17 Principles of Ultra-Wideband Sensor Electronics

Jürgen Sachs

Technische Universität Ilmenau, Ilmenau, Germany

17.1 Introduction

The ultra-wideband (UWB) technique will open a new perspective in microwave moisture sensing. This chapter intends to introduce the readers to this technique and to familiarize them with the appropriate electronics and related aspects.

Microwave moisture sensors take advantage of the strong influence of water on the propagation of an electromagnetic field. Thus the moisture content can be determined by measuring the propagation and development of an alternating electromagnetic field. However, not only water but also other substances, the material constitution (homogeneity, granular size, particle inclusions, material texture, etc.) and the geometry of the arrangement under test (size and shape of the test object, surface properties of the test object, object sensor distance, etc.) influence the electromagnetic sounding fields. One tries to exclude these troubling effects in classical microwave moisture sensors by calibration referring to the material under test and by a fixed geometry of the exposed volume. But increasing claims referring to application flexibility and robustness against troubling factors require improved sensor principles.

The measurement of moisture by microwaves is an indirect method. The moisture content is deduced based on the behavior of the electromagnetic field which is generated by a known source and influenced by an unknown object within its propagation path. That means it can be concluded from knowledge (measurement) of the resulting field as to the reason (i.e., the moisture) which made the field behave as it did. By generalization, this kind of question is called an inverse problem. It is known that the solution of such problems gives ambiguous results under most experimental conditions. From a mathematical point of view, inverse problems are ill-conditioned and ill-posed, since the quantity and quality of the available data are not sufficient to solve the problem in an unambiguous way. But practical experience gives rise to some optimism. The probability of finding an adequate solution of an inverse problem increases with the availability of better information (measurement data) about the test objects. Furthermore, a situation-specific model (adapted to the present test scenario) is

needed in order to be able to extract the required information from the measurement data.

Translated to the task of microwave moisture sensing, the message from the above details can be summarized by the following three points:

1. *Need for a sufficient quality and quantity of measurement data*: Starting from a general viewpoint, the information content gained from a pure microwave measurement is determined by the frequency diversity, by the diversity of sensor locations, and by the polarization diversity. That means, for example, that by measuring over a large bandwidth there is a better chance to separate different substances due to their specific dispersion or to separate volume effects (moisture, grain size, etc.) and boundary effects (size and shape of test arrangement, surface structure). Or it could be useful to gather data from several sensor positions (i.e., sensor arrays) under different polarizations in the case of inhomogeneous objects, objects with limited dimensions, or a rough surface. Consequently a sophisticated sensor arrangement might be a wideband system and it should cover lots of measurement channels to operate with sensor arrays which could be fully polarimetric.
2. *Need for an adequate model of the test scenario*: In order to be able to extract the required information from the measurement data, a model of the test scenario is necessary. The model can be a purely formal or physical-based system of mathematical relations. The key issue of modeling is a set of (desirably few) parameters which should describe the behavior of the real object as it was measured. The values of these parameters are the information searched for, i.e., moisture content, material density, etc. Thus, the task of the model is to link the available (measurement) data with the model parameters of interest.
3. *Need for methods to extract the model parameters from the measurement data*: UWB models are of course more complex than the usual narrowband models. With the further improvement of algorithms to solve inverse problems and the permanent growth of computing power, both the methodical and technical prerequisites are given as a matter of principle to extract the required information from a huge amount of data. The selection of an appropriate solver depends on the data throughput, the model structure, and the restrictions regarding computing power (costs, size, and power consumption of the system) and the available computing time (e.g., the control of a fast industrial process).

In summary, the improvement in sensor performance is based on the technical improvements of measurement and computing electronics as well as the improved capabilities of modeling and parameter extraction. Currently, the major handicap in the volume application of the UWB method for moisture sensing is seen by the lack of economical solutions for the appropriate electronics. In what follows, this chapter will concentrate on this topic in order to give a perspective for future sensor developments.

The UWB term is defined by the fractional bandwidth b, which has to be between 25% and 200% for a UWB system

$$b = \frac{f_u - f_l}{\frac{1}{2}(f_u + f_l)} 100\% . \tag{17.1}$$

Here f_u and f_l refer to the upper and lower 10 dB bound of the occupied spectrum for the stimulus signal. Following FCC rules [17.1], the UWB term (cf. Eq. (17.1)) is limited to an instantaneous spectrum. This fact may be important for regulation issues but not for the function of moisture sensors. Hence it will disregarded in what follows.

The classical device which covers a large bandwidth is the network analyzer. For moisture sensing, it is a useful device for research and development purposes but it is useless in industrial volume applications due to its cost, size, weight, and power consumption. The goal of this chapter is to show that there are some interesting concepts which can replace this device with economically interesting solutions and to give the readers the opportunity to select the best principle for their purpose.

Before some of these device concepts are introduced, a short review of UWB system theory will be given. It will be followed by a summary of practical constraints which UWB arrangements underlie. The basic concepts of UWB systems architecture will be explained and three of the most suitable (concerning volume applications) device concepts will be presented. Finally some examples will give an impression of the utility of the UWB technique for moisture sensing.

17.2 Basics of UWB System Theory

The basic parts of an arrangement for microwave moisture sensing are built from the material under test (MUT), the applicators – the real sensor elements providing and capturing the sounding fields – and the RF-electronics as well as the signal processing unit. There are lots of different measurement arrangements and applicator principles in use. But in all cases, the MUT applicator arrangements may be considered as a one-, two-, or n-port system which we call the device under test (DUT). Figure 17.1 schematically symbolizes such an arrangement.

At any port, two different signals are needed to describe completely the energetic interaction with the DUT. These signals are either a voltage and a current or an ingoing and outgoing wave each referred to a defined location (measurement plane). Half of these signals can be chosen independently. They are called input signals, stimulation signals, or perturbation signals. The ingoing waves a_i are mostly considered as these stimulation signals in the microwave technique. The remaining signals, the so-called system response (usually the outgoing waves b_j), depend naturally on the stimulation signals but also on the behavior of the DUT. Thus there is some possibility to determine the properties of the DUT by measuring these signals.

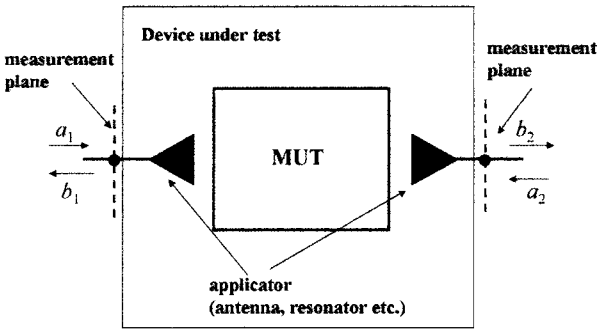

Fig. 17.1. Schematic arrangement of a microwave moisture sensor forming an electrical two port. The port signals applied here are the normalized waves a and b as usual in microwave techniques.

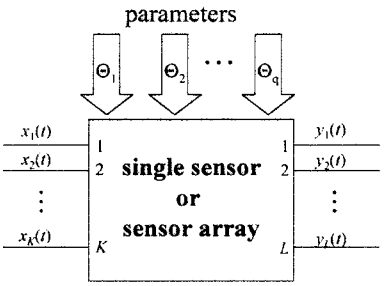

Fig. 17.2. Sensor arrangement symbolized as MiMo system having K inputs (stimulation points) and L outputs (measurement points). In the case of microwave moisture sensors, the input variables x_k usually correspond to the ingoing waves a_k and y_l correspond to b_l. The parameters Θ_q refer to the quantities which influence the properties of the DUT. Examples of such parameters are the moisture, the material density, the temperature, the salt content, etc.

Mathematically, the relation between the stimulation signals and the response signals via the DUT behavior is described by a multiple-input – multiple-output (MiMo) system as symbolized in Fig.17.2. In order to abstract from a specific physical meaning of the stimulation signals, they are described as $x_i(t)$ and the response signals as $y_j(t)$. An arrangement of n applicators provides a maximum of n stimulation and n response signals. But all these signals must not or can not be used so, in general, one has to deal with an $L{\cdot}K$ system having L output and K input variables.

In general, every output signal $y_l(t)$ of a MiMo system depends on every input signal $x_k(t)$ and the behavior of the DUT which is influenced by the parameters Θ_q. This can be expressed by the following generic relation:

$$y_l(t) = f_l\left(t, x_1(t), \cdots, x_k(t), \Theta_1 \cdots \Theta_q\right). \tag{17.2}$$

Equation (17.2) simplifies greatly if time invariance and linearity of the DUT can be supposed. Time invariance means that the dynamic behavior of the system merely depends on an "internal propagation time" but not on the ("absolute") observation time. Of course, this is not a completely exact assumption but for practical purposes it can be accepted as long as the DUT does not (only slightly) change its behavior during the measurement time. Therefore, highly variable phenomena require a short measurement time. Linearity holds if the superposition of signals is valid. This is generally the case as long as limiting or hysteresis effects do not occur and if no semiconducting pn-transitions are involved within the DUT. For the sensors we consider, the use of low peak power stimulus signals usually prevents non-linearities.

Under the constraints mentioned above, an arbitrary signal $x_k(t)$ may be decomposed into a series of standard functions such as Dirac pulses, sine waves, or step functions. Thus the stimulation effect of every sub-component of every variable $x_k(t)$ on the DUT may be considered independently of each other. In the case of a Dirac-pulse decomposition – often described as "time domain consideration" – Eq. (17.2) transforms into a set of convolutions which may be expressed in matrix form:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_L(t) \end{bmatrix} = \begin{bmatrix} h_{11}(t, \Theta_1 \cdots \Theta_q) & h_{12}(t, \Theta_1 \cdots \Theta_q) & \cdots & h_{1K}(t, \Theta_1 \cdots \Theta_q) \\ h_{21}(t, \Theta_1 \cdots \Theta_q) & h_{22}(t, \Theta_1 \cdots \Theta_q) & \cdots & h_{2K}(t, \Theta_1 \cdots \Theta_q) \\ \vdots & \vdots & \ddots & \vdots \\ h_{L1}(t, \Theta_1 \cdots \Theta_q) & h_{L2}(t, \Theta_1 \cdots \Theta_q) & \cdots & h_{LK}(t, \Theta_1 \cdots \Theta_q) \end{bmatrix} * \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_K(t) \end{bmatrix} \tag{17.3}$$

$$\mathbf{y}(t) = \mathbf{h}\left(t, \Theta_1 \cdots \Theta_q\right) * \mathbf{x}(t).$$

Here the asterisk * symbolizes the convolution product and $h_{lk}(t, \Theta_1 \cdots \Theta_q)$ is called the impulse response function (IRF). The IRF $h_{lk}(t, \Theta_1 \cdots \Theta_q)$ corresponds to the system response $y_l(t)$ at output l by stimulating the input k with $x_k(t)$. Consequently, a MiMo system is characterized by a set of $L \cdot K$ time functions that have to be measured.

