per piconet.

Node density,

Refers to the number of nodes per area. Only a few nodes would be needed in a sewage plant, but distributed over a larger surface area. Conversely, in a production or robot cell, many communication devices are required in a small area.



Each wireless application field has its specific demands

Energy per Bit
It is a measure of the energy efficiency of a system. Combined with a data rate, the total energy consumption per time unit can be calculated, which allows an estimation of the energy supply. Wireless systems can be scaled down to energy self-sufficient systems. Self-sufficient systems harvest their energy requirement from the environment.

// Range and mobility

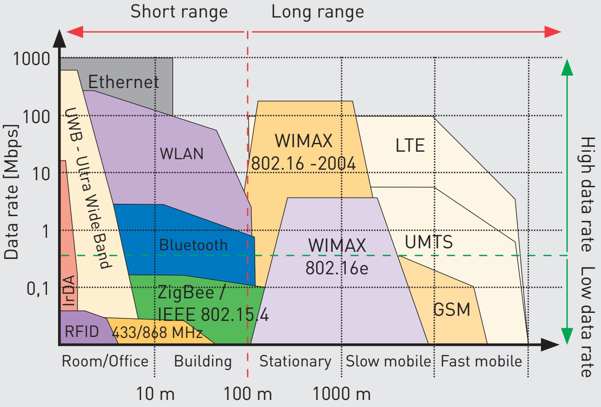
Wireless or radio technology is strongly influenced by the need to be mobile. Still, this freedom of movement is relative. Therefore, it is necessary to clearly define this requirement. How long must the transmission range be to allow uninterrupted communication? How fast will the object receiving communication be moving, and how will this affect the fluency of wireless communication? These are deciding factors in making an informed choice of system.

In practice, pragmatic technology clusters have been defined using two parameters. First, one should differentiate between a high and low data rate. The limit between high and low is somewhat randomly set between 100 kbps and 1 Mbps. The differentiation between long range and short range is also pragmatically set at the 100 metre mark.

Four fields can be differentiated with the following classifications:

1. Short range/low data rate
2. Short range/high data rate
3. Long range/low data rate
4. Long range/high data rate

One of the first applications using EnOcean technology in industry: A wireless multifunction handle switch for machine tools.



Differing systems are employed, depending on application and use.

The mobility factor is also relevant. Especially in the long range solutions, one should consider at least these three categories:

1. Stationary – Fixed stations that are immobile, such as those used in infrastructural networks.
2. Slow/pedestrian mobility – Slow mobility between transmitter and receiver with a typical speed of up to 6 km/h.
3. Fast mobility – Transmission to receiver at high speeds, though this depends on the definition of »high speed«: In Europe this can be up to 300 km/h, while in the USA 80 mph is termed »fast«.

// Basics of wireless technology: Techniques and standards

08

IEEE 802.15.4

ZigBee

Wireless HART

Bluetooth

ISM Band	1	2	3	4	5
Frequency range	26.957 ... 27.283 MHz	40.660 ... 40.700 MHz	433.050 ... 434.790 MHz	868 ... 870 MHz	2.400 ... 2.483 GHz
Bandwidth	326 kHz	40 kHz	1.74 MHz	2 MHz	83 MHz
Advantages	Low energy consumption	No continuous interference carriers, low energy consumption	Relatively broad band, good propagation, moderate costs	Relatively broad band, low occupancy, maximum switching time limit of 10 %.	Large bandwidth, worldwide approval
Disadvantages	Higher interference level due to CB radio	Insufficient bandwidth, too large antenna	Highly occupied, amateur radio, 70 cm	More expensive than 434 MHz	Propagation problems

Frequency fields and characteristics in the ISM band

// Introduction

Fundamental differences to non-industrial wireless applications

Since its discovery, wireless technology has been used to bridge significant distances. For this reason, it was and still is primarily intended for long distance communication. Radio services with their coordinated frequencies are therefore legally protected against interference. In Germany, this is carried out by the radio monitoring and measuring services of the Federal Network Agency.

By contrast, industrial applications using Short Range Devices (SRD) have increased considerably over the past years. With transmission over short distances and use of publicly available frequencies, they are not protected by law. Users of these frequencies have to negotiate and coordinate use in a decentralised manner. These conditions result in a series of policies and specifications designed to ensure contractual co-existence (compatibility) or harmonious cooperation (interoperability) of local networks using SRD. These policies are the focus of the chapter on co-existence starting on page 139.

From the previous reports, it is clear that a single technology fulfilling all requirements is not achievable. The following chapter therefore provides an overview of relevant technologies, demonstrating their primary functions as well as limitations.

// Sub-GHz band

If radio communication is needed within a contained structure, a frequency below 1 GHz is recommended. With lower attenuation as at 2.4 GHz, for example, a better link budget of between 15 and 25 dB can usually be expected. With a similar transmission power, considerably higher ranges are possible. Moreover, a very low transmission power with this frequency range enables the use of devices running on maintenance-free batteries or using harvested energies.

Region	Standard	Transmit Power
USA, Canada	FCC 15.231 (e) RSS-210 FCC 15.240	Field strength: 4400 uV/m @ 3m 10 sec blink rate Field Strength: 55,000 uV/m @ 3m (deployment restrictions apply)
Europe, Africa	ISM Band EN 300 220	Max ERP: <10 mW @ 10 % or <1 mW @ 100 % duty cycle
China	SRRC Regulation	Max ERP: 10 mW, occupied bandwidth < 400 kHz
Australia, New Zealand	AS/NZS 4268:2003	Max ERP: 15 mW

Regulatory requirements in the 433 MHz band

Wireless applications in sub-GHz band are popular and extremely versatile. A classification according to a specific area of application is not possible. The devices serve as wireless modems, alarm systems, weather stations or are used for general telemetric tasks as well as for smart-metering.

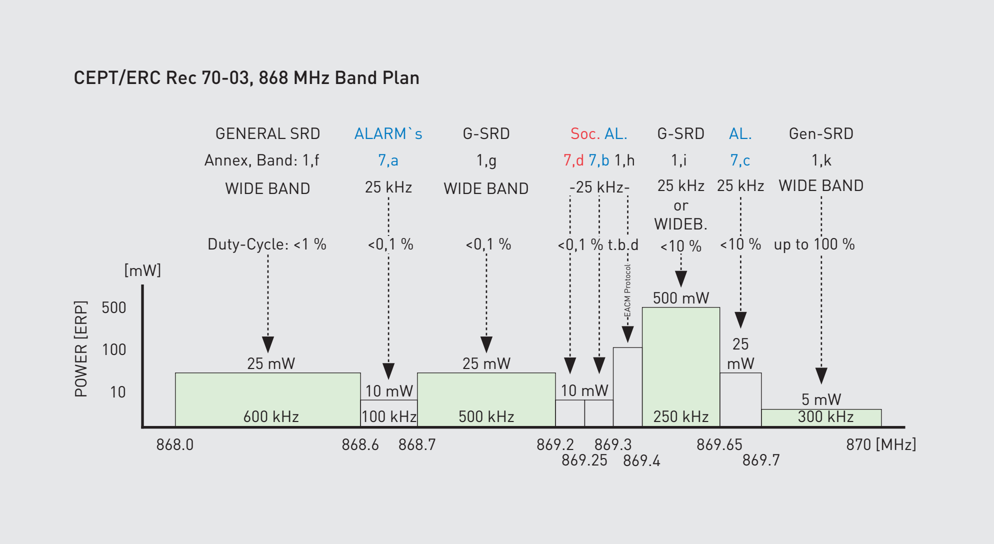
If only sporadic data communication is needed, these devices can be very energy efficient – one can extend the battery life almost indefinitely by the clever management of transmission and reception. This is a significant advantage that enables their use in all in the described fields of application.

433 MHz

The 433 MHz ISM band uses the frequency range of between 433.05 MHz and 434.79 MHz. It is predominantly implemented in client-specific, individual system solutions with short to medium range coverage. In the case of maximum transmission power of up to 10 mW, the duty cycle of up to 10 % is permissible. In the case of transmission power <1 mW, the duty cycle is not restricted at all.

This frequency band is especially attractive because the available bandwidth is unrestricted on a spectral and time level. Also, the type of modulation is not prescribed, allowing for a variety of procedures to be implemented. This advantage, however, incorporates the disadvantage that special measures regarding interference or the prevention of interference need to be undertaken. To achieve reliable communication, the potential for interference should be analysed and the channel coding adjusted accordingly. This challenge having been resolved, cost-effective, energy-saving wireless solutions can be successfully implemented.

The 433 MHz band is also attractive for future innovation and sustainability. To achieve consistent global harmonisation, the IEEE is motivated to allow this frequency as an alter-



Regulatory requirements in the 868 MHz band

native MAC-layer for wireless personal area networks according to IEEE 802.15.4f. The key advantages are the reduction of path loss, as well as the simplified implementation of wireless chips.

In Europe the 433 MHz band is used in technical applications predominantly for cordless microphones and headphones, baby intercoms, access control systems, door openers, motion sensors and medical systems.

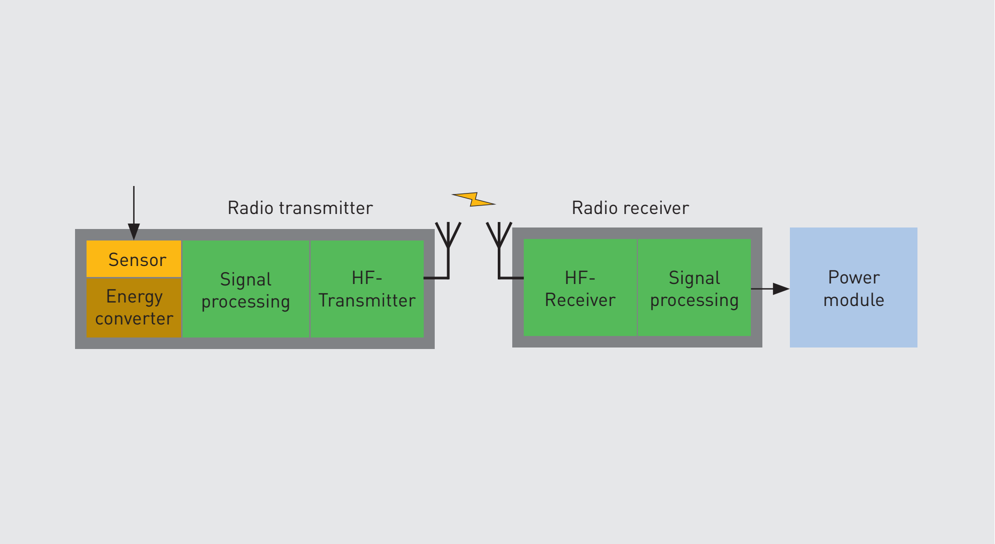
868 MHz

One of the most attractive frequency bands in Europe is the 868 MHz-SRD band. Short Range Devices (SRD) are part of our daily lives. An almost unimaginable number of applications, from a remote key for one‘s car through chest belt monitors in sports to headphones and wireless alarm sensors, are evidence of the variety of applications. Unlike the 433 MHz-band, the 868 MHz-band was considered for various applications at an early stage and the available frequency bands were divided into sub-bands applicable for these various uses. The applications, duty cycles and maximum transmission capacities are specifically regulated in these sub-bands in order to ensure the compatible coexistence of devices.

Transmission capacities of between 5 and 500 mW with a typical duty cycle of 0.1 % to up to 100% are allowed. Combined with ideal conditions for propagation, coverage in the kilometre range is possible. A good reception within buildings is also guaranteed.

Non-standard global application

A disadvantage is the non-standard global application of frequency bands. The 868 MHz-band is allowed in countries classified as group 1. In the USA, these frequency ranges are used in the application of CDMA mobile technology. The frequency range 902 ... 928 MHz is available for SRD applications in the USA. In Europe, this frequency range is used for GSM mobile com-



The EnOcean technology facilitates battery-free sensor nodes.

munication. Fortunately, these frequency bands are situated so closely that dual-mode devices can be developed. Almost every 868 MHz wireless chip manufacturer allows for the 915 MHz-band as well.

So too, in the 868 MHz-band, the modulation and transmission can be freely selected. A multitude of cost-effective system solutions are available on the market. The securing of data from intrusion is not regulated. This means that encoding or application security has to be incorporated in the implementation of the respective protocols. The data rate of modules in typical industrial applications is 38400 baud. There are, however, variations with up to 115200 baud capacity possible.

As only the physical layer and the data link layer are defined, the 868 MHz-ISM-module does not provide a standardised application interface or have standardised protocols. Ad hoc networks too cannot be adapted without a great amount of effort. Usually a discrete micro-controller is needed, which is responsible for the system's protocols, encoding and other sundry security mechanisms. The 868 MHz-SRD system is therefore a typical system for low-cost and low-end sensor/actuator networks.

»Energy harvesting«: The EnOcean standard

EnOcean, with its patented technology, has made its mark in sub-GHz wireless technology. EnOcean is a spin-off company of Siemens which was made independent in 2001 with the idea of establishing authentic zero-power wireless technology in the market.

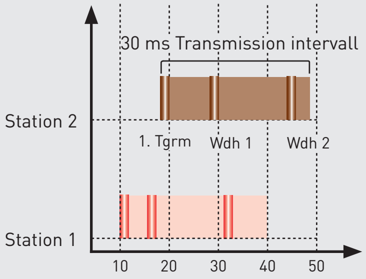
This idea rapidly found a market. In building services engineering in particular, the costs involved in battery exchange of a large number of switches, even with a battery life of several years, is simply too high. EnOcean follows the key concept of generating and harvesting energy, either from the switching process or the environment, using an extremely efficient

Description of serial data structure	
Bit 7	Bit 0
SYNC_BYTE1 (A5 Hex)	
SYNC_BYTE0 (5A Hex)	
H_SEQ	LENGTH
ORG	
DATA_BYTE3	
DATA_BYTE2	
DATA_BYTE1	
DATA_BYTE0	
ID_BYTE3	
ID_BYTE2	
ID_BYTE1	
ID_BYTE0	
STATUS	
CHECKSUM	

With a data rate of 120 kbps, a telegram takes less than a millisecond to transfer.

wireless system operating on harvested energy. This concept has won a number of technology awards since its first implementation. EnOcean is now an established authority in ultra-low power wireless technology.

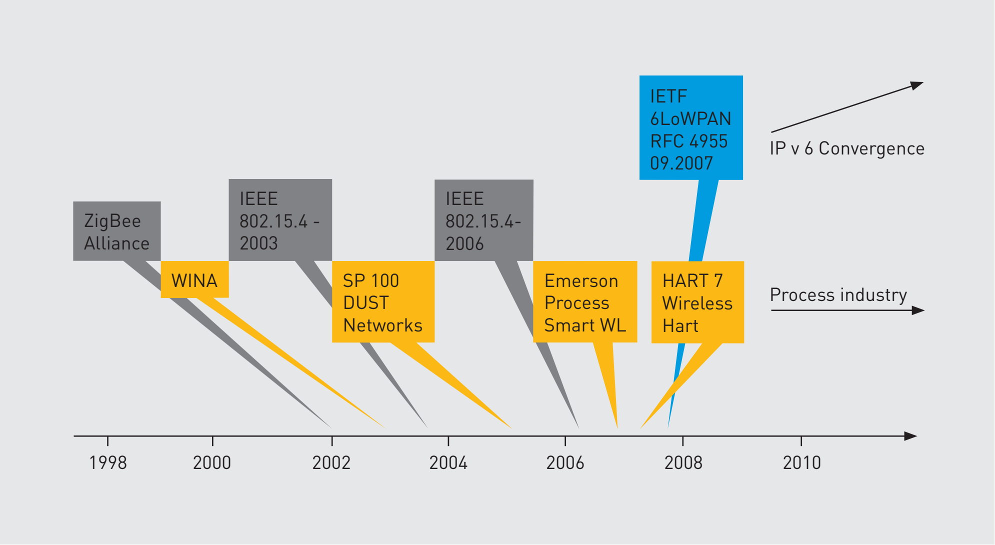
EnOcean uses the 868 MHz-SRD band for its transmission or the 315 MHz-ISM band for applications in countries outside group 1. The high receiver sensitivity of better than -95 dBm, a transmission power of up to 10 mW (10 dBm) and an energy efficient ASK (Amplitude Shift Keying) modulation results in a complete wireless command using a mere 50 mW of energy. A typical coverage of 30 metres within buildings and 300 metres free field can be achieved. With a 32 bit-ID, which the wireless receiver can learn from its transmitter, a unique device assignment is possible. Typically, 4 bytes of user data is transmitted. In serial mode, a permanent data link with up to 9.6 kbps is possible.



Transmission scheme of unidirectional links

One needs to distinguish between two operating models.

1. Unidirectional communication is used for basic sensor operation. Generally, a piezo element or a miniature energy generator gains as much energy from the switching process as necessary to implement the entire transmission process. Accordingly, the control command is transmitted three times 30 milliseconds. Due to the high data rate and small data amount, the transmission path is only briefly occupied. This means the probability of collision is optimally reduced. Data feedback to the sensor is not possible. To enable maintenance-free autonomous sensors, energy converters that permanently draw on and store harvested energy in a



The IEEE 802.15.4 standard is the basic technology for various wireless networks.

buffer are employed. Small solar cells, vibration or thermal converters are ideal for this purpose.

2. In bidirectional operation, the transmitter and receiver data exchange is permanent. With the Dolphin Platform, EnOcean has made bidirectional communication available since 2010.

EnOcean wireless technology is currently the leader in zero power technology. More than sixty companies offer hundreds of sensors that are largely compatible with each other. An increase in bidirectional communication has resulted in new models in which feedback between sensor and actuator is possible.

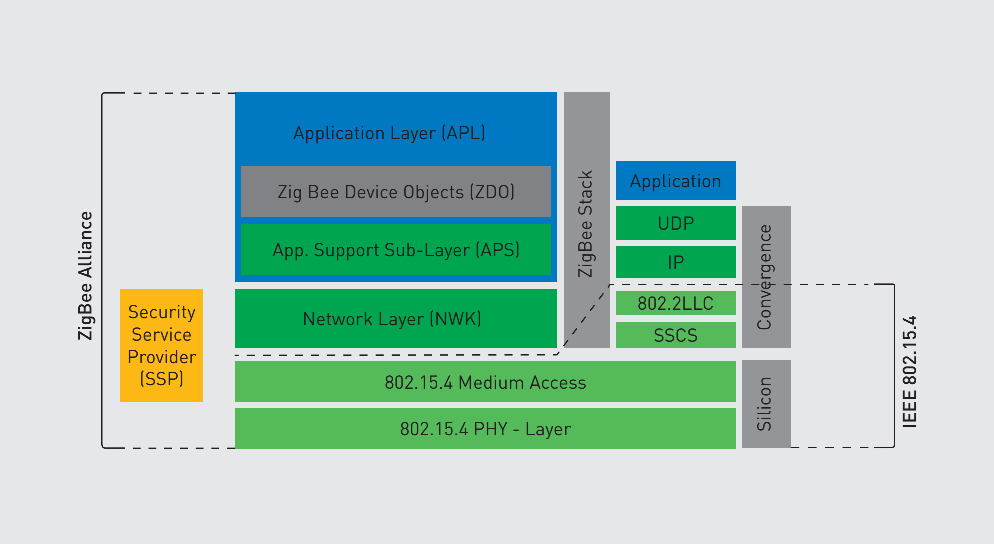
// IEEE 802.15.4

The terms IEEE 802.15.4 and ZigBee are often used synonymously incorrectly. The IEEE standard only defines the lower layers (PHY and MAC) in ISO-OSI models for WPANs (Wireless Personal Area Networks). The higher protocol levels are regulated by other organisations, for example the ZigBee Alliance or the HART Foundation.

The standardisation of the physics and the Mac-layer was defined in the WPAN working group IEEE 802.15 as sub-group 4. As of February 2003, the first specification IEEE 802.15.4-2003 was released. In the 2006 version, supplementary data rates were added in sub-GHz frequency.

Further enhancements in the working **group a:** (additional physical layers namely UWB Ultra Wide Band and CSS Chirp Spread Spectrum), **c:** (314-316 MHz, 430-434 MHz and 779-787 MHz for China), **group d:** (950-956 MHz for Japan), **e:** (additional functions for industrial applications), and **f:** (active RFIDs), are proof of dynamic standardising activity. These developments are pioneering, especially with regard to sensor networks (WSN – Wireless Sensor Networks).

A development of particular interest is seen in the very universally applicable IEEE 802.15.4 base



ZigBee and IEEE 802.15.4 are not identical.

layer. A variety of different technologies use this base layer, expanding the protocol stack only at the higher layers. Two trends are currently apparent:

1. In the field of process automation, the implementation of an IEEE 802.15.4 layer in the wireless HART in HART 7 standard has been established on a large scale.
2. With the trend towards IP-based networks, the IETF 6LoWPAN standard has created a good basis for IP-based sensor networks.

Besides these two trends, further solutions will certainly be created in the future on this basis with a great variety of application options.

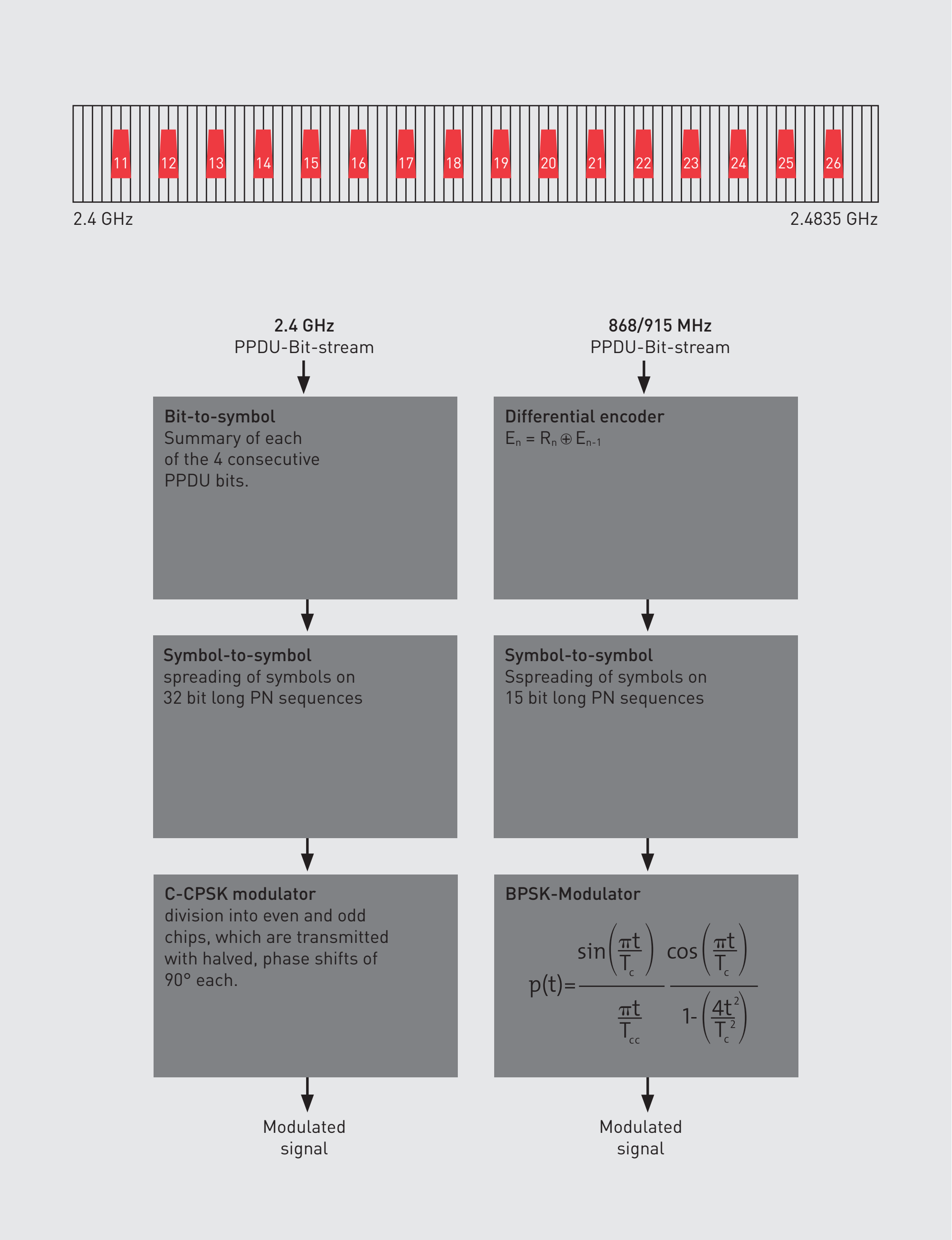
Basics

The IEEE 802.15.4-2003 defines a wireless transmission for the 2.4 GHz-ISM band, the European 868 MHz-SRD band and the American 915 MHz band.
In the 2.4 GHz band, there are 16 channels available worldwide with a maximum bandwidth of 250 kbps. Due to limitations of the SRD band, in Europe only channel 868.3 MHz with 20 kbps is allowed. On the North American continent, however, 10 channels with a data rate of 40 kbps are allowed.

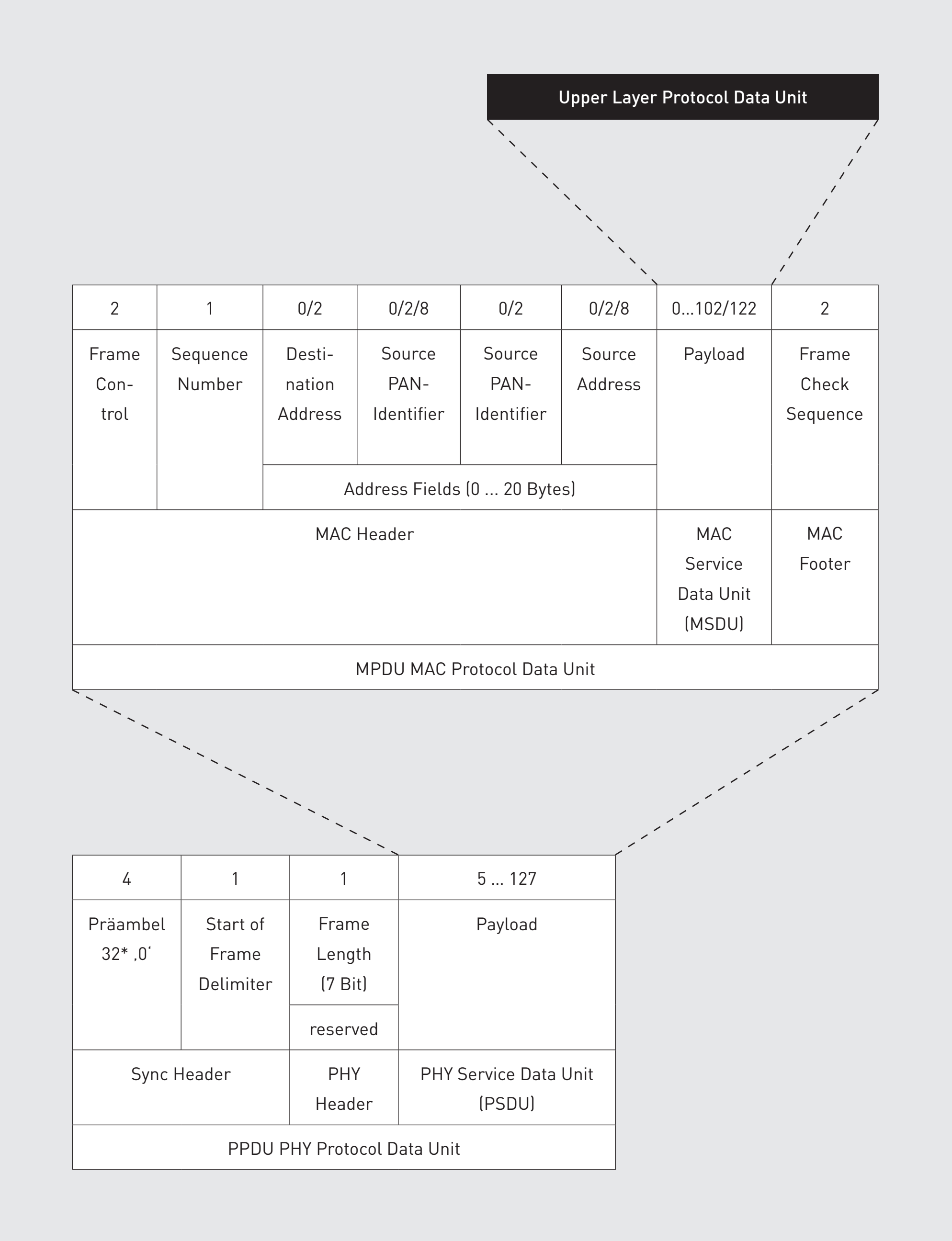
Currently three trends have surfaced:

1. The sub-GHz band is being adopted in standard use according to the opportunities of the respective countries to make full use of the propagation features (802.15.4 c, d).
2. The 2.4 GHz range is particularly attractive in technical applications and is also provided with application interfaces.
3. Alternative physical layers are available.

