

Autonomous driving and simulational challenges

For more than one hundred years the automotive industry has been a warrant for technological innovation, sustainable employment and prosperity. With the introduction of seat belts and airbags vehicles got increasingly safer. The development of an electronic fuel injection system lead to severely diminished emissions due to lower fuel consumption. Moreover, and despite the rising technological complexity prices dropped considerably with regards to the Consumer Price Index. As a summary it may be stated that motor vehicles got cleaner, safer, and more affordable. The automobile as a symbol of freedom, indepenence, and prosperity lead to a continually increasing demand. In conjunction with additional political measures the automotive sector has become the leading industry in many industrialized nations like the United States, China, and Germany, to name just a few.

In spite of all this, time does not stand still and so car manufactures are facing great upheavals nowadays. For a long time period California has been a sole pioneer for tightening emission regulations and thus increased the demand for electric vehicles. While senior industry experts kept saying lithium-ion technology is still several years away Tesla introduced the first all-electric car based upon lithium-ion battery cells in 2008. Even though new competitors are struggling with numerous difficulties till date their way of working as well as their influence to current and future development projects in the automotive sector is already considerable. Economy and society are on the brink of another new technological revolution: autonomous driving [1].

1.1 Safety validation and simulative test drives

The almost one hundred year old idea of self-driving cars might solve several intractable problems related to everyone's daily mobility. These include, among others, traffic congestions, availability of public transport, medical or age-related mobility constraints as well as the often quoted number of road fatalities. The particular interest of many companies in this technology, however, is especially based on associated disruptive business models [2].

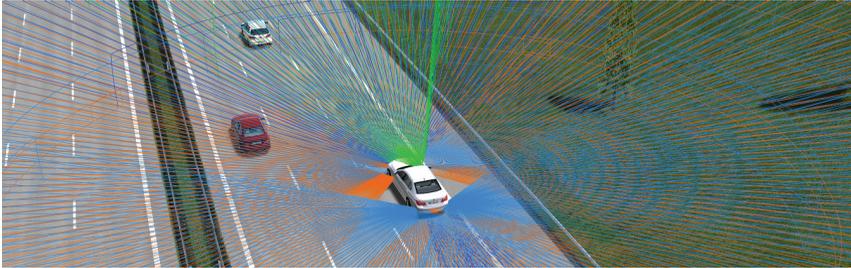


Fig. 1.1. Schematic, exemplary view of the various sensors around a vehicle with a driver assistance system. The picture is taken from a highway scene within a virtual environment. The colored sensor cones show each sensors approximate field of view. The different colors indicate different types of sensors, e.g., radar, lidar, and camera sensors (see also [3]).

However, the almost 25-year-long development of adaptive cruise control (ACC) has demonstrated that innovation development up to the stage of production readiness requires a certain amount of effort. On the one hand, there was uncertainty about the customer features and the technology to be used (pulse-width modulation or frequency modulation), and on the other hand, there was a lack of suitable suppliers. From today's point of view, it is not surprising that numerous product releases failed due to the extensive efforts required for integration and deployment, as well as the difficulty in communicating the resulting benefits to the customers. Nevertheless, adaptive cruise control has matured to become a mass product these days, paving the way for the further development of advanced driver assistance systems.

The development of autonomous driving and modern advanced driver assistance systems always focuses on controllability by the driver and constitutes the basis for safety validation of individual functions. This applies in particular to conditional automation, the so-called Level 3 of driving automation according to the SAE J3016 standard. Here, all aspects of the driving task are performed by an autonomous driving function, whereby the driver is always expected to be able to intervene on demand in case of doubt. Typically, the driver is expected to be ready to take over the driving task within a few seconds. This strategy leads to enormous demands on the view of environmental sensors. However, particularly due to the fact that even the simplest tasks are covered by the autonomous driving function, even the very simplest but rarely occurring events must be safe guarded prior to a vehicle's release. In case of events that occur only after an average of 50,000 kilometers or more (see 1.2), such as an emergency break event, several million test kilometers have to be performed beforehand. Additional requirements for a particular test concept, such as weather conditions or configuration options, may again drastically increase the number of test kilometers to be covered [4, 5].