The determination of the whole number of IRFs requires a minimum of K steps. Successively, within every step j, only one of the input channels will be stimulated by a short pulse $x_{kj}(t)$. The corresponding output signals are recorded. If they are additionally indexed by the number j of measurement, a new matrix equation (17.4) can be built, which provides all the required IRFs of the MiMo system. Particularly if the stimulation signals are close to Dirac pulses $\delta(t)$, the time shape of the IRFs directly correspond to the measured output signals $y_{lj}(t)$ (see Eq. (17.4), where **I** is the identity matrix):

$$\begin{bmatrix} y_{11}(t) & y_{12}(t) & \cdots & y_{1K}(t) \\ y_{21}(t) & y_{22}(t) & \cdots & y_{2K}(t) \\ \vdots & \vdots & \ddots & \vdots \\ y_{L1}(t) & y_{L2}(t) & \cdots & y_{LK}(t) \end{bmatrix} = \begin{bmatrix} h_{11}(t) & h_{12}(t) & \cdots & h_{1K}(t) \\ h_{21}(t) & h_{22}(t) & \cdots & h_{2K}(t) \\ \vdots & \vdots & \ddots & \vdots \\ h_{L1}(t) & h_{L2}(t) & \cdots & h_{LK}(t) \end{bmatrix} * \begin{bmatrix} x_{11}(t) & 0 & \cdots & 0 \\ 0 & x_{22}(t) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{KK}(t) \end{bmatrix}$$

$$\mathbf{y}_m(t) = \mathbf{h}(t) * \mathbf{x}_m(t) \tag{17.4}$$

$$\mathbf{y}_m(t) \propto \mathbf{h}(t) \qquad \text{with} \quad \mathbf{x}_m(t) = \delta(t)\,\mathbf{I}\,.$$

The set of $L \cdot K$ IRFs covers all accessible information about the behavior of the DUT, since a Dirac pulse is able to stimulate all system states which may be externally stimulated. These functions form the database to determine the required parameters, i.e., the moisture etc. via an appropriate model and inversion algorithm. However, since this chapter is mainly directed toward measurement problems rather than inversion problems, the parameter dependence of the IRFs will be omitted in what follows (and even in Eq. (17.4)) in order to simplify the notation of the equations. But it should be stressed that the actual goal is directed toward the extraction these parameters.

A second possibility to characterize the MiMo behavior is to decompose the time signals at the different ports into a series of sine waves whose amplitude and phase are given by the complex values

$$\underline{X}_k(f) = X_k(f)\mathrm{e}^{j\varphi_k(f)} \text{ and } \underline{Y}_l(f) = Y_l(f)\mathrm{e}^{j\gamma_l(f)}$$

for each frequency component. In equivalence to Eqs. (17.3) and (17.4), this results in a simple matrix product omitting any convolution:

$$\underline{\mathbf{Y}}(f) = \underline{\mathbf{H}}(f) \cdot \underline{\mathbf{X}}(f) \tag{17.5}$$

and for the representation of the K measurements

$$\underline{\mathbf{Y}}_m(f) = \underline{\mathbf{H}}(f) \cdot \underline{\mathbf{X}}_m(f) \tag{17.6}$$

which simplifies to

$$\underline{\mathbf{Y}}_m(f) \sim \underline{\mathbf{H}}(f) \quad \text{if} \quad \underline{\mathbf{X}}_m(f) = X_0\,\mathbf{I}\,. \tag{17.7}$$

Equation (17.7) indicates the usual measurement procedure, in which every input channel is individually stimulated by a sine wave of constant power. Its phase is considered as a reference phase and its frequency is stepped over the band of

interest. The function $\underline{H}_{lk}(f)$ is called the frequency response function (FRF). It includes all electrically accessible information about the DUT as does the IRF. In microwave measurements the FRFs often relate to the scattering matrix.

Equations (17.3) and (17.5) as well as (17.4) and (17.6) can be transformed mutually by the Fourier transform. Thus there is no difference between both measurements concerning the information content about the DUT. Finally a third method is partially in use, which applies a decomposition into step functions. Correspondingly the characteristic system function is called the step response function (SRF). It can be measured in an analogous way to the IRF. But it can also be calculated from the IRF by integration; thus the SRF also does not provide any new information on the DUT.

In summary, the starting point to determine the parameters of the MUT like the moister etc. is knowledge of the characteristic functions of the DUT. These are the IRF, the FRF, and the SRF. Theoretically there is no preference of any kind with these functions since they all include the same information. But sometimes it is usual to have them all, because some effects can be better interpreted by the IRF while others are easier to see with the FRF or SRF. In practice this means measuring one of them and providing the others by transformation. This is a straightforward method. But some attention has to be paid in order to avoid errors. Some problems which can arise will be discussed shortly in the following section.

Finally it should be mentioned that a sensor arrangement having n applicators can provide n^2 characteristic functions in total, from which $(n - 1)n/2$ functions are redundant because of the reciprocity of the DUT.

17.3 Practical Constraints of UWB Measurements

The previous considerations are based on very theoretical assumptions which usually cannot be fulfilled exactly in reality. In order to minimize errors, some basic rules should therefore be followed. In what follows, they will be considered using a single-input – single-output DUT for simplicity.

Assume a bandwidth B_{DUT} of the DUT and a length of its IRF of T_{DUT}. Note that the B_{DUT} value in our sense does not refer to the 3 dB value, rather it means the highest frequency component which is not yet buried by noise. This also holds for T_{DUT}, which implies that the IRF cannot be considered to be terminated before all signal components have died out, i.e., they are covered by random noise. Under the following measurement conditions, the characteristic functions will be well represented:

$$B_x \approx \frac{1}{\tau_x} \geq B_{\mathrm{DUT}} \tag{17.8}$$

$$f_{\mathrm{s}} \geq 2\, B_{\mathrm{DUT}} \tag{17.9}$$

$$T_y \geq T_{\mathrm{DUT}}\,. \tag{17.10}$$

Here B_x is the 3 dB bandwidth of the stimulus and τ_x refers to its impulse width, f_s is the sampling rate of the receiver, and T_y the record length of the captured signal. From this it can be seen that the data from any the characteristic functions covers

$$N = T_y\, f_s \geq 2\, T_{\mathrm{DUT}}\, B_{\mathrm{DUT}} \tag{17.11}$$

data points. Thus the IRF or the SRF is a "chain" of N (real-valued) data points sampled at a time interval $t_s = 1/f_s$,whereas the FRF represents a series of $N/2$ complex-valued (non-redundant) data points determined at a regular frequency grid spanned from DC to $f_s/2$ by increments of $\Delta f = 1/T_y$.

For time domain measurements within the microwave range, the sampling theorem Eq. (17.9) burdens the electronics with extremely high measurement rates. But, fortunately, the system response can also be measured piecewise by repetitive stimulation with the same signal as long as the DUT does not change its properties. Such a procedure is described as under-sampling where only a subset of data points is gathered by one stimulation. Referring to n data points captured at every stimulation, the actual sampling rate f_{sa} is given by

$$f_{\mathrm{sa}} = \frac{f_s}{u_{\mathrm{sf}}} \quad \text{with} \quad u_{\mathrm{sf}} = \frac{N}{n}\,. \tag{17.12}$$

The value u_{sf} represents the under-sampling ratio and f_s is now called the equivalent sampling rate. In order to avoid aliasing, the equivalent sampling rate must respect the sampling theorem Eq. (17.9). However, the real sampling rate f_{sa} may be quite low and easy to handle with technical systems. This will expand the time T_{obs} which is necessary to observe the DUT, i.e., to complete the data set. Supposing a repetition rate f_0 of the stimulus, the observation time will be

$$T_{\mathrm{obs}} = T_0\, u_{\mathrm{sf}} \quad \text{with} \quad T_0 = 1/f_0 \geq T_y \geq T_{\mathrm{Dut}}\,. \tag{17.13}$$

Note that the period length T_0 of the stimulus must be larger than the settling time T_{DUT} of the DUT, otherwise time aliasing arises. The maximum value of the under-sampling ratio u_{sf} is limited by the non-stationarity of the DUT. The most critical non-stationarity in moisture measurements is certainly given by fast-moving MUTs. The wideband ambiguity function is one way to evaluate the sensor performance under such conditions. Corresponding considerations are showing that the measurement arrangement behaves nearly stationary, i.e., time invariant if

$$4\,T_{\mathrm{obs}}\,|v_{\max}| < \frac{c}{B_{\mathrm{DUT}}}. \tag{17.14}$$

Thus the under-sampling factor u_{sf} must be less than

$$u_{\mathrm{sf}} < \frac{c}{8|v_{\max}|N}. \tag{17.15}$$

Here $v_{\max}$ and c stand for the maximum speed of the MUT and the speed of light respectively. Equation (17.14) holds independently of the measurement principle. That means the sine wave sweep technique underlies the same requirements. Here, T_{obs} corresponds to the time needed for one sweep over the frequency band of interest.