Within the frequency bands, different modulation processes are used, though always using a DSSS (Direct Sequence Spread Spectrum) process with 15, or 32 bit range.
To access the channel, 802.15.4 uses a CSMA-CA algorithm comparable to that of WLAN according


11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
2.4 GHz
2.4835 GHz
2.4 GHz
PPDU-Bit-stream
868/915 MHz
PPDU-Bit-stream
Bit-to-symbol
Summary of each of the 4 consecutive PPDU bits.
Differential encoder
$E_n = R_n \oplus E_{n-1}$
Symbol-to-symbol
spreading of symbols on 32 bit long PN sequences
Symbol-to-symbol
Sspreading of symbols on 15 bit long PN sequences
C-CPSK modulator
division into even and odd chips, which are transmitted with halved, phase shifts of 90° each.
BPSK-Modulator
$p(t)=\frac{\sin\left(\frac{\pi t}{T_c}\right)}{\frac{\pi t}{T_{cc}}}\,\frac{\cos\left(\frac{\pi t}{T_c}\right)}{1-\left(\frac{4t^2}{T_c^2}\right)}$
Modulated signal
Modulated signal

Top: Use of the 2.4 GHz band in IEEE 802.15
Bottom: Modulation process for IEEE 802.15.4

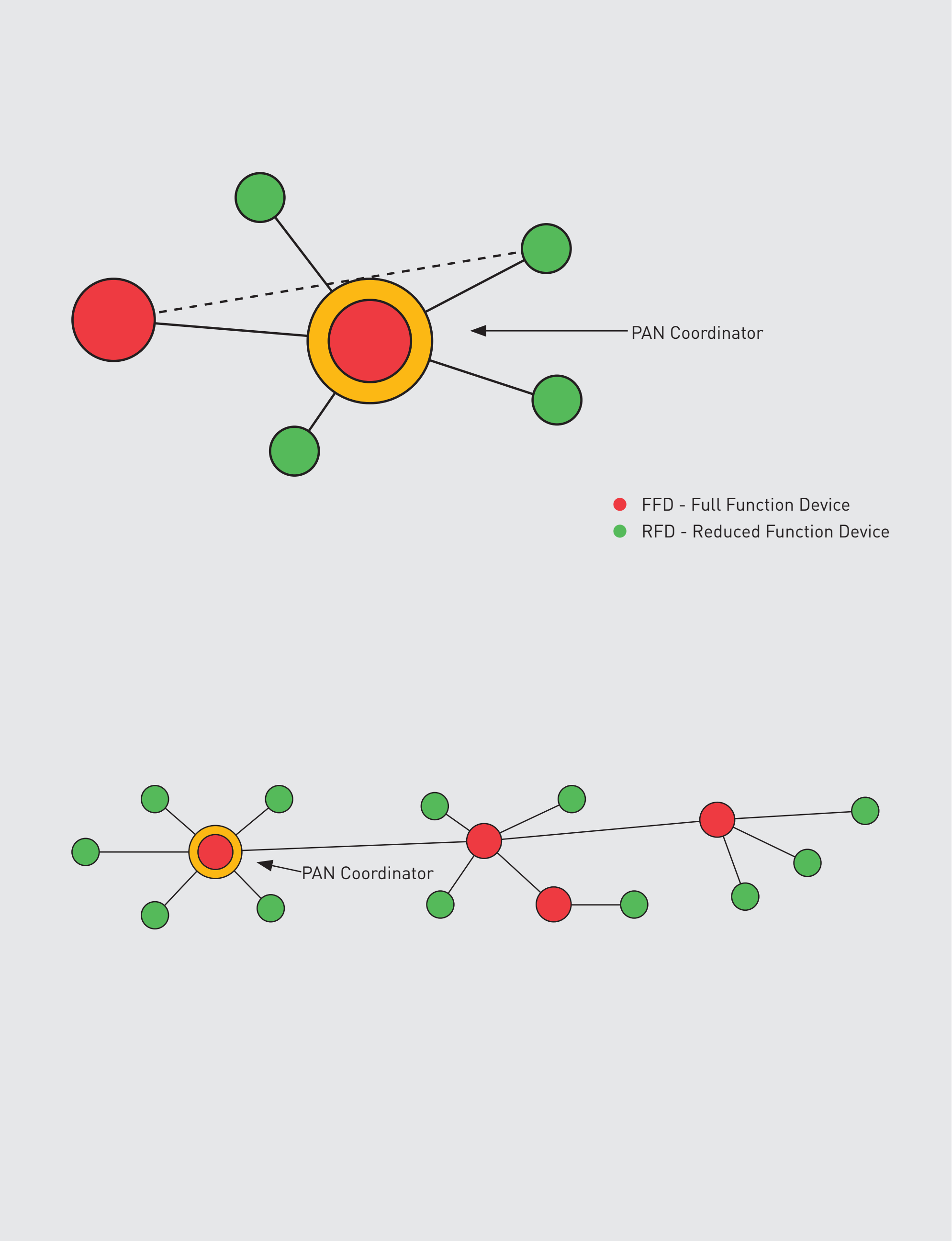


Upper Layer Protocol Data Unit

2	1	0/2	0/2/8	0/2	0/2/8	0...102/122	2
Frame Control	Sequence Number	Destination Address	Source PAN-Identifier	Source PAN-Identifier	Source Address	Payload	Frame Check Sequence
		Address Fields (0 ... 20 Bytes)					
MAC Header						MAC Service Data Unit (MSDU)	MAC Footer
MPDU MAC Protocol Data Unit							

4	1	1	5 ... 127
Präambel 32* ‚0‘	Start of Frame Delimiter	Frame Length (7 Bit)	Payload
		reserved	
Sync Header		PHY Header	PHY Service Data Unit (PSDU)
PPDU PHY Protocol Data Unit			

IEEE 802.15.4 Protocol parameters



Top: RFDs and FFDs can span a complex network
Bottom: FFDs can be operated in even more complex networks.

to IEEE 802.11. For optimal channel access, the so-called super frames can also be used, for which time slots can be reserved for time-critical applications. To make these time slots known, a network coordinator (PAN coordinator, see illustration on page 105) transmits signals at selected intervals, by which registered stations can be recorded. A channel switching mode similar to Bluetooth is not implemented, but can be done if needed by the network layer.

A special feature of the IEEE 802.15.4 standard is the network topology itself (see illustration, page 105). A complex mesh star typology is supported that enables either low power or high availability.

A differentiation is made between Full Function Devices (FFD) and Reduced Function Devices (RFD) to allow for cost-effective sensors and actuators. RFDs can only be enabled as slaves in FFDs and communication between the two is not possible. The result is that complexity is significantly reduced, enabling RFDs to run on a mere 8 bit processor with very little memory. FFDs are central network nodes that can operate on their own and need to prevail over a much wider range of functions. FFDs are able to connect with each other or with RFDs. In a network segment, the FFD adopts the role of the PAN coordinator.

These nodes are responsible for the connection management in a PAN cell. All nodes that belong to the cell, in both FFDs and RFDs, must register on the cell. A PAN coordinator is also able to manage communication between two RFDs. With the ability to link FFDs to each other, one can establish almost any complex network structures, with a routing of messages across the segments made possible. These are also referred to as mesh networks. To address each node, the MAC frame provides both destination and source addresses for the subnet (PAN-2-byte identifier) as well as for the node itself (address 2/8 bytes). Theoretically, a network structure can include up to 216 times 264 nodes.

// ZigBee

ZigBee contributes a service layer to the IEEE 802.15.4 standard. It focuses on the configuration and administration of communication nodes and the routing of messages between devices. The implementation of an application layer is not directly defined in the ZigBee standard, though there are definitions of application scenarios and suggestions for use in its profiles. The focus is therefore on a specific application field.

Definitions for building automation, housing technology and industrial control as well as for remote control of consumer devices are available.

The standardised device specifications mean that different providers are able to develop compatible devices. A complete ZigBee protocol stack also includes an added security layer, an adaptation layer as the intermediate application layer and device objects (ZDP ZigBee Device Objects) and their application definitions in the various profiles.

// HART

HART (Highway Addressable Remote Transducer) is a communication standard that was developed in the 1980's under the direction of the company, Rosemount. In 1989, what was initially a proprietary technology became a standard and the HCF (HART Communication Foundation) was launched. HART is an extension of the classical 4 ... 20 mA analogue technology, which is widely used in processing industries. Since its debut, more than 24 million HART-enabled field devices have been installed. HART extends the analogue interface to a digital instrumentation bus. For transmission of digital data, the analogue current signal is superimposed on a floating-free digital

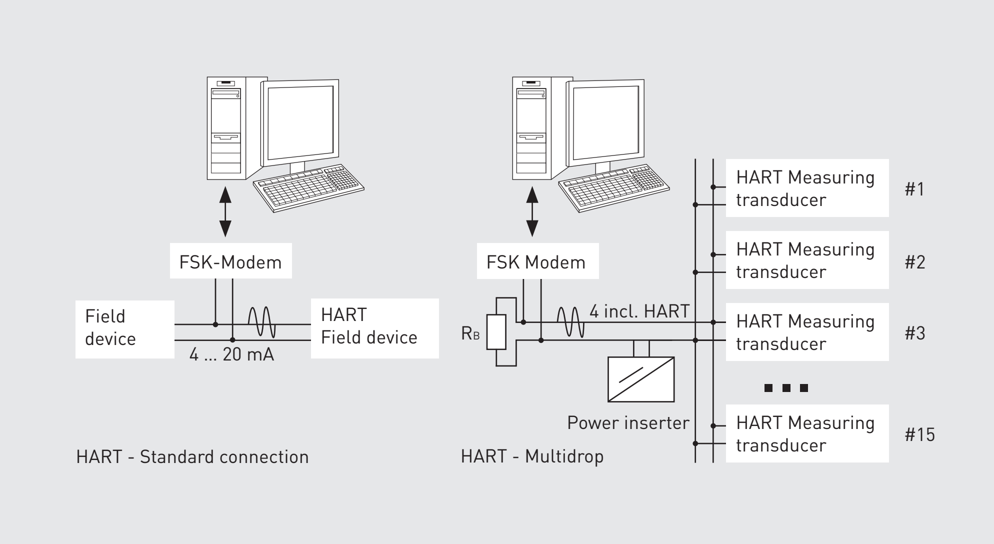
Mandatory (Direct & Indirect Addressing Clusters)

Cluster Name	Clusters[1]	SB[2]	Attrib Name	Attribid		TransDataTyp		Units	Note
				b15 b6	b7 b0				
Input: OnOff DRC	0xXX	X	OnOff	0000 0000	0000 0000	Unsigned Integer 8-Bit	0x1	Data	
			Data Definitions				Value	Units	
			On				0xFF	Data	
			Off				0x00	Data	
			Toggle Output (Used for 3 Way Switches)				0xF0	Data	
Input: DimBright-DRC	0xXX	X	DimBright	0000 0000	RRRR RRRR	Unsigned Integer 8-Bit	0x1	Data	5
			Data Definitions				Value	Units	
			Dim				0x00	Data	
			Bright				0xFF	Data	
Input: Preset DRC	0xXX	X	Store Preset	0000 0000	0000 0000	No Data	0x0	N/A	6
			Preset	0000 0001	RRRR RRRR	Unsigned Integer 8-Bit	0x1	%	5,6

Optional (Direct Addressing Clusters)

Cluster Name	Clusterid[1]	SB[2]	Attrib Name	Attribid		Trans Data Typ		Units	Note
				b15 b6	b7 b0				
Input: Adj DRC	0xXX		Current Level	0000 0000	RRRR RRRR	Unsigned integer 8-Bit	0x1	%	5
			Previous Level	0000 0001	RRRR RRRR	No Data	0x0	N/A	5
			Stop	0000 0101	0000 0000	No Data	0x0	N/A	
			Min Dim Level	0000 0110	0000 0000	Unsigned Integer 8-Bit	0x1	%	
			Max Bright Level	0000 0111	0000 0000	Unsigned Integer 8-Bit	0x1	%	
Input: Light Level LSM	0xXX		Current Level	0000 0000	0000 0000	Semi-Precision	0x08	Lux	
Input: Occupancy OS	0xXX		Current State	0000 0000	0000 0000	Unsigned Integer 8-Bit	0x01	N/A	7
			Data Definitions				Value	Units	
			Occupied				0xFF	N/A	
			Unoccupied				0x00	N/A	
			Wait: occupancy sensor is stabilizing (power up cycle)				0xF0	N/A	

ZigBee‘s definition of a»virtual« dimmer for building automation


FSK-Modem
Field device
4 ... 20 mA
HART Field device
HART - Standard connection
FSK Modem
RB
4 incl. HART
Power inserter
HART Measuring transducer
#1
HART Measuring transducer
#2
HART Measuring transducer
#3
HART Measuring transducer
#15
HART - Multidrop

Wireless HART can be used in existing peer-to-peer connections or even to establish complex networks

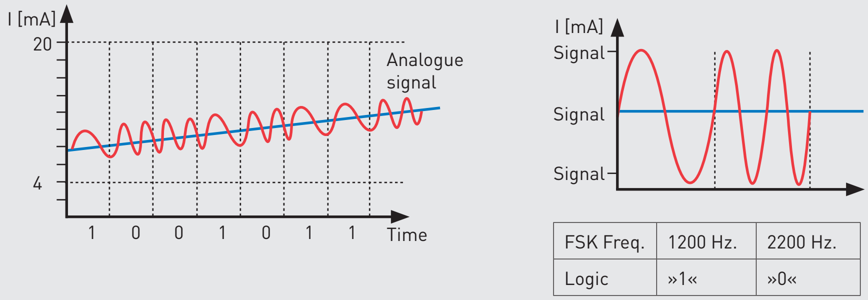
signal. Connected with a master-slave communication, this allows for bidirectional data exchange. HART allows up to two masters. Generally, the engineering console at the routing station along with a secondary device on location, for example a hand-held terminal or laptop, are used as masters.

HART was developed for configuration of field devices. This means that a superordinate station (master) has access to a passive field device. For data preparation, the digital signal is activated on a two-wire conductor using a HART-FSK modem and the control commands are superimposed on the current signal.
In addition to this point-to-point mode, a multidrop operation is possible, with up to 15 field devices able to operate as in field bus applications. To enable this, however, it is necessary to disable analogue data transmission of the measuring transducer. The multi-drop mode is in any case only used for transducers and not for position signals due to the slowness of the bus used. Position signals are usually transmitted as 4 ... 20 mA standardised signals.

HART is represented by the HCF (HART Communication Foundation), a non-profit organisation, which aims to promote the implementation of HART and its specifications independent of the manufacturer.

Cable-linked HART communication

The data transmission of the HART components is based on the Bell 202 specification. The analogue signal is superimposed on a floating-free digital signal in FSK code (Frequency Shift Key). A logic »1« is modelled as a 1200 Hz signal, a logic »0« as a 2200 Hz signal, with an amplitude of +/- 0.5 mA. The specification stipulates that the master sends a voltage signal, while the slaves deliver the communication using load-independent currents. The total resistance load of the current loop should be between 230 and 1100 ohms to ensure proper operation. According to specification, unshielded 0.2 mm two-wire cable is ideal



FSK Freq.	1200 Hz.	2200 Hz.
Logic	»1«	»0«

HART superimposes an FSK-coded digital telegram on the analogue signal.

for short distances. Distances of up to 3000 m are possible with twisted and pair-shielded cables.

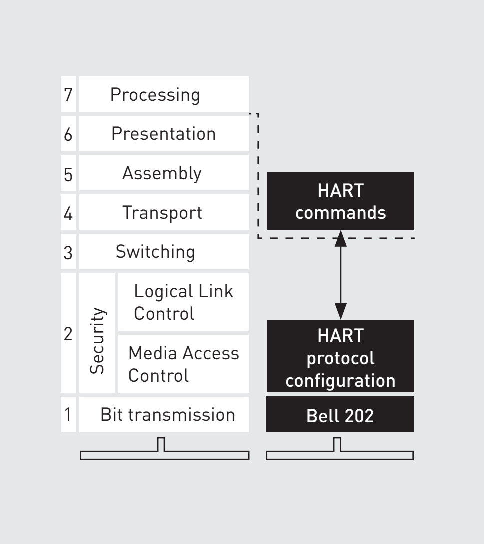
Different manufacturer-specific enhancements enable special features; for example, the FSK bus (Hartmann & Braun) uses HART as a device bus with up to 100 devices. By implementing buffer amplifiers and because of the resulting decoupling of the analogue signals, it is possible for the controller and actuator to continue working together in analogue mode without being influenced by other devices. If the buffer amplifiers are configured ideally, implementation is also possible in areas with a risk of explosions.

Data link layer

HART uses a pure master-slave protocol. All activities start from the master. A primary and a secondary master are permitted. The primary master is usually the guiding system; the secondary master could be a service device on location. All HART field devices are passive slaves.

In standard master-slave communication, a master broadcast and a slave burst mode are supported communication structures. In a standard communication, a master telegram immediately follows the answer received from a slave with the relevant data. The burst mode enables a slave to cyclically deliver communication telegrams. The number of possible telegrams is thereby doubled to four telegrams per second. A HART telegram transfers each byte as a UART character of 11 bit length and at a speed of 1200 bps.

In a typical telegram with a 10 byte protocol and 25 byte user data, a master-slave transaction including all synchronisation mechanisms needs an average of 500 ms. It is clear that HART‘s time performance is no match for state-of-the-art field buses and is not suitable for control tasks.



HART can be classified within the ISO/OSI scheme. .

Application protocol

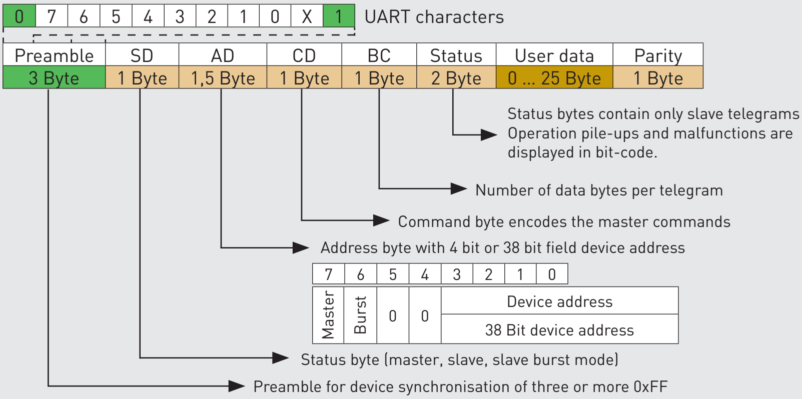
HART's application protocol provides a relatively simple command interface. Pre-defined commands allow the master specifications or communication to be delivered to a field device. For a universal and cross-device communication, the HART commands are divided into different groups for field devices (slaves) and display and control devices. Each type of device with its specification has to conform to respective classes of commands or conformity classes. Universal commands are understood by all HART field devices. Standard commands are unique to certain device groups and the device-specific commands enable access to specialised device features.

To allow for uncomplicated configuration of the device-specific properties and to ensure an essential basis for the interoperability of devices, HART devices are described in DDL (Device Description Language). The description of the device is often stored in the device in binary form so that the configuration tools have access to the database during a configuration process, thereby allowing them the same capabilities as field devices.

// Wireless HART

In 2005, the company DUST Networks, established a low-power mesh network based on the 802.15.4 standard. With the development of TSMP (Time Synchronised Mesh Protocol), a stable base was developed for a secure and especially reliable low-power network with a long-life battery. In 2006, the Emerson Process Management decided to integrate the protocols developed by DUST Networks initially in a 912 MHz band into their products, making the first step towards wireless sensor networks in the processing industry. The standardisation efforts that followed led to the integration of wireless technology - now based on the IEEE specification 802.15.4-2006 in 2.4 GHz band - into the Wireless HART standard in September 2007. Wireless HART is part of the HART 7 specification, according to IEC 62591:2010.

The special feature of Wireless HART is the TSM protocol. Based on the IEEE 802.15.4 standard for the PHY and MAC layers, an upgrade of the access layer to a chronologically synchronised network followed. The entire signal transmission is scheduled in a 10 ms grid. Each station is assigned a specific time slot in which communication with the relevant partners can occur. During the period when no communication is to occur, the respective stations are in sleep mode. This leads to an extended battery life. In general, update cycles for the entire network takes from a few seconds to a few minutes.

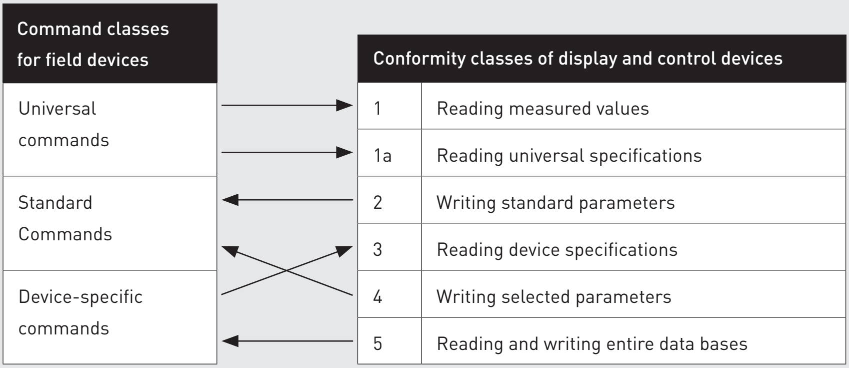


0	7	6	5	4	3	2	1	0	X	1

Preamble	SD	AD	CD	BC	Status	User data	Parity
3 Byte	1 Byte	1,5 Byte	1 Byte	1 Byte	2 Byte	0 ... 25 Byte	1 Byte

7	6	5	4	3	2	1	0
Master	Burst	0	0	Device address 38 Bit device address			

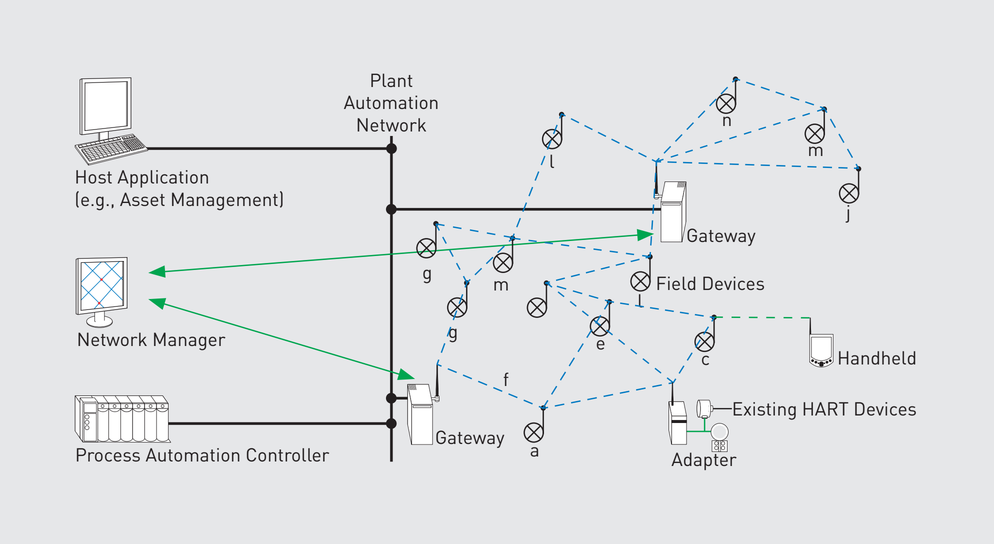
UART characters
Status bytes contain only slave telegrams
Operation pile-ups and malfunctions are
displayed in bit-code.
Number of data bytes per telegram
Command byte encodes the master commands
Address byte with 4 bit or 38 bit field device address
Status byte (master, slave, slave burst mode)
Preamble for device synchronisation of three or more 0xFF



Command classes for field devices
Universal commands
Standard Commands
Device-specific commands

Conformity classes of display and control devices	
1	Reading measured values
1a	Reading universal specifications
2	Writing standard parameters
3	Reading device specifications
4	Writing selected parameters
5	Reading and writing entire data bases

Top: Structure of a HART telegram
Bottom: Command and conformity classes enable interoperability with a HART device.



Structure of wireless HART systems as a mesh network.

To improve interference behaviour, each data exchange is allocated a different frequency in a pseudo-random pattern so that top reliability is ensured. A total of five mechanisms are responsible for reliable operation:

1. Planned, chronologically synchronised communication
2. Channel switching after each transmission-related event.
3. Automatic linking of communication nodes and construction of the network structure
4. Redundant mesh data transmission with channel monitoring
5. Data encryption for high-level security.

To meet the diverse requirements of devices, there are three different types of Wireless HART terminals:

Adapter

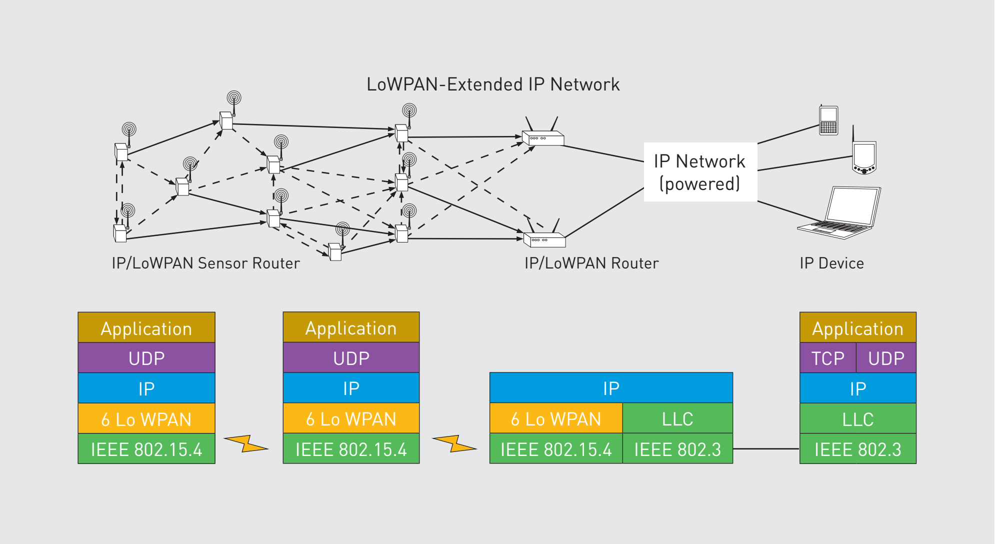
A Wireless HART adapter uses the conventional HART interface according to the Bell 202 standard and serves as a data hub for the wireless network. Several devices can be connected via power supply lines.

Terminal

A Wireless HART terminal is a device usually used for measuring and monitoring. In general, these are intelligent sensors integrated as independent measuring stations in a system environment. Devices are used where, for cost reasons, no conventional wiring can be used. Especially in the processing industry, extremely high installation costs are common, particularly in outdoor areas.

Gateway

At the heart of a Wireless HART infrastructure is the gateway. According to the specification, it consists of a wireless unit that serves as an access point for wireless devices, the network manager (which is responsible for controlling transmissions) establishing the network and managing the safety certificates of terminals and ultimately, the gateway which acts as a system interface between the host system (typically an



6LoWPAN enables the design of IPv6 (sensor) networks, based on IEEE 802.15.4 (see page 114).

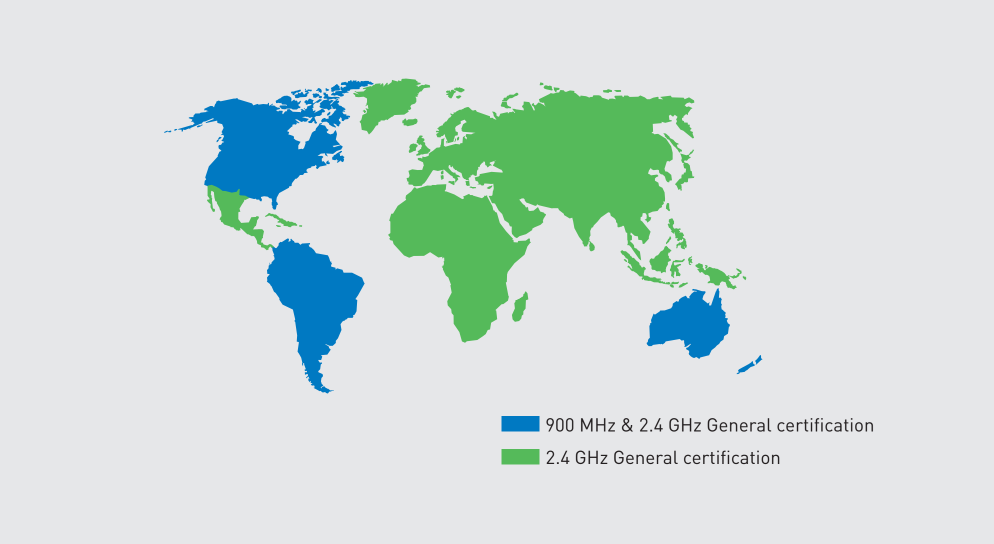
Ethernet network) and the Wireless HART network. By using multiple gateways, added redundancy can be incorporated to achieve higher availability.

6LowWPan

Another interesting communication technology based on the IEEE 802.15.4 specification is 6LoWPan. The abbreviation is for »IPv6 over low power WPAN (Wireless Personal Area Network)« The goal of the specification RFC 4944 launched by IETF (Internet Engineering Task Force) is the efficient communication of IPv6 data volumes in IEEE 802.15.4 networks.

Of central interest in 6LoWPAN is the implementation of an IP-based sensor network with a minimum of effort. The advantages are clear: IP networks are now the most prevalent and well-defined, and the necessary protocol stacks are easily available. This also renders particular expectations on the technology, which has to enable more embedded integration than IEEE 802.11 WLAN can provide, while delivering lower energy consumption capabilities by a factor of 100. Particular challenges arise with the need to adapt the expansive IPv6 header to the requirements of IEEE 802.15.4 and also to segment data volumes, so that the MTU of 802.15.4 data volumes of 127 bytes are exceeded. The RFC 4944 provides guidelines on how the respective headers can be compressed in order to achieve interoperability between technologies.

The first non-commercial and commercial 6LoWPAN protocol stacks are now available. The idea of extending »all-IP« to sensor networks as well is very attractive and is sure to play an important role in the future.


🟦 900 MHz & 2.4 GHz General certification
🟩 2.4 GHz General certification

Only the 2.4 GHz band offers the advantage of worldwide general approval

2,4 GHz-Technologien

The 2.4 GHz band is probably the most interesting frequency band for all areas of technical wireless systems. This is so for various reasons. First, sufficient capacity is available with a bandwidth of 83.5 MHz. Second, this is the only band that is largely harmonised globally. If the creation of global products is the intention, this is a manufacturer‘s first choice.

The popularity of its technology, however, does not come without disadvantages. Almost every device designed to transmit run in the 2.4 GHz band, ranging from established wireless technologies like Bluetooth and Wi-Fi to the afore-mentioned ZigBee and Wireless HART to finally video transmission systems and baby monitors.

In Germany, there is a general allocation of the 2.4 GHz band for almost all applications with transmission capability of up to 10 mW without limitation on bandwidth or occupation of frequency. Moreover, devices with up to 100 mW used in the application, »Local Wireless Networks« (WLAN Wireless Local Area Networks), are allowed to be operated, where using frequency hopping transmits a maximum power density of 100 mW / 100 kHz or has a spread spectrum output of 10 mW/1 MHz.

The current frequency allocation is valid until 31 December 2013. An amendment is very likely as other systems, for example, digital remote control used in model airplanes and other devices in this frequency range are pressing for allocation.

Proprietary technologies: steute wireless

Many proprietary wireless systems use the 2.4 GHz technology because of the international acceptance. As an example, steute‘s wireless system, specifically developed for the requirements of medical technology, are described in detail on page 182.

// Bluetooth

Bluetooth is the most successful wireless communication technology to date. No other wireless consumer technology in the last ten years has achieved similar growth in such a short time. An assembly of over 10,000 members, actively supporting the implementation of Bluetooth technology, form the Bluetooth Special Interest Group (SIG). Hardly any other interest group could develop so fast so quickly. At the same time, Bluetooth chip manufacturers have registered a market growth of 20% per year. In a nutshell: Bluetooth is without doubt a state-of-the-art technology.

History

In the 90's, the mobile phone manufacturer Ericsson conducted research on how different communication terminals accessed their networks. It soon became clear that cable connections would be abandoned in favour of cheaper, more efficient radio interfaces. The project was so well-received that in May 1998, the companies Ericsson, Nokia, Intel, IBM and Toshiba joined forces in a Special Interest Group (SIG) to develop a common standard. Evidently, a wireless standard could only be successful if it gained sufficient interest from the main supporters. Three objectives were pursued by the SIG:

1. Development of a standard specification for the purpose of standardising hardware and software interfaces.
2. Creation of a developer forum with possibilities for product certification to secure interoperability
3. Disclosure of the standard for zero-cost licences, accessible by anyone interested in its specifications.