When deriving the number of test kilometers, a Poisson distribution is usually assumed. For example, if a certain event occurs n times on average for a distance length s_0 , then

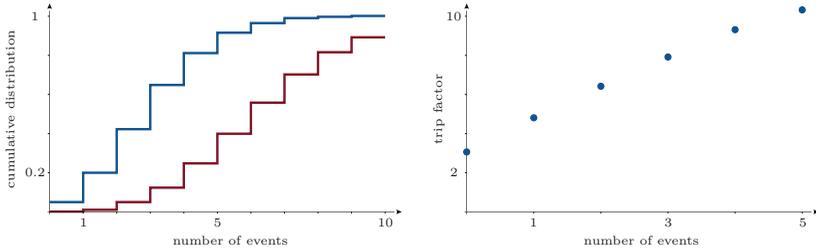


Fig. 1.2. Effects on the number of test kilometers assuming an underlying Poisson distribution. On the left, the cumulative probability for the number of occurrences, if $\lambda = 3$ (blue) or $\lambda = 6.3$ (red) events are to be expected during the test. On the right, the stretch factor for the test drive in case of a given number of events (see also [4, p. 1175 ff.]) for further details.

it is expected that in case of a trip length s the event occurs $s/s_0 \times n$ times. However, this does not guarantee that the event will occur that often. Instead, this value forms the expectation value λ of the Poisson distribution. For instance, if an event occurs on average only once at a distance s_0 , then at the three-fold distance, i.e., $\lambda = 3$, the event does not occur at all in 5% of the cases (see 1.2). This is referred to as significance level and is usually tolerated up to a probability of 5%. Therefore, for a given event, at least three times the distance must be completed without an accident. Alternatively, 6.3 times with a maximum of one accident.

Such a safety validation strategy rapidly leads to a coverage of up to multiple hundred million test kilometers and exceeds the technical, personnel and financial capabilities of today's companies. Even if this effort was ventured for an initial system under limited circumstances, one has to consider that the test would have to be performed again with at least one third of the initial effort after each modification of the system, which is obviously not economically justifiable [4]. Therefore, the prevailing opinion among automotive manufacturers, suppliers and also within the research community tends towards a decisive role for simulation-based safety validation in the future [6].

1.2 Principles of automotive radar sensors

The first radar systems originated in the field of military and aviation technology. For about two decades, these sensors have also been increasingly integrated into automobiles. Today, upper-class vehicles already include numerous radar sensors that provide data for various driver assistance functions. In general, however, vehicle manufacturers only specify certain sensor requirements and subsequently either purchase these sensors or develop them in collaboration with suppliers.

Regardless of whether the design of models is performed by the sensor manufacturer himself or by automotive or simulation manufacturers themselves, it is crucial to understand

the fundamental principles of radar sensor technology and to be familiar with the definition of the interfaces. Essentially, to determine the distance of an object, a radar sensor usually measures the time of flight t taken by an electromagnetic wave to travel back and forth between sensor and target. Hence the distance R is given by

$$R = \frac{c_0}{2}t, \quad (1.1)$$

whereby c_0 is the speed of light. For automotive applications, however, so called pulsed radar systems, which measure the runtime directly, have numerous disadvantages. Thus, targets can only be distinguished from each other if the distance between them is sufficiently large, i.e., depending on the pulse duration t_p the distance must be at least

$$\Delta R \geq \frac{c_0}{2}t_p. \quad (1.2)$$

Due to legal restrictions and the fact that the -3 dB bandwidth is inversely proportional to the pulse duration, t_p cannot be chosen arbitrarily small. Therefore, the frequency modulated continuous wave (FMCW) measurement principle is widely accepted and currently de facto standard for automotive purposes. Therefore, a short introduction to the essential principles of modern automotive radar sensors will be given in the following. The subsection focuses primarily on signal processing and concludes with a general description of the radar point cloud interface.

1.2.1 Static targets and their radar signal

Frequency modulation is characterized by the transmission of a continuous wave, with periodically increasing or decreasing frequency. The following explanation deals with the common case of sawtooth frequency modulation, i.e., the time dependent transmit frequency f_{Tx} is given by

$$f_{\text{Tx}}(t) = f_c + \frac{B}{T} \left(t - \left\lfloor \frac{t}{T} \right\rfloor T \right). \quad (1.3)$$