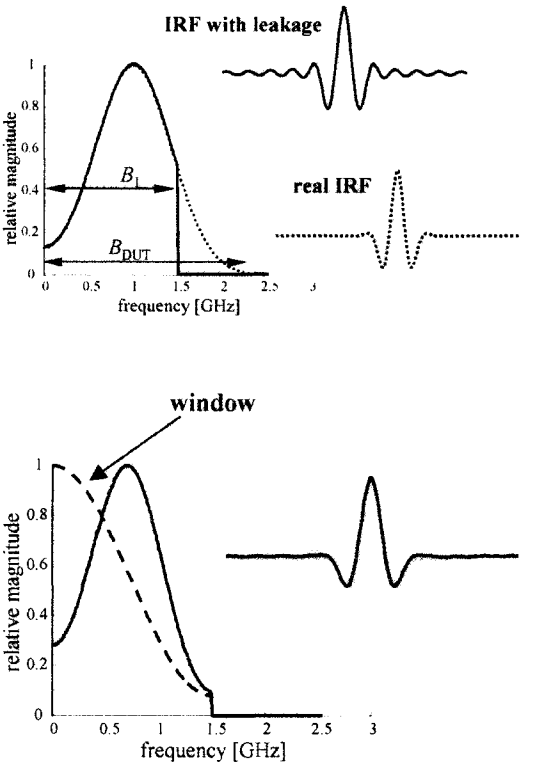

Fig. 17.3. Leakage effect. *Top*: ringing on the IRF by leakage (dotted line: real FRF and IRF of the DUT; solid line: measured part of FRF and corresponding IRF gained from the Fourier transform). *Bottom*: reduced ringing by windowing (dotted line: applied window function; solid line: windowed measurements and resulting IRF)

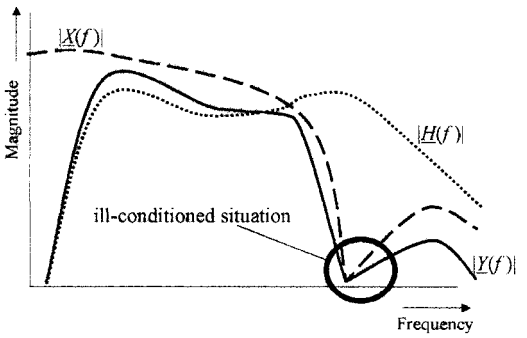

Fig. 17.4. Ill-conditioned deconvolution caused by zeros in the stimulation spectrum (dotted line: real FRF; dashed line: spectrum of stimulus; solid line: spectrum of measured signal)

By disregarding the condition in Eq. (17.8) two problems arise. Consider first a sine wave, i.e., a frequency domain measurement which covers the bandwidth $B_1 < B_{DUT}$ (see Fig. 17.3). Thus some data of the FRF missed. By transforming this reduced data set to the time domain in order to get the IRF, the abrupt cut will be "interpreted by the Fourier transform" as a steep filter flank causing a sin x/x ringing in the IRF. However, a window function can be applied to the frequency data to smooth the cut resulting in a more friendly IRF. But note that windowing cause data loss, reducing the time resolution of the IRF by widening the main lobe. So a reasonable compromise must be found.

The second problem arises if the deviations of the stimulus signal referring to a Dirac pulse cannot be tolerated. Then the stimulus $x_m(t)$ must be deconvolved from the measurement signal $y_m(t)$ in order to gain the IRF $h(t)$. The easiest way to do this is to transform the stimulus and measurement signal into the frequency domain, to divide their (complex) spectra, and to re-transform the quotient back to the time domain. Thus,

$$h(t) = \mathrm{IFT}\left\{\frac{\mathrm{FT}\{y_m\}}{\mathrm{FT}\{x_m\}}\right\}. \tag{17.16}$$

FT{} represents the Fourier transform and IFT{} the inverse Fourier transform.

But unfortunately deconvolution often represents an ill-conditioned problem. Figure 17.4 demonstrates the difficulties. Suppose that the spectral energy of an (arbitrary) stimulus is not regularly spread over the frequency band of interest (i.e., B_{DUT}). Obviously it is difficult to divide the spectra at those spectral segments where their values are close to zero. Mathematically the term 0/0 is indefinite, which practically means that noise will determine the result. Although no energy appears in the process at these frequencies, there may be huge values in the FRF caused by the uncertainties of the division. By transforming such an FRF back to the time domain, usually those quantities dominate the shape of the whole IRF. They must be suppressed by appropriate methods.

17.4 Measurement Errors

Electronic measurement devices are always associated with errors. One distinguishes several types of errors which differ by their temporal behavior. Additive noise and jitter are short-term perturbations, drift refers to medium-term variation of the system properties, and systematic errors are stable deviations from the desired conduct. Particularly due to the large bandwidth, the real challenge of the UWB system design is the error handling. In what follows, the different error types will be discussed and some general measures will be given to suppress them.

17.4.1 Additive Random Noise

Random noise is considered as a stochastic short-term variation of the amplitude values of a signal. Caused by the large bandwidth, the noise rejection of UWB systems is very poor. That is why proper design and signal processing must suppress the noise as much as possible. The first measure is to restrict the bandwidth to the actual needs by short time integration of the measurement signal (see Fig. 17.5). The integration window either continuously slides over the waveform (low-pass filter) or is moved in steps (S&H circuit with an appropriate sample time). A short time integrator reduces noise at the expense of bandwidth. In an ideal case it should limit the bandwidth of the receiver to the bandwidth of the DUT. The remaining variance of the captured waveform is approximately

$$\sigma_T^2 \approx \frac{N_x}{T} \approx N_x B. \tag{17.17}$$

Here σ_T is the standard deviation of the captured signal, and N_x the power spectral density of the noise (which was supposed to be white). The bandwidth B of the receiver is given by the inverse of the integration time.

A second measure particularly in the case of a periodic stimulation is synchronous averaging as demonstrated in Fig. 17.6. It does not affect the resulting bandwidth of the recorded waveform as long as only additive noise is perturbing the signal, though in the presence of jitter or time drift, the resulting bandwidth will decrease (see Chap. 17.4.2). The improvement of the noise performance is proportional to the square root of the averaging number m:

$$\sigma_{\mathrm{av}}^2 \approx \frac{\sigma^2}{m}. \tag{17.18}$$

Note that the observation time T_{obs} is stretched by the same factor which slows down the measurement speed.

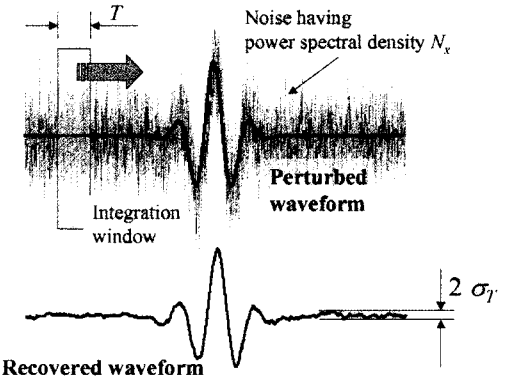

Fig. 17.5. Noise reduction by short time integration. An integration window is moved over the waveform smoothing out short noise peaks

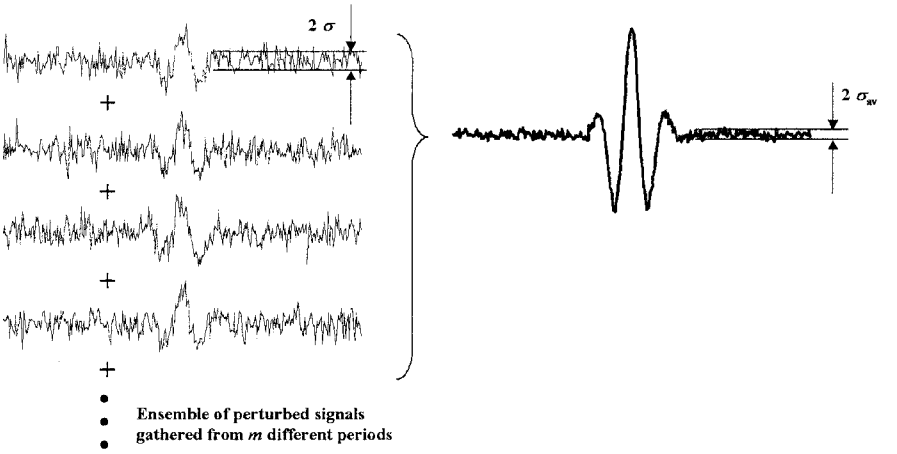

Fig. 17.6. Noise reduction by synchronous averaging

The problem of random noise should, however, not be reduced to the question of how to suppress it. Rather the question is more how to maximize the signal-to-noise ratio *SNR*. This quantity represents a central point. It influences all the other aspects of the electronic design and behavior. The *SNR* value is given by

$$SNR = \frac{E}{N_x}. \tag{17.19}$$

This states that the receiver must accumulate a certain quantity E of coherent energy connected with the measurement process under consideration. The accumulated energy depends on the average signal power P_{av}, the accumulation time τ_{ac}, as well as on an efficiency factor η of the receiver:

$$E = \eta P_{\mathrm{av}} \tau_{\mathrm{ac}}. \tag{17.20}$$

The higher the average signal power P_{av} of the stimulus signal, the shorter the accumulation time τ_{ac} can be for a given *SNR*. Accumulation time in that sense is another expression for measurement and observation time T_{obs}. Among other things, the receiver efficiency η depends on the under-sampling ratio $\eta \sim 1/u_{\mathrm{sf}}$. Only a small part of the available signal energy is really included in the measurement data if under-sampling is applied. Thus a high under-sampling ratio may have a negative impact on the *SNR* value.

Whereas the average power accounts for the *SNR* value, the peak power P_{peak} is usually the limiting factor due to the electronics. The crest factor *CF* of a signal is a measure to characterize the uniformity of the instantaneous power distribution:

$$CF = \frac{\hat{X}}{X_{\mathrm{rms}}} = \sqrt{\frac{P_{\mathrm{peak}}}{P_{\mathrm{av}}}}. \tag{17.21}$$

Here $\hat{X}$ represents the peak value of the signal and X_{rms} is its RMS value. The lower the crest factor, the more evenly the instantaneous signal power is distributed. This means that fewer the systems are stressed by high-power "shocks." Thus, thanks to the low crest factor signal more energy can be handled by the systems without going beyond their limits.