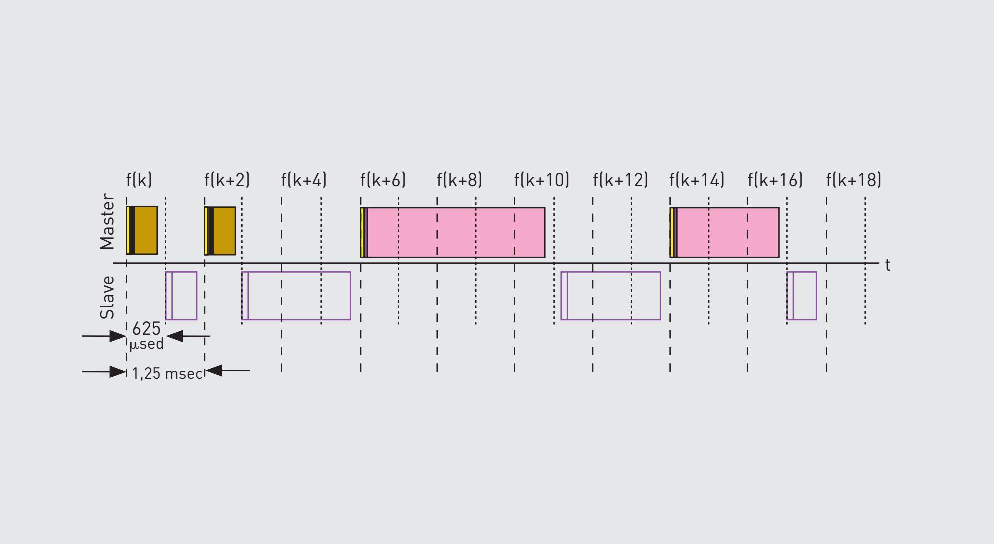
The first working version (V0.7) was published on the 19 October 1998, in which only the rudimentary functions of the base band and the link manager were specified. In January 1999, version 0.8 followed, which included a detailed specification of wireless communication and the fundamental software layers for data transport. The first official launch of the Bluetooth core specification version 1.0a took place on the 26 July 1999 on the SIG website.

The first version was in need of amendment. By the end of the year (12/1999), version 1.0B that particularly included interoperability requirements was published. The breakthrough came in February 2001 with version 1.1. Products that work quite well with this version are still on the market.

Bluetooth is categorised as a PAN (Personal Area Network), which is the focus in the IEEE 802.15. In 2002 the adaptation of parts of the Bluetooth V1.1 standard to the IEEE standard as IEEE 802.15.1-2002 took place. Traditionally, the IEEE 802.x working group specifies only the lower protocol layers of the MAC and PHY, while the higher protocol layers and profiles are not covered in IEEE standards.

Due to the increasing use of Bluetooth and high penetration of WLAN systems (IEEE 802.11b/g WLAN), it was necessary to optimise the co-existence of largely complementary wireless technologies. The 2.4 GHz ISM band ultimately has limited resources and the persistent methods used by Bluetooth wireless FHSS (Frequency Hopping Spread Spectrum) with 1600 hops per second, resulted in a significant impairment of the WLAN links.

To account for these factors, version 1.2 released in November 2003 supplied the first complete overhaul of Bluetooth. The focus of the new specification was an improved connection setup, adaptive frequency hopping for a better co-existence, improved mechanisms for synchronisation and data flow control, and data trans-



Master and slave communication in a piconet.

mission via the synchronous SCO (Synchronous Connection Oriented) links.

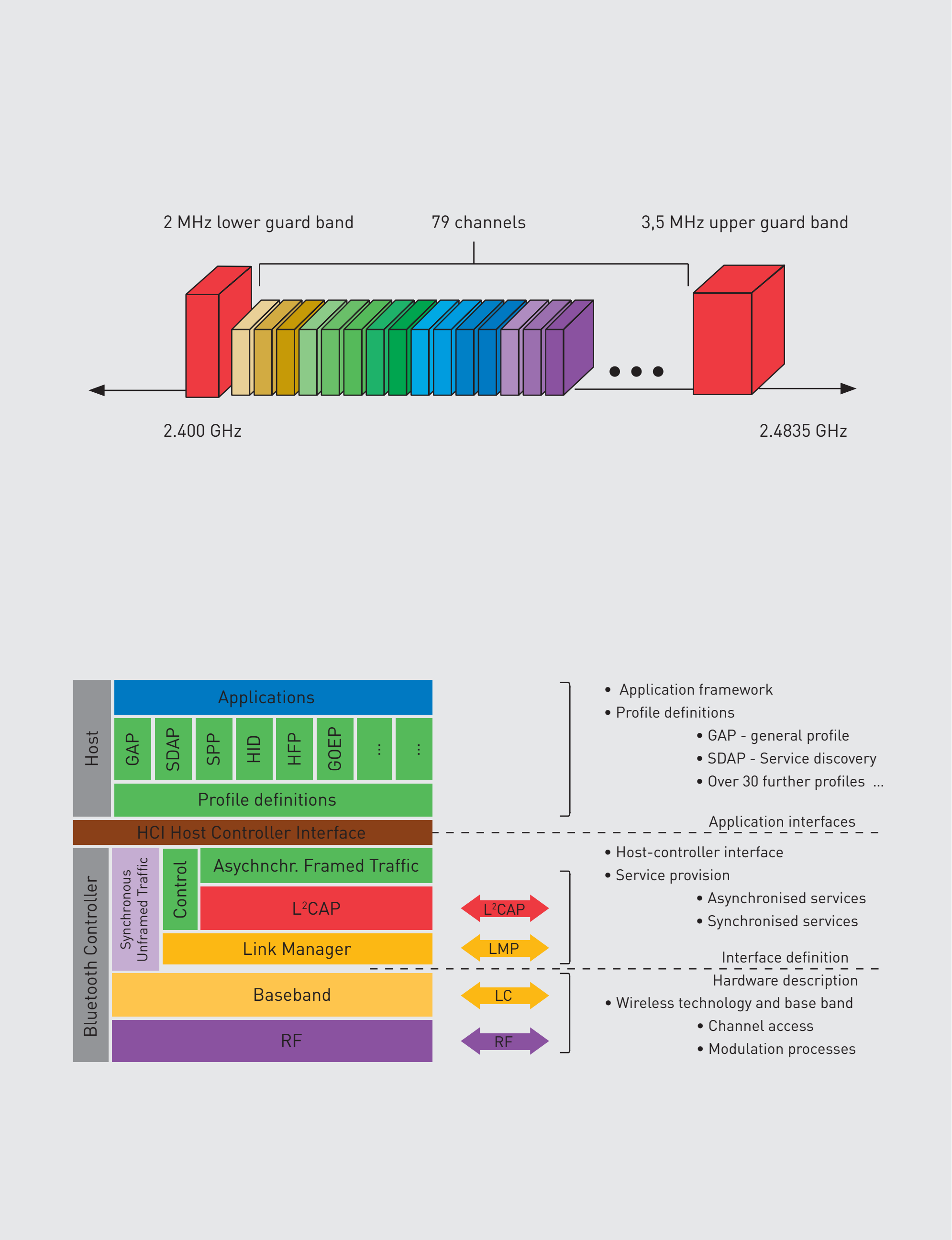
These ongoing improvements actually warranted a fully reformulated version, which resulted in version 2.0 dated 1 August 2004. An important revision of version 1.2 to version 2.0 is the addition of »EDR«, Enhanced Data Rate. While the classic Bluetooth piconet only allows 1 Mbps gross data rate, the high-speed modes enable gross rates of 2 or 3 Mbps.

In version 2.1 + EDR of July 2007, numerous new features were added, such as »Secure Simple Pairing«, which essentially provides Bluetooth better manageability and faster access times.

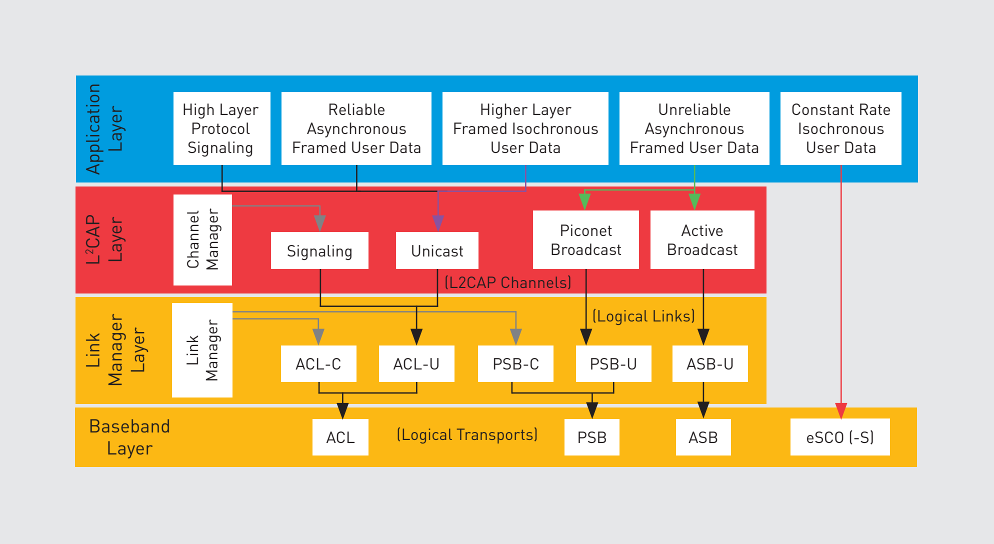
Current Bluetooth specifications

Bluetooth version 3.0 + HS (High Speed) was ratified in April 2009. The high-speed channel is based on WLAN 802.11 and allows for a faster data channel via a second wireless technology. Establishing a connection and the provision of service as well as security can continue to be implemented via the Bluetooth channel. Integration of WLAN allows a data rate of up to 24 Mbps. The high-speed channel was initially planned under the use of UWB (Ultra Wide Band – ECMA 0368), which was, however, abandoned later.

The current specification is Bluetooth 4.0 EDR. Its purpose is further optimisation, allowing particularly fast connection (<5 msec) and higher security levels. Furthermore, the reduction of energy use in newer versions is a significant milestone. The objective is the operation of a Bluetooth device on a remote energy source for many years.



Top: Bluetooth uses 79 channels with 1 MHz bandwidth in the 2.4 GHz ISM band
Bottom: Overview of Bluetooth protocol stacks



Overview of the Bluetooth data channels in version 2.1

Technological overview

In technological terms, Bluetooth is significantly different from other wireless technologies, as embedded integration is already taken into account in the specification. This means that a TCP/IP-based starting point for data transmission has not been chosen. Rather, a service-oriented framework provides a scaled provision of services. More than 30 different profiles have already been defined, or can be found in the specification to implement specific solutions. The multitude of different applications corresponds with the numerous technical peripheral conditions.

Bluetooth technology can be divided into three logical sections:

Radio technology and base band

provide a low-level functionality of the system. Principal operation rules are relegated, which are applicable to practically every device and invariably must function identically.

The link manager and the host controller interface

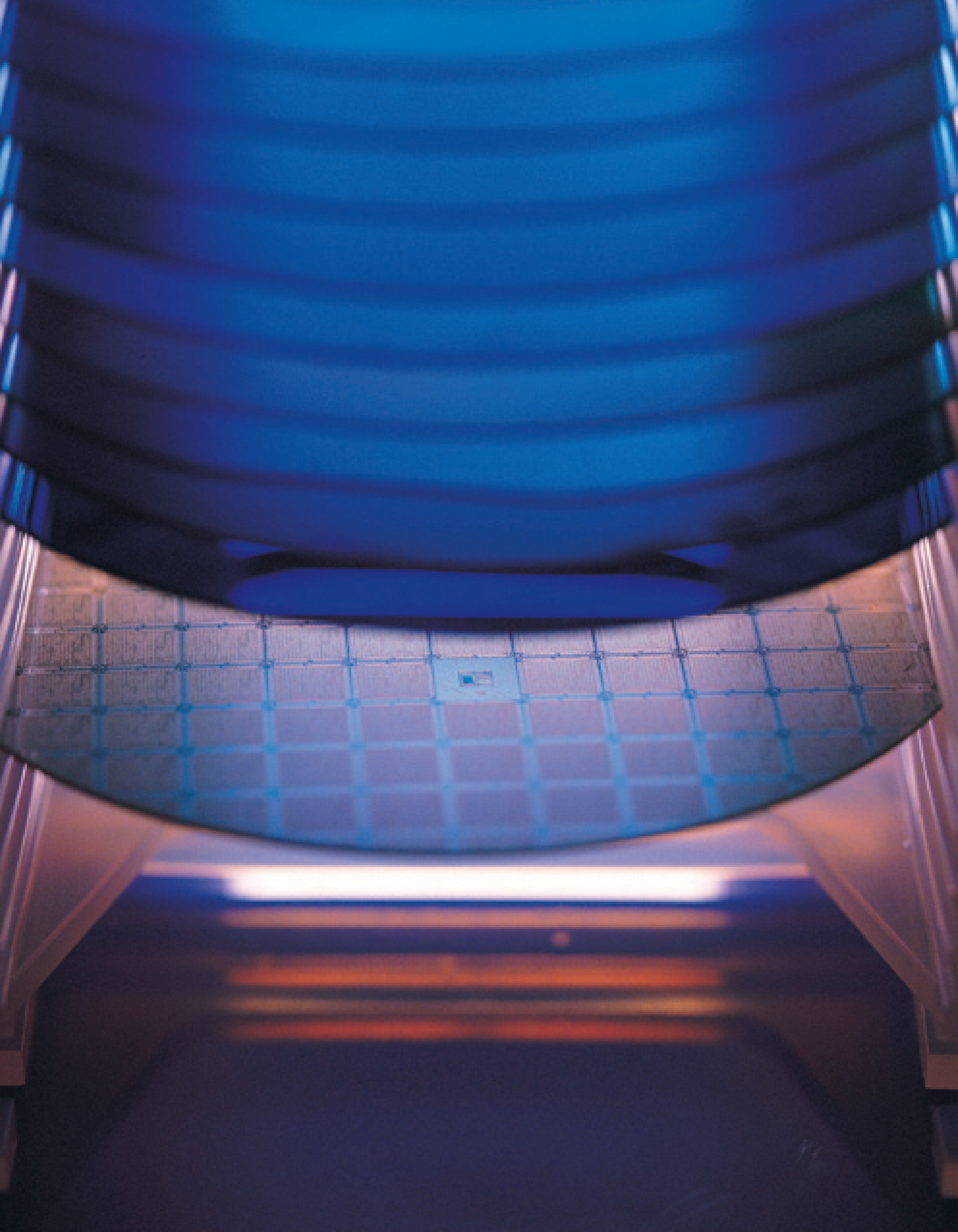
link the application and wireless technology. The link manager is responsible for connection management, the HCI (Host Controller Interface) provides a standardised interface to the host computer.

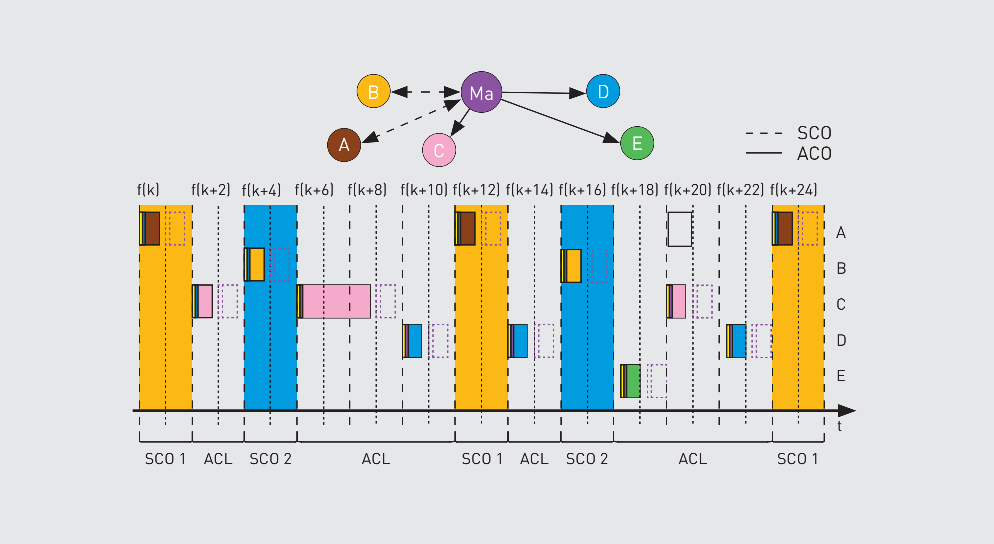
Application layer

Differentiated applications require defined Bluetooth usage models, the so-called profiles. These form the application interface of Bluetooth.

Bluetooth protocols: Wireless technology and base band

Bluetooth, being a wireless technology, also uses the 2.4 GHz ISM band. Unlike WLAN, however, a frequency hopping method with up to 79 channels of 1 MHz bandwidth has been selected. The channels are pseudo-randomly traversed at a hopping rate of 1600 hops/second. The hopping sequence is characteristic of a piconet and basically dependent on a clock and the device





Master and slave communication in a piconet.

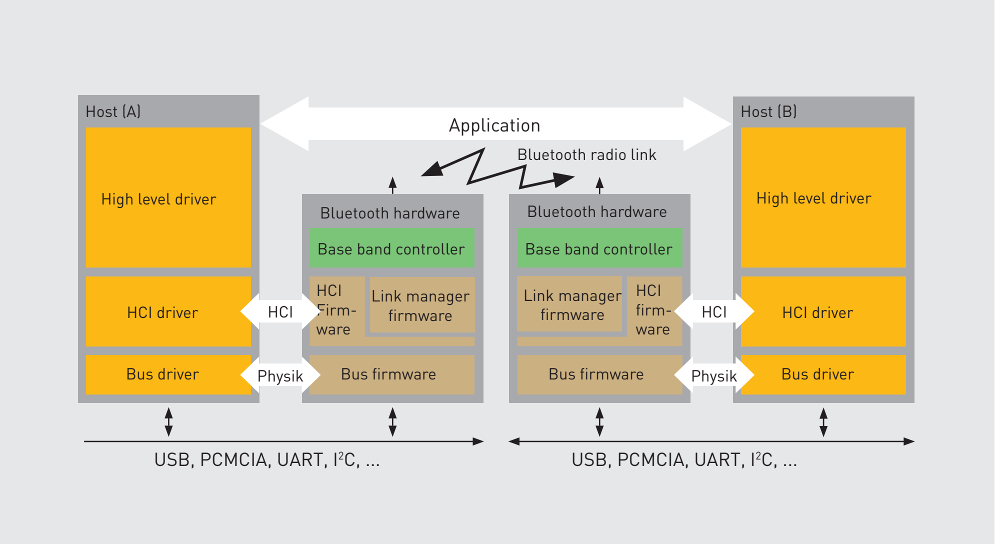
address (Bluetooth MAC-ID). Due to the hopping process, compared to other wireless systems, Bluetooth is very robust in terms of narrowband interference. Since version 1.2, adaptive frequency hopping has been added, which explicitly maps the frequency ranges, making the operation free of interference and enabling conflict-free co-existence with other wireless systems in the same frequency band.

Three power classes

Bluetooth defines three power classes for wireless modules. Class 3 modules characterise the lower power end, with a transmission capacity of 1 mW and a typical range of 10 metres. Class 2 modules have a transmission capacity of 2.5 mW and a range of about 30 metres. High-end application is achieved with 100 mW modules with ranges of up to 100 metres. Class 2 and class 1 modules are equipped with power controls so that transmission is carried out with only the minimum required power. The actual achievable range depends significantly on the HF design, the antenna used and the overall quality of the components.

A Bluetooth network forms what is called a piconet. A node takes on the role of a network master and transmits its Bluetooth ID and timing of the hopping sequence to the connected Bluetooth slaves. Up to seven slaves can actively communicate with the master. A further 255 slaves can be included as parked members of the piconet, depending on system implementation. This form of network formation can theoretically include up to 20,000 participants, coexisting without conflict in a radio field, in 79 piconets. Bluetooth therefore provides an extremely high density of communication nodes.

Through the master-slave communication scheme, a precise timing given by the master can be achieved. A communication time slot is 625 microseconds. A complete data frame consisting of master request and slave response takes 1.25 ms. To increase the usable data, 3-slot



The HCI interface reads out the Bluetooth hardware to the host.

or even 5-slot data packets are possible. This means that up to V 1.2, maximum data rates of 433 kbps symmetric, or 723.2/57.6 kbps asymmetric can be achieved. According to the EDR Bluetooth V 2.0 definition, a maximum data rate of asymmetric 2178.1/177.1 kbps is possible.

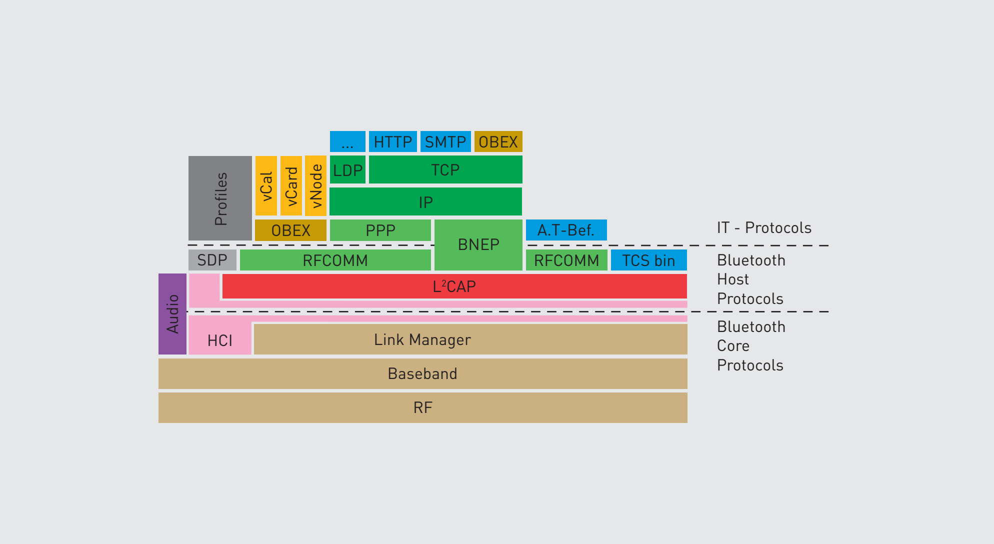
Versions up to 1.2, Bluetooth uses a Gaussian Frequency-Shift-Keying process (GFSK) for modulation. After Version 2.0, for the transmission of EDR data packets, a π/4-DQPSK (2-Mbps) – and 8DPSK (3-Mbps) modulation was added. Only the user data portion of the data is sent via PSK modulation.

Physically, the synchronous Bluetooth controller (SCO - Synchronous Connection Oriented) and the asynchronous (ACL – Asynchronous Connection Less) provided data channels. Synchronised data channels are used for sending streaming data. SCO channels allow scheduled data transfer without resending. Extended SCO channels (eSCO) allow resending to some extent. The asynchronous channels are essentially subject to an ARQ scheme (Automatic Repeat Request).

Link manager and link controller

The link manager is responsible for setting up a connection and the allocation of channel bandwidth. Set timeslots are reserved for SCO, or eSCO channels. The surplus timeslots can be used for asynchronised communication.

In addition to managing the data channels, the link manager provides services for synchronisation of the data packets, for the definition and monitoring of service levels (QoS – Quality of Service) as well as for security.



Bluetooth profiles define the complex interaction between protocols and applications

The security model

Bluetooth supports a scalable multi-level security model, which uses three different security modes in addition to various safety mechanisms. All mechanisms are integrated in the Bluetooth controller, so no additional security hardware or firmware needs to be used.

The weakest security mode is security mode 1. No authentication is needed and the data is usually transmitted without encoding. Consumer devices are mostly found in mode 2. The security procedures are carried out by the host so it is always necessary to connect to the Bluetooth controller. The advantage of scalable application-level security is contrasted with potential security vulnerabilities. All known vulnerabilities in Bluetooth-based attacks are due to this mode-2 problem.

A module-level security is ensured in the security mode 3. In this mode, all the security features of Bluetooth chips are used and there are no additional measures needed from the host. Mode 3 will ensure that only reliable sources gain access to the system. The security mechanisms provided by Bluetooth offer in addition the establishment of invisible point-to-point connections, authentication by the master and slave, and the change of the link codes that can take place according to each data packet. An encryption with up to 128 bits and the corresponding key management are also integrated into this technology.

HCI – Host Controller Interface

The connection of the Bluetooth controller to the host system is via the Host Controller Interface. The HCI layer provides an abstract interface, which provides all relevant base band and link manager functions in an API to the host. Typical physical interfaces at this level are USB (Universal Serial Bus) and UART.

High level host software

Bluetooth offers a very powerful protocol stack compared to other wireless systems. Not only the data transport, but the quality of service and security are also implemented by the Bluetooth controller. Also, the overlaying protocol stack on the host driver level and application level are defined. Basic protocols for this are SDP, TCP, and BNEP AVCTP and AVDTP.

The Service Discovery Profile (SDP)

It provides basic services for finding and delivering service features available on a computer. This is necessary because not all Bluetooth devices need to provide all services and some measure of leeway is available in the implementation of the various forms. Figuratively speaking, SDP assumes the function of the »Yellow Pages« in a Bluetooth device and is mandatory for all devices.

TCP is the Telephony Control Protocol,

which is responsible for the administration and use of telephone services. Only devices with telephone services need this protocol.

BNEP (Bluetooth Network Encapsulation Protocol)

provides basic services for the delivery of layer 3 protocols via Bluetooth. This protocol is necessary in access points and for connecting, for example, to the Ethernet.

AVCTP (Audio Video Control Protocol) and AVDTP (Audio Video Distribution Transport Protocol)

are newer protocols for the transport and control of video and audio data.

Besides these Bluetooth protocols, established IT protocols like OBEX or TCP/IP can be connected and make the respective application layers available. How the individual protocols are used, is regulated in the profiles.

Bluetooth profile

The use of the transport layer of the link manager is controlled by profiles. Profiles define exactly which role and task a master device and a slave device have in a specific application. This makes a very sophisticated scaling of the functionality of Bluetooth devices possible, making the best use of this wireless standard.

The GAP and SDAP profiles must be implemented in all Bluetooth devices. GAP (Generic Access Profile) defines general access to a Bluetooth device with the security modes, the visibility and inquiry. The SDAP (Service Discovery Application Profile) controls access to the service database of a Bluetooth device from which the main communication parameters and supported services and profiles can be identified. The basis for many other profiles is the SPP (Serial Port Profile), which is also used as a universal serial interface. SPP can transmit multiple virtual serial RS-232 channels via a Bluetooth interface, making it the base for cable replacement.

Bluetooth Low Energy (LE)

Bluetooh Low Energy (LE) rounds off the Bluetooth wireless range. Especially for small, low-energy sensors, Bluetooth is much too powerful. With reduced functionality and a streamlined protocol stack, LE only uses 10-30 % of the energy of a conventional Bluetooth module. There are no differences in the use of the 2.4 GHz ISM band frequency hopping and the simple GMSK (Gaussian Minimum Shift Key) modulation is possible with an almost unaltered radio signal. Thereby net data rates of up to 1 Mbps can be reached. Depending on technical specifications, ranges of 2 to 300 m are possible.

The differences are found in the much simpler protocol stack from the link manager. The master and slave roles are clearly defined. Role switching is not possible. The slave is categorised as an advertiser. Periodic broadcasts send a service

A2DP	Advanced Audio Distribution Profile	Transmission of high quality stereo audio data
AVRCP	Audio Video remote Control Profile	Remote control of audio/video devices
BIP	Basic Image Profile	Transmission of image data
BPP	Basic Printing Profile	Printing
CIP	Common ISDN Profile	ISDN connections via CAPI
CTP	Cordless Telephony Profile	Cordless telephony
DUN	Dial-Up Networking Profile	Internet dial-up connections via PPP
DI	Device Identification	Identification of devices by specified manufact./providers
ESDP	Extended Servoce Discovery Profile	Extended service detection for uPnP
FAX	Fax Profile	Fax services
FTP	File Transfer profile	File transfer
GAVDP	Generic Audio Video Distribution Profile	Transfer of audio/video data
GOEP	Generic Object Exchange Profile	Data exchange of OBEX objects
HCRP	Hardcopy Cable Replacement Profile	Printer applications
HFP	Hands Free Profile	Wireless connection of mobile phones in a vehicle
HID	Human Interface Device Profile	Input devices such as a keyboard and mouse
HSP	Headset Profile	Speech output via headset
OPP	Object Push Profile	Exchange of calendars and business cards
PAN	Personal Area Network	Virtual network connection
PBAP	Phonebook Access Profile	Access to the telephone book of a mobile phone
SAP	SIM Access Profile	Access to the mobile phone SIM card and deactivation of GSM module in vehicle use
SDAP	Service Discovery Application Profile	Tracing service signature and profiles
SPP	Serial Port Profile	Serial data transmission
SYNCH	Synchronisation Profile	Synchronisation of OBEX objects
VDP	Video Distribution Profile	Distributing video data

Bluetooth profiles and their application

Physical data rate	1 Mbps
Typical range	typ. 10 m (2 ... 300 m)
Frequency band	2.4 GHz (ISM band)
Power consumption	stand-alone BT x 0.1 ... 0.25 Dual mode BT x 0.8
Modulation	GFSK h= 0.5 (BT)
Access method	TDD (Time Devision Duplex)
PDU user data	variable length 1 ... 31 Byte
Net data rate	Unidirect. 350 (816) kbps Bidirect. 2 x 259 (2 x 449) kbps

Performance data of Bluetooth LE

notification, to which a master responds. Masters are also called scanners that look for service advertisements. When a service indicator is found, the master takes on all the connection management. The scanner can manage up to 255 slaves in a TDD (Time Division Duplex) process.

Since only 3 call frequencies are used instead of 32, a much faster connection is possible. Connection is therefore made within only ten milliseconds.

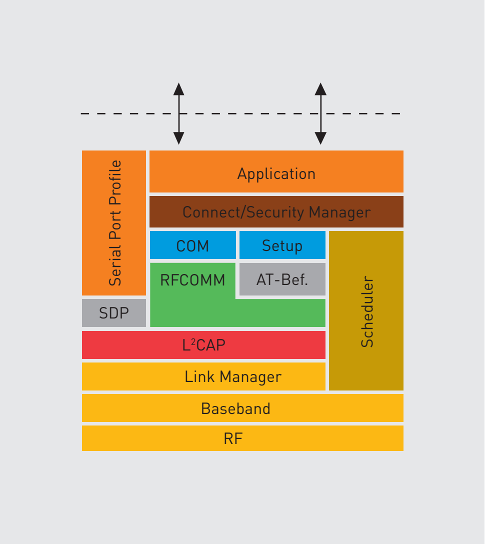
The wireless transmission on both conventional and LE Bluetooth is largely identical, so the logical step of constructing dual-mode Bluetooth devices is imminent. The additional cost of a conventional Bluetooth system is extremely low.

Embedded Integration

Due to the powerful protocol stacks, Bluetooth often seems complicated. However, that is not true. First, as profiles are established, clear definition of functions are usually very easily achieved. Second, many providers now allow the integration of the upper-layer protocol stack in the Bluetooth controller. This enables the entire application to be housed in the Bluetooth controller, the embedded device requires only a simple UART interface with the processor for the integration of a corresponding wireless connection. Preferred SPP and HID profiles for easy embedded devices are possible.