Each sawtooth, also called chirp, passes through the frequency range from f_c to $f_c + B$. Thereby, f_c indicates the carrier frequency, B the bandwidth and T the chirp or sweep time. In case of a reflective target at a distance R , the signals time of flight is

$$\tau_0 = \frac{2R}{c_0}. \quad (1.4)$$

This results in a time dependent receive frequency

$$\begin{aligned} f_{\text{Rx}}(t) &= f_{\text{Tx}}(t - \tau_0) \\ &= f_c + \frac{B}{T} \left(t - \left\lfloor \frac{t}{T} \right\rfloor T - \tau_0 \right). \end{aligned} \quad (1.5)$$

Now, if the difference frequency Δf between the transmit and receive frequency is measured, the distance R to the target can be determined, i.e.,

$$R = \frac{c_0 T \Delta f}{2B}. \quad (1.6)$$

To be more specific about digital signal processing and the frequency measurement principles, it is necessary to examine the actual transmit and receive signal more closely. The signal of an unmodulated wave at frequency f is given at the transmitting antenna as

$$\begin{aligned} u_{\text{Tx}}(t) &= u_0 \cos(\varphi_{\text{Tx}}(t)) \\ &= u_0 \cos(2\pi ft). \end{aligned} \quad (1.7)$$

Thereby, the factor u_0 denotes the amplitude and φ_{Tx} the phase of the transmit signal. To derive a more profound interrelation between frequency and phase of a frequency modulated signal the term instantaneous frequency must first be explained. Since frequency is the number of revolutions (or simply repetitions) per time, the formula can be expressed in terms of the phase angle as well. Therefore, the average frequency \bar{f}_h in a small interval $[t, t+h]$ of length h is given as

$$\bar{f}_h(t) = \frac{1}{2\pi} \frac{\varphi(t+h) - \varphi(t)}{h}. \quad (1.8)$$

Now, the limit $h \rightarrow 0$ defines the instantaneous frequency. Thus, it can be concluded that frequency is the time derivative of the phase, i.e.,

$$f(t) = \frac{1}{2\pi} \frac{d}{dt} \varphi(t). \quad (1.9)$$

Conversely, the phase can be determined by integrating the time dependent frequency. Hence the phase is given by

$$\varphi_{\text{Tx}}(t) = 2\pi \int_0^t f_{\text{Tx}}(u) du. \quad (1.10)$$

Provided that all chirps start in phase, it suffices to simply integrate from 0 to $(t - \lfloor t/T \rfloor T)$ to determine $\varphi_{\text{Tx}}(t)$. An analogous formula applies to the phase of the receive signal φ_{Rx} at $t \geq \tau_0$, i.e.

$$\begin{aligned} \varphi_{\text{Rx}}(t) &= 2\pi \int_{\tau_0}^t f_{\text{Rx}}(u) du \\ &= 2\pi \int_0^{t-\tau_0} f_{\text{Rx}}(u + \tau_0) du \\ &= 2\pi \int_0^{t-\tau_0} f_{\text{Tx}}(u) du \\ &= \varphi_{\text{Tx}}(t - \tau_0). \end{aligned} \quad (1.11)$$

The intermediate frequency signal u_{Δ} of u_{Tx} and u_{Rx} is a nearly continuous wave with frequency Δf . Technically, the differential signal is generated by means of a multiplicative mixer and a subsequent low-pass filter. The mixer multiplies both signals with each other. According to the trigonometric product-sum identity, the resulting signal u_* has the form

$$\begin{aligned} u_*(t) &= u_{\text{Tx}}(t)u_{\text{Rx}}(t) \\ &= u (\cos(\varphi_{\text{Tx}}(t) + \varphi_{\text{Rx}}(t)) + \cos(\varphi_{\text{Tx}}(t) - \varphi_{\text{Rx}}(t))). \end{aligned} \quad (1.12)$$

In general, since the receive signal passes through an amplifier before, there is an additional constant $c > 0$ such that $u = u_0 u_1 c$. In the context of signal processing, the exact value of the amplitude can usually be neglected anyway. Now, the low-pass filter removes all high-frequency components, hence only the difference frequency remains present. Consequently, the intermediate frequency signal is given by

$$u_{\Delta}(t) = u \cos(\varphi_{\text{Tx}}(t) - \varphi_{\text{Rx}}(t)). \quad (1.13)$$