Measurement devices based on sine the wave technique respect this fact. Pulse measurement systems do not. This is why they usually have a poor noise performance. One way out is to spread the energy of a wideband signal over a large time and to compress it again in order to get the IRF of a DUT. The convolution of Eq. (17.3) with the time-inverted and transposed stimulation vector permits such an approach:

$$\begin{aligned} \mathbf{y}(t) * \mathbf{x}^{\mathrm{T}}(-t) &= \mathbf{h}(t) * \mathbf{x}(t) * \mathbf{x}^{\mathrm{T}}(-t) \\ \mathbf{C}_{yx}(t) &= \mathbf{h}(t) * \mathbf{C}_{xx}(t) \\ \mathbf{C}_{yx}(t) &\propto \mathbf{h}(t) \quad \text{if } \mathbf{C}_{xx}(t) = \delta(t)\,\mathbf{I}. \end{aligned} \tag{17.22}$$

Taking $\lim_{T\to\infty} 1/T$ over the first line in Eq. (17.22) results in two matrices of correlation functions $\mathbf{C}_{xx}$ and $\mathbf{C}_{yx}$. The matrix $\mathbf{C}_{xx}$ forms a diagonal matrix of functions like a Dirac pulse if the stimulus signals at the different input ports are mutually un-correlated and if they occupy a large spectrum. Thus the set of IRFs is approximately given by the matrix of the cross-correlation functions $\mathbf{C}_{yx}$ which must be determined by the measurement system. Obviously, the actual signal

shape of the stimulus is of no further interest. This opens up a great deal of different device architectures.

The determination of the cross-correlation functions $\mathbf{C}_{yx}$,i.e., the IRF, is performed by a correlator and a matched filter. Figure 17.7 schematically demonstrates the procedure. A wideband waveform of any shape stimulates the DUT. Its response is processed by a matched filter or correlator having the stimulation waveform as reference. The compressed waveform represents the IRF of the DUT as long as the bandwidth of the stimulus is larger than the bandwidth of the DUT under consideration.

For estimation of the *SNR* improvement, let us consider a DUT of infinite bandwidth. Referring to Fig. 17.7 and Eq. (17.19), and taking into account the energy conservation Eq. (17.23) by the matched filter (respectively correlator)

$$E = x_{\text{rms}}^2 \, T_0 \approx \hat{Y}^2 \, t_m \approx \frac{\hat{Y}^2}{B}. \tag{17.23}$$

the improvement of the *SNR* value of the compressed waveform can be approximated by

$$SNR_{\text{out}} = \frac{\hat{Y}^2}{n_{\text{rms}}^2} = \frac{E}{N} = \frac{x_{\text{rms}}^2}{n_{\text{rms}}^2} B T_0 = SNR_{\text{in}} \, B T_0. \tag{17.24}$$

Here B is the bandwidth of the matched filter (respectively the waveform), and $n_{\text{rms}}^2 = NB$ is the total noise power. The improvement factor BT_0 is also called the correlation gain.

The aim of impulse compression is to load critical system components only by low peak power signals and to regain high peak power signals by impulse compression at less critical points within the systems. The intention is to put as much energy as possible into the measurement arrangement. Impulse compression is performed if the receive filter has an IRF which is a time inverse replica of the transmit waveform. A second possibility is to perform a cross-correlation which applies a reference signal corresponding to the desired input waveform. The output waveform is as short as the bandwidth $B \approx 1/t_m$ of the input signal is wide and the energy E stored in the filter is as big as the time T_0 of the input signal is long. Both factors determine the compression and the correlation gain.

It should be noted that impulse compression by analog circuit principles is always limited in the dynamic range because of saturation effects of the components. A compression in the software domain is less critical since an appropriate number representation can always be used. Furthermore, the quantity of energy which can be stored in an analog matched filter for the purpose of compression is also quite limited, particularly if it is spread over a wide spectrum. Finally, analog correlator principles must always fight against the inadequacy of programmable wideband delay lines. As may be seen in the subsections below, a digital principle has benefits also with respect to these two aspects.

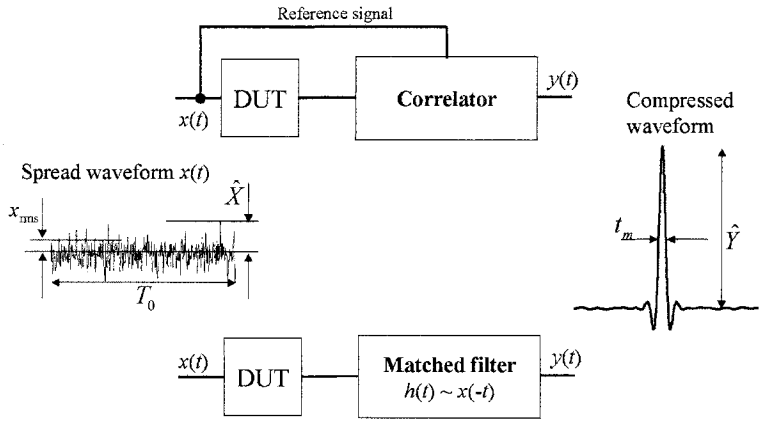

Fig. 17.7. Improvement of the *SNR* by impulse compression. If the signal $x(t)$ is misshapen by a linear system (DUT) before it enters the impulse compression system, its output signal will represent the IRF of the DUT as long as the bandwidth of $x(t)$ is wide enough

17.4.2 Jitter

Jitter is a stochastic short-term instability of the temporal position of a signal and of a sampling moment. Jitter affects signal flanks by increasing the noise but not flat signal parts. Averaging of jittered waveforms flattens impulse flanks, which means it reduces the available bandwidth (see Fig. 17.8 and Eq. (17.25), where t_r is the rise time).

$$t_{r,av}^2 \approx t_r^2 + t_j^2 \tag{17.25}$$

Jitter is caused by additive random noise within timing control circuits or oscillators. Figure 17.9 demonstrates one possibility of its origin. The quantity of trigger jitter can be estimated by Eq. (17.26). Obviously it can be largely avoided by steep flanks.

$$t_j \approx \frac{n_{rms}}{V_0} t_r. \tag{17.26}$$

A second countermeasure is to omit VCOs for clock or stimulus generation, since monochromatic frequency sources provide less phase noise, i.e., jitter, than frequency tunable circuits due to their high-quality resonators.

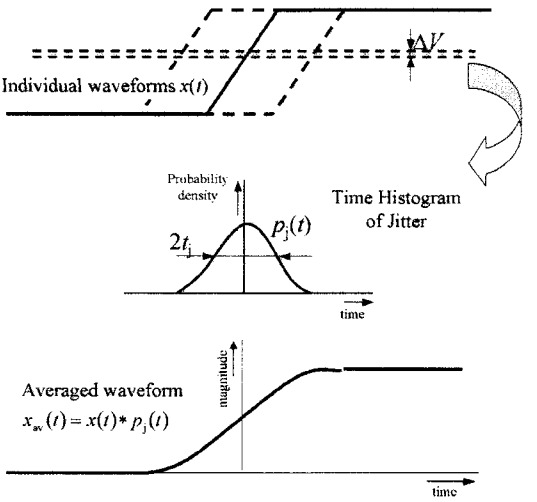

Fig. 17.8. Effect of averaging on jittered waveforms. The jitter behavior may be determined by recording a histogram over a small voltage gap

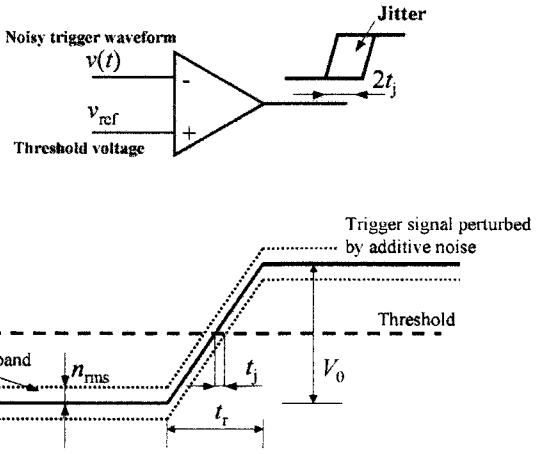

Fig. 17.9. Generation of jitter by triggering at a threshold

17.4.3 Drift

Drift is a medium-term instability which acts on both amplitude and time values. Drift limits the maximum usable number of averages and prevents correction of systematic errors in the sensor arrangement. Thus the drift behavior is a critical

parameter of UWB systems. Usually it is caused by variations of threshold or offset values and gain due to changes in temperature or supply voltage or by aging. Three types of drift can be distinguished as follows.

Time drift: Comparable to the jitter performance, time drift can be largely suppressed by the steep flanks of the control signals (compare Fig. 17.9) and a temperature-stabilized high-Q oscillator.

Offset drift: This may be critical in connection with additive sampling (see Fig. 17.14). The stability of the threshold value as well as the amplitude stability of the sampling generator translate immediately into amplitude errors. Since the threshold and sampling pulse must be larger than the measurement signal, even small, relative variations of their values cause considerable additive errors. Homodyne receivers are also sensitive to drift caused by self-mixing effects (see Fig. 17.11 and the following text).

Gain drift: Gain errors are proportional to the actual measurement value. They can be eliminated by referring the measurement signal to a reference signal (usually the stimulus). It is important that the reference signal is captured by an identically constructed receiver, because then it can be supposed that they behave in the same way.

17.4.4 Systematic Errors

Systematic errors can be classified into linear and non-linear effects. Since such errors do not change with time, they can be removed theoretically from the measurement result. The success of such an error correction largely depends on the stability of the measurement system (drift, noise, jitter), the error model, and the quality of the calibration standards.

The causes of systematic errors are manifold (mismatched ports, frequency-dependent receivers, crosstalk, limited directivity, saturation effects of amplifiers, etc.) The simplest and most promising way to deal with such errors is to consider them by a formalized strategy, i.e., to model the real measurement system and sensor arrangement by an ideal device cascaded by an error box including all errors independently of their cause. Following Fig. 17.10, a sensor arrangement having n applicators represents an n-port system resulting in an error box having $2n$ ports. Such a system is characterized by $4n^2$ characteristic functions as long as the errors behave linearly. However, one of them is redundant, thus $4n^2 - 1$ error terms may appear in total. If the error terms are known, one can remove them from the measured values (a_{ei}, b_{ei}) by matrix inversion.

The determination of the error terms is based on a calibration routine which uses DUTs with a known behavior – the so-called calibration standards. The simplest way is to use commercial calibration standards. But they usually refer only to standardized RF connectors. Unfortunately, that kind of calibration excludes the influence of the applicators on the systematic errors. In order also to deal with them, either a numeric applicator model is needed or the applicators must be included into the calibration routine by using appropriate calibration materials.