WLAN – WIFI – IEEE 802.11

The different transmission power levels used in the encoding process and the bandwidth of a channel essentially determine the range and data throughput in wireless technologies. It is in these points that the available technologies differ significantly. A discussion of whether either Bluetooth or WLAN should be chosen, in this sense, is inconclusive. A good perspective of the whole problem is assumed when one sees the situation in terms of standardisation. The IEEE standardisation institution plays an important role in this matter. Under the working group 802,



Implementation of embedded devices with embedded stack

there are different interest groups working on the standardisation of wireless and wireline communications. The WLAN specifications in the working group 802.11 reflect only a portion of the cable-free communication. Lest we digress too far, the following classifications should be sufficient:

802.11 WLAN

The main activities for a wireless Ethernet are summarised in this group. The standardisation is underway and various PHY and MAC layers for high-speed communications at 2.4 GHz and 5 GHz band are defined. The dynamics of work groups is documented in alphabetical increments, now already at 802.11n.

802.16 WMAN

This denotes the Wireless Metropolitan Area Network. This type of network describes a standard for wide area communications, which is covered, at least in Europe, by GSM and UMTS wireless technology.

802.15 WPAN

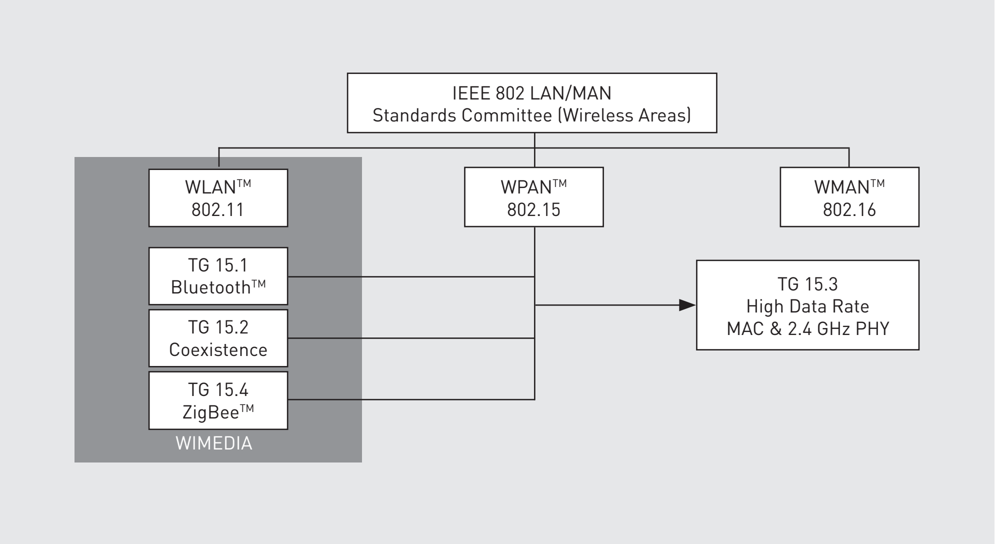
Wireless Personal Area Network describes networks for communication at close range. WPANs play a unique role for the ad hoc networking and the exchange of data between mobile devices such as PDAs (Personal Data Assistant) or mobile phones. This band is now covered by Bluetooth. In the 802.15.1 working group, efforts are being made towards the integration of Bluetooth into an IEEE standard. Bluetooth MAC and PHY are already specified in the IEEE. Due to the high complexity and the associated less than optimal low power, ZigBee 802.15.4 provides for the wireless communication in the lower range. ZigBee is a pure master-slave system with data rates up to 128 kbit optimised for low power and the transmission of even the smallest of packets. Furthermore, the working group deals with the interoperability of different wireless technologies in 802.15.2 and the optional increase of data rate in the subgroup 802.15.3.

The frequently encountered opinion that WLAN, Bluetooth and ZigBee technologies are interchangeable, therefore, is not correct. On the contrary, these standards complement each other due to their different features, or they are suitable for specialised applications.

The history of IEEE 802.11

What is usually meant by WLAN is Wireless LAN according to IEEE 802.11. Behind this abbreviation is an entire menagerie of different and sometimes not even compatible technologies.

Historically, WLAN was developed from the early 90s, with the admission of the 2.4 GHz ISM band and the efforts of the IEEE to develop a standard for the MAC and PHY layers. In 1997, the 802.11 standard was approved, the addition of a diffuse infrared transmission that used a less demanding technology to full frequency hopping (FHSS - Frequency Hopping Spread Spectrum) and allowed a data rate of 1 or 2 Mbps. A modified

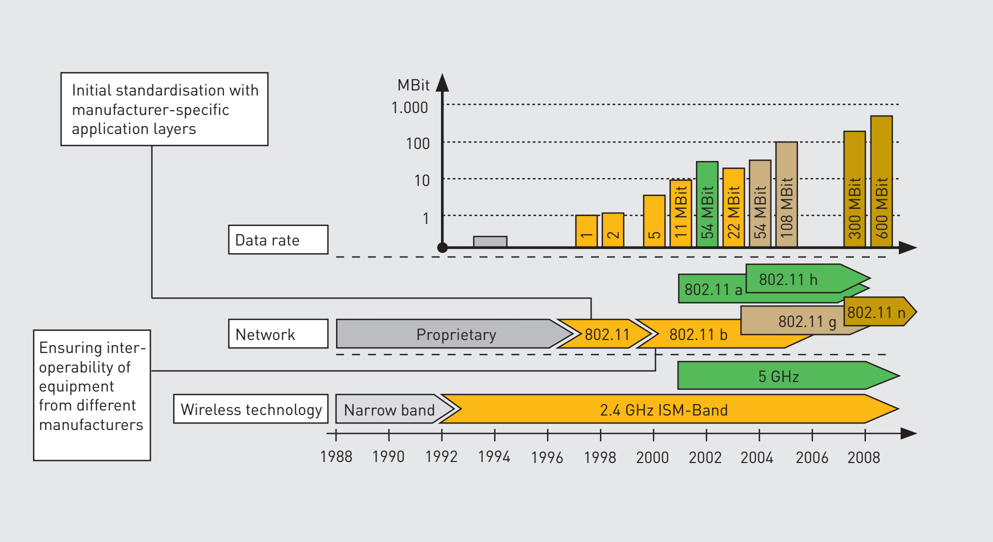


Wireless technologies according to IEEE 802

modulation method, the DSSS (Direct Sequence Spread Spectrum), was finally developed in 1999 under the 802.11b, which was a very successful alternative to wireless with data rates of 11 Mbps in the 2.4 GHz band. This standard is now the mother of all WLAN implementation models.

It quickly became clear that the attainable net data rate for high-quality applications was not sufficient and that the 2.4 GHz band offered too low a channel capacity for a large number of wireless nodes. In Europe one worked on the Hyper-LAN standards and America answered with the 802.11a standard. By using the 5 GHz ISM bands combined with OFDM (Orthogonal Frequency Division Multiplexing) modulation, data rates of 54 Mbps could be typically achieved. Due to the difficult harmonisation of the 5 GHz band, the modulation method in the 2.4 GHz band was transferred and codified as the 802.11g standard. 802.11g represents the state of the art in wireless technology of the 802.11 group.

In the following section the main aspects of the 802.11 alphabet are described in brief. The sequence of letters, however, has no chronological significance. Today, the standard 802.11g and h pertain to transmission, 802.11i to security, and 802.11d, e and f deal with the coordination of WLAN systems.



Development of technologies according to IEEE 802.11

Protocols

First of all, some basic channel access methods for 802.11 networks have to be made clear in order to understand the fundamental rules of operation. IEEE 802.11 networks can implement different MAC requests.

WLAN standards include the provision of asynchronised data services. Pending data packets are sent according to the »best effort« principle. Theoretically, broadcast as well as multicast are possible. For asynchronous data services, a DFWMAC-DCF or CSMA/CA method for channel access is implemented.

A ready-to-transmit station observes the air interface to determine whether a carrier is present. If a period DIFS (Distributed Coordination Function Interframe Spacing) channel is not occupied, it can immediately begin transmission. If the medium is occupied, the stations wait until a channel becomes available. For a destruction-free arbitration, competing stations wait for SIFS (Short Interframe Space), which are reserved for arbitration messages, such as confirmation messages. Then a time-bound service using PCF (Point Coordination Function) can be started within the PIFS (Point Coordination Function Interframe Space). Thereafter, after the DIFS has terminated, the competition phase of the low priority stations can begin. In order to avoid simultaneous transmission, a concurrence meter is started (back-off algorithm), which guarantees random access.

In addition to this standard access method, even RTS-CTS-driven CSMA-CA access, or polling methods (MAC-PCF) can be optionally implemented, which are not available with all hardware.

The DFWMAC DCF method with RTS-CTS control is used to avoid the hidden station problem. If a station wants to transfer data, it first sends out an RTS signal (Ready To Send). Within the SIFS, a ready-to-receive node sends out a CTS signal

Function		Layer
802.11	1–2 Mbit, FHSS, DSSS, IR	MAC, PHY
802.11a	Up to 54 MBit, OFDM, 5 GHz	PHY
802.11b	Up to 11 Mbit, DSSS, 2.4 GHz	PHY
802.11c	MAC-Bridge	MAC
802.11d	International Roaming	MAC
802.11e	QoS Quality of Service	MAC
802.11f	Interaccess Point Protocol	MAC
802.11g	Up to 54 Mbit, DSSS/CFDM, 2.4 GHz	PHY
802.11h	Transmit Power Control/Dynamic Frequency Selection	MAC
802.11i	Security	MAC
802.11k	Radio Ressource Management	PHY
802.11m	Maintenance of Standart	MAC, PHY
802.11n	High Troughput (>100 Mbit)	MAC, PHY

Summary of the significant 802.11 technologies

(Clear To Send), whereby all stations that are located in the range of the two nodes are informed about the data transfer. After receipt of CTS, the stations go into the NAVCTS mode by waiting for an acknowledgment (ACK) of the communication and then become active again.

MAC-PCF (Point Coordination Function) is an optional method for the processing of time-critical services with prioritised access. PCF is only available in the infrastructure mode. The access point coordinates the access to the medium at fixed time intervals. Within the CFP time-frame, the access point works through a polling list of different explicitly registered nodes within one of theses priority time-frames. For the duration of the polling cycle, the other stations are excluded from the active communication. Upon completion of the PCF, the competitive nodes can resume participation in data exchange. The MAC PCF function is not supported by all devices and must be used explicitly.

IEEE 802.11a

The standard 802.11a for up to 54 Mbps in the 5 GHz band was ratified in 2002. The first products were available from the middle of 2003. The particular advantages of 802.11a are in its utilisation of the seldom used 5 GHz band, which is far less occupied than the 2.4 GHz band, as well as the OFDM modulation method which permits up to five times higher symbol rate than DSSS. This, however, is at the cost of interference protection in the coexistence of the systems.

The standard offers many more parallel existing data channels than the 802.11b standard with only three non-overlapping channels. Still, here too there were some problems. In the main markets of the USA and Europe, significant differences in product characteristics within each product zone have to be managed as the channels and transmission power levels are not harmonised.

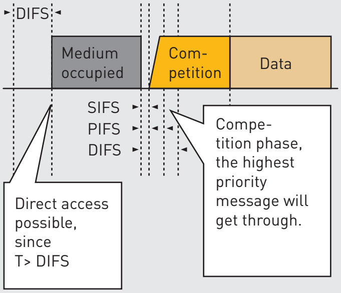
Moreover, one must remember that the data rate is strongly dependent on distance. A gross data rate of 54 Mbps can only be achieved within a range of a few metres. Longer distances (greater than 70 m) usually only allow a data rate of a few Mbps. Lastly, not only the gross data rate should be considered: 802.11a has a maximum net data rate of 15 Mbps in high speed mode.

Considering the development, 802.11a is clearly in competition with the high-speed 802.11g option and the European HyperLAN2 that allows significantly higher net throughput of up to 40 Mbps with the same symbol rate.

IEEE 802.11b

802.11b describes the de facto standard for wireless LAN and was published in 1999. Today, one can find WLAN components according to the 802.11b as a standard in mobile terminal devices with Centrino technology, or in many compact devices up to PDA and mobile phones. The data is transmitted in up to 14 overlapping, 25 MHz-wide channels in the 2.4 GHz ISM band. Depending on the modulation method, data rates of up to 11 Mbps can be achieved. Extensions of the specification also allow transfer rates up to 22 Mbps.

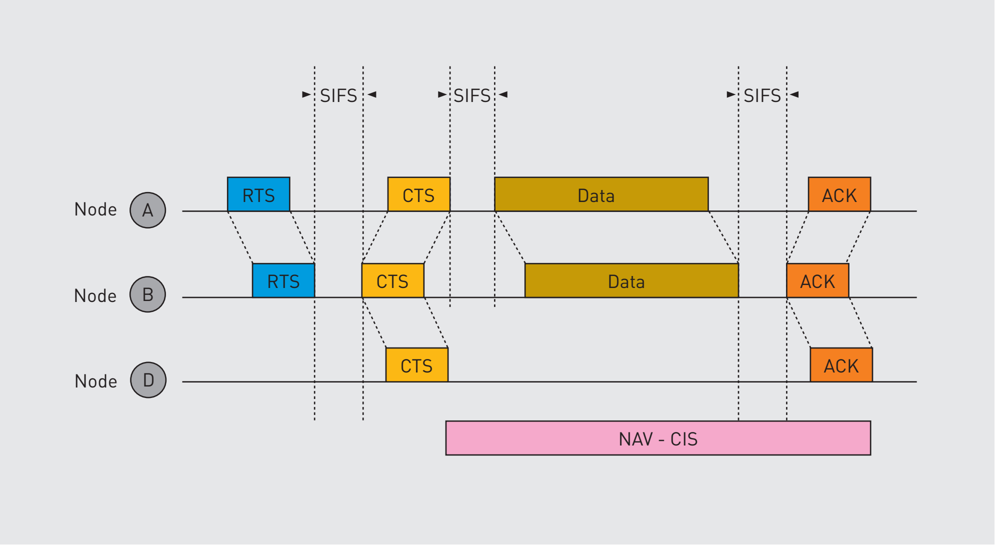
To achieve a high noise immunity, the signals are transmitted in a DSSS process. This means that the application information can be dispersed in an 11-bit pseudo-random sequence, so that potential interference is suppressed at the receiver by 10.4 dB. The advantages of the channel spread is offset by the limitation of only three non-overlapping channels.


DIFS
Medium occupied
Com-petition
Data
SIFS
PIFS
DIFS
Direct access possible, since T> DIFS
Compe-tition phase, the highest priority message will get through.

Medium access with CSMA/CA processes

Frequently mentioned advantages of WLAN are a high data rate and broad propagation range. However, caution should be observed. Even with WLAN 802.11b, the data rate is very strongly influenced by the modulation method used and the distance. In addition, the net data rate is essentially dependent on the modulation.

Another major advantage of 802.11b is the possibility of certifying the interoperability of devices by different manufacturers. For this purpose, several wireless manufacturers formed the WECA (Wireless Ethernet Compatibility Alliance). The WECA enables well-defined test scenarios to demonstrate interoperability and indicates it by the WiFi logo (WiFi – Wireless Fidelity). WiFi is often used interchangeably with 802.11b, but this is not correct. Wikipedia contradicts the definition that WiFi means wireless fidelity. Quote: Wi-Fi is an invented artificial concept for marketing purposes, it does not stand for Wireless Fidelity, as one could assume, as in the popular analogy for Hi-Fi. However, it



Avoidance of hidden station problems by RTS-CTS signals

was well-adopted by the Wi-Fi Alliance as a pun on Hi-Fi.

Specific developments are defined in the 802.11b + standards. These include reverse-compatible transmission standards with a data rate of up to 108 Mbps.

IEEE 802.11d

In order to take national peculiarities into account in terms of frequency division and the transmission power, the d-standard was formulated. The aim is to provide worldwide access points with the same parameters by country code, so that just by a software modification, the regional operating parameters are adapted.

IEEE 802.11e

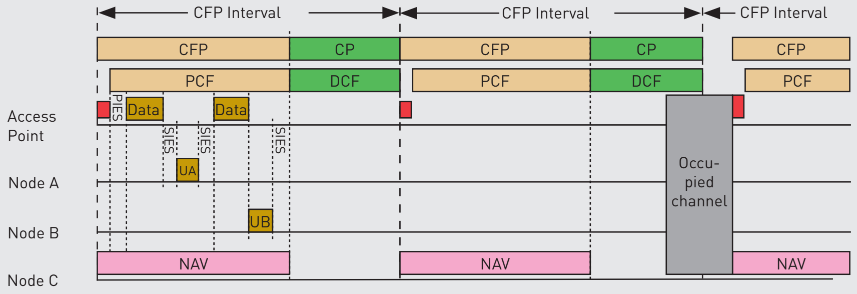
The 802.11e specification is a proposed extension for the provision of defined qualities of service for Voice over IP (VoIP) and other time-critical services that require an isochronous data transfer. The 802.11e MAC extension lets you extend PCF functionality, whereby the stations are assigned a designated bandwidth. Essentially, the adaptation of the ETSI-Hiperlan/2 specification occurs in the IEEE 802.11 standards.

IEEE 802.11f

The finding of mobile stations in a network infrastructure has long been reserved for implementation by individual manufacturers. Only the 802.11f standard allows the definition of an Inter-Access Point Protocol (IAPP), the standardised data exchange between the infrastructure components.

IEEE 802.11g

A significant development was characterised by the adoption of the 2003 802.1g standards. It is compatible with the existing 802.11b standard. Reverse compatibility with the existing infrastructure components and the use of the globally available 2.4 GHz ISM bands are part of this standard. In 802.11g, the signal encoding process



PCF enables a processing of prioritised cyclic messages.

from the 802.11a (OFDM) is adopted and thus provides the additional 22 to 54 Mbps operation modes.

IEEE 802.11h

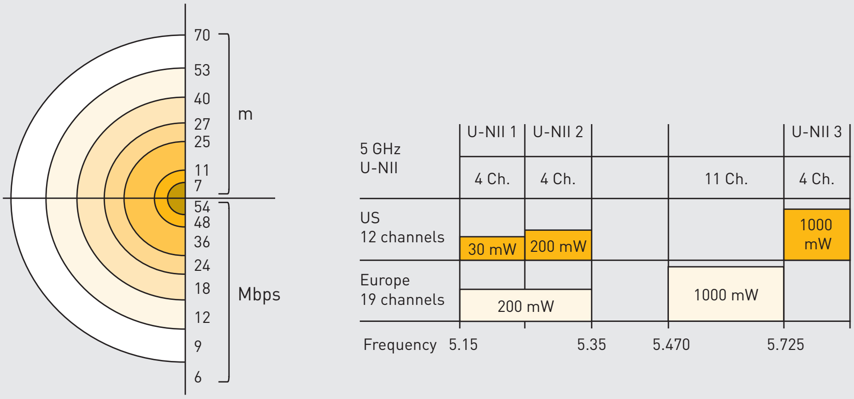
The 802.11h standard is essentially an extension of the existing 802.11a standard in the 5 GHz band with the technical characteristics of the European ETSI HiperLAN/2-Specification. First and foremost, additional modulation methods with higher efficiency of channels, a spectrum management for masking radar frequencies, and the management of transmission power are intended.

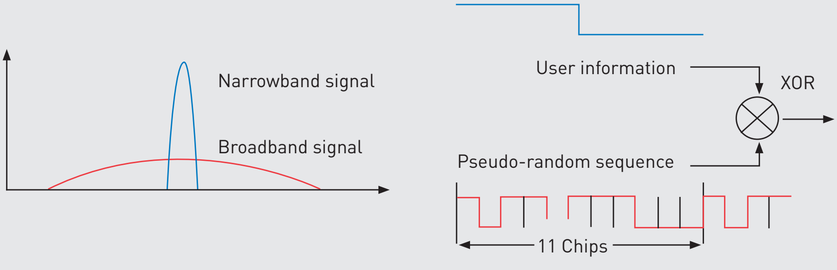
IEEE 802.11i

One could devote an entire book to the security of WLAN. There is so much to be discussed about vulnerabilities and countermeasures. In general, we find that all is not well regarding WLAN security. Two aspects significantly contribute to this:

- Access control is uncertain because the MAC addresses are transmitted unencrypted and therefore can be identified by a unique sender-receiver assignment in the communication. Attackers can sneak into a network pretending to be »real« MAC addresses in a network.
- WEP coding can be cracked or sniffed out with relatively little effort, since they are static. By recording with the appropriate tools, the key is easily reconstructed within a few minutes. .

The 802.11i describes an amendment of the 802.11 a, b and g, to close the gaps in the WEP algorithm (Wired Equivalent Privacy). This includes the authentication of stations using EAP (Extensible Authentication Protocol). EAP uses a variety of authentication mechanisms that require a RADIUS server (Remote Authentication Dial In User Service). To secure the key, the use of a TKIP (Temporary Key Integrity Protocol) is recommended in 802.11i. Here, the static key is circumvented by generating a new key for each data packet. Interception is still possible, but not

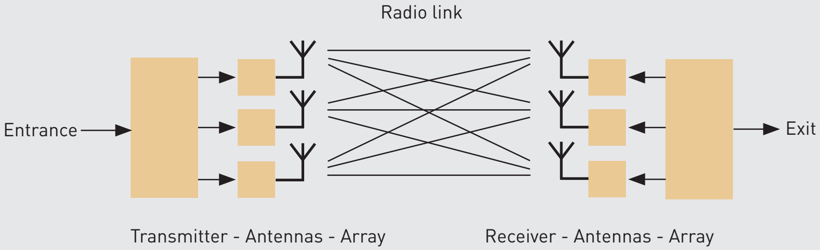




Top: Frequency bands and data rates in IEEE 802.11a networks
Bottom: IEEE 802.11b uses a 22 MHz DSSS signal

1 MBps	DBPSK Differential Binary Phase Shift Keying
2 MBps	DBQSK Differential Quadrature Phase Shift Keying
5.5 MBps	CCK Complementary Code Keying
11MBps	CCK+DQPSK
22 MBps	OFDM Orthogonal Frequency Division Multiplexing

	11 Mbps	5 Mbps	2 Mbps	1 Mbps
Free field	150 m	250 m	300 m	400 m
Buildings	30 m	35 m	40 m	50 m
Net rate	5.04 Mbps	4.33 Mbps	1.59 Mbps	0.82 Mbps
Efficiency	46%	62%	79%	82%



Top: Modulation and net data rate in IEEE 802.11b
Bottom: IEEE 802.11n uses a MIMO system for better correlated signal strength

the reconstruction of the key. WiFi-compliant systems use a backup based on 802.11i. However it must be turned on because backup is always optional.

IEEE 802.11n

The most up-to-date technology in IEEE 802.11 is the standard n. The first draft was adopted in January 2006 and immediately incorporated in a variety of products, even before the standard was ratified. This led to the so-called draft-n products. The final standard was published in October 2009 and has since allowed for a transmission speed of up to 600 MBps. A significant improvement over the previously established 802.11g- or a-standard was made by the use of so-called MIMO antenna arrays. MIMO (Multiple Input Multiple Output) configurations allow for multiple use of a data channel in the same frequency range, resulting in a broader range or a higher data rate.
For modulation, the n-standard uses OFDM (Orthogonal Frequency Division Multiplex) with up to 58 OFDM carriers. The individual carriers can, depending on line quality, be modulated with BPSK (Binary Phase Shift Keying), QPSK (Quadrature Phase Shift Keying) or QAM (Quadrature Amplitude Modulation). While conventional wireless systems amount to a spectral efficiency of up to 5 bits per Hz, IEEE 802.11n allows up to 20 bits per Hertz. With the option of bundling on channel 40 MHz bandwidth, a gross data rate of up to 600 Mbps is achievable. The attainable net data rate is substantially lower. Due to the need for additional redundancy, a realistic net data rate of a maximum of 74-248 Mbps is achievable.

In addition, with increasing spectral efficiency, vulnerability to interference sources similarly increases, but the number of simultaneous similar networks that coexist in the same frequency band will decrease.

C50
L12
C53
37
C65
C157
R41
R47
R46
C77
C78
R54
R55
R56
C80
102
153
R112
35
R57
C88

// Coexistence, safety and security

09

Coexistence in the 2.4 GHz band

Real time in wireless networks

IEEE 802.11 – WLAN

WLAN sideband behaviour

Bluetooth® and WLAN

Coexistence management

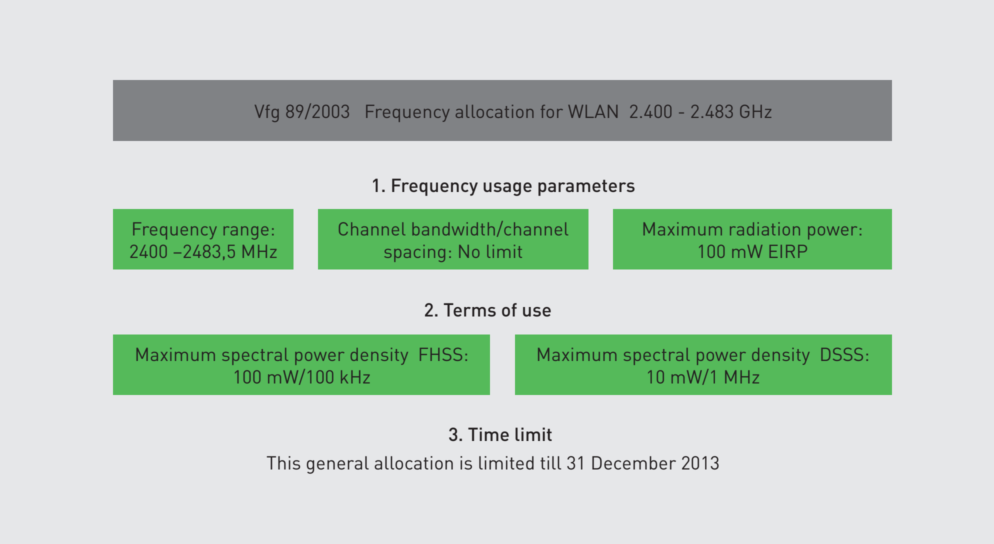
Performance parameters

Coexistence planning

Safety

Functional safety

Information security



Vfg 89/2003 Frequency allocation for WLAN 2.400 - 2.483 GHz

1. Frequency usage parameters

Frequency range: 2400 –2483,5 MHz

Channel bandwidth/channel spacing: No limit

Maximum radiation power: 100 mW EIRP

2. Terms of use

Maximum spectral power density FHSS: 100 mW/100 kHz

Maximum spectral power density DSSS: 10 mW/1 MHz

3. Time limit

This general allocation is limited till 31 December 2013

Frequency allocation for WLAN

// Coexistence in 2.4 GHz band

In general, coexistence means the conflict-free presence and trouble-free operation of various systems in one place. It is often also referred to as the concept of compatibility. As far as wireless technology is concerned, this concept is not so easy to implement. Wireless signals are not detectable with human senses, and above all, there is a wide range of different systems on the same frequency band in the air. To ensure a conflict-free coexistence, the operator of the wireless equipment has the responsibility to evaluate the use of different systems in advance and must plan accordingly. It means that instead of wireless coordination by the state Network Agency, the obligation to coordinate lies with the operator.

Simultaneous presence of the twelve wireless services in the same frequency range listed in the table may not necessarily and directly lead to errors. Different measures can be taken to use the spectrum several times. Many techniques already use very different mechanisms. The possibilities of the diversification of frequency, time and code as well as the spatial or temporal separation were already introduced. Reference is made to the special aspects of WLAN.

With Regulation 89 / 2003, the Federal Network Agency has adopted special rules for such WLAN wireless networks that use frequency hopping or frequency spreading procedures.

Based on these procedures, SRD technologies for short range devices were developed and standardised, which compete now with each other in the frequency band 2400-2483.5 MHz.

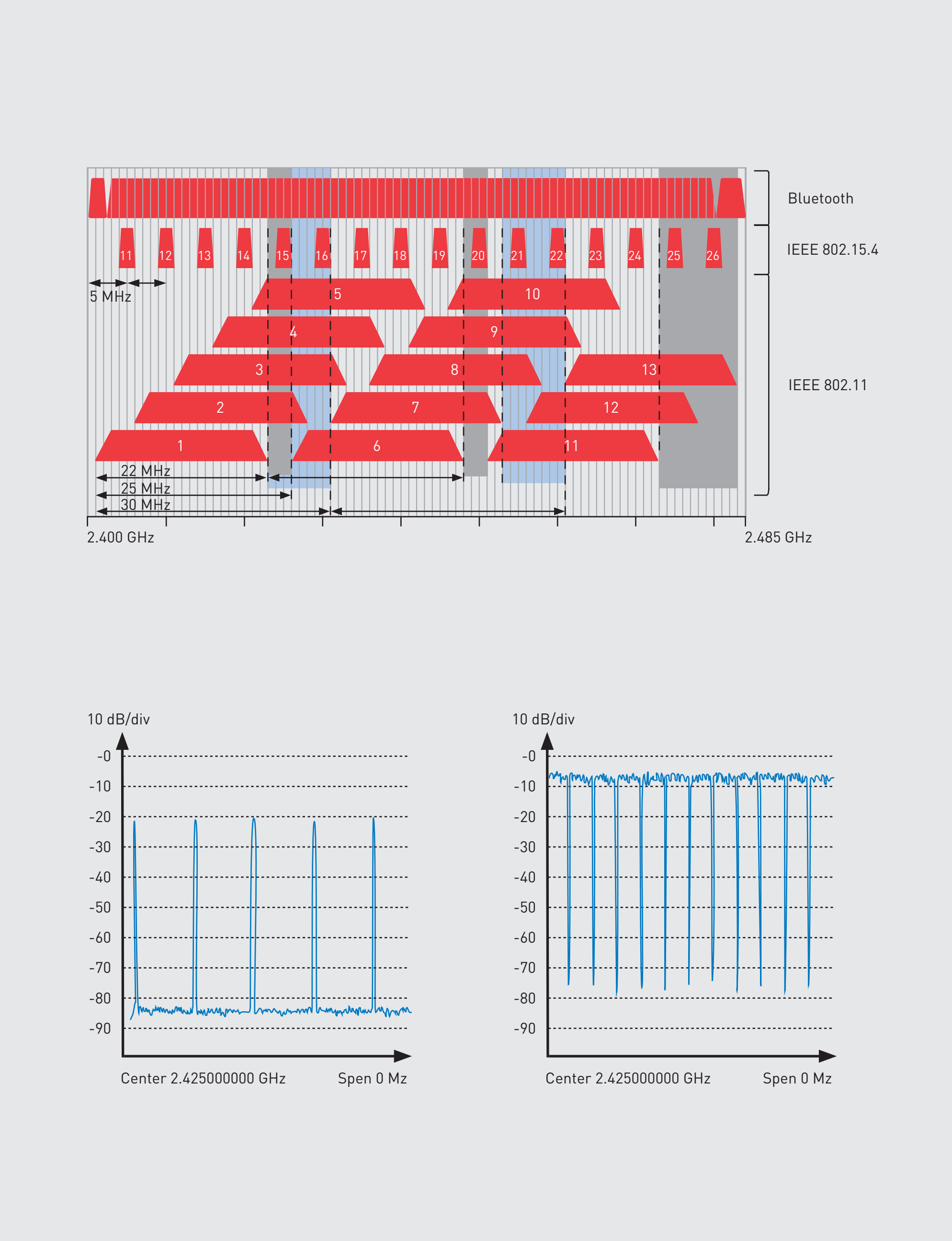
Frequency hopping procedures prevail, because they have on one side a relatively high spectral power density and on the other side frequency agility. Preferably, they can use free vacancies in the partly occupied 2.4 GHz band. With growing numbers of networks, the FHSS networks can

Wireless services	Maximuml admissive equivalent radiation power	Conditions
General wireless short range applications;	**10 mW EIRP**	
Wireless demonstration for educational institutions	5 W ERP	
Wireless remote control Transmission of data signals	10 W ERP	Channel spacing 2.5 MHz Channel bandwidth 2.5 MHz
WLAN **Broadband data transmission**	**100 mW EIRP**	**in relation to the band 2400-2483.5 MHz**
Wireless amateur service		Wireless Amateur Act § 6 Clause 1
Wireless amateur service via satellites		Wireless Amateur Act § 6 Clause 1
Military mobile services		
New radio Transmission of image and sound	27 dBW EIRP (500 W);	13 dBW transmitter output (usually) (20 W)
Wireless low range motion detector, speed and distance measurement	25 mW EIRP	
Military non-navigational wireless service		
Identification purposes	500 mW EIRP outside 4 W EIRP within buildings	
Public trains Automatic automobile identification	500 mW EIRP	B= 1.5 MHz

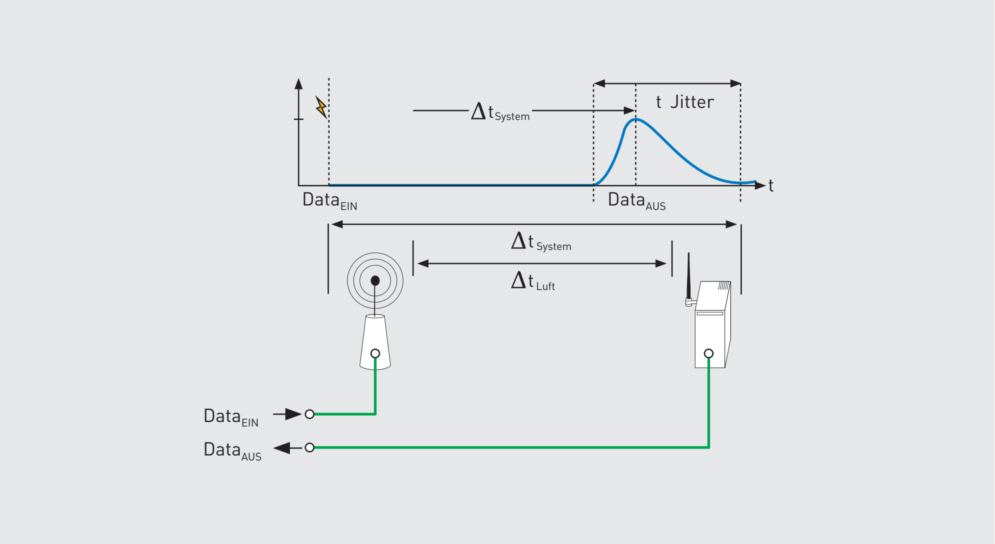
Use of frequency in 2.4 to 2.5GHz area according to FreqNPl, segment plans 276 and 277

however disrupt each other, because they have to reckon with increasing collisions due to their stochastic frequency and channel selection. Usually, this is clearly felt at a capacity of 40 % in the 79 theoretically possible pico networks.