According to equation (1.10) and (1.11) one can derive a detailed expression for the intermediate frequency signal. However, since this leads to an extremely elongated expression, the signal is usually only given for the range of constant intermediate frequency. This range extends from τ_0 to $T - \tau_0$ and accounts for considerably more than 90% of the signal. For example, a signal with $T = 20\mu\text{s}$ contains 99% constant frequency components for a target at a distance of 15m. For the previously mentioned range the temporal signal is given by

$$\begin{aligned} u_{\Delta}(t) &= u \cos\left(2\pi\left(\frac{B}{T}\tau_0\left(t - \left\lfloor\frac{t}{T}\right\rfloor T\right) + f_c\tau_0 - \frac{B}{2T}\tau_0^2\right)\right) \\ &\approx u \cos\left(2\pi\left(\frac{B}{T}\tau_0\left(t - \left\lfloor\frac{t}{T}\right\rfloor T\right) + f_c\tau_0\right)\right). \end{aligned} \quad (1.14)$$

Since a window function is usually applied to the intermediate frequency signal anyway, it is sufficient for almost all applications and models to consider a signal of constant intermediate frequency. For consideration of more complex effects due to the time ranges $[0, \tau_0]$ and $[T - \tau_0, T]$, the phase must be calculated for the full range. In addition to a different phase, and depending on the window function, there is also a time-dependent amplitude. The following neglects these effects and proceeds with the formula given in equation (1.13).

1.2.2 The Doppler effect and its influence

For the derivation of the frequency modulated continuous wave radar signal only static targets have been considered so far. However, if the target moves the frequency of the received signal is shifted due to the Doppler effect. For the receive frequency f_{Rx} of a moving target at the radial speed v_r , the following applies

$$f_{\text{Rx}} = \frac{c_0 - v_r}{c_0 + v_r} f_{\text{Tx}}. \quad (1.15)$$

Since an exact relativistic consideration is usually not necessary at low speeds, the above formula is often simplified by a first-order Taylor approximation. Therefore, equation (1.15) can be written as

$$f_{\text{Rx}} \approx \left(1 - \frac{2v_r}{c_0}\right) f_{\text{Tx}}. \quad (1.16)$$

As is the case with a primary radar, the above formula given in equation (1.15) considers two relativistic Doppler shifts. The first shift occurs during the reception at the target.

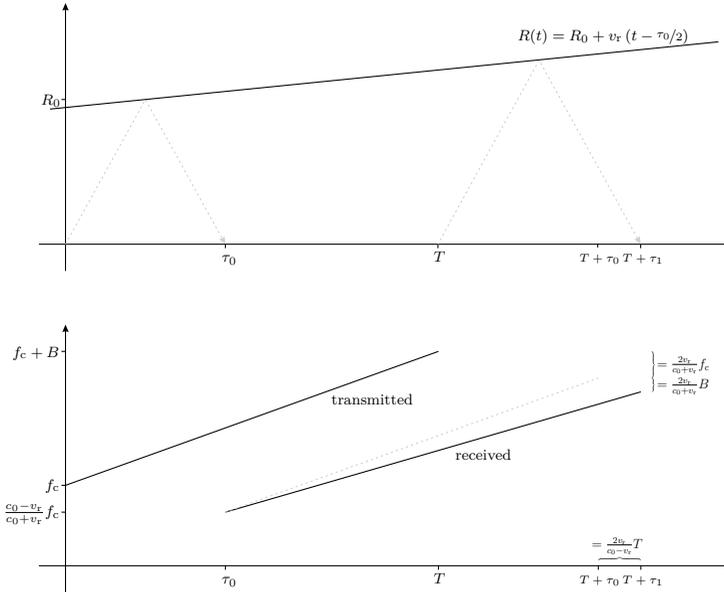


Fig. 1.3. Illustration of the relativistic Doppler effect and its influence on a single chirp. During the transmission of a chirp, the target vehicle continuously moves and as a result, a frequency shift occurs. Since the chirp duration is also subject to a Doppler shift, the graphs of transmitted and received frequencies are not perfectly parallel. However, these nuances are generally negligible.

The second one when the signal reaches the transmitter after reflection. For example, if the target moves radially away from the sensor, $v_r > 0$ applies and thus $f_{Rx} < f_{Tx}$. In order to finally reproduce such a frequency shift within the phase of the time signal u_Δ , an exact relativistic consideration of the received phase is necessary. In numerous publications and contributions one can find simpler derivations, which only consider a time-dependent runtime τ in equation (1.11). However, this reproduces the Doppler shift only approximately and neglects certain effects due to frequency modulation (see the following remark).