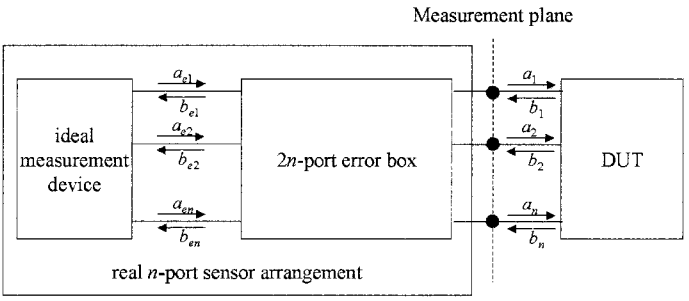

Fig. 17.10. Generalized modeling of systematic errors

Finally it should be mentioned that except for the simplest error correction methods, the full n-port matrix must be measured. Nowadays, this is not usual in microwave moisture sensors. For future UWB moisture sensors, however, it is recommended to equip their electronics with a sufficient number of measurement channels in order to be able to determine the complete n-port matrix of the DUT.

Even if it seems theoretically possible to remove all systematic errors by the above-mentioned procedure, only the most important one can be actually eliminated in practice, because the more error terms are respected the more (different) calibration standards are needed and the more reference measurements have to be done. Since every additional measurement increases the random error level of the final result, there exists an optimum number of systematic error terms that should be removed.

17.5 Architecture of UWB Systems

Concerning the electronics, a UWB sensor system is built from three parts – an RF unit which provides the stimulus signals, an RF unit which captures the measurement data, and a usually digital unit which is charged by the data processing.

The stimulation of the test objects can be done either sequentially by a swept or stepped narrowband (sine wave) source, or instantaneously by a periodic or random wideband signal. The nature of the stimulus signal has a determinant influence on the architecture and the behavior of the system.

The task of the measurement unit is to determine the response of the DUT due to its stimulation. In the case of moisture sensing, the time constants which fix shape and duration of the response function are to be found in the nano- and sub-nanosecond range. Thus a typical record length T_y of the measurement data is in the order of tens of nanoseconds. The gathering of such data streams in real time requires an extreme technical effort. Fortunately the characteristic parameters change only slowly, thus signal recording may be distributed over a longer time. This fact can be used to reduce the data gathering rate which enables less critical

and hence more cost-effective technical system implementations. There are three basic approaches to reduce the real-time requirements, which will be summarized below shortly.

Nowadays, sensor measurements include a more or less sophisticated level of digital signal processing. The burdens of the digital unit and hence its complexity and costs depend on the extent of the algorithms which extract the desired information from the measurement data as well as from the data rate. Though a low data rate system often permits cost-effective technical implementations, it should be noted that the stochastic confidence in the measurement results will degrade if data rate reduction is simply done by losing data (as in case of under sampling). Consequently, an optimum must be found for every application between the performance of the UWB system and its costs. Such an optimum solution mainly depends on the overall architecture of the UWB system and the available technology for the electronic components.

An assessment of the usability of a specific UWB principle for a certain task should cover not only technical parameters (bandwidth, dynamic range, stability, measurement speed, etc.) but also features concerning power consumption, aspects of technical implementation (size, costs, weight, level of integration) as well as the capability to form multi-channel sensor systems or to permit systematic error corrections. Certainly, in deciding in favor of a specific principle, the various features weight differently. But this consideration must be left to the user, corresponding to the actual task.

In what follows, some examples for a UWB system layout will be given including some general assessments.

17.5.1 Down-Conversion

Down-conversion is a method to shift the RF measurement signals into a lower frequency band where they can be easily handled. However, a simple down-conversion is not able to reduce the data rate. Down-conversion is often applied in connection with sine wave excitation, which finally results in a considerable reduction of measurement speed. The idea is to determine the real and imaginary parts of the FRF $\underline{H}(f_x)$ of the DUT referred to the applied excitation frequency f_x. In order to record the full FRF the sine wave sources must be (slowly) stepped and swept over the entire band of interest.

There are two basic methods – the homodyne and the heterodyne approach (see Fig. 17.11). In the case of the technically simpler homodyne receiver, the response of the DUT is mixed by the two channels of a quadrature modulator with the replica and a 90°-shifted version of the stimulus sine. If the stimulus source keeps its frequency constant, both output channels of the quadrature demodulator will provide a DC voltage whose values correspond to the real part (inphase component) and the imaginary part (quadrature component) of the FRF of the DUT. The noise suppression and hence the receiver sensitivity is performed by the low-pass filters included in the quadrature demodulator. At first glance, the homodyne concept seems to be ideal for volume applications due to its simplicity

and the absence of bulky off-chip components which are amenable to monolithic integration. However, the homodyne approach exacerbates a number of issues [2]:

- The required 90° phase shift of the reference signal within the IQ demodulator can only by handled for narrow band sweeps.
- The measurement values are superimposed by a varying DC offset originating from self-mixing of the reference signal due to leakage of the mixers.
- Even-order distortions also initiate DC perturbations.
- Any tolerance between both branches of the quadrature modulator leads to an I/Q-mismatch causing systematic errors.
- Flicker noise ($1/f$) may have a profound effect since the down-converted spectrum is located close to DC.

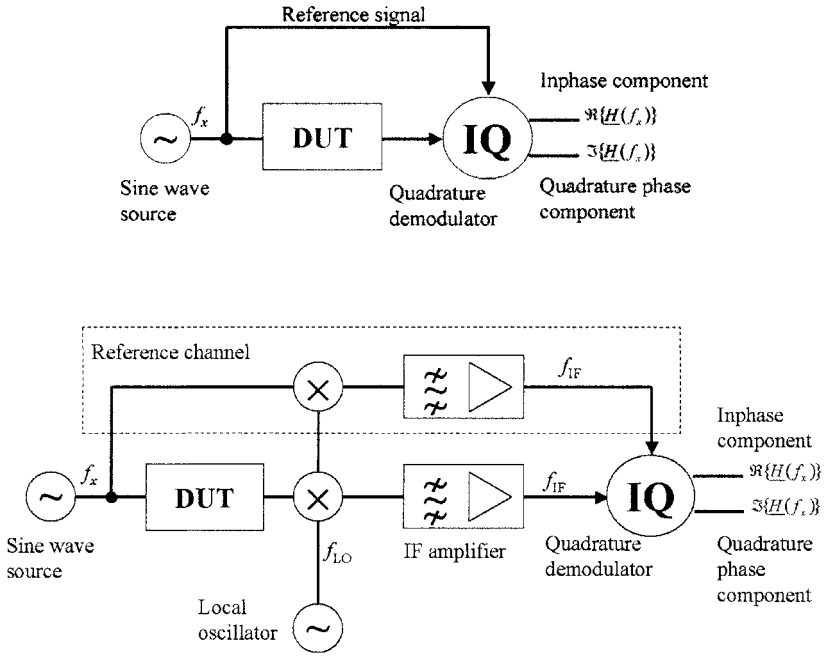

Fig. 17.11. Down-conversion by a homodyne (top) and heterodyne (bottom) approach. Note that the heterodyne concept usually includes an image rejection filter in front of the mixers. It can be dropped only if single tone signals exists as shown

The heterodyne concept avoids most of the inadequacies of the homodyne receiver by introducing an intermediate stage (see Fig. 17.11). The RF signal is mixed by the local oscillator to a fixed intermediate frequency $f_{IF} = |f_x - f_{LO}|$.

Thus narrow band filtering can be performed which effectively rejects noise and distortion. The quadrature demodulator must be designed only for a fixed frequency.

High-precision, multi-purpose UWB measurement devices as network analyzers are based on such heterodyne principles. However, their excellent features must be paid for by system complexity. Narrowband systems as IF amplifiers require high-Q elements which prevent monolithic integration. But the most critical parts in this sense are the sine wave sources (stimulus generation, local oscillator), which must be sophisticated synthesizers in order to gain both a wide sweep range and an excellent frequency stability (accuracy and phase noise).

Even if the down-conversion in connection with a swept/stepped sine wave excitation is used to leave the microwave range and to have a modest data rate, the measurement rate is often unacceptably low. Indeed, most of the time cannot be used for data gathering in waiting for synthesizer and the narrowband filters to settle down.

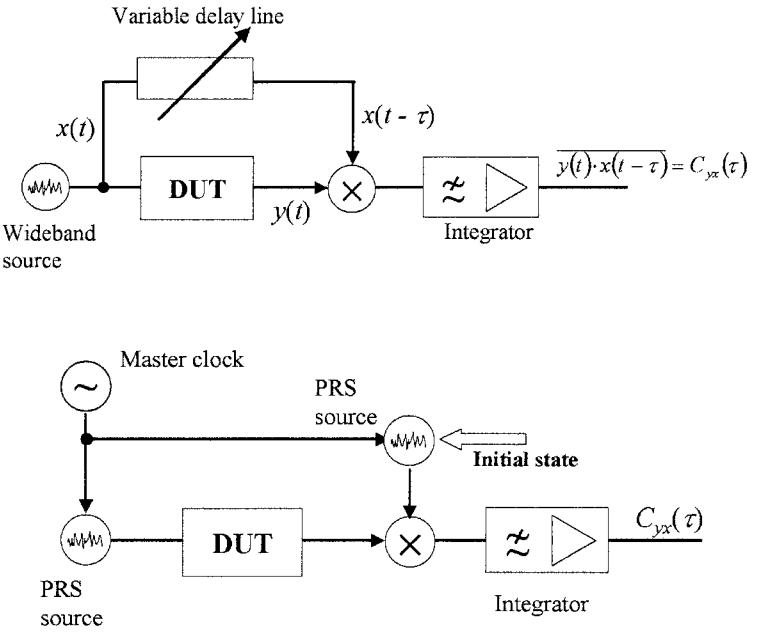

Fig. 17.12. Sliding correlator with variable delay line (top) and a reference source (bottom) which is shifted by its initial state

17.5.2 Sliding Correlation

Sliding correlation is a further method to record successively the shape of a response function of the DUT from DC values. Following Fig. 17.12 (top) and supposing a fixed delay τ_0, the correlator forms the mean value of the product between the DUT response $y(t)$ and a delayed version of the stimulus $x(t - \tau_0)$. As demonstrated by Eq. (17.22), this value corresponds to the cross-correlation $C_{yx}(\tau_0)$ at τ_0 which is equal to the IRF $h(\tau_0)$ if the DUT is stimulated by any wideband source. In order to capture the whole IRF, the delay time must be successively swept over the whole record length T_0. This is the bottleneck of the principle, since no satisfactory method is known that does this in an automatic and cost-effective way. An alternative is to use two identical pseudo random signal sources. One of them successively shifts its initial state, which has the same effect as a delay line (Fig. 17.12, bottom).