In contrast, the advantage of the direct sequence spreading method is that it suppresses all interfering signals in the same spreader ratio when despreading their useful signals in the 22MHz broad channel. With the usual spreading procedure, the spreading gain is 11 to 32 chips per useful bit values of 10.5 to 15 dB. Therefore, DSSS networks interfere with each other only when the sum level of the »noise floor« is 10.5 to 15 dB above the user receiver's reception level. The



Top: Frequency occupation of some popular techniques within the 2.4 GHz band
Bottom: Oscillograms of 10 % (left) and (right) almost 100 % channel use (duty cycle)



In wireless systems, the whole system and not just the air interface should be considered.

DSSS networks are therefore considered very robust as far as interferences are concerned. The selection of a favourable spreading code also plays an importation role here. At the same time it helps to suppress similar interference signals via the cross- and auto-correlation of digital signal processing.

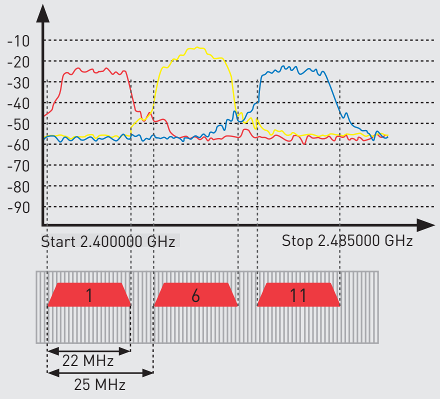
In an interference scenario FHSS against DSSS, FHSS will be only marginally affected due to the 20dB higher spectral power density. FHSS will also only slightly interfere with DSSS. Here, the spreader gain favourably affects the relevant frequency, the lower channel count of 22MHz (attributable to FHSS) and its retention period. A robust evaluation of the interference in both scenarios depends on further unfavourable basic conditions, first of all on the actually present scenario on site and its wireless field load. However, one may assume that in general, these two systems, FHSS and DSSS, will coexist with each other. We will now take a closer look at this. First, it seems alarming that overlapping is inevitable. Even within a technology, such as WLAN (according to IEEE 802.11), the available channels overlap.

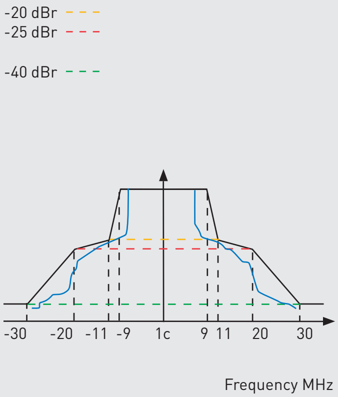
Overlapping is only disruptive if the channels are really used for data transfer. Pure existence of potentially overlapping channels is not harmful.

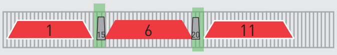
An important criterion is the channel load, which is also described by the duty cycle. Purely statistically, with a very low channel load, the likelihood of interference between different wireless systems in the same frequency range is very low. Only when there is a large channel load will a significant interference on the relevant wireless technologies occur. The following observations will make that even clearer.

// Real time in wireless networks

Besides the duty-cycle view, the entire system plays an important role in evaluating the performance of wireless distances. Here, the







Top: WLAN canals 1.6 and 11 in the frequency scan
Bottom left: WLAN channel mask for the IEEE-802.11g standard
Bottom right: Interference IEEE 802.15.4 on channel 15 and 20 parallel to IEEE 802.11g channel 6

system transit time comprises of the transmission via the air interface and the processing of data bundles by the technical system. Also, there will always be a system-related delay and a transmission that is jitter-afflicted. The jitter is caused primarily by repeated transmission through the air interface and will increase in an environment exposed to interferences. As two asynchronous processes work together, the distribution function will likely follow a Raleigh distribution.

In a system for deterministic transmission, it must be ensured that the jitter is very small compared to the system time t_{System}. In addition, the system time should be deterministic and be ignored in relation to the system dynamics.

// IEEE 802.11 – WLAN

A previous figure demonstrates clearly that only three of the thirteen WLAN channels do not overlap. Those are either channels 1, 6 and 11, 2, 7 and 12 or 3, 8 and 13. A corresponding frequency plan would ensure that the channels theoretically do not interfere. In reality, it looks slightly different. The following oscillogram of real wireless systems shows significant overlapping of the frequency bands.

In the process, 22 MHz channels (that are DSSS spread) do not interfere with each other, but the other users do. These spread channels are well isolated by digital cross-correlation and autocorrelation.

It is obvious that the 22 MHz bandwidth of a WLAN channel is pure theory. In reality, the channel mask is much wider and also depends on the type of modulation. The 802.11b channel has therefore a significantly different mask than an 802.11g channel. We can take advantage of these different properties – or alternatively, we have to live with the limitations.

// WLAN sideband behaviour

Looking at the theoretical channel use in the 2.4 GHz band, it should be possible to operate IEEE 802.15.4 channels 15 and 20 completely without breakdowns and interference to the WLAN Channel 6 (see figure page 144 centre). In a reference measurement, the round trip times were operated from a control system with a cycle time just above a millisecond over an IWLAN distance on Channel 6 in operating mode g – i.e. with OFDM modulation. The statistics show a very short response time.

The response behaviour changes significantly once IEEE 802.15.4 systems run parallel with different duty cycles. With a duty cycle of 10 %, a change is barely noticeable; however, a 50 % duty cycle clearly reveals an increase of the round-trip time. When IEEE 802.15.4 systems transmit with a maximum network load, the WLAN system can go down.

In the previous table, the course leading to a channel breakdown can be clearly seen. However, the presented results are not representative for all WLAN and all IEEE 802.15.4 systems. Generally, this approach just indicates a trend. A large network load, also in the side band, can lead to a breakdown of a connection that theoretically should not interfere. This is the bad news. The good news is that a low network load <10 % has virtually no effect at all on the transmission. Also overlapping channels with a small duty cycle will (purely statistical) not cause any dramatic mutual interference.

To conclude that a larger minimal distance between the wireless channels should be defined as a consequence would not be helpful. This is demonstrated by another extreme example. When WLAN channels 5 and 6 are operated at the same time, one might think that only inadequate communication is possible.

Active interferer	-	802.15.4	802.15.4	802.15.4
Channel	-	15	15	15
Channel load	-	Max	50 %	10 %
# sent	4500	4500	4500	4500
# error	0	0	0	0
# lost	0	4500	0	0
min [ms]	1.48	-	1.5	1.5
max [ms]	2.59	-	12.67	5.89
avg [ms]	1.57	-	2.7	1.69

Statistical analysis for different duty cycles

That is only partly true. As long as the two wireless distances have only a sufficiently small duty cycle, a drop in communication is barely noticed. But just increasing the communication load to a 50 % duty cycle (10mbps net @ 54 mbps gross) leads to doubling of the communication jitters.

These values cannot be generalised. However, they demonstrate that even with overlapping channels and moderate duty cycle, error-free communication is possible. However, if the communication load would increase further bus errors and connection breakdowns can increase by about 75 %.

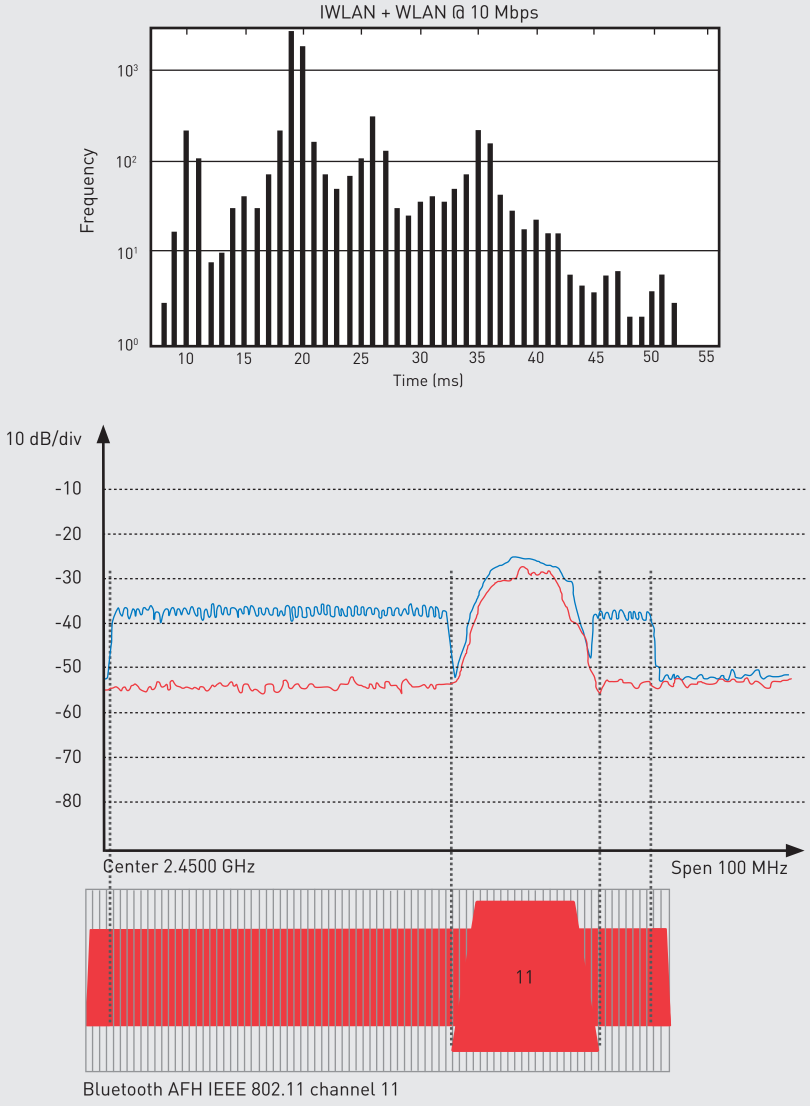
// Bluetooth® and WLAN

Bluetooth uses the modulation process FHSS (Frequency Hopping Spread Spectrum). Each Bluetooth station changes its frequency with 1600 hops per second. These frequency changes occur in a pseudo-random sequence, which is determined by the 48 bit Bluetooth address and an internal 23 bit number and has a period of almost 24 hours.

Frequency hopping avoids interferences

Due to the channel definition, 78 pico networks could co-exist without conflict within the wireless range. Networks that hop stochastically over the 78 channels are in fact interfering with each other increasingly in the utilisation of up to about 40 %, corresponding to 32 pico networks. Nevertheless, Bluetooth has a very large communication density, and due to the fast frequency changes, it has great robustness compared to the narrowband interferers.

Once the Bluetooth pico network is heavily used, it has an effect on the other wireless systems like a stochastic broadband interferer. At worst, it has an effect like a 78 MHz wide wireless system with a duty cycle of around 1.28 %. Any additional pico network increases the duty cycle also by 1.28 %. Bluetooth will prevail in this scenario, but the



Top: With channel load, the jitters and the round-trip time also rise with overlapping channels
Bottom: Bluetooth AFH allows the automatic masking of occupied channels. Here, the WLAN channel 11 is clearly hidden.

other wireless technologies will suffer interferences (see page 147).

To avoid this, AFH (Adaptive Frequency Hopping) is implemented in every Bluetooth device from version 1.2. If Bluetooth notices through an increase of package errors that a channel is permanently occupied, AFH ensures that the occupied frequency spectrum is completely hidden in the hop schema. Important is the permanent concept. If a wireless distance is only sporadically occupied, the AFH algorithm will not apply, and thus, there will be no improvement of the communication. To ensure a greater predictability, industrial Bluetooth solutions also allow blacklisting (manually hiding potentially occupied channels) or white listing (manually allowing certain frequency ranges).

Reference measurements show that there is hardly any loss of communication with parallel operation when using Bluetooth, thanks to the adaptive frequency hopping on a fully loaded WLAN channel.

// Coexistence management

The previous explanations clarify that a management of frequencies and systems is necessary especially in the free ISM bands and in the 2.4 GHz band in particular. Wireless should not be left to itself. Each access point and each wireless station should be known in its application context and surrounding.

Duty cycle as an important criterion

Another determining factor is the channel load. The previous sections demonstrated how strongly the duty cycles influence the transmission behaviour of overlapping and adjacent frequency ranges. Here, we generally determine that a channel load up to 10% can be considered non-critical. Typically, influence of the other communication party is then hardly measurable. With a duty cycle of 50%, it gets critical. Here, a significant increase of jitter time can be almost always noted. This is a sure indication of collisions, even if there is a remedy for the communication in the form of transmission repetitions. In case of bigger channel load, attention should be paid to exclusive channel allocation and side bands with low communication.

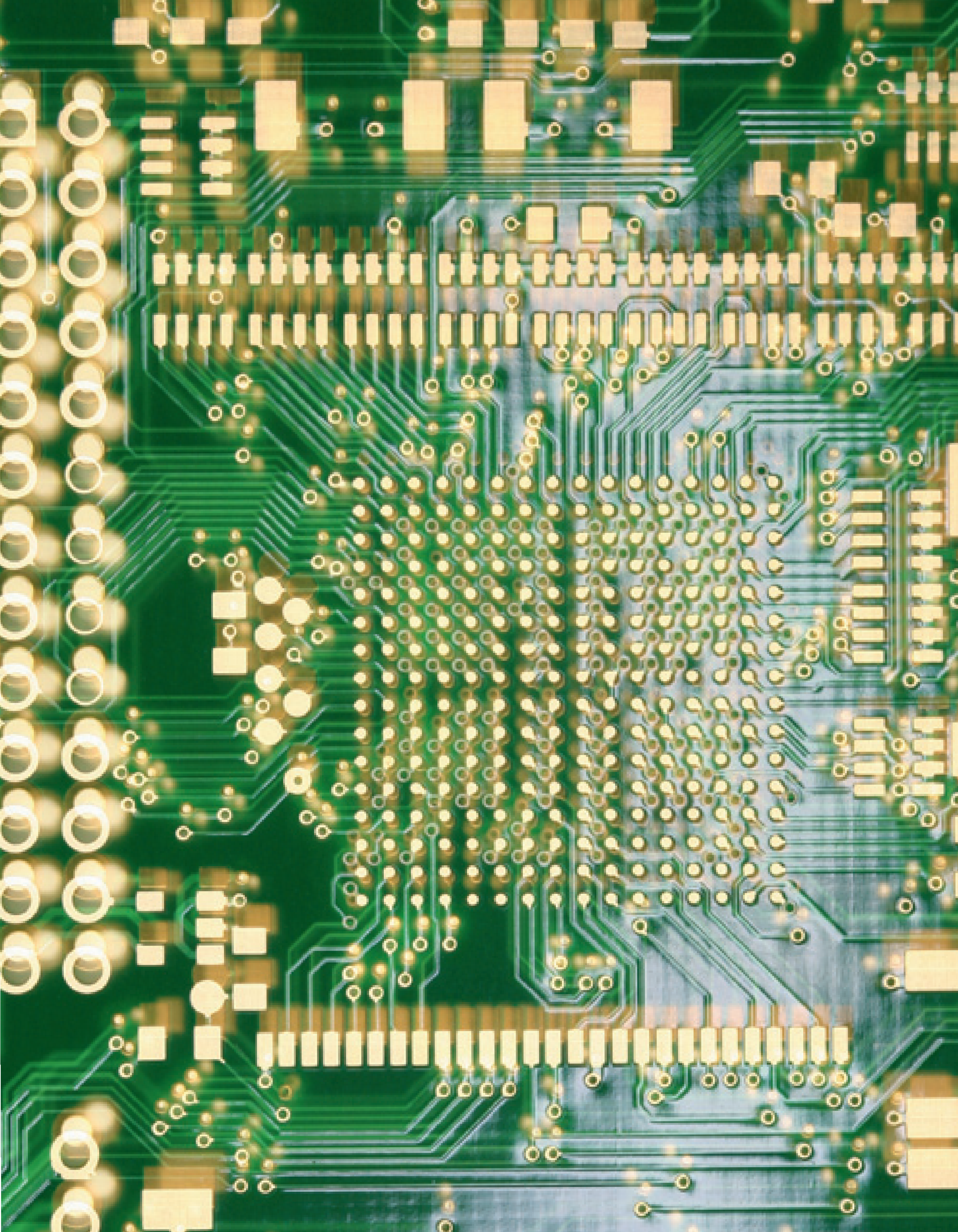
The fundamental issues of coexistence management

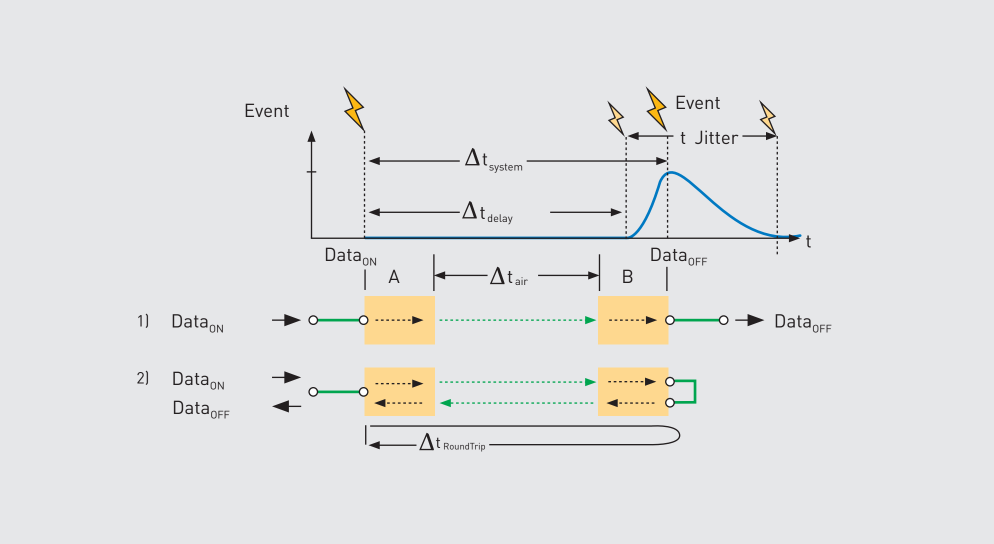
The main considerations in wireless planning are the following:

- What wireless systems exist in the area?
- Are there wireless systems that are used by other divisions and/or units?
- Are wireless consumer devices (laptop, mobile, Smartphone, headset, etc.) allowed in this area, and if yes, what services and functions are being used?
- What frequency ranges are used by the identified systems?
- Are there relevant overlaps of the various frequency ranges?
- What data is transmitted via the relevant channels and how often?
- Is it possible to determine the duty cycle for the considered systems?
- Are there minimum requirements with regard to the response time, round trip time or data rate?
- Are there jitter limits within the considered systems?

// Performance parameter

Performance parameters play an essential role in the assessment of the radio links. That is, the requirements imposed on a wireless communication system. Two essential cases of application have to be distinguished here: radio bridges and bidirectional devices.





Response times in a time-variable radio link system.

Radio bridges:

The term radio bridge refers to a system in which node A of any event, such as an input signal, receives DATA ON, which is transferred via the air range to node B and is placed there as an event on an output. Most of the time, this will be (hopefully) the unmodified data stream DATA OFF. The time it takes for receiving the result after an event is important for a real-time system. This time is also called the reaction and/or system time. It is composed of the processing times of nodes A and B as well as of the transmission time of the air interface. In the best case, the transfer is complete after a minimum delay. As explained in the previous section, with a real system the system time will be however flawed by a jitter. This is generated by the variable processing time in the communication nodes and through possible signal propagation delays and transmission repetitions on the air interface. Temporal requirements for the radio bridges are determined by the process. System times up to 100 milliseconds are for example acceptable for the controls of human-machine communication. For machine-to-machine applications, only a few seconds of system times are sufficient in many cases.

Bidirectional devices:

In bidirectional devices based on an event, node A initiates a communication to node B, which is attached via the air interface. It processes this event, calculates a result and transmits it back to the requesting node A, which makes this a response event. The complete processing time is known as round trip time. Round trip times always play a role when there is a need to communicate bidirectionally with controllers, sensors or actuators.

Besides the mentioned values, other performance parameters are also of importance. They are:

1. Period

A period is the gap between two communication events. It is a position indicating system; the period must be strictly observed. If there is a stochastic and event-driven system present, no period can be specified but just a frequency per time unit.

2. Amount of data

The amount of data describes the size of a telegram that must be transmitted when an event occurs. In an automation system, these can be a few bits or bytes. Also large data bundles from several 100 to some 1000 bytes are possible.

3. Data rate

Each transmission distance has a specific data rate. A high data rate provides short transmission times for small amounts of data. Small data rates reach long distances.

With the described data rates, the performance parameters of a channel can be described really well. We can immediately recognise the potential for optimisation. If small amounts of data are to be transmitted and the frequency band is already well occupied, it will be helpful to use a very fast transmission technique. In that way, the transmission time on the air interface can be clearly reduced and the duty cycle is kept small.

// Coexistence planning

The previous explanations demonstrate clearly that the use of wireless should not be left to chance. A comprehensive knowledge of processes, system dynamics, wireless infrastructure and used frequency ranges are definitely required. Different applications must be also taken into consideration. A generally applicable approach to achieve full coexistence of all systems is not realistic. As a rule, a compromise that describes an acceptable solution must be found in the technology and frequency selection. A differentiated plan of measures is helpful. It must be created on the basis of the collected performance parameters, system requirements and framework conditions, which also includes a prioritisation of the objectives to be achieved. Some other aspects of coexistence can be found in the VDI Directive 2185 / page 2.

// Safety

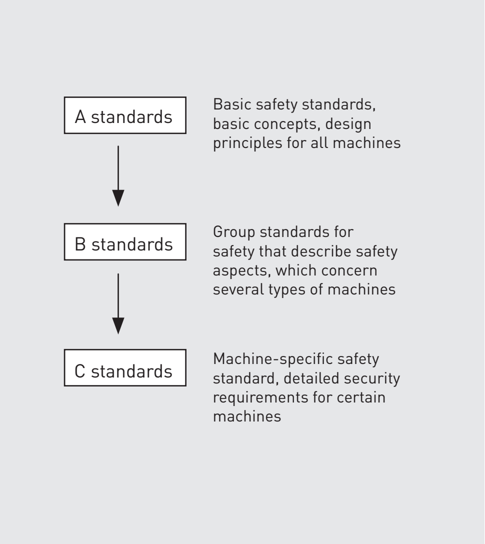
The safety of technical systems is essential for their operation. Generally, the definition of safety is the absence of danger. However, the hazard potential can come from various directions.

Safety and security - a complex topic

The term »safety« describes the safety of the person operating technical equipment or the machine. Life and health of the person must be protected. As a rule, a safety system has to make sure that the equipment or machine is safe, so that there is no danger to health and life. For this purpose, the term »functional safety« was established.

In contrast, if a machine, device or a building needs to be protected, the activity is called »security«. We are talking about measures that protect technical systems against attack by third parties. Most of the time, it is all about data security: attackers from the outside should have no influence on the data, or control the machines.

This means that safety and security are very complex issues. Rules and legal regulations of many different disciplines must be observed in order to provide a secure system. Safety and security are both about the identification of hazard potential and measures taken in order to keep the system in a secure state. We must be aware that maximum security can only be achieved with maximum effort. This requires differentiation of safety and security levels. In the following sections, the aspects safety and security will be discussed separately.


A standards
Basic safety standards, basic concepts, design principles for all machines
B standards
Group standards for safety that describe safety aspects, which concern several types of machines
C standards
Machine-specific safety standard, detailed security requirements for certain machines

Safety standards are hierarchically structured

Functional safety

Programmable electronic safety systems (PES, safety controllers) are now widely used. Initially, they were just electro-mechanical solutions for machine safety. Interesting software-based and fieldbus-based safety concepts were already developed in the 1980s. Today, we find a variety of different system solutions for safety problems. They are used in control and automation technology and also in drive, automotive and medical technologies. Thus, the trend towards configurable safety systems on an electronic basis is obvious.

Machinery directive 2006/42/EC

Another important step towards Europe-wide harmonisation of safety policy took place with introduction of the new Machinery Directive (MD) on Safety of December 29, 2009. In addition, several new standards for functional safety were established at the same time with the MD. Not just the deterministic consideration of failure scenarios are in the foreground, but also a comprehensive assessment of the likelihood of (probabilistic) error scenarios.

Modifications to the previous version of MD 98/37/EC are not dramatic but there are significant extensions in some essential points. For example, provisions were amended for assemblies and partial devices that perform no specific function by themselves and which are only complete after installation and addition of protective equipment.

Now, the directive applies also for safety components, such as light barriers, safety mats, automatically moving safety guards and interchangeable components, which alter the function of a machine (interchangeable tools, attachment parts). CE labelling applies to all these systems in accordance with the Machinery Directive (2006/42/EC), which requires a risk assessment.

Hierarchical structure of the standards

For the implementation of the directives, the applicable harmonised standards are published in the Official Gazette of the European Union. These alone are decisive for the implementation. Once a directive is replaced, as the directive 98/37/EC at the end of 2009, only the new harmonised standards are relevant for further use. This means more than 600 existing standards need to be partly revised. The standards are adapted to international circumstances and are standardised worldwide. A good example of this is EN 292 »Safety of Machinery, Basic Concepts, General Principles for Design«, which is replaced by ISO Standard EN ISO 12100-1. The standards are hierarchically divided into A, B and C standards. The A standards, also known as basic safety standards, deal with the basic terms and general guidelines, design principles and general aspects that apply to all machines. They are binding and form the basis for all further standardisation.

EN standard	Title
EN ISO 12100-1: 2003/ A1: 2009	Safety of Machinery - Basic concepts, General Design Principles - part 1: Fundamental Terminology, Methodology
EN ISO 12100- 2: 2003/ A1: 2009	Safety of Machinery - Basic concepts, General Design Principles - part 2: Technical Principles
EN ISO 14121- 1: 2007	Safety of Machines – Risk evaluation – part 1: Principles

Basic safety standards

Among the B standards are the safety groups standards. They deal with the application-independent standards that apply to different machine types and are aimed at the developers and manufacturers of safety-related components. There is a distinction between B1 and B2 standards. B1 standards have to do with special safety aspects, such as safety distances, temperatures and noise levels. B2 standards deal with safety-related processes and equipment, such as shut down, two-hand controls and interlocking devices.