In order to determine the phase of the reflected signal, in case of a single dynamic target, a twofold coordinate transformation must be performed. For this purpose, the transmitted spacetime-signal (for simplicity limited to only one chirp)

$$u_{Tx}(t, x) = u_0 \cos \left(2\pi \left(f_c \left(t - \frac{x}{c_0} \right) + \frac{B}{2T} \left(t - \frac{x}{c_0} \right)^2 \right) \right) \quad (1.17)$$

is transformed into the reference frame of the target. According to the Lorentz transfor-

mation, the coordinates (t, x) are given in the reference frame of the target as

$$\begin{aligned} x' &= \gamma (x + v_r t) \\ t' &= \gamma \left(t + \frac{v_r x}{c_0^2} \right), \end{aligned} \quad (1.18)$$

whereby $\gamma = (1 - v_r^2/c_0^2)^{-1/2}$ is the Lorentz factor [7]. Plugging (t', x') into equation (1.17) one gets the transmit signal from the perspective of the target vehicle. Now, a reflection is equivalent to the instantaneous emission of the signal received from the sensor. Due to the principle of relativity, instead of a moving target and a fixed sensor, one can just as well assume a stationary target and a moving sensor. Therefore, the reflected signal in the sensor reference frame is obtained by another Lorentz transformation. Just with exchanged roles of (t', x') and (t, x) . If one considers the additional time of flight τ_0 , this results in the following overall transformation

$$\left(t - \frac{x}{c_0} \right) \rightsquigarrow \frac{c_0 - v_r}{c_0 + v_r} \left(t - \frac{x}{c_0} \right) \rightsquigarrow \frac{c_0 - v_r}{c_0 + v_r} \left(t - \tau_0 - \frac{x}{c_0} \right) \quad (1.19)$$

for the transition from u_{Tx} to u_{Rx} . The phase of the reflected signal in the reference frame of the sensor at $x = 0$, i.e., the phase of the received signal, is therefore given as

$$\varphi_{\text{Rx}}(t) = 2\pi \frac{c_0 - v_r}{c_0 + v_r} \left(f_c (t - \tau_0) + \frac{c_0 - v_r}{c_0 + v_r} \frac{B}{2T} (t - \tau_0)^2 \right). \quad (1.20)$$

REMARK. Deviding equation (1.20) by 2π and performing a time derivative, one obtains the instantaneous receive frequency

$$f_{\text{Rx}}(t) = \frac{c_0 - v_r}{c_0 + v_r} \left(f_c + \frac{c_0 - v_r}{c_0 + v_r} \frac{B}{T} (t - \tau_0) \right)$$

Now, one clearly sees that the Doppler shift affects the current transmit frequency on the one hand, but also the chirp frequency $1/T$ on the other. Consequently, the duration of the chirp also changes (see Fig. 1.3). The latter two facts are usually neglected.

To obtain a concise, closed form for the intermediate frequency signal u_{Δ} , again several first-order Taylor approximations are necessary. An extensive calculation then yields

$$u_{\Delta}(t) \approx u \cos \left(2\pi \left(\frac{B}{T} \tau_0 t + f_c \tau_0 - \frac{B}{2T} \tau_0^2 + \frac{2v_r}{c_0} f_c (t - \tau_0) + \frac{4v_r}{c_0} \frac{B}{2T} (t - \tau_0)^2 \right) \right) \quad (1.21)$$

for a single chirp. In the limiting case $v_r \rightarrow 0$, one obtains the already known formula for the static case presented in equation (1.14). In practice, an additional approximation is usually performed, and hence one usually considers

$$u_{\Delta}(t) \approx u \cos \left(2\pi \left(\left(\frac{B}{T} \tau_0 + \frac{2v_r}{c_0} f_c \right) t + f_c \tau_0 \right) \right) \quad (1.22)$$

only. For example, this approach is also used in [8, p. 49 ff.] but derived differently (see the following remark). The intermediate frequency signal u_Δ contains not only a time of flight dependent part, but also an additional frequency shift due to the Doppler effect¹. Therefore, the objective of radar signal processing is to separate the time-of-flight component from the Doppler component as precisely as possible in order to enable simultaneous measurement of distance and velocity. Subsequent to the basics of angle estimation, the signal processing of modern radar sensors is explained taking this derivation into account.

REMARK. The formula given in equation (1.22) is usually derived by simpler methods than a Lorentz transformation, see [8, p. 49 ff.] for example. However, from the physical point of view it is not entirely correct. Nevertheless, since it is common practice, a possible procedure is briefly presented here.