The bandwidth of the integrator affects the sensitivity, i.e., the noise rejection, and limits the measurement speed. Note that at every delay step the integrator must have settled down before the next measurement value can be captured. Furthermore the sliding correlator is sensitive to DC offset (by self-mixing and even-order distortion) and flicker noise as in the homodyne principle. Some examples of sliding correlators are to be found in [3] – [5].

17.5.3 Bandwidth Compression by Under-Sampling

Under-sampling is strictly limited to periodic (repetitive) signals. Without violating of the Nyquist theorem, the gathering of a complete set of data samples may be distributed over several periods. Hence the data gathering rate can always be broken down to a speed which does not overload the more or less slow and cost-effective digital signal capture. The sampling method reduces the number of analog components and opens up the full flexibility of the digital world as close as possible to the RF measurement ports.

Theoretically, under-sampling is nothing but a multiplication of the input signal having a repetition rate f_0 with a pin-pulse train having the rate f_{sa}. If properly done (see Fig. 17.13) this results in a signal with reduced bandwidth which can be sampled at a lower speed without disregarding the Nyquist theorem. Suppose the bandwidth of the measurement task requires an actual sampling rate of f_s. In the case of real-time sampling this would give $N = f_s/f_0$ samples per period of the measurement signal. However, it can be shown that, corresponding to Eq. (17.27) any other sampling ratio f_{sa} may also be used to collect the same samples as long as N and the under-sampling factor u_{sf} do not have a common divider.

$$u_{sf} = \frac{f_s}{f_{sa}} = N\frac{f_0}{f_{sa}} \quad (u_{sf}, N \text{ integer numbers}) \tag{17.27}$$

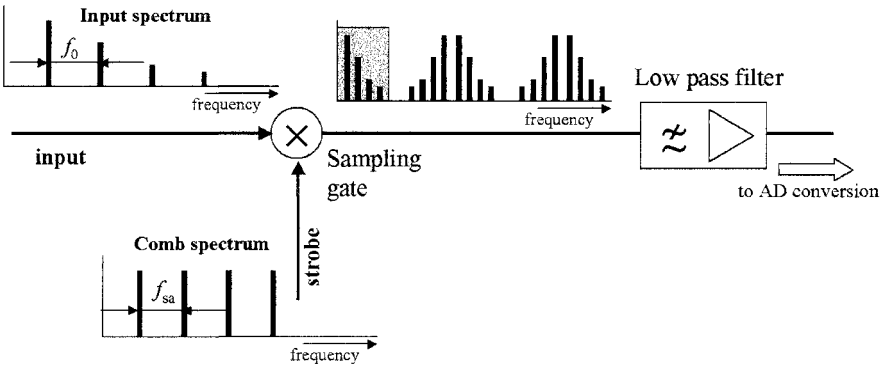

Fig. 17.13. Bandwidth compression by convolving the input spectrum with a comb spectrum. (A comb spectrum is generated by a train of short pulses)

However, the original order of spectral lines and the samples will only be respected for sequential sampling, i.e.,

$$u_{sf} = n\,N + 1 \quad (n = 1, 2, 3, \cdots). \tag{17.28}$$

In all other cases, the samples must be reordered before further processing.

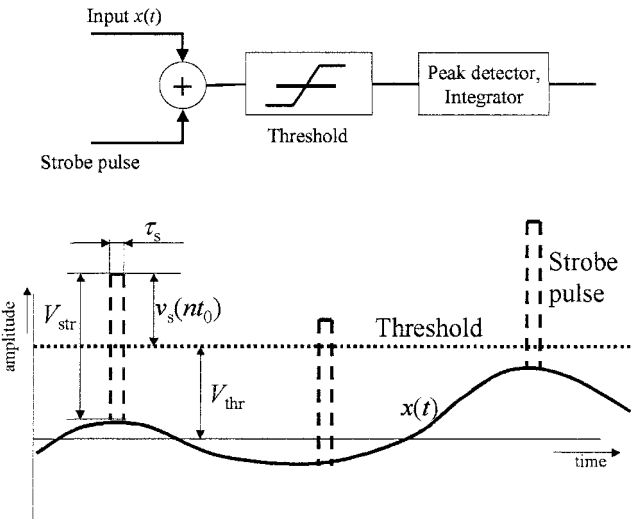

Fig. 17.14. Additive sampling. Either the peak value $v_s(nt_0)$ or the signal area (charge) beyond the threshold provides the value of interest

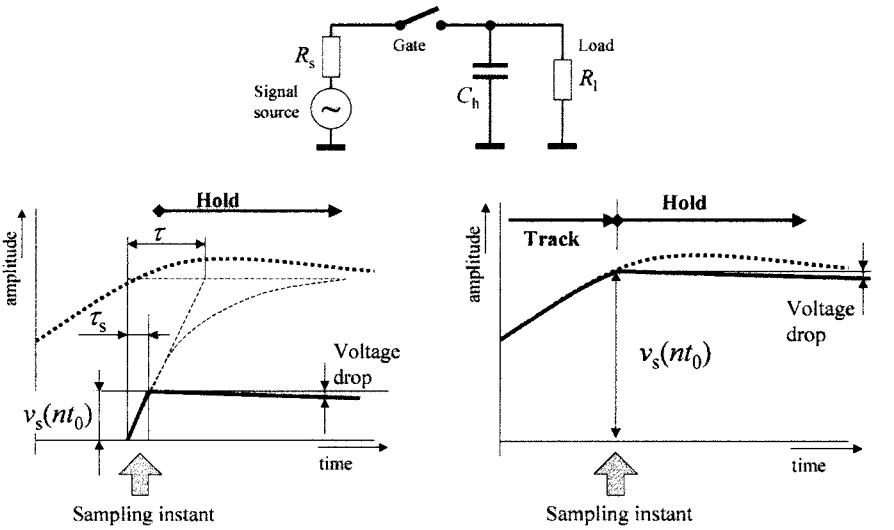

Fig. 17.15. Multiplicative sampling. Circuit schematic (top) and typical signals in the sample and hold mode (bottom left) and in the track and hold mode (bottom right)

The probing itself is done by two approaches – additive or multiplicative sampling. For additive sampling, the measurement value results only from that signal part which goes beyond a threshold (see Fig. 17.14). The threshold is chosen in such a way that the pure input signal can never pass it. An additive sampler can be implemented by very simple and power-efficient circuits [6]. However, they are strongly sensitive to threshold drift and variations of the strobe pulse magnitude.

Multiplicative sampling (Fig. 17.15) provides the measurement value by charging a hold capacitor via the sampling gate, which is controlled by the strobe pulse. Two modes of charging the capacitors are in use. In the S&H mode, the gate is only closed for a very short time τ_s. The capacitor will only be charged to a small fraction of the input amplitude. Thus the sampling efficiency will be low (only a few percent) but the operational bandwidth will be high. The opposite behavior is to be found for the T&H mode. The gate is closed over a long time so that the voltage across the capacitor follows the input signal. This will result in a higher measurement value (higher efficiency). But with increasing frequency, the signal source is no longer able to drive the capacity even if a drive amplifier is inserted. If the gate is open, the circuit is in the hold state. The capacitor voltage drops only slowly and a simple ADC is able to capture the value.

Multiplicative sampling circuits are more stable than additive sampler ones but they are also more power consuming. Sampling circuits can be built for a bandwidth from DC up to 100 GHz approximately.

17.5.4 UWB Principles for Volume Application

Costs, size, robustness, and power consumption of the sensor electronics will finally decide against or in favor of a volume application of UWB principles in moisture sensing. With the exception of laboratory tasks, sophisticated classical wideband systems such as network analyzers or TDR-sampling scopes will hardly find use in field applications. However, pushed by the progress in semiconductor technology, there exist some UWB principles which will have the power to transfer UWB moisture sensors into large-scale applications. Three of them will be discussed shortly in what follows.

FMCW principle: The FMCW method permits a simple and cost-effective system implementation. It uses a frequency-modulated continuous wave (FMCW) signal to stimulate the test objects. For that purpose a VCO is swept linearly over the frequency band of interest. The system response is transformed by the homodyne concept (see Fig. 17.11) into the low-frequency range where the signals can be easily captured by general purpose digital hardware. The commercial availability of different VCOs provides a great deal of different operational frequency bands. Due to the large bandwidth, the FMCW method has to waive the quadrature demodulation, so only the real part of the FRF can be measured. This greatly limits the signal processing since only part of the information about the DUT is available.

Let us consider for simplicity the wave transmission between two frequency-independent antennas. The complex FRF of such an arrangement looks like a screw thread line ($\underline{H}(f) \sim \exp(j2\pi f\tau)$) if visualized in 3D coordinates (real axes – imaginary axes – frequency axes). The "thread lead" depends on the traveling time τ which can be a measure of moisture, for example. The projection of such a line into the real-axes – frequency-axes plane results in a sinusoid as function of f having a period proportional to τ. Since in the case of FMCW the VCO sweeps linearly in time, frequency and time can be interchanged. Thus the measured signal is finally a "time-dependent" sinusoid whose frequency gives access to the traveling time τ of the wave. However, each deviation from a linear frequency sweep causes remarkable errors which must be suppressed by additional means [7, 8].

Pulse method: The pulse method represents another principle which profits from the simplicity of the electronic circuits. The DUT is excited by a short, i.e., wideband, pulse and the response of the DUT is usually obtained by sequential sampling (see Fig. 17.16). The pulse generation is typically based on avalanche principles, step recovery diodes, or tunnel diodes. The peak power of the stimulation pulse largely determines the *SNR* value. Very simple, power, and cost-efficient pulse modules can be built up to a few gigahertz bandwidth, particularly if they are equipped with additive samplers [6]. But such simple solutions usually suffer from some drawbacks. As a matter of principle, the strobe pulse generation tends to drift and produces jitter (compare Fig. 17.9 and Eq. (17.26)) and a non-linear ramp results in timing errors. Thus additional and expensive measures must be taken to improve the behavior. Sophisticated pulse systems are available up to 100 GHz bandwidth.