EN standards:
Some safety standards groups

GSA 1.2.1 Safety and Reliability of Control Systems

EN ISO 13849-1: 2008 Safety of Machinery – Safety-related Parts of Control Systems, Part 1: General Design Principles

EN ISO 13849-1: 2008/AC: 2009 Safety of Machinery - Safety-related Parts of Control Systems, Part 1: General Principles for Design

EN ISO 13849-2: 2008 Safety of Machinery – Safety-related Parts of Control Systems,
Part 2: Validation

EN 62061: 2005 Safety of Machinery – Functional Safety of Safety-related Electrical, Electronic and Programmable Electronic Control Systems

GSA 1.2.3 Start-up

EN 1037: 1995 + A1:2008 Safety of Machinery – Prevention of Unexpected Start

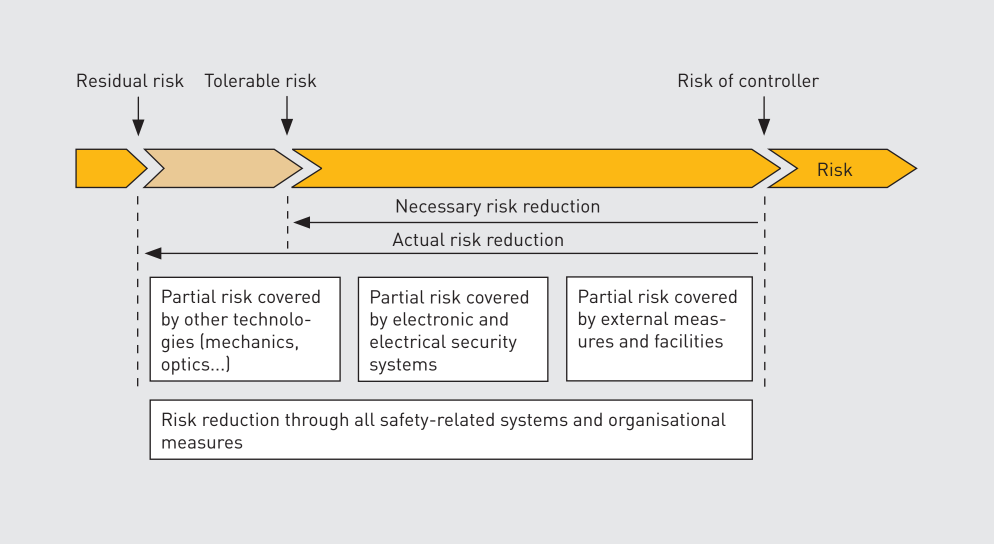
GSA 1.2.4.3 Shut-down in an Emergency

EN ISO 13850: 2008 Safety of Machinery – Emergency Stop – General Design Principles

GSA 1.4.2 Special Requirement of Safety Guards

EN 953: 1997 + A1: 2009 Safety of Machinery – Safety Guards – General
Requirements for Design and Construction of Fixed and Movable Separating Safety Guards

GSA 1.5.1 Electrical Energy Supply


Residual risk
Tolerable risk
Risk of controller
Risk
Necessary risk reduction
Actual risk reduction
Partial risk covered by other technologies (mechanics, optics...)
Partial risk covered by electronic and electrical security systems
Partial risk covered by external measures and facilities
Risk reduction through all safety-related systems and organisational measures

Principles of risk mitigation in accordance with IEC 61 508

EN 60204-1: 2006/A1:2009 Safety of Machinery – Electrical Equipment of Machines – Part 1: General requirements

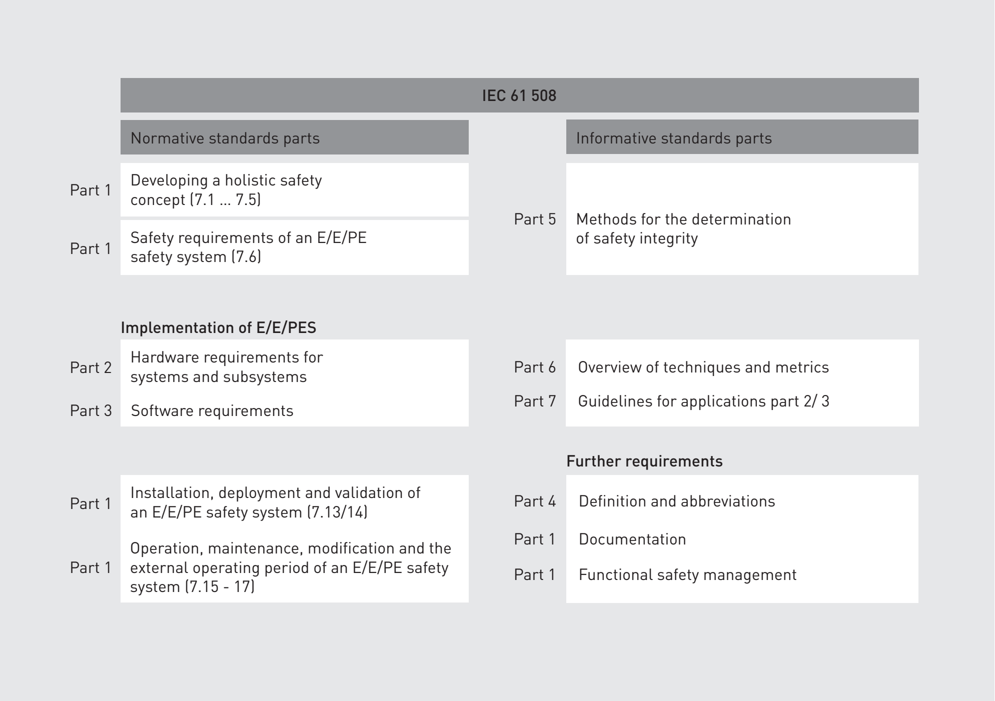
EN 60204-11: 2000 Safety of Machinery – Electrical Equipment for Machinery – Part 11: Requirements for High-voltage Equipment for Voltage Above 1000 VAC or 1500 VDC but Not Over 36 kV

C standards are specialised safety standards that are valid only in specific application areas and/or with specific classification of machinery. They are divided into the so-called subject areas (ICS). The relevant standards are only valid in the subject area and accordingly highly specialised. In case there is no C standard for a type of machinery, the A and B standards must be applied to prove compliance with the directive.

The levels of machine safety

Since the introduction of the machinery directive, only machines with risk assessment can be sold in Europe. In the process, the IEC 61508 »Functional Safety of Electrical/Electronic/Programmable Devices« plays an important role. This standard calls for a quantitative proof of safety.

Institutions or individuals who market machines and equipment must conduct risk assessments and have to keep such proof for 10 years. The directive is through implemented self-certification, and the EU relies on manufacturers' personal responsibility. If there is no qualified risk assessment, it is a violation of applicable law. In case of an accident, there will be an investigation whether the security concept complied with the state of the art or whether the accident was based on a design error, or if an inadequate warning had occurred. If negligence can be proven, it generally leads to criminal prosecution.



IEC 61 508

Normative standards parts

Part 1 Developing a holistic safety concept (7.1 ... 7.5)

Part 1 Safety requirements of an E/E/PE safety system (7.6)

Implementation of E/E/PES

Part 2 Hardware requirements for systems and subsystems

Part 3 Software requirements

Part 1 Installation, deployment and validation of an E/E/PE safety system (7.13/14)

Part 1 Operation, maintenance, modification and the external operating period of an E/E/PE safety system (7.15 - 17)

Informative standards parts

Part 5 Methods for the determination of safety integrity

Part 6 Overview of techniques and metrics

Part 7 Guidelines for applications part 2/ 3

Further requirements

Part 4 Definition and abbreviations

Part 1 Documentation

Part 1 Functional safety management

Contents of IEC 61508

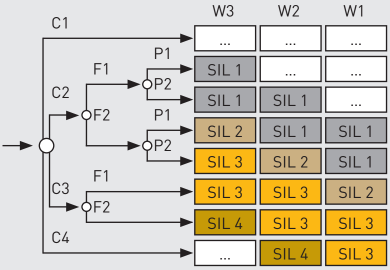
Operational steps that need to be applied in conformity with the standards are:

1. Identification of the system's possible hazard potential and assessment of the potential risk. (IEC 61508)
2. Risk elimination and reduction by design changes; application of measures during the development / construction of the machinery according to EN ISO 12100-2, clause 4
3. Reduction of risk through protection measures in accordance with EN ISO 12100-2, clause 5
4. Reduction of risk by means of training/ warnings; description of residual risk by means of training and displaying warnings on machinery in accordance with EN ISO 12100-2, clause 6.

// Functional safety

IEC 61508 applies to all applications that use electrical or electronic equipment and assemblies to perform safety functions. The standard calls for quantitative proof of the remaining risks based on error probabilities. All calculations are to be evaluated for a complete safety chain from sensor through control system up to the actuator. Therefore, the probability of failure (PFD – Probability of Failure on Demand) is a measure of evaluating the safety.

IEC 61508 is a basic standard that can be directly applied. It contains all key aspects and implementing regulation for the execution and operation of secure electrical / electronic / programmable electronic devices. Devices and equipment are considered »safe« if they do not cause any



W3	W2	W1
...	...	...
SIL 1	...	...
SIL 1	SIL 1	...
SIL 2	SIL 1	SIL 1
SIL 3	SIL 2	SIL 1
SIL 3	SIL 3	SIL 2
SIL 4	SIL 3	SIL 3
...	SIL 4	SIL 3

	Extent of damage
C1	Small, slight injury to person Smaller harmful acts against the environment
C2	Severe, irreversible injury of one or more persons, death of a person or temporary major harmful environmental influences
C3	Death of several persons or continuation or larger harmful environmental influences, e.g. after an accident regulation
C4	Catastrophic impact, many people dead
	Length of stay
F1	Rarely till often
F2	Often till always
	Averting of danger
P1	Possible under certain conditions
P2	Hardly possible
	Probability of occurrence
W1	Very low
W2	Low
W3	Relatively high

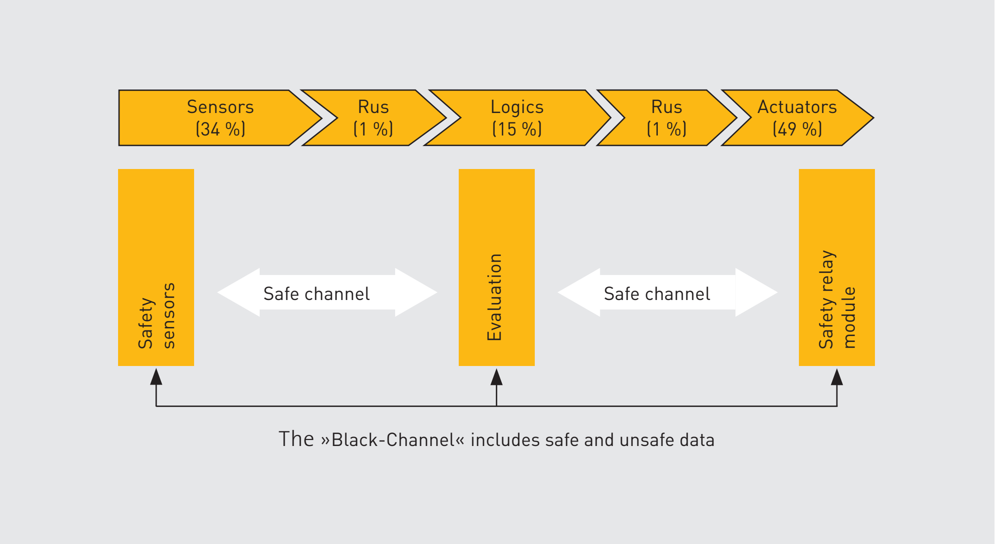
Risk graph according to IEC 61508 for the assessment of safety levels

unacceptable injury risk to a person or damage to the environment. Functional safety refers to the function of safety control itself and to the likelihood of the safety element functioning in case of a failure. This standard describes four Safety Integrity Levels (SIL 1 to 4). SIL 4 provides the highest protection.

In compliance with the relevant standard, safe systems can be created. It is the decision of the manufacturers and users as to what solution they want to use. It is assumed that currently about 10% of all sensor and actuator components in the manufacturing technology are safety-related and a significant increase is expected in the next few years. User organisations and leading manufacturers react accordingly and offer appropriate solutions for the leading field buses, such as Profibus, Interbus, CAN, and AS interface – however, currently only for wired systems.

The trend: integrated safety concepts

Traditionally, safety solutions have been implemented on the basis of standard controls and electro-mechanical safety control devices. These allow the use of proven technology and low-cost components. But they require also a complicated circuitry technology and high (error-prone) wiring costs. Additionally, the emergency stop



A black channel transmits safe and unsafe data via a medium

switch, sirens and light grid are connected with the control computer. This allows feedback to the user on the side of the operating elements and the visualisation.

Current machinery solutions have a clear trend towards integrated safety concepts. Hardware components become more expensive. This is contrary to the advantage of perfect integration through the fieldbus systems and engineering tools of their respective manufacturers. Safety connections can be planned and parametrised with appropriate tools and an automatic validation is also possible. Based on higher communication performance and complete networks, an integrated solution can implement a response time much faster and handle shut downs more intelligently than conventional systems. Integrated solutions also enable the functional safety required by IEC 61508 over the entire life cycle of a machine or plant.

Safe communication

Integrated safety solutions require communication systems that ensure a consistent, transparent and accurate communication. Initially, it was assumed that this is only possible with special safety buses. However, safety-related communication is now built on proven standard field buses. Safety protocols on higher application levels allow a logical dual-channel nature and the protection of communication through software measures alone. The goal is the establishment of availability, which meets the requirements of the relevant technical standard. Ultimately, the standard makes sure that all errors are detected. Measures to take may depend on the bus system:

- Sequential numbering of the safety telegrams
- Time expectations with acknowledgement of the telegram
- Identification of transmitter and receiver to ensure the authenticity of the telegram
- Additional backup (CRC Cyclic Redundancy Check).

Performance Level PL	Average probability of risk-bearing downtime within an hour	SIL EN 61 508-A
a	> 10^{-3} … <10^{-4}	No special requirements
b	> = 3 x 10^{-6} … 10^{-5}	SIL 1
c	> = 10^{-6} … 3 x 10^{-6}	SIL 1
d	> = 10^{-7} … 10^{-6}	SIL 2
e	>= 10^{-8} … 10^{-7}	SIL 3

Functional safety required by IEC 61508

According to IEC 61508, the influence of security system components is divided into 35 % sensors, 15 % control logics and 50 % actuators. In this approach, communication is not taken into account. To be fair to distributed systems with bus communication, one percent of the influence factor is accredited to the bus system. Considering the error influence according to the performance level of a secure system, a performance level e may have 10^{-9}/h hazardous breakdowns within one hour as a maximum.

As far as the transmission distance is concerned, it means a reliable error probability that is better than 10^{-9}/h with SIL3. It implies that it is equivalent to a data transport for over 100,000 years without undetected errors. This also defines the requirement of a safe (within the meaning of the directive) wireless standard.

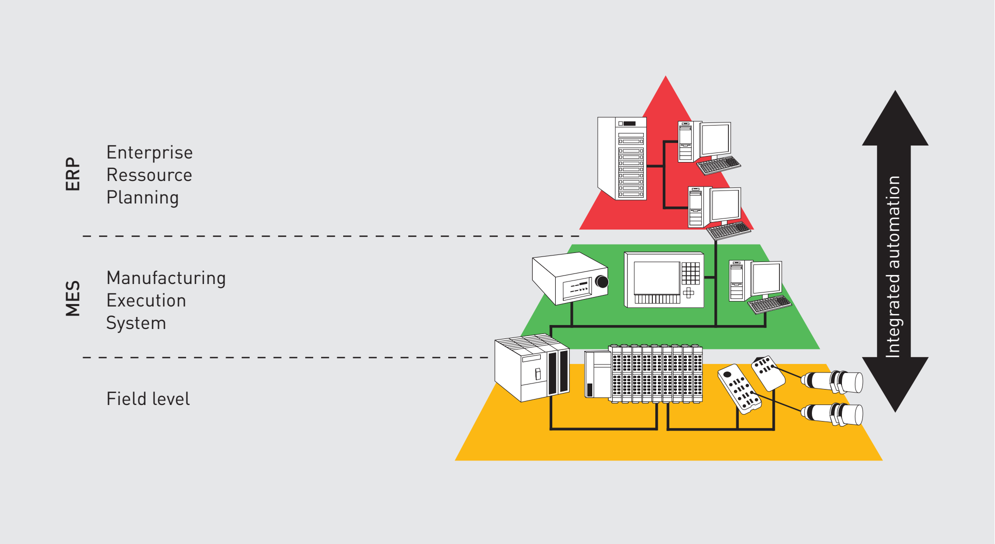
Requirements for »Wireless Safety«

The keyword here is »unknown error«. Secure systems require no error-free communication, only error detection. This opens up the possibility to use virtually any bus system as a black channel. Today, all industrial bus systems such as AS interface, Profinet, Profibus Ethercat, Sercos III and Ethernet IP have the possibility to transmit safe telegrams. The safety aspect also needs to be applied to the transmission distance. As no specific requirements are to be met, only the reliability of the bus system has to be verified. This is implemented in many solutions, and it allows the transmission of secure data via Bluetooth, WLAN or any proprietary wireless systems without further problems. Initial approaches and even real cases exist already. Several manufacturers introduce safety-oriented wireless standards for functional safety. Control of lorry loading cranes and industrial cranes is a field in which the safety-oriented transmission path has already prevailed on a broad level.

EMC Directive and ATE X Directive

Other safety-related requirements are described in the EMC Directive (2004/108/EC). Its goal is the avoidance of an electromagnetic disturbance through a (new) equipment. It dictates how the electromagnetic compatibility of electrical equipment in the EU Member States should be designed. This directive requires the setting up of an adequate level. Details are deliberately left undefined.

The Ordinance on Explosion Protection (11th GPSGV) transforms the European ATEX Directive 94/9/EC into federal law for Germany. The purpose of the directive is the protection of people who work in potentially explosive areas. Appendix II contains the basic health and safety



The desire of vertical integration comes with new security challenges

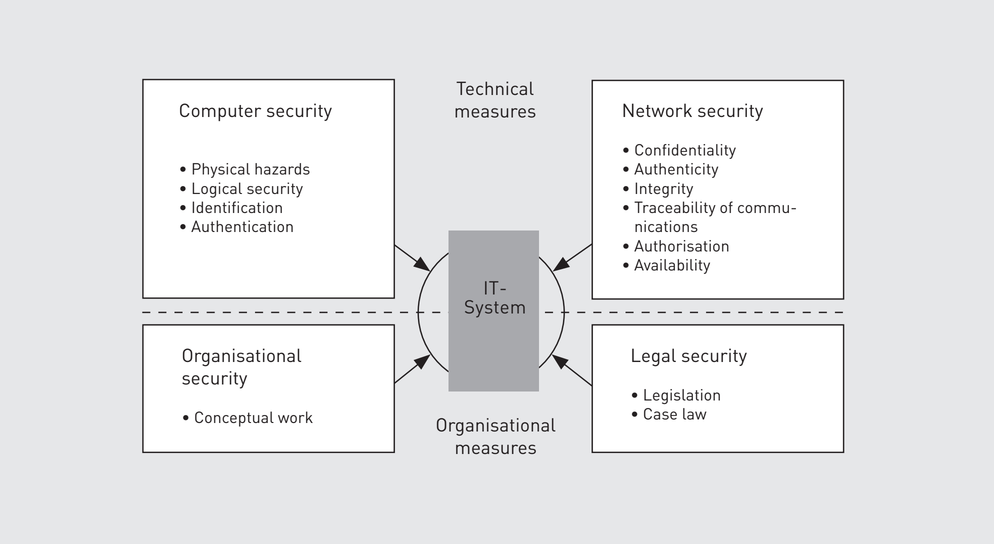
requirements that the manufacturer has to observe. For more information, see the book, »Explosive topics« by steute.

// Information security

Information security within automation technology is a fairly new topic. In autonomously working, conventional industrial field buses, unprotected data exchange was previously completely unproblematic. But vertical integration, e.g. continuous data exchange on a company level, makes new protection requirements necessary.

Overall, we see three trends, which also quickly demonstrate the importance of data information security in industrial plants.

1. Remote maintenance and remote monitoring of systems and equipment impose new demands on the communication: with Internet connection, devices are exposed to comparable risk. Virtually the same requirements apply as those that apply to the public Internet.
2. The second determinant is the use of Ethernet technology in automation, which can be compared with common Internet technology. At least as far as asynchronous communication is concerned, the TCP/IP protocol family is used above the Ethernet, which makes it not just easier to find and manage devices, but it also offers a comfortable base for parameter setting and documenting of devices. Simple web servers, even miniature controls and embedded devices, are today state of the art. Use of Internet technologies in an internal context of a company (so-called extranets) lead to security problems into the company. If a facility is integrated from an executive level up to the control, comparable security requirements (as in the Internet) must be implemented.
3. Wireless technology is increasingly used in industrial systems. System developers often operated from the viewpoint that what cannot be seen is not dangerous. Thousands of attacked WLAN networks prove that these



Information security is achieved through technical and organisational measures

networks in various industrial sectors need to be protected. In the field of automation technology or technical systems, the attack is even more tragic if technical processes or functions are corrupted by outside devices. Sensitive data is possibly exchanged via the air interface, which are worthy of protection. Here, there is only one solution: comprehensive state-of-the-art protection.

These three trends explain that IT has entered the area of automation technology. Even if this requires radical rethinking, information security is at the very top of the agenda in terms of company-wide use of Ethernet-based or wireless-based systems.

Information must be protected

Information security means generally the protection of any kind of information and its origin. The protection of data in people's minds and on paper has to be solved organisationally. IT information security takes care of the electronic data protection in computer systems. In this scope of application, computer systems can be very broadly defined and range from mainframe to desktop computers, mobile phones, industrial controllers and intelligent sensors in automation technology.

Basic needs: CIA

The classic basic needs of information security are summarised in the abbreviation CIA: Confidentiality, Integrity and Authenticity. Data must be treated confidentially. Data may only reach a person who is truly authorised to use it. The data must be reliable. In other words, it has to be received by the customer exactly as it was sent. Unauthorised modification must be prevented in any case, or should be at least traced. Authenticity means that the user of the data is really the person who he says he is.

We must bear in mind that there are not only intentional threats, such as computer viruses, Trojans, interception of data lines or theft of hard drives or entire computers. The database is also threatened by force majeure, such as fire, lightning, water, electricity surges, erroneous updates and human error.

Various measures address this security problem. Network security plays a major role in wireless systems. Besides confidentiality and integrity, communication should be verifiable and be used only by authorised systems. Another important aspect in wireless systems is the availability.

Usually, the technical measures taken are to be supported by organisational measures. This includes first of all the organisational security (worked out in the form of rules and directives), how IT security is implemented and ultimately, also the legal security. Minimum requirements for IT systems have also been defined in different standards, which are comparable to the safety standards. And finally, there are also infrastructural security measures that have to be observed. They prevent the physical access of unauthorised persons to the IT device with the intention of stealing or manipulating. Structural safeguards and access controls are part of these measures.

Norms and standards

Just as with the technical requirements, the standardisation work in the area of IT security is also in full progress. The standards grow steadily because of the higher risk potential. At this point, it is not possible to refer to all the rules and directives. Standard series ISO 27000 ff dictates the relevant standards for the information security of a management system.

ISO 27000: This standard provides a general overview of information security management systems. It describes the relationship of the various standards in the 2700x family as well as the basic principles, concepts and terms for information security management systems.
ISO 27001: ISO 27001 standard »Information Technology – Security Techniques – Information Security Management System Requirements Specification« describes general recommendations for the introduction and operation of appropriate management systems.

ISO 27002: »Information Technology – Code of Practice for Information Security Management« is a framework that describes how a functioning safety management can be built and how it should be anchored in the organisation. The recommendations are rather intended for the management. They are not precise technical references.

ISO 27005: »Information Security Risk Management« contains standard guidelines for risk management. This directive replaces ISO 13335-2.

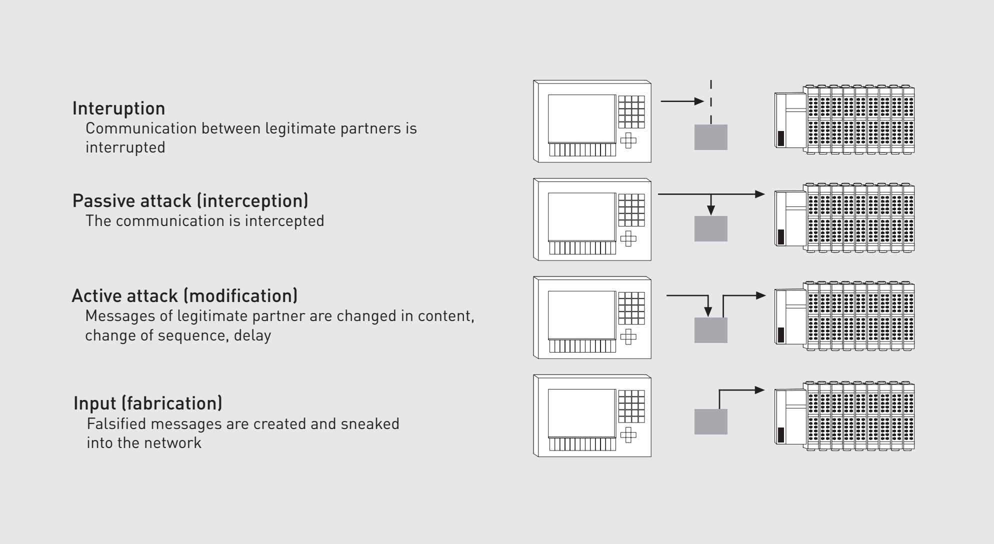
ISO 27006: Requirements »Information Technology – Security Techniques – Requirements for the Accreditation of Bodies Providing Certification of Information Security Management Systems« specifies the accreditation and certification bodies.

Other standards: ISO 2700X is growing steadily and will be developed into a long term standard. The certifiability will cause awareness just like the ISO 9000, and it will become a quality feature for an organisation.

Contrary to the technical standards in the area of safety, the security standards situation is essentially related to the organisation. Security is a holistic challenge. It cannot be realised simply with a technical system.

Attack strategies

In addition to the threats posed by own employees (e.g. indiscretions), technical possibilities



Interuption
Communication between legitimate partners is interrupted

Passive attack (interception)
The communication is intercepted

Active attack (modification)
Messages of legitimate partner are changed in content, change of sequence, delay

Input (fabrication)
Falsified messages are created and sneaked into the network

The four basic principles of attacks on IT networks

of a potential attack plays a major role. Here, two opposing trends are notable. On one hand, the complexity of the tools to spy out technical information increases. On the other hand, the attacker needs less and less technical understanding because there are lots of software and tools available on the Internet. This availability enables the attacker to prepare and carry out a highly effective attack, without the attacker requiring much know-how. These trends lead to a rising risk potential and certain vulnerability, if we do not intensively deal with the subject of information security.

In principle, we distinguish four basic types of attack on technical systems.

Interruption

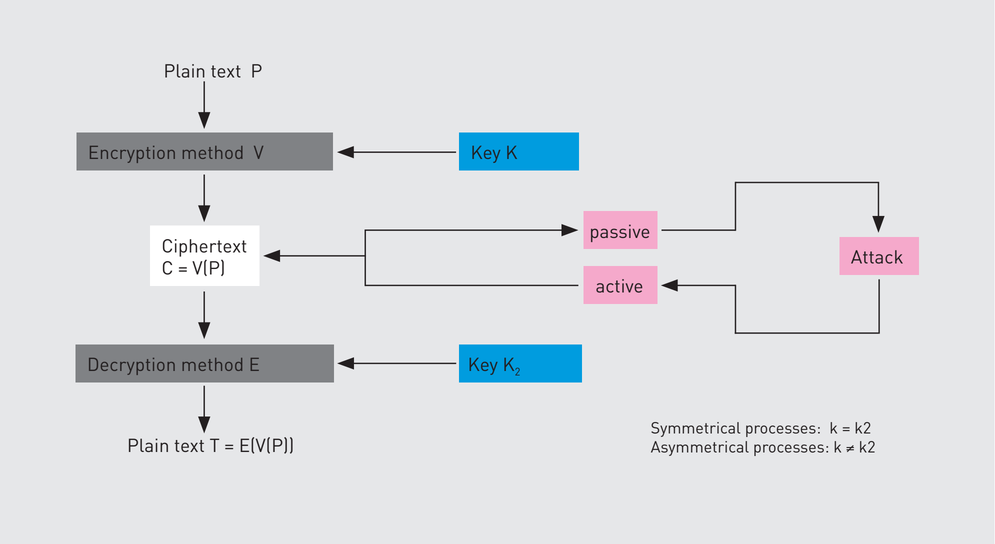
The communication between two or more legitimate communication partners is interrupted. In case of wired communication, the interruption can be achieved by the capping, for example, of a connection. In case of wireless systems, the disturbance of the wireless channel through jammers is frequent enough.

Interception (passive attack)

Eavesdropping on the channel where the operator often does not notice it. The attacker is just interested in the transmitted data. A passive attack on wireless systems can be conducted completely unnoticed, because it is enough to position a receiver in wireless range. Only encrypted data transfer can guarantee sufficient protection.

Modification

In this active attack, the message content is modified by an attacker who interferes without disclosing his own identity. A typical representative of this type is the »Man in the middle attack«. Security can only be achieved through well-functioning encryption and additionally through communication- and telegram-related security mechanisms.



Different attack methods and encryption processes

Fabrication

If an attacker has sufficient information about a communication through a passive attack, he can, in a second step, supply the target system with incorrect information by sending messages. In doing so, the attacker usually uses the identity of a legitimate partner. Here, only an authenticated and validated communication sequence can provide security.

Encryption - Cryptography

In a technical system, confidentiality requirements can be implemented only by encryption. Various methods with distributed or common keys are state of the art. What is used depends on the vulnerability of data. Purely technically, the procedures are basically similar. Original data in plain text P is encrypted using an encryption method V, encrypted with a key k. The longer and more complex the key, the higher the security, because a potential attacker must put more computing power into decoding the key. 128 bit keys are state of the art today. Hopefully, the transmitted ciphertext C is so well protected that the attacker cannot decrypt the data. The receiver on the other side knows the decryption method E, which is an inverse function of V. With the knowledge of a key, the original plain text can be restored.

Authentication is the process in which a participant proves his authenticity. He has to authenticate himself to protect sensitive areas against abuse, deception, phishing and intrusion.

Processes as Challenge Response Authentication and Zero Knowledge are being used to exclude the risk of disclosure of the authenticated subject‘s own identity. With these methods, the authenticated subject himself no longer transmits the identification data but transmits only an evidence that he undoubtedly has these identification data. In addition, it is always a good idea to encrypt the authentication.

Security in wireless communication

Wireless communications systems offer a potential for attacks. Attackers can communicate with jammers and interrupt the communication or quietly listen in on the air interface. Secure wireless communications is hardly possible without appropriate measures. Technologies already include most diverse techniques to make attacking difficult for potential attackers.

Which technology is used for the respective wireless technology often depends on the type of application. Measures for IT security are in many cases scalable, and the operator can decide on the level of selected security. However, we must take into account that each of the processors needs processing time for the selected security mechanisms, in other words for the calculation of the keys. This is an essential consideration particularly in embedded applications. Furthermore, a secure question and response communication requires additional entries and mechanisms. Devices need to be configured, keys are to be exchanged, and the behaviour after a power failure or system change (due to a service event) must be investigated. Sometimes, this complicates the use of security techniques.

In order to make wireless safe, the following basic principles have proved their worth.

1. Frequency change methods allow suppression of narrow-band interferers. Wireless systems change the frequency after every data packet. The organisational effort is often implemented in the link manager of each technology.
2. Configuration of devices. Based on the applied hardware, only trusted devices should communicate with each other. For this purpose, a single binding process is necessary. It connects the sensors, actuators and other devices to a controller or an access point. Various mechanisms, e.g. exchange of MAC addresses, individual protected exchange of connection codes and other mechanisms are possible.
3. Encryption is a must. Only encrypted data is protected against eavesdropping. The effort is paying off with a higher security. 128-bit key and AES procedures are standard today and offer adequate protection. Other measures are possible too, which lead to an increase in security. Finally, some characteristics of proven wireless technologies should be discussed.

Bluetooth

It provides all mechanisms for secure data transmission. In terms of hardware, each device has a unique Bluetooth MAC address. Based on this and a 32-bit counter, a hopping sequence for the frequency hopping method (FHSS frequency hopping spread spectrum) is individually negotiated for each pico net. With that, each pico net is hopping in its own individual frequency within the entire allowable frequency band so that interfering narrow-band sources can be very well suppressed. The authentication uses a challenge-response process. The transmitter sends a 128-bit key. This key is processed by the receiver with a link key and his 48-bit address and sent back as result. A symmetric key with 8-128 bit and the AES algorithm is used for encryption. Security according to the state of the art is thereby fully provided - if not all security options were disabled by default. The user is also responsible for proper security.