Suppose the target moves with constant radial speed, then the distance R changes linearly and so does the runtime τ_0 . Hence the time-dependent runtime is given by

$$\tau(t) = \tau_0 + \frac{2v_r}{c_0}t. \quad (1.23)$$

If one proceeds analogously to the static case as described in equation (1.11) and uses the above relation during integration, one obtains the following approximation

$$\begin{aligned} \varphi_{\text{Rx}}(t) &= 2\pi \int_{\tau(t)}^t f_{\text{Rx}}(u) du \\ &\approx 2\pi \left(\left(1 - \frac{2v_r}{c_0}\right) \frac{B}{2T} t^2 + \left(\left(1 - \frac{2v_r}{c_0}\right) f_c - \frac{B}{T} \tau_0 \right) t - f_c \tau_0 \right), \end{aligned} \quad (1.24)$$

if terms of order v_r^2/c_0^2 , v_r^3/c_0^3 and τ_0^2 are neglected. After calculating the phase difference as indicated in (1.13) and neglecting terms of order t^2 , one obtains the the same intermediate frequency signal for the dynamic case as presented in equation (1.22).

1.2.3 Angle estimation with multiple receivers

Measuring the speed and distance of a target vehicle is still not sufficient for an environmental perception. In addition, an angular measurement in azimuthal direction is also necessary. In order to evaluate time-of-flight dependent differences and thus determine the direction of incidence, radar systems require at least two receiving antennas. Assuming a reflective target is located in polar coordinates at (R_0, θ) with $\theta > 0$, the reflected signal will reach the left receiving antenna first. Although the time difference is minimal due to the speed of light, the phase of the received signal changes (see Fig. 1.4). The ratio of the phase difference $\Delta\varphi$ and 2π matches the ratio of the additional path length ΔR and the wavelength λ , i.e.,

$$\frac{\Delta\varphi}{2\pi} = \frac{\Delta R}{\lambda}. \quad (1.25)$$

¹Using equation (1.22), the non-parallelity of transmitted and received chirps (see Fig. 1.3) is neglected.

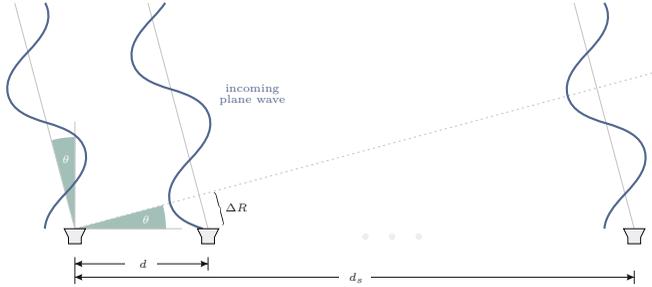


Fig. 1.4. Illustration of the angle measurement principle. The antenna distance d_s of the fictitious receiving element on the right side is chosen in a way that a phase difference of 2π , respectively λ arises compared to the first receiving antenna. This argument is used within the signal processing subsection.

The additional path depends in good approximation (plane wavefront) only on the angle of incidence θ and the spacing between the antennas d . Thereby

$$\Delta R = d \sin \theta \quad (1.26)$$

applies. Since the phase difference must be in the range from $-\pi$ to π , the antenna spacing is also directly related to the unique angular range of the sensor. For a common antenna spacing of $\lambda/2$, a maximum range of -90° to 90° is obtained due to

$$\alpha = \sin^{-1} \left(\frac{\lambda \Delta \varphi}{d \cdot 2\pi} \right). \quad (1.27)$$

Due to the non-linearity, however, a significantly larger angular error results at the boundaries of the range of uniqueness even in case of slightest phase measurement errors.

This simple principle is already sufficient for determining the angle of a single target. In practice, though, often more than two antennas are used. The exact way in which the phase difference is measured is discussed in the following section.

1.2.4 Radar signal processing and point clouds

The purpose of signal processing is to determine the distance, the velocity and the incidence angle of a target. Essentially, the signal processing is based on the frequency analysis by means of Fourier transformations, which is explained step by step in the following. The subsequent detection is generated with a constant false alarm rate algorithm, which will be briefly discussed afterwards.

Frequency analysis with Fourier transformations. If the signal from equation (1.21), resp. (1.22) is received by the receiving antenna, the analog signal is sampled at a frequency f_{AD} using an analog-to-digital converter. Therefore, a chirp of length T