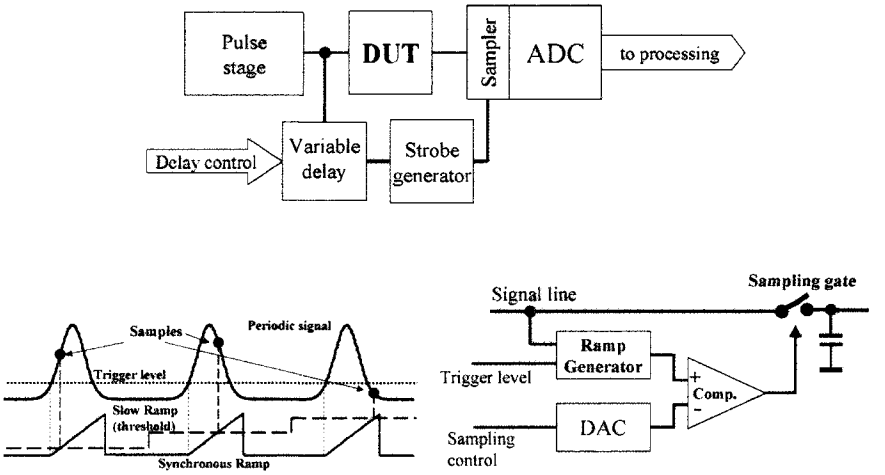

Fig. 17.16. Typical structure of a pulse system (top). The example shown uses sequential sampling for data gathering in which the strobe pulse is triggered from the stimulus via a variable delay. The variable delay is performed by sweeping a threshold (dashed line) over a voltage ramp which is synchronized with the stimulus (bottom)

MLBS approach: The direct use of a maximum length binary sequence (MLBS) with microwave frequencies in connection with under-sampling is a further approach promising superior performance at a bandwidth up to the 10 GHz range. Mainly its stable operation gives access to sophisticated data processing which is the key issue of all UWB measurement methods.

Figure 17.17 demonstrates the basic concept and a practically implemented RF part. A fast digital shift register with an appropriate feedback is able to provide an MLBS having a bandwidth up to several gigahertz. This signal stimulates the DUT. The MLBS is a periodic signal consisting of numerous pulses having a variable width. They are quasi-randomly distributed over the whole period. In contrast to the classical impulse excitation, the MLBS distributes the pulse energy over the complete measurement time. Thus the signal amplitude may be comparatively low even if a large energy is required in order to gain a certain *SNR*. Signals with low amplitudes are easy to handle and they promote monolithic circuit integration, resulting in an improved RF behavior. As in the pulse method the wideband measurement signal is captured by under-sampling. But now it is controlled by a binary divider. This is a very effective method since drift and jitter are largely suppressed due to the steep flanks of the divider and furthermore it absolutely avoids nonlinear sample spacing.

The period length of an MLBS corresponds to $N = 2^n - 1$ clock cycles in which n is the length of the shift register. If the equivalent sampling rate f_s is to be chosen equal to the clock rate f_c the usable bandwidth will cover the range from DC to $f_c/2$ and the measured system response functions will consist of N data points. Moreover it can be seen from Eq. (17.27) that any under-sampling factor of a

power of 2, $u_{sf} = 2^m$, is applicable because $2^n - 1$ can never be divided by 2. See [9 – 12] for more information on the MLBS principle.

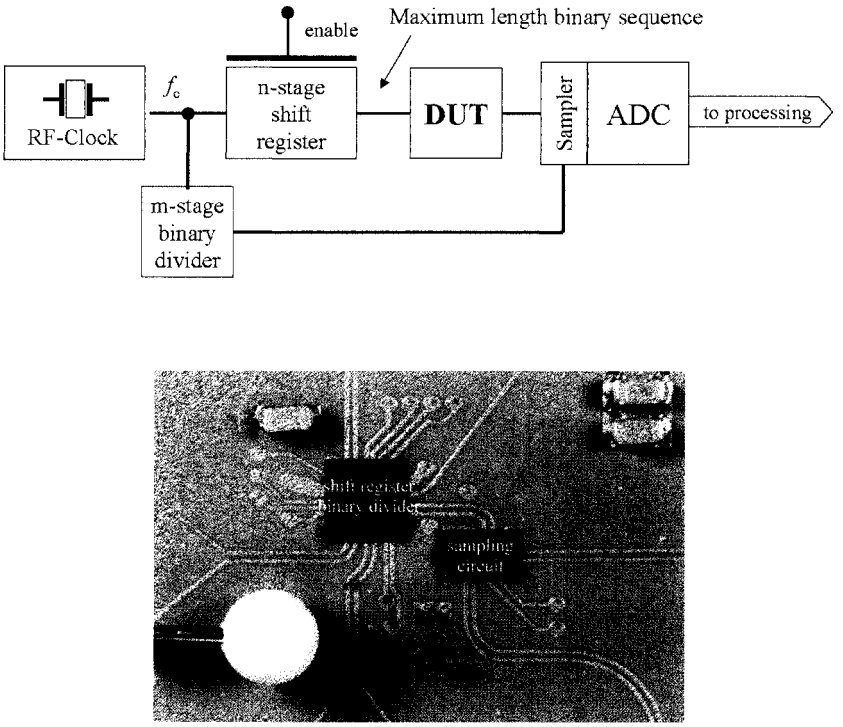

Fig. 17.17. Basic structure of MLBS baseband UWB system (top) and an example of the RF front end (except the clock generation) of a version for 10 GHz clock rate (bottom, courtesy of MEODAT GmbH). The pin head on the left is for comparison of dimensions

Admittedly the MLBS principle requires a more complex circuit structure compared to the basic versions of the above-mentioned pulse and FMCW methods. Nevertheless it can be economically manufactured thanks to modern low-cost RF and semiconductor technologies. In connection with its outstanding features the MLBS approach open up a great deal of future applications in UWB moisture sensors.

As expected, the power consumption of the MLBS RF part is higher than for simple pulse systems with additive samplers. However, concerning the total power consumption, it plays only a secondary role since the required signal processing usually has a greater power need. Furthermore it should be underlined that the MLBS approach enables in a simple way the building of fast-operating multi-channel arrangements [9]. It is much less expensive with this technique to equip every measurement channel with its own signal source (shift register) and receiver

circuit (T&H or S&H) than to use RF switches. The shift register can be controlled by a simple TTL signal, so no additional RF hardware is required to build a multi-channel arrangement.

UWB-modulated carrier: It is often desirable to work in a certain frequency band. The FMCW method gives this possibility. The pulse- and MLBS approaches are restricted to the baseband. By combining them with the homodyne concept, the situation can be changed. Figure 17.18 shows a possible circuit structure using an MLBS to modulate the carrier. Note that it is more suitable to modulate by an MLBS than by a pulse, because it does not burden the modulator with short high-power shocks.

The advantage compared to the FMCW method is mainly to be seen in the possibility to gather completely the information content of measurement data due to quadrature demodulation. In contrast to the FMCW method a 90° phase shifter for the quadrature demodulator can now be built since the device is working with a fixed carrier frequency.

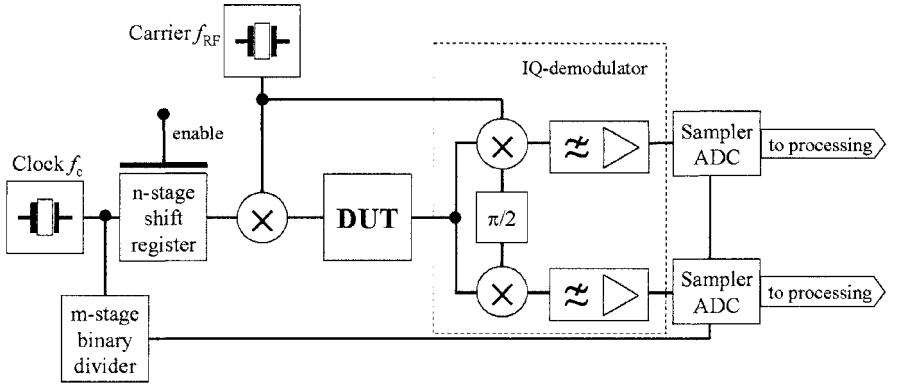

Fig. 17.18. Basic structure of a carrier frequency MLBS system. The measurement signal is divided into three parts: (1) the carrier which is known a priori; (2) a baseband signal describing the even part of the spectrum – detected by the upper channel; (3) a baseband signal describing the odd part of the spectrum – detected by the lower channel

17.6 Examples

This section refers to some simple laboratory experiments which demonstrate the versatility and the usability of the UWB technique. It is not the intention of the author to deeply interpret the measurement results with respect to quantitative statements. Furthermore, the demonstrations are limited for simplicity to

SiSo sensors (single-input – single-output) even if it is an easy task to upgrade the measurement head to a multi-channel system. All experiments are based on MLBS RF front end, as shown in Fig. 17.17, which was used in the measurement arrangements as indicated in Fig. 17.19.