WLAN IEEE 802.11

Used by default in office networks. Even in automation networks of today, WLAN is state of the art. Various security mechanisms are also implemented here, which can be used optionally. WLAN offers no frequency hopping procedure but uses previously configured channels. However, the established DSSS and OFDM modulation processes inherently protect well against narrow-band interferers. Various methods are implemented during the encryption process. This has historical reasons and ensures backwards compatibility with newer devices – however, this takes place at the expense of security. WEP (Wired Equivalent Privacy) can be hacked within a few minutes using the correct Internet tools. The WPA (WLAN protected access) offers a slightly higher level of security. RC4 uses the same encryption algorithm as WEP, but is changed with each 10 kbyte of data. This change of key is also referred to as TKIP (Temporal Key Integrity).

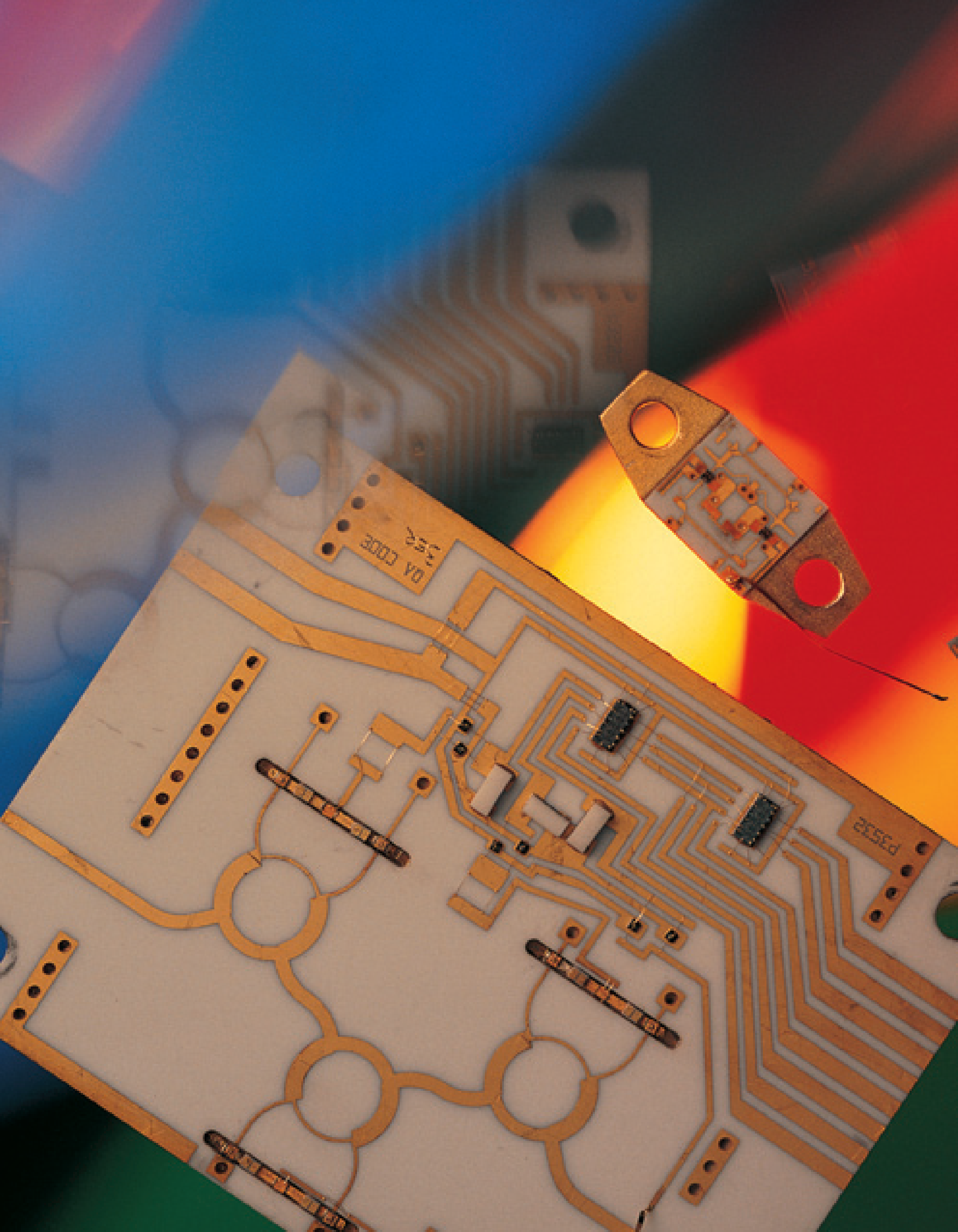
WPA2 is the state of the art in current wireless products of today. It uses the very secure AES algorithm. AES implementation in WPA2 is also known as CCMP (Counter Mode with Cipher Block Chaining Message Authentication Code Protocol). The name does not necessarily contribute to a better understanding.

IEEE 802.15.4

This is based on the ISO OSI layers 1 and 2 of different wireless networks, such as ZigBee or Wireless HART. Proprietary solutions are also possible with the chips that are currently on the market. IEEE 802.15.4 also provides security mechanisms on the MAC layer by means of a message integrity check and symmetric encryption. The keys are separately determined through the higher protocol layers for each communication party and managed by the MAC layer. In principle, various encryption algorithms such as CCM and AES are intended. The type of encryption used depends on the implementation. Support for the implementation sometimes takes place in the chipset. In other cases this is done by the host processor.

Overall, we observe that all the latest wireless technologies now offer a secure transmission. Typically, security is not preconfigured but can be selected as an option.

The level of security that can be reached depends ultimately on the security levels available in the application and the user.



QA CODE
P3532

// Antenna technology

010

Historic information

Antennas are necessary to transmit (emit) as well as to receive electromagnetic waves. A transmitting antenna converts electromagnetic waves into free space waves, and a receiving antenna converts the space waves back into electrical waves. Antennas have been used since the beginning of wireless technology: Heinrich Hertz used stretched wire as an antenna, which became known later as the so-called Hertz dipole.

By the way, the term antenna derives from the Italian term »tent pole« or »sailing pole«, because in 1897 Guglielmo Marconi, a pioneer of wireless technology used a wooden tent pole during his first attempts in constructing an antenna. In the early days of wireless technology, a simple dipole was first used as receiver and the loop antenna came later. After the First World War, people also used antenna arrays, horn radiators and parabolic antennas.

How does an antenna work?

In principle, an antenna is nothing other than an open electric oscillating circuit, in which alternating voltages and/or alternating current oscillations are induced. In other words, a simple dipole antenna generates and/or receives electric fields and also vertical standing magnetic fields. Once the circuit goes into electrical resonance, the capacitor and the coil form an oscillating circuit that sends out closed electric field lines.

The magnetic fields that run vertically to the electric field form closed circles around the conductor. In the near field, the field strength decreases proportionally to the cube of distance r. In the far field, it is proportionally reduced only by 1/r. The power density of the radiated electromagnetic waves is proportional to the product of electric and magnetic field strength and is therefore reduced by 1/r (inverse square law).

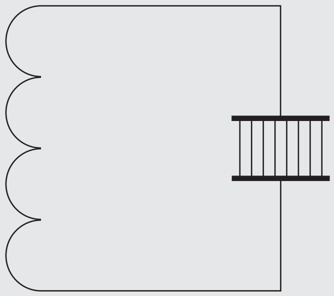
Transmitting antennas must be installed high up.

Important factor: the polarisation

Polarisation is an important factor when evaluating antennas. This refers to the orientation of the emitted electric field lines when compared to the earth‘s surface. A linear horizontal or linear vertical polarisation occurs when the orientation of the electric field remains the same over time. If the strength of the electric field is the same over time but the direction of field lines are time-dependent, we speak of circular and/or right-handed and left-handed polarisation. This classification is important for the operation of the antenna; if receiving and transmitting antenna do not have the same type of polarisation and direction, the signal transmission is heavily attenuated.

Resonance and resonance frequency

An even more important factor is the response. It is the tuning to the wave length that the antenna should receive or transmit. The resonance frequency is determined by the antenna‘s dimensions. Through special constructive measures,



In the end, a dipole antenna is nothing more than an open oscillating circuit.

$$Z_{w0} = \sqrt{\frac{\mu_0}{e_0}} \approx 120\pi\ \Omega \approx 376{,}73\Omega$$

The characteristic impedance of free space

one can reach the base point resistance (impedance) over a large frequency range, which remains relatively the same, and thus is able to transmit a wide frequency band. Such antennas have the name broadband antennas.

Brief description of additional parameters

The parameters that also have impact on the efficiency and functionality of an antenna and/or which result in its efficiency include:

- Absorption area
- Antenna gain
- Aperture
- Equivalent isotropically radiated power
- Bandwidth
- Impedance
- Side-lobe rejection
- Directivity
- Radiation resistance
- Front-to-back-ratio
- Front-to-sides-ratio
- Efficiency.

Absorption area (active area): effective antenna surface

Antenna gain: ratio of radiant intensity emitted in the main direction related to a loss-free antenna of same feed power with an antenna gain of zero (in dBd and dBi)

Aperture: free opening or their diameters, which are emitted or received through the radio waves

Equivalent isotropically emitted power: EIRP; product of the power feed to a transmitting antenna and the antenna gain (related to an isotropic radiator)

Bandwidth: frequency range, in which the impedance changes only slightly

Impedance: resistance at the connecting terminals with the used frequency (ohms)

Side lobe rejection: ratio of the main lobe gain in dB to the level of the highest side-lobe.

Directivity: directive efficiency of an antenna with preferred direction; the value that lies in terms of the preferred direction above the average value of all directions. Specifically: the ratio of the square of the maximum electric field strength (or magnetic field strength) in the main radiation direction to the square of the field strength of an assumed isotropic emitter (dBi) or a Hertzian dipole (dBd).

Radiation resistance: relation between the antenna current in the feed point and the converted power in the wave type (ohms).

Front-to-back-ratio (also front-to-rear-ratio): measured ratio level 180° dB ratio level of the main lobe to the level of the back lobe.

Front-to-sides-ratio: measured ratio level 90° or 270° dB of the main lobe to the level of the side lobe.

Efficiency: ratio of actual radiated power, to the total power (in percent).

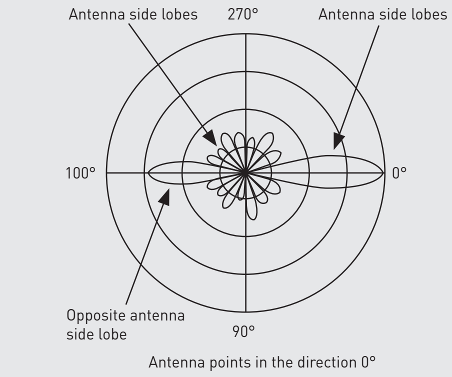
Shorter waves need visual contact

In the early days of wireless, for long and medium waves operation, the spatial location and geometric shape of the antenna were not really critical. The waves followed the earth's surface also far beyond the visual horizon. Only with the FM radio broadcasting service was experience won in regarding quasi optical propagation of wireless waves, and radio antennas were installed as high above the earth as possible: transmitting antennas were installed on high mountains, the receiving antennas on the roofs of houses.

The higher the frequencies and shorter the wavelengths became, the more the wireless propagation approached optical conditions. For a good wireless link, both antenna heights had to be optimised. This also means that the antennas needed visual contact with each other through the decimetric waves. In this regard, the laws of physics have often been disregarded in practice, or were forgotten – also, and particularly, in matters of industrial applications!

Different tasks – different systems

Antennas have the task to emit the generated admitted transmission power with as little loss as possible in the area (transmitting antennas) or to capture as much transmission power from the area (receiving antennas). This already could result in different structures, like a high transmitting mast on one side and a simple wire aerial on the other side. This applies particularly to unilateral wireless connections, such as the radio. With two-way wireless connections, similar antennas are more the rule. In practice, however, this reciprocity is limited by factors, such as the transmitting power. An antenna that is designed to receive is not necessarily suitable for use with the high electrical capacities of the transmitter.

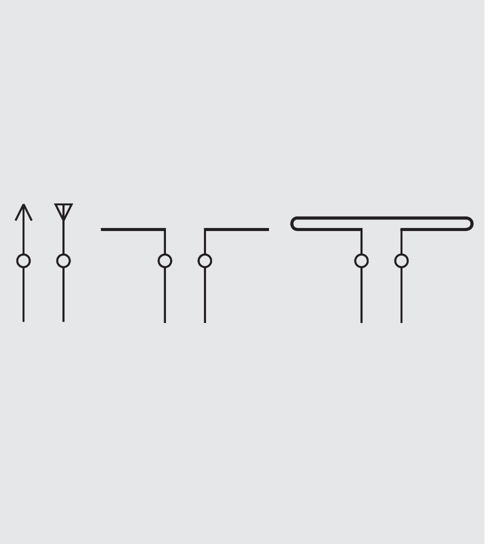

Antenna side lobes
270°
Antenna side lobes
100°
0°
Opposite antenna side lobe
90°
Antenna points in the direction 0°

Through the directivity, one can increase the range of a wireless connection.

Incidentally, the free space also has a characteristic impedance. More precisely, it is a fundamental physical constant, being the quotient of two fundamental physical constants. It has approximately 377 ohms as shown in the picture on page 171 above.

In other words, the antenna fits the characteristic impedance of the feed line (such as 75 ohms) to

the free space, so that the high-frequency energy can pass through without reflection and that means, without loss. In order to do so, the resonance conditions must be met as well. That is why we speak of »periodic« antenna coupling. On the other hand, there is the »aperiodic« antenna coupling, where a non-tuned piece of wire is coupled to the tuned input circuit, such as the telescopic antenna of a transistor radio. However, periodic antennas are the rule.



Circuit symbols: left, single pole long-wire and rod antenna, right, dipole antenna

Types of antennas

The diversity of the respective wireless services and also their different tasks lead to a correspondingly large number of designs. Here we name just a few typical designs. The best known and the most used antenna is the tuned dipole, used by Heinrich Hertz, precisely the half-wave-length dipole.

The ideal antenna:
The isotrop or isotropic radiator

This dipole is often the reference dipole for antenna comparisons as specifying antenna gain or directivity, because it emits uniformly in all directions. However, strictly speaking, it applies only to the level on which this dipole stands vertically, its equator level. Vertically, it has significant drops on the poles in the direction of the wire. This is the reason, why today, as theoretical reference, a preconceived spherical wave radiator (isotropic radiator) is used, which would emit equally intense HF energy in all directions. By definition, it has the antenna gain 0 dB. Antenna gains are related to these two antennas and specified in dBd for the half-wave dipole and dBi for the isotropic radiator. The value of a Hertz dipole lies at 1.8 dBi.

Here is a possible classification of antennas, according to their type:

- Linear radiator (linear antennas – the dipole is one of them)
- Group antennas (such as transmitting antennas for FM and TV)
- Panel radiator (panel antennas) in two types:
- Aperture radiator
- Reflector antennas (e.g. satellite dish and directional wireless parabola)

Other antenna designs, which can be classified under the above mentioned types, are for example:

- Helical antennas (emission in the direction of the axis of a wire or strip spiral, circular polarisation)
- Vivaldi antenna (two-dimensional exponential funnel at the end of a slot-conductor)
- Aantennas, created through slots in waveguides (emitting direction transversely or longitudinally to the waveguide)
- Spiral antennas, emission on both sides perpendicular to a spiral that is composed of strip lines, circular polarised
- Fractal antenna
- Sleeve antenna
- Patch antennas and PIFA on conductor plates
- T2FD (type similar as a folded dipole, through a terminating resistor but without resonance

effects). Some commonly used types are explained below

Linear antennas

Linear antennas convert a conducted standing wave in free space waves and vice versa. This includes all types of long wire antennas as well as dipole antenna and folded dipoles. Practical examples for linear antennas are radio transmitter masts for long and medium waves, wire antenna for amateur and ship radio, λ/2 dipoles as emitter in Yagi antenna and λ/4 dipoles in rod antennas for wireless services, wireless phones, CB radio, etc.

Shortened linear antennas

We can insert an inductance near the feed point to break through the fixed relation of the antenna length to the wavelength and to enable the construction of significant smaller antennas. These have an effect like an »electrical extension«. This way, antennas can be constructed that are much smaller than a quarter of the wavelength. Additional capacity at the end of the shortened element achieves the same effect. Such antennas are often characterised by higher losses, smaller antenna gain and smaller efficiency.

Half- and quarter-wave dipole

The length of this dipole is equal to half of the wavelengths. As there is a voltage maximum and a minimum electric potential at the feed point, the antenna shows a low impedance of 73.2 Ω. It is called a folded dipole if the direction of the current of a half wave dipole is divided into two ways. With it, the impedance of the feed point quadruples to approx. 240-300 Ω. It can be grounded and attached to the antenna carrier and can be used as a cheap symmetrical two-wire line.

The quarter-wave dipole is a special form, in which only one branch of the half-wave dipole is used as an antenna rod. The function of the other half serves for example as an electrically conductive surface. A user‘s body can take over this function, for example, with hand-held devices.

Magnetic antenna

Magnetic antennas do not use the electric field but a magnetic field to generate radiation and/or they receive primarily magnetic field components of the electromagnetic radiation. They consist of coils (in the simplest case with only one winding) that have a directivity (eight characteristics when the coil is standing) and can be very small compared to the wavelength if the coil consists of several windings.
Magnetic antennas also include loop antennas, DF antennas and ferrite antennas consisting of a rotating coil. By the way, it was a magnetic antenna that led to the discovery of radio waves. It was the second spark coil that was in the same laboratory room of Heinrich Hertz and sparked.

Phase array and dielectric antennas

Another type is the »phased array antenna«, which can mainly be classified as an aperture radiator. Their directivity can be generated electronically. It means that they are able to change and pivot their directivity pattern without the antennas as such having to move or change mechanically. At the same time, they offer an ability to suppress interference by placing a directional zero on their indicative chart on the interferer. And that is why they are able to suppress it so easily.

Particularly the so-called dielectric antennas (e.g. ceramic antennas) with decimetre and centimetre waves are now covering the market with great success. Its dielectric constant allows smaller dimensions. This promotes their practical applicability if compact dimensions are desired. However, losses are reported often on them and they only tolerate low voltages.

SW H. 12
FEED
SW H. 11
SW H. 13
Antenna
SW H. 15
SW H. 14
VERT
POL Y
SW
POL X
FEED
Antenna
Control

steute's standard antenna for wireless devices

Alternative antenna with detuning sleeve

In terms of SRD applications, it is not a significant hindrance. Protruding dipoles cause more problems.

Connection of antennas via coaxial cables

Special cables, the so-called coaxial cables, are used for the connection of antennas. These are twin-pole cables with a concentric structure, consisting of inner and outer conductors. Between them is an insulator, which may also consist of air. Regular coaxial cables have an outer diameter from 2 mm to 15 mm, and special forms have diameters up to 100 mm. There is also a coaxial design of overhead power lines.

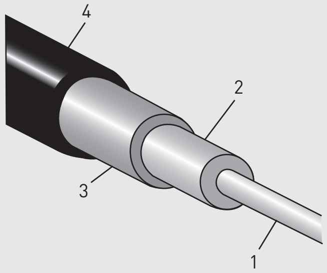
Most of the time, the inner conductor of flexible coaxial cables is made of braided or stranded copper wires. Also the cable shield consists of such copper wires and can be completed by a foil. Coaxial cables have defined characteristic impedance. It is usually 75 ohm for radio and TV reception technology and 50 ohm for other applications.

Antenna simulation

Besides calculation and/or metrological determination, computer-based simulation becomes increasingly important when setting up antennas or determining antenna parameters. In this manner, we can determine exactly the desired values by considering the other influencing factors in the vicinity, such as poles, metallic emission surfaces, etc. Parameters such as directivity characteristics for each polarisation direction and the antenna impedance can also be calculated, e.g., the emission characteristic for a defined frequency as well as the impedance values over the entire (or a defined) frequency range. Moreover, the programmes also specify the electricity distribution along the antenna. This data can be used as a basis for the optimisation of existing antennas.

Antennas: a central theme in wireless technology

Practical errors are often made because many technicians have to first familiarise themselves with the topic of wireless in automation and do



Coaxial cable cutaway model:
1. Core or inner conductor
2. Insulation or dielectric between inner conductor and cable shield
3. Outer conductor and shielding
4. Protective jacket

Design of a coaxial cable

not give enough consideration to the correct selection of the antenna. A comprehensive overview of antenna technology and the various antenna designs would go beyond the scope of this book. More information can be found on Wikipedia under the heading »Antenna technology«. In addition, there are specialist books that deal specifically with antenna technology.

// Wireless technologies and products from steute for medical equipment

011

First applications in the medical industry

For more than ten years the steute's development department has been working on wireless technologies. The initial applications were for the medical industry, where steute has for many years designed, developed and produced foot controls for medical equipment. The function of these foot controls plays a critical role in the faultless operation of the equipment. Free positioning of the foot control on the operating room floor, the absence of cables which can, for example, be a tripping hazard and hygienic requirements all give rise to the need for wireless solutions.

Start with infrared

The first wireless switch components that steute developed for the medical sector used infrared as the transmission medium. For example, a foot switch that allows wireless control via an IR signal was developed for an oral camera. However, this process has limits, because the transmission relies on a focused beam.

Engineers at steute evened out this disadvantage by using a high-energy signal. It is a method that exploits the reflections and emissions that come, for example, from the walls of the operating room. Moreover, the switches send signals in various frequency ranges to achieve a high transmission reliability.

The first steps with 868 Mhz technology

The first applications with 868 MHz technology were made with a foot control device of the type MKF-MED GP23. However the 868 MHz frequency cannot be used worldwide. Additionally, the device was operated under very strict conditions: The foot controls were used exclusively in MRI rooms which are hermetically sealed against radiation. There were no problems with coexistence or transmission reliability.

Next step: new standard based on Bluetooth

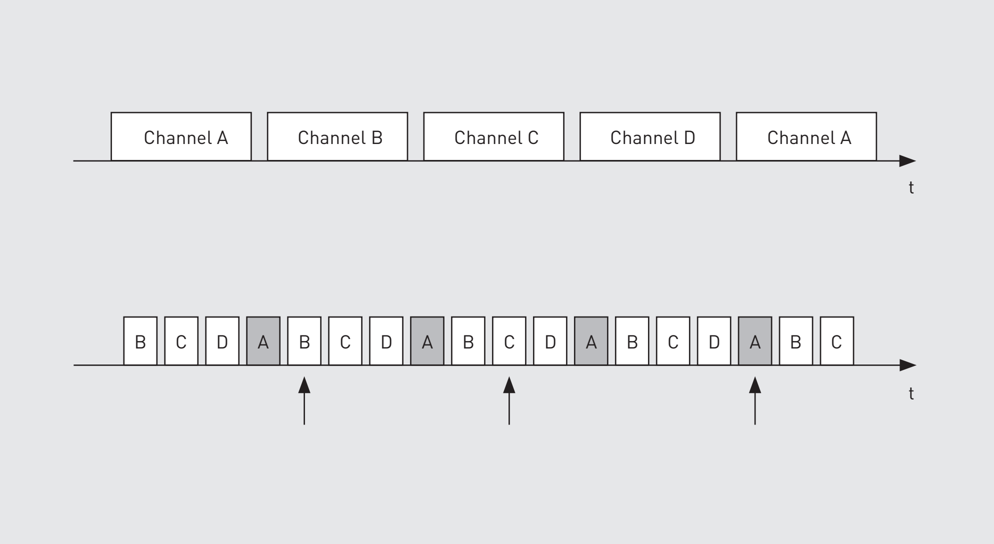
The next step in 2003 used the Bluetooth standard within the license-free 2.4 GHz ISM frequency band range for the development of a new generation of wireless control units for medical applications. The decision for Bluetooth was based on several reasons:

- The frequency band range is royalty-free and used uniformly all over the world.
- With a range of 10 to 100 metres, the standard for use and replacement of conventional cables in the medical and industrial sector is very well suited.
- Multiple devices can be operated at the same time.
- Due to the spread spectrum of frequency hopping, the signals are resistant to interference and thus are ideally suited for critical operations.
- The system can co-exist with other wireless networks (such as the WLAN networks used in hospitals).

Enhancement to meet the high safety requirements

However, the Bluetooth standard had to be modified to meet the high safety requirements of the medical industry. A fundamental feature of Bluetooth systems is that any two Bluetooth devices can communicate with each other. It is exactly this that must be prevented for medical foot controls. In this case there must be a very clear assignment of the foot (or hand) control with the specific medical device it is communicating with. This is why steute further developed the Bluetooth protocol with a customised configuration block.

The universal RF/BT module with the Bluetooth connection can be integrated in various types of final control equipment.



Channel hopping during the connection process

Individual configuration for each Bluetooth module

Origin and functions of the device can be entered into the module. Besides module classification data, such as the name of the manufacturer, device type and the device class (e.g. foot switch), configuration data of the module are stored here – and so are the numbers and functions of digital and analogue inputs and outputs. In addition, operational settings, such as alarm and warning signs, and the transmission speed of the serial interface can also be set. These configuration data allow the transmitter and receiver to verify already during the connecting process whether they are compatible and if a safe operation is possible.

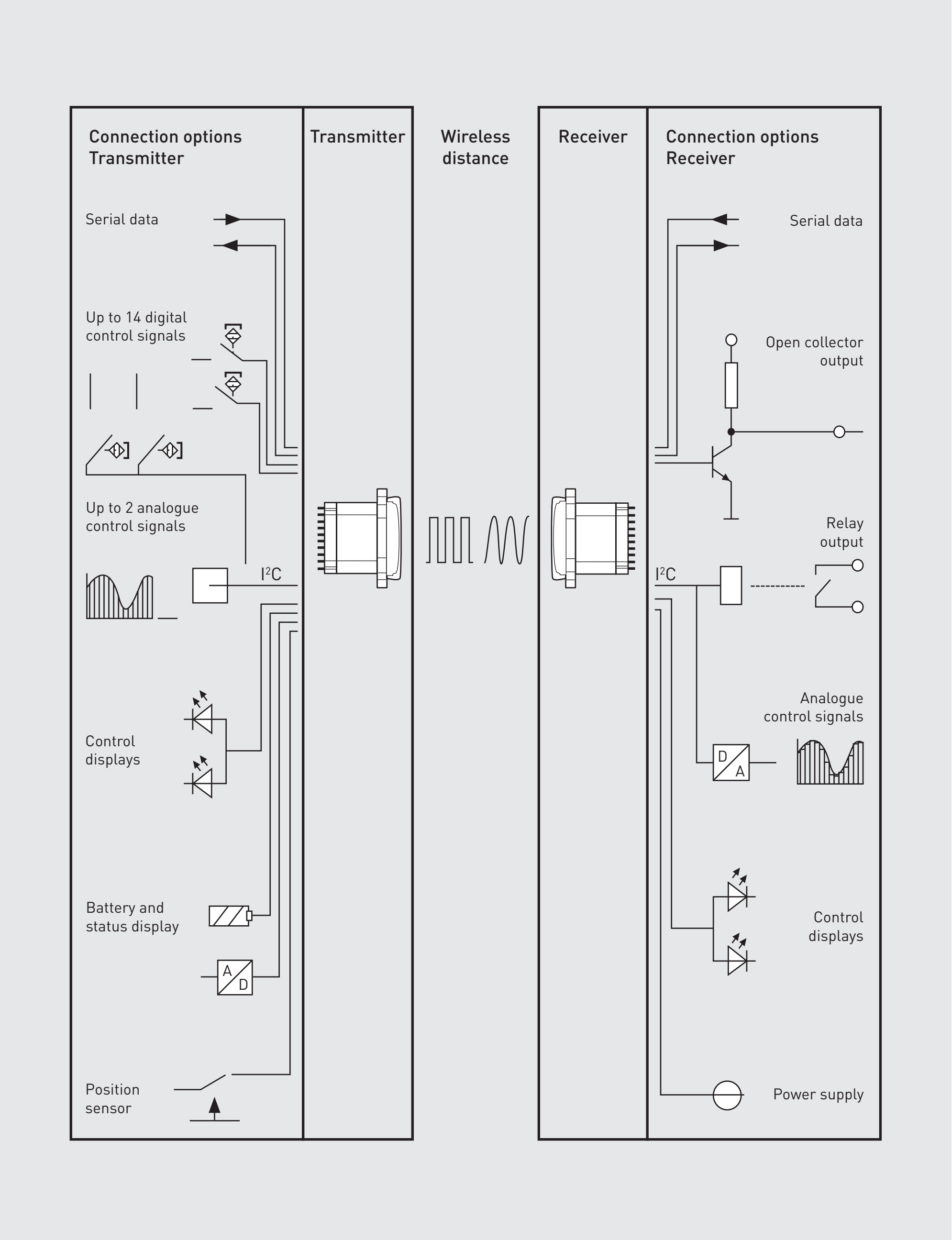
Registration process verifies switch and device compatibility

steute integrated several other innovative features to increase the operational safety even further and to ensure the compatibility of foot switch and device. A special registration procedure was designed to couple the device with the foot switch for initiation. steute developed various procedures for this purpose. A second data channel, for example, is created via an infrared route, which allows transmitter and receiver to handshake. Or, a backup cable is used for a on-off conformation. Alternatively, the pairing can be initiated by pressing a key on the medical device. Or a unique identification number can be used, the transmitter and receiver also exchanging the manufacturer and configuration data.

Sensor technology eliminates operating errors

Wireless foot switches are no longer tied to one location, and that means that the operator must take special precautions to avoid accidental operating errors. Because of this, sensors continuously monitor the status of the foot switch.

Once the switch leaves the ground, all switch functions are locked and a warning signal will be generated. If the switch is then not reset over a



The wireless standard steute uses today offers various input/outpu options for the transmitter and the receiver.

longer time period, it will stop the wireless transmission automatically.

Universal use with RF/BT module

The fact that the medical technology programme developed by steute is designed as a modular system has most certainly contributed to the success of the wireless switching devices, which steute developed for the medical technology sector. The user can always choose between a wired and a wireless version. In the wireless options, the cables are replaced by a compact module that contains the wireless board, commercially available batteries and the antenna.

All input and output signals of the RF/BTmodule are fed to the 26-pin connector on the edge of the board. Various digital switch functions can be carried out with minimal external circuitry. A compact receiver module is usually integrated into the customer's system.

Versatile: the periphery of the RF/BT module

Additional peripheral components can be connected to the RF/BT module via the I²C bus. Up to fourteen digital and two analogue control signals can be transmitted this way. Maximum response time of this transmission is 50 ms. In addition, a serial interface with adjustable transmission speed is also provided. Depending on the external circuitry, this wireless interface allows the replacement of asynchronous wired interfaces (such as the RS 232) through modules such as the RF/BT.

Power consumption: need for improvement

steute developed a technology which is establishing itself across the medical technology sector. This technology is used not just for relatively simple foot controls, such as single and double pedal foot switches, but is also for very complex surgical equipment. However, steute engineers and operators noted that there was still room for improvement with regard to one aspect: the power consumption of the wireless control devices is relatively high, in spite of intelligent power management. As the Bluetooth system takes a relatively long time to establish connection, the connection must be permanently maintained. This costs energy and is often unacceptable. This was the starting point for a new development.

New development: Energy-efficient 2.4 GHz wireless technology

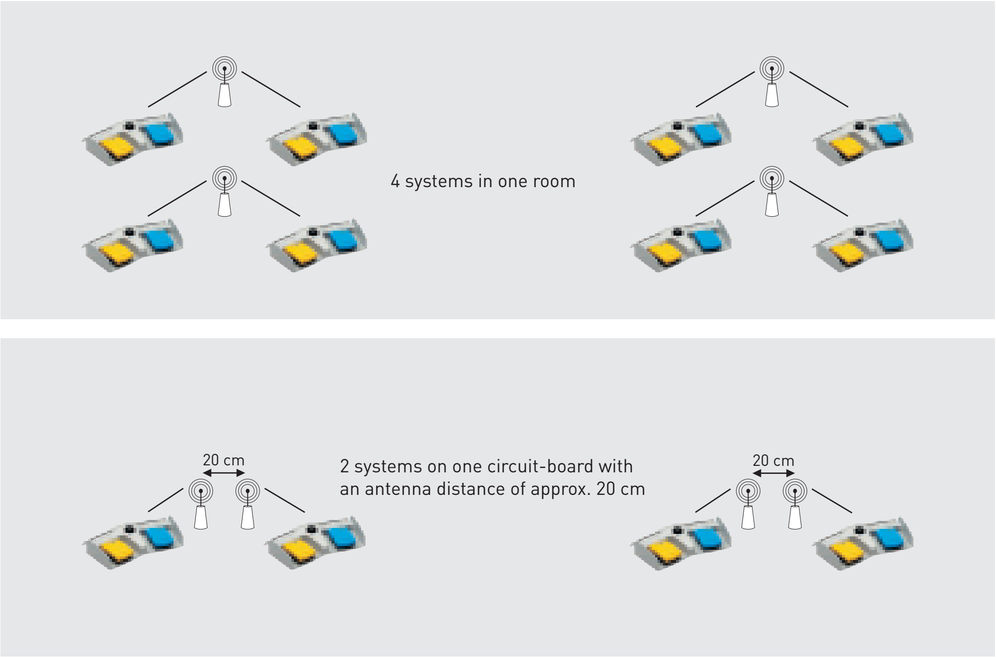
Among others, the »steute wireless« standard for medical technology was developed with the goal of achieving significantly lower energy consumption and to improve battery life compared to conventional Bluetooth systems. This is achieved through a system that uses - just like Bluetooth - the world-wide available and licence-free public 2.4 GHz band. The system was optimised for the communication of one or two foot switches (slave) with a receiver (master). The initiative to connect always starts from the foot switch. The receiver (with its non-critical energy supply) is constantly active and waits for a request to connect via the foot switch.

Fast response time – battery life of up to one year

According to this principle, a connection is made with channel B after 25 ms and with channel A after 80 ms. In each case 40 ms for sending and receiving a first telegram are added. At the latest after two passes, the connection is set up. Based on that, the connection time is 25 ms to a maximum of 200 ms. This creates the conditions to be able to operate the control devices through an energy saving sleep mode. Once in the sleep mode, the device can be reactivated after a maximum of 200 ms.

Wireless controlling equipment can be operated with a battery life of up to one year by means of these additional features of steute's wireless technology. Ranges of 10 metres are possible, and an error-free transmission and

Ophthalmology is one of the medical disciplines for which steute designs complex, customised controlling equipment.


4 systems in one room
20 cm
2 systems on one circuit-board with an antenna distance of approx. 20 cm
20 cm

With steute Wireless one device can be equipped with two receivers and up to four wireless systems can be operated in one operating room.

reception is guaranteed, even when multiple 2.4 GHz WLAN networks operate within the same area.

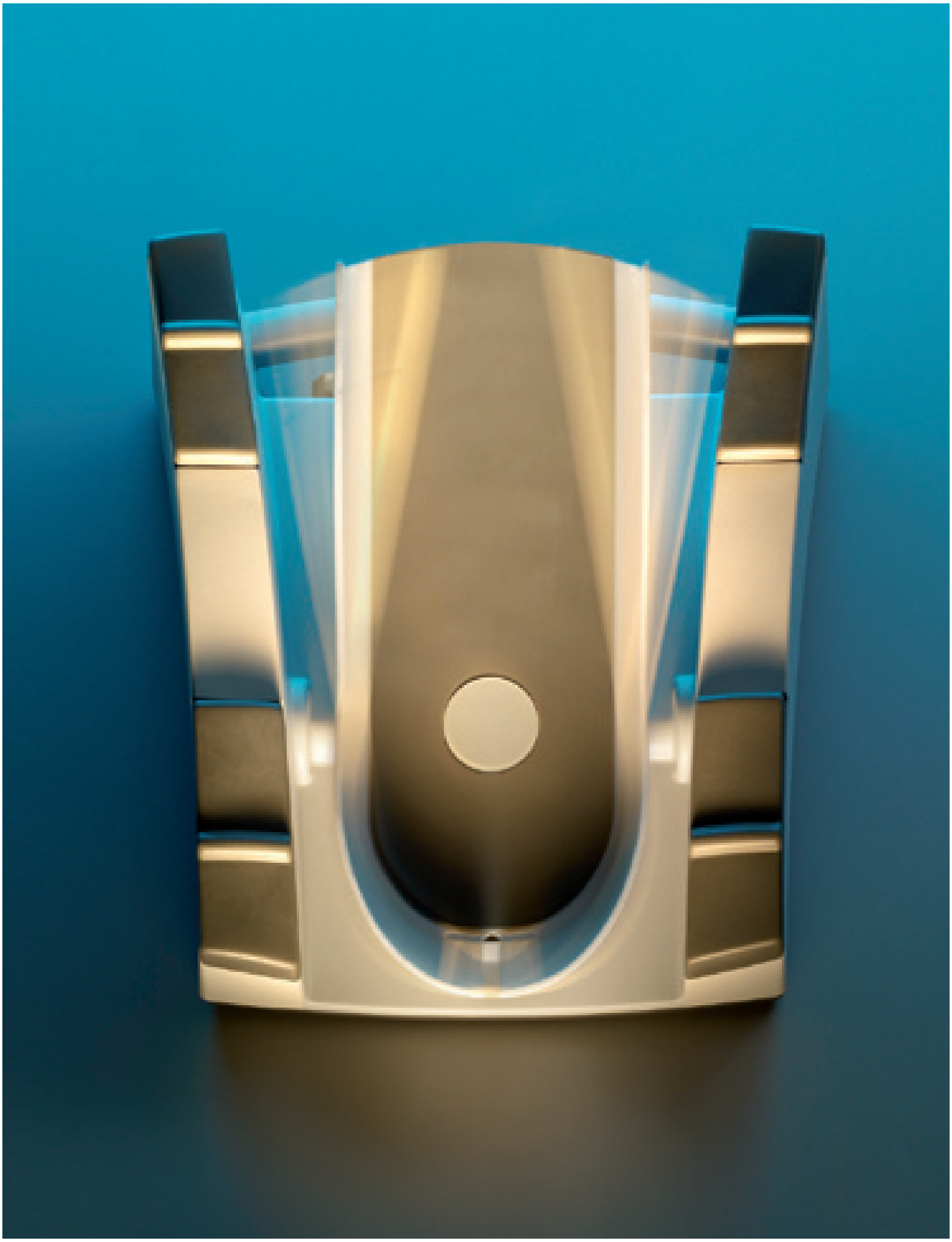
The operation of multiple receivers in a medical environment is possible without any interference provided that the antennas of the receiver are maintained at a minimum distance of 20 cm. In this way up to four transmitter/receiver pairs can be placed in an operating room without them interfering with each other (see image on page 186).

This system is already successful in the field and from the perspective of steute, it is the technology of choice for wireless medical technology applications.

Best solution: final controlling equipment for phacoemulsification

Controlling equipment for ophthalmology (eye care) demonstrates for example the functions of the steute foot switch for medical apparatus as well as the benefits of wireless technology. The world‘s most widely performed surgery is in the treatment of cataract: clouding of the natural eye lens leads to blindness. It can be cured by removing the clouded lens and replacing it with an artificial lens.

For that purpose, the surgical procedure called phacoemulsification was developed: the ophthalmologist makes a cut of just a few millimetres in the lens of the eye and extracts by suction the emulsified lens using ultrasonics. He injects a rolled up artificial lens through the cut opening,



The final controlling equipment for the phacoemulsification transmits the signals of the operator by wireless to the control of the surgical device.

which then »unfolds«. This gentle and minimally invasive procedure requires maximum precision and concentration by the surgeon. Since both hands are needed to control the instruments, all essential functions are controlled by the foot through the complex controlling equipment.

In the centre of the operation unit is a pedal that can be adjusted in four degrees of freedom. Four rocker switches are installed on the right and left side of the pedal. Each can be operated in two directions, and this means that the switches can be used for up to eight different functions.

The operation unit is linked to the control system of the phacoemulsification device so that the individual switches can take over different roles in the various process steps and menus. In doing so, the user can control complex processes with a manageable number of operating controls. It creates the precondition for ergonomic and intuitive work: the surgeon can concentrate entirely on the patient. Possible movements that the foot switch offers are not restricted by cables.

High demands - growing distribution

Finally, one can say that the wireless solutions developed by steute Meditec for medical apparatus are always specified to the customers‘ particular requirements. This applies to the wireless standard based on Bluetooth as well as to the steute wireless system.

Regardless of the selected wireless standard, a bi-directional signal exchange is always required - the receipt of a signal transmission is not sufficient due to the high demands by the security of transmission. Two channels are therefore very important. Analogue systems are also in demand here, because the controlling equipment often executes continuous movements and transmits signals accordingly.

steute‘s experience with wireless technology within the medical sector continiues to grow: acceptance is high and the benefits with regard to ergonomics and hygiene are evident. Use of wireless standards that are particularly designed for the medical technology sector and the long battery life for the latest wireless technologies will accelerate this development even more.

// Wireless technologies and products from steute for industrial applications

012

Industrial applications

Explosion protection

Industrial applications: analysis of technologies and market potential

Parallel to the development of the first wireless foot switch for the business division Meditec, steute analysed the market and the market potential of wireless technologies for industrial applications. In this area, standards were already established, and based on customer requirements, steute had already developed its first generation of products that function on the basis of 2.4 GHz technology.

It was however clear, that the wireless systems currently available on the market provided only half the solution: Signal transmissions were indeed conducted by wireless but the unit still depended on a wired power source if they wanted to avoid the maintenance issues inherent with battery technology.

Unique characteristic: energy self-sufficient switch devices for industrial automation

In addition, the well occupied market of the conventional »wireless automation« offered little opportunity to differentiate from the switching equipment manufacturer in the sector of automation technology. This possibility arose in 2004, when the newly founded company EnOcean GmbH used the idea of energy production through »energy harvesting« for switching devices and presented several options to extract energy from the environment. A solar module with miniaturised energy storage was among the innovations that EnOcean presented at that time, and the company won the prestigious Hermes Award 2004 award for these.

Core technology: low-energy wireless standard

The platform that EnOcean developed created the technical possibility to abandon cables and also the battery. It is not only modules that draw energy from the environment which belong to the EnOcean core technologies. The wireless standard that EnOcean founder Frank Schmidt developed is just as important. This wireless standard needs only very little energy and still ensures high transmission reliability.

Rapid success in building technology sector

steute was one of the first companies that recognised the importance of this technology. It entered into a contract with EnOcean and began with the development of wireless and energy self-sufficient switching devices. Other companies such as AEG OSRAM, Funkstuhl, Peha, Thermokon and Warema also continued on this path. However, all other EnOcean partners were active in the sector of building automation. »Energy harvesting« provides quite obvious advantages here: no cables are required between switches and electric consumers (lights, blinds, shading systems, ventilation flaps, etc.) at all, and thus can be considerably more flexible when building alterations are required, which are often necessary in administrative buildings as well as in industrial buildings. In this case, the electrician can simply place the light switch elsewhere on the wall, without having to lay cables or pull cables through cable ducts.

Use for automation technology

steute adapted this technology for industrial automation and control and used the EnOcean solar module for the first generation of the energy self-sufficient switching devices, whose integrated accumulator also allows a wireless operation during dark periods of up to 48 hours. Among the users who took advantage of this technology from the very beginning is, for example, a well-known German manufacturer of machining centres. The working area of the unit is secured by large doors that can be moved sideways. To connect the door handle's release button and the other switch functions with the controller, the designers had to provide a complex flexible cable harness. A door handle switch from the TGF programme does not require such complex types of cable feeding: the energy

Example of the first applications of »energy harvesting« in industrial automation: A multifunction handle from the TGF programme on sliding doors eliminates the need for flexible cable systems.

The solar cell generates energy to activate the switch.

Command devices belong to the programme of energy self-sufficient switching devices.

that the switches require is generated by the solar module. It is integrated in the appropriately-shaped door handle together with multiple controls.

Other innovations: the electrodynamic power generator

EnOcean presented another innovation, the electrodynamic power generator that generates the energy that is necessary to transmit the wireless signal through the actuation of the switch. Ultimately, the operator who actuates the switch generates the power, turning kinetic energy into electric energy. This principle is particularly elegant, because the energy is generated exactly at the moment in which it is needed. There is no need for a memory device, provided that a presence signal is not required.

Comprehensive range

In the meantime, a wide range of wireless switching devices based on EnOcean technology are available. Included in this programme are position switches, foot switches, pull-wire and pull switches, various command devices design types and the aforementioned door handle switches. There are, of course, also various solutions for the receiver. In addition to single- and multi-channel wireless receivers for the EnOcean standard, a repeater to increase the transmission range is also available.

Solution from the case

In the sector of industrial automation, the principle of »energy harvesting« is gaining more and more acceptance and steute actively promotes this development. »Starter packs« are available with all necessary components for a trial system.

In a few steps the assignment of transmitter and receiver via automated teach-in is possible. These cases are offered with various types of switching devices (such as pull switches and position switches).

»Plug and play«: the wireless case contains all components needed to operate a wireless switching device.

Creativity of the users

Practice shows that energy self-sufficient switching device applications are not only very diverse in the building technology sector but also in industrial automation, because the designers are very creative in finding new areas of application.

There are for example numerous energy self-sufficient pull switches where their pull wires are used for opening and closing rolling doors. The installation of a wired switch is often very expensive as long cable lines must be laid on high indoor walls and ceilings.

Another area of application is energy self-sufficient pull wires on the assembly lines in the automotive industry. Here, the switch functions as a »rip cord« in case of quality problems. It is used to send a signal to a control centre, or – according to the famous »Toyota production system« – to stop the whole line in order to resolve the detected problem immediately.

There are recent application cases that are more attributed to building technology than control technology: wireless switching devices with En-Ocean technology are used, for example, to monitor functions for solar systems and fire extinguishing systems.

Generally, energy self-sufficient switching devices allow flexible solutions on movable or remote equipment and a simpler, more cost effective installation within the entire sector of mechanical engineering and plant construction. It also applies to retrofitting existing machines with wireless and energy self-sufficient switching devices. This approach often helps to increase the productivity and ergonomics of the unit, to avoid parts that can wear and to increase availability.

Further development to meet the needs of industrial automation

Despite the fact that the use of energy self-sufficient switching devices is so innovative in

Energy self-sufficient pull switches are used to open and close rolling shutters

industrial automation, there is also a limitation: EnOcean's power generator was developed for use within the building technology sector and therefore designed for 50,000 operations. It is quite enough for this application, but industrial machines might generate up to 5000 operations per day. For steute, this was the reason for a new development. It uses the same board, but different and considerably more durable mechanics.

Own system in 868 MHz technology

Furthermore, early users had additional requests, which, among others, related to the size of the system, the range and the bi-directionality of the wireless communication. steute's declared goal was to also meet these requests.

However, the development goals could not be achieved on the basis of EnOcean technology. Therefore, in 2010, a completely new wireless system was developed by steute on the basis of the 868 MHz technology. The circuit board is considerably smaller than that of EnOcean. At the same time, the wireless communication is bidirectional which ensures higher transmission reliability. The range is also much higher: 700 to 800 metres are reached in the free field, while EnOcean technology reaches a maximum of only 300 metres.

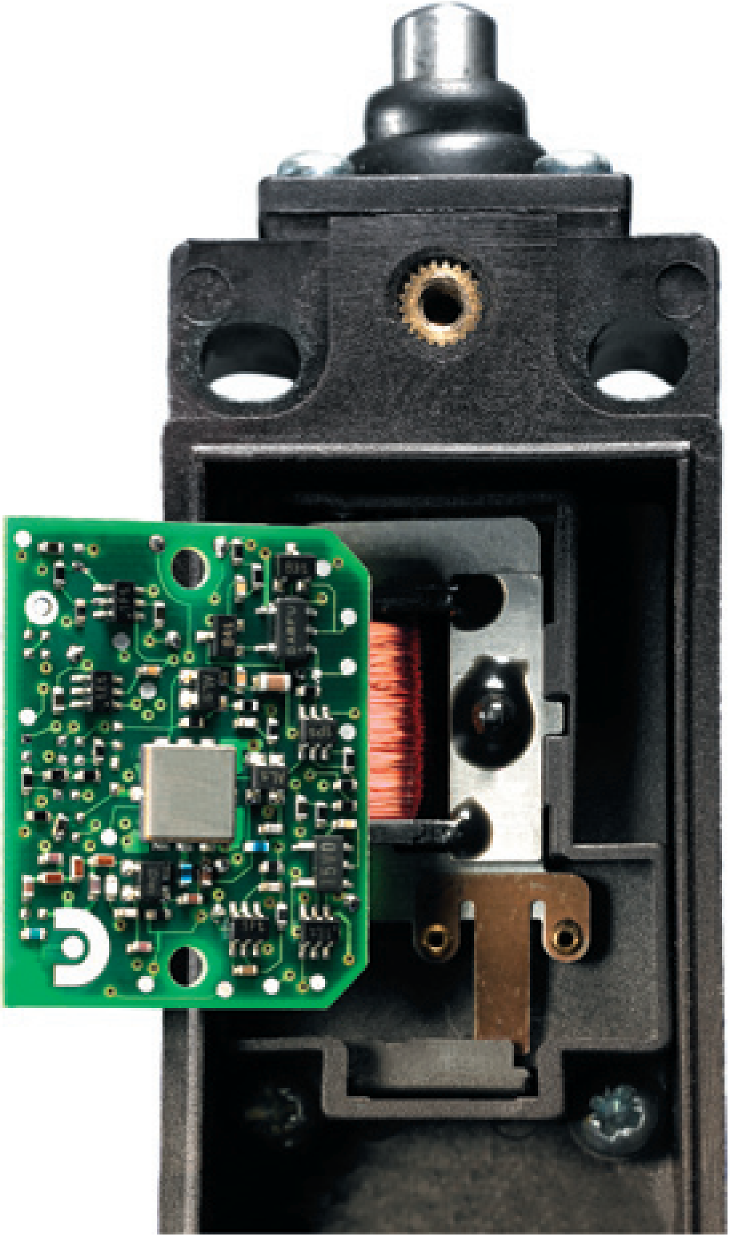
Important: bidirectionality and high range

Bidirectionality was not within reach until recently with energy self-sufficient switch components, but bidirectionality was also very important to the customers of steute. With the original EnOcean standard, the recipient does not transmit a receipt acknowledgement as confirmation. A bidirectional EnOcean standard was recently created with the Dolphin system. However, the energy-autonomous 868 MHz system developed by steute combines the bidirectional exchange of data with a higher range.

It is quite challenging to achieve such bidirectionality with a low energy system and no storage medium. steute developers mastered this challenge by sending the signal and receiving the confirmation while operating the electrodynamic power generator. The feedback is then stored in an EPROM transmitter and sent along with the next transmission. Thus, a storage medium for energy can be omitted, while enabling the transmission and guaranteeing of the acknowledgement of receipt.

New ways also with battery-based wireless standards

At first glance, it is not clear why a user of EnOcean technology wants a battery-based system. However, there are good reasons for it. An acknowledgement of receipt is needed after each transmitted signal, or the user might need a regular status signal. Since power consumption of the system is very low due to the power content of the signal, the user can insert very long-lasting batteries, in many cases even



Wireless position switch with electrodynamic power generator

Wireless switches in unusual form: the wireless cube RF 10

lifetime batteries. A key factor in this kind of wireless switch is the reliability of the input and output circuitry. steute has developed a series of special motherboards to achieve this aim.

The »Wireless Cube RF 10«

A good example of a newly developed battery-powered wireless switch is the Wireless Cube RF 10. It is built extremely compact and can be integrated easily in hard to reach places on a machine. The typical ranges that EnOcean standard allows are 300 metres in the free field and 30 metres in buildings. The electronics of the switches is powered by energy through a standard battery, which can be replaced with a simple tool. A spring serves as actuator of the switch device, and activates a micro switch. The actuation is a short process and the necessary actuating power is also low. Alternatively, the device can be also fitted with a magnetic switch and a reed contact. In addition to the compact design and easy assembly, longevity is one of the specific characteristics of the remote switch. The mechanical life extends to over a million switching cycles and also with a very high switching frequency the compact device is in its element. It can run up to 1,800 switching cycles per hour.

With these features, the new Wireless Cube RF 10 is the ideal switching device for poorly accessible and confined installation situations in industrial automation machines and equipment. It can be an economical alternative to conventionally wired position switches.

// Explosion protection

Wireless technologies in potentially explosive areas

steute has always had expert knowledge of the development and production of switching equipment for potentially explosive areas. It was a logical step for steute to use the EnOcean industrial automation technology also for ex-switching devices – particularly since the energy self-sufficiency in this application area offers additional benefits, as every wired connection is a potential risk. And when switching devices »spark« in hazardous areas, it makes sense to use conventional and significantly lower cost receivers outside the hazardous area. For these reasons, steute exploited at an early stage the use of EnOcean technology with the electrodynamic power generator for ex-switching transmitters. In order to achieve this, the devices had to be adapted to the demanding requirements of explosion protection. This applies to the mechanical as well as to the electrical construction.

Hurdles had to be overcome with the certification of devices under the ATEX Directive 94/9/EC, because first of all, appropriate test standards had to be developed. With these hurdles overcome, steute now offers a comprehensive portfolio of »wireless ex« switching devices that are certified for use in gas-ex zones 1 and 2 and as

Various wireless control devices are available for potentially explosive areas well in the dust-ex zone 21 and 22. The range includes among others command devices, position switches, pull switches and foot switches.

Also in the field of energy self-sufficient ex-switching devices, one can already find interesting examples of applications, e.g. position monitoring of valves on a gas pumping station. No power supply is located in this station area. Position switches capture the position of the valves, which are opened and closed manually via hand wheels and which pass the signals to a control centre.

Another typical area for application is the unloading of wagons in a chemical company. A trolley drives underneath the wagon to be unloaded. It hooks up underneath and pulls the wagon by rope into the desired position. The confirmation that the train is actually hooked up is verified via a wireless ex-position switch. These switches are used among others also in a refinery. There they are used to monitor the position of loading arms during the loading of tank lorries.

steute's wireless range

steute has a comprehensive portfolio of wireless switching devices for three major application areas: automation, explosion protection and medical technology. In the automation and medical technology sectors, the user can even choose between multiple wireless standards, depending on the ambient temperature and the conditions of the application.

To represent this diversity economically, steute uses a modular system of switching devices and wireless solutions. Wireless standards are often modified for unique customer-specific applications.

Command device with Wireless Ex technology

// Prospects

013

Energy harvesting draws energy from the environment

Prospects: growing markets for wireless switching devices

In steute‘s opinion, wireless solutions will expand in the industry and in the medical sector. One major reason for this is the increased flexibility of devices that can be used without cables.

Here, wireless technologies will develop similarly as in the consumer goods market, where TV remote controls, mobile phones, WLAN and Keyless Go are now part of everyday life. Well-known wireless techniques, such as television and radio, go the opposite way and are now frequently transmitted through wired devices but that is not contrary to the wireless concept. Moreover, wireless technology will also benefit from further development of battery technology, in which many invested recently: it promotes longer service life of wireless devices.

steute estimates that »energy harvesting« is just beginning to develop. Already established procedures that generate energy from the environment, such as sun and movement (such as the electrodynamic energy generator) will advance further. Research and development laboratories of universities and companies are also working on other processes to remove the need for external energy supply and batteries. They concentrate for example on temperature and vibration. Mobile phones could be equipped with a practical example in the near future: with devices that recharge through the movement of people (walking).

In short, there are many indications that wireless technologies within the industry and the medical technology sector are just at the beginning and that there is a lot more potential. steute will help to exploit this potential. Its goal is to optimise the operation of machines and equipment, to simplify their constructions, to reduce installation and assembly costs and to achieve greater flexibility within industrial production and also for medical equipment that help to treat diseases. Here, proven wireless technology can and will make important contributions in the future.

ANS
RING
BUSY
ATND
HOLD
1
2
3
4
MJ
MN
1
ABC
2

// Appendix

Glossary

Image source directory

Authors

Imprint

Concept/statement/page

DMT	Discrete Multitone Transmission/Digital modulation process 75
DSSS	Direct Sequence Spread Spectrum/Direct sequence spread process 63
EDGE	Enhanced Data Rates for GSM Evolution/Multi-layer digital modulation process 75
EIRP	Effective Isotropic Radiated Power 42
ETSI	European Institute for Telecommunication Standards 54
FDMA	Frequency Division Multiple Access/Multiplex process for multiple usage of frequencies 56
FEC	Forward Error Correction/Error backup process 80
FFD	Full Function Device 106
FHSS	Frequency Hopping Spread Spectrum/Frequency hopping process (interference protection process) 63
FreqBZP	Frequency Band Range Plan 51
FreqNP	Plan for Frequency Use 51, 54
FSK	Frequency Shift Keying/Digital modulation process 75
HART	Highway Addressable Remote Transducer 108
IEEE	Institute of Electrical and Electronics Engineers 126ff
ISM	Industrial Scientific Medical/Licence-free frequency range for devices with low transmission power 53
ITU	International Telecommunication Union 50
LBT	Listen Before Talking/Access regulation for the frequency use 56
LOS	Line of Sight/Visual contact between transmitter and receiver 64
MANET	Mobile Ad Hoc Networking/Ad-Hoc networks 70
MIMO	Multiple Input Multiple Output 135
MTU	Maximum Transmission Unit 113
NLOS	No Line of Sight /Visual contact without visual contact between transmitter and receiver 64
OFDM	Orthogonal Frequency Division Multiplex/Interference protection process 64, 135
PCM	Pulse Code Modulation 78
PSK	Phase Shift Keying/Digital modulation process 75
QAM	Quadrature Amplitude Modulation/Digital modulation process 75, 135
QPSK	Quadrature Phase Shift Keying/Multi-layer digital modulation process 75, 135
RFD	Reduced Function Device 106
RTS	Ready to send-Signal 130
SAR	Synthetic Aperture Radar 64
SRD	Short Range Devices/Wireless device for the ISM band 52, 96
TDMA	Time Division Multiple Access/Multiplex process for multiple usage of frequencies 57
WECA	Wireless Ethernet Compatibility Alliance 131
WPA	WiFi Protected Access 165

Page/Image source

25 »Heinrich Hertz«, source: Wikipedia.de

25 »Experimental Arrangement for Experimental Detection of Electromagnetic Waves« source: The Development of Wireless Technology. From Post Office Worker Concert in 1920 to Digital Radio in 1993 by Siegfried Hermann (author), Wolf Kahle (author), Joachim Kniestedt (author), publishing house R. v. Decker Verlag, Heidelberg

27 »The First Radio Relay Link: Experimental Set-up by Heinrich Hertz« source: German Experimental Wireless Message Systems until 1940 by Fritz Trenkle, publishing house Hüthig Verlag, Copyright: Telefunken Systemtechnik GmbH

28 »Pluggable Crystal Detector with Pyrite Crystal and Lace« source: Wikipedia.de, Creative Commons

32 »Spark Extinguishing Transmitter from Telefunken« source: Stiftung Deutsches Technikmuseum Berlin, Historical Archives

33 »View of a Spark Gap with Tesla Transformer«, source: Wikipedia.de

34 »View of a Flashing Arc Transmitter after Poulsen«, source: Wikipedia.de by Clemens Pfeiffer, Creative Commons

35 »Machine Transmitter by Alexanderson, the Generator on the Left, the Multiplier on the Right«, source: Wikipedia by Gunther Tschuch, Creative Common

36 »Different Reception and High-performance Tubes« source: Wikipedia by Stefan Riepl (Quark 48), Creative Commons

37 »Tube type RE 144 by Telefunken (in the 20s)« source: Wikipedia.de by Hihiman, Creative Commons

Jörg F. Wollert

Wolfram Gebhardt

Jörg F. Wollert is a professor of software engineering and computer networks at the Technical University of Bielefeld. After completing his electrical engineering studies with focus on telecommunication, he graduated from the RWTH Aachen, Faculty of Mechanical Engineering, in the subject of distributed object-oriented systems in automation technology. He gained industry experience as a project manager for complex automation systems by working for a market-leading company in the logistics sector. In 1999, he was appointed at the University of Bochum. Since 2000, he studied distributed real-time systems with focus on Ethernet-based field buses and wireless systems. With a group of about 10 people, he edited various R&D projects system solutions in the area of wireless networks and convergence of wireless and wired real-time systems.

Wolfram Gebhardt had already discovered his passion for wireless technology at his scientific secondary school. He studied aerospace in the German Air Force with telecommunication as the main subject. He initiated the development of cheaper antennas for aircraft radio already during his first troop deployment. In a NATO command agency, he was responsible for the establishment of directional radio and shortwave networks. As department head for wireless systems of the Air Force, he ordered the radio networks and frequency allocations of the Air Force and was responsible for all aspects of radio coordination.

As head of the Air Force's communications systems at the German Ministry of Defence, he was involved in future technology and participated in the research and development of new systems. The civilian characteristics of these systems are used today in mobile communications as well as in the general applications of Bluetooth and WLAN.

Imprint

Issued by
steute Schaltgeräte GmbH & Co. KG, Löhne, Germany, Wireless business division

steute's editorial team authored Chapters 10, 11 and 12 and are also responsible for the coordination and editorial work. This team consists of the following members:
Andreas Hofmann, Development; Andreas Schenk, Product Management
Dipl.-Ing. Marc Schmidt, Development; Dipl.-Ing. Georg Töpler, Head of Marketing
Dipl.-Ing. Ralf Twellmann, Head of the Technical Division

External editors: Gerald Scheffels, Wuppertal, Germany

Concept and design: Stephan Degens, Büro Longjaloux GmbH, Wuppertal
Implementation of the graphics Mark Muchowski, Büro Longjaloux GmbH, Wuppertal

Translation: C.E.T. Central European Translations GmbH, Cologne

Total circulation: Printing Office Westfalia-Druckerei, Löhne, Germany

First edition, November 2010, 1000 wd

ISBN 978-3-00-038951-1

steute
Schaltgeräte GmbH & Co. KG
Brückenstraße 91
32584 Löhne, Germany
Phone + 49 (0) 57 31 7 45-0
Telefax + 49 (0) 57 31 7 45-200
E-mail info@steute.com
www.steute.com

ISBN 978-3-00-038951-1