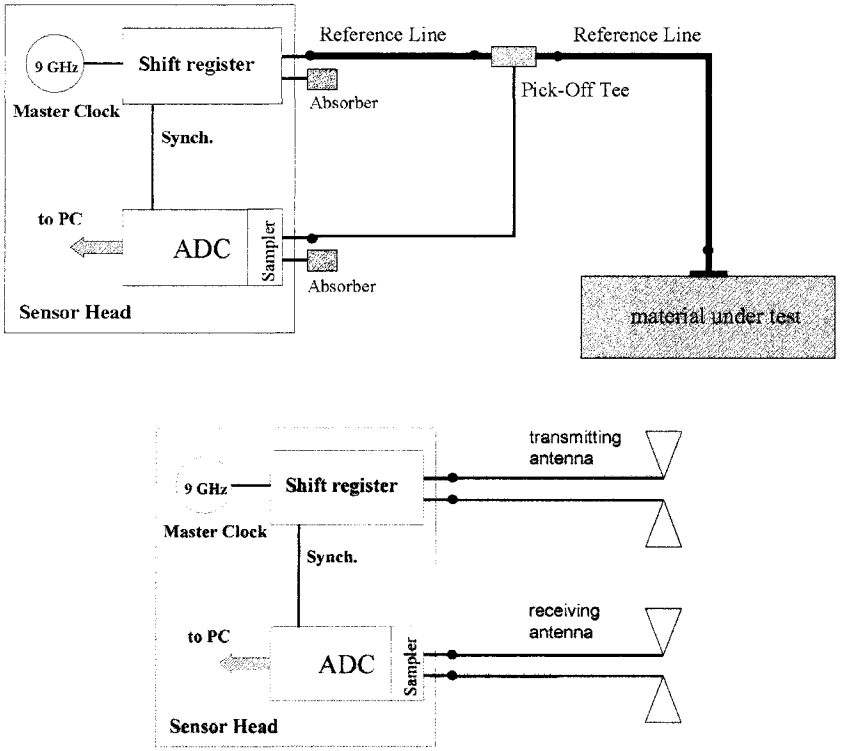

Fig. 17.19. Typical UWB measurement arrangement. Top: Time domain reflectometry (TDR) arrangement for measuring the reflection of an open coax line. Bottom: transmission arrangement for imaging purposes or if waves should cross the object under test

Water diffusion in a brick: Corresponding to the schematics on the top of Fig. 17.19, an open coaxial line was placed on the top side of a brick. Four minutes after the start of the measurement, the bottom side of the brick was exposed to fresh and in a second experiment to salty water. The temporal changes in the reflection coefficient of the probe were captured. They are presented for both experiments in Fig. 17.20. As expected, no effect appears when the water was added to the bottom side. But after approximately 20 min, the water arrived at the upper surface by diffusion and the reflection coefficient of the probe changed in both amplitude and phase. To make the presentation in Fig. 17.20 more clear, the same situation is pictured in Fig. 17.21 representing the data in the complex plane.

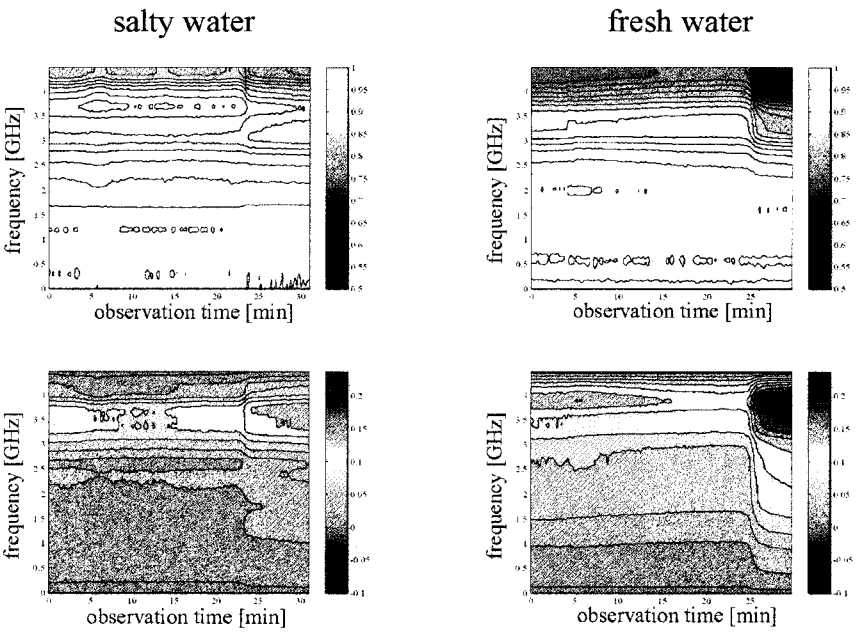

Fig. 17.20. Contour plots showing the temporal variation of the magnitude (top) and the phase (bottom) of the normalized reflection coefficient of the open coax line gained from TDR measurements. The considered frequency band covers DC to 4.5 GHz

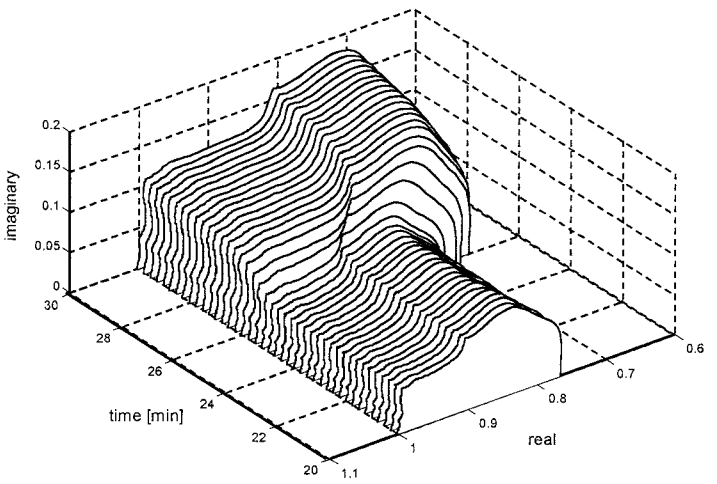

Fig. 17.21. Variation of the normalized complex reflection coefficient. The data shown refer to the fresh-water example

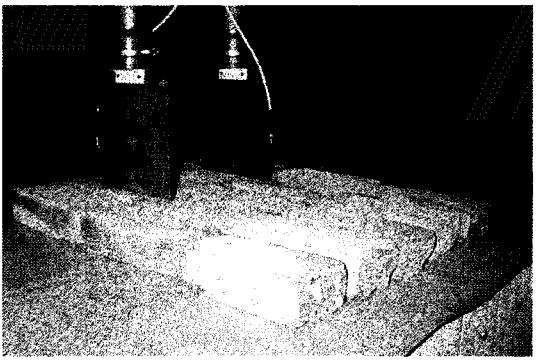

Fig. 17.22. Laboratory test arrangement using 3 by 5 bricks. The antennas used were of the Vivaldi-type

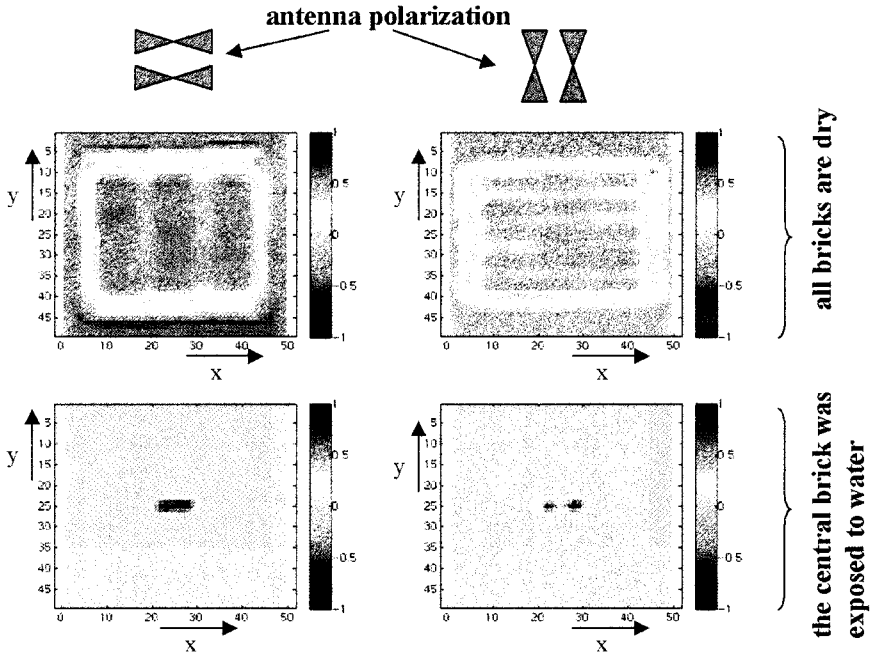

Fig. 17.23. C-scan representation of the scattered waves for different polarization. It should be noted that in the lower images the gaps between the brick are not visible because of a bigger scaling factor. By applying the same scaling factor to all images they will appear once again

Microwave imaging: The UWB technology offers interesting possibilities to build hand-held or automatic scanners in order to capture "moisture images" or moisture profiles respectively. The high bandwidth of the measurement electronics as well as of the antennas results in a range and cross-range resolution of a few centimeters. The following pictures demonstrate an example in which the signal processing is limited to a minimum in order to give an impression of the quality of the measurement data. It is interesting to note that by deploying array processing or synthetic aperture algorithms the image resolution can be greatly increased.

The goal of the experiment was to investigate the influence of dry and moist bricks on the scattering of electromagnetic waves. For that purpose, an arrangement of 3 by 5 bricks (see Fig. 17.22) was automatically scanned by a grid of 2 by 2 cm. Figure 17.23 shows some radar images. Obviously, dry and moist stones differ and it can also be seen that the polarization of antennas emphasizes either the lengthwise or cross-gaps between the bricks. A C-scan is also often called as time slice. It represents the reflectivity at a certain depth.

Time – frequency analysis [13, 14]: As mentioned above, all accessible information about the sensor behavior is included in the IRF or FRF. The problem which often arises is to extract it from these functions. Propagation effects such as scattering at boundaries or similar can be easily interpreted in the time domain. However, the frequency domain is preferred in connection with wave dispersion caused by a frequency-dependent permittivity as for water, for example. A new approach to data analysis should help to overcome these drawbacks. It is the so-called time – frequency representation (TFR) which joins the advantage of the pure IRF and FRF representation and makes it possible to analyze signals in the joint time – frequency domain.

One distinguishes between linear and non-linear TFRs. The well-known Wavelet transform and the short time Fourier transform count among the linear types. Linear TFRs have a time – frequency resolution which does not meet our requirements. Non-linear TFRs have a better resolution but they provoke cross-terms which complicate the data interpretation. One of the best known non-linear TFRs is the Wigner distribution.

The following example is intended to give an impression of how TFR works. The experiential set-up was built from a sandbox having a footprint of about 50 cm by 50 cm and a height of 20 cm. The sandbox was placed on sheet metal of the same size as the footprint of the box. This arrangement was scanned by the same UWB radar scanner as mentioned in Fig. 17.22. The resulting radargram is given in Fig. 17.24. A radargram or B-scan displays the IRF at every scanner position whereby the magnitude of the IRF is represented by a color. In the central region of the radargram, it can be clearly observed that signals arise by reflection at the surface and the ground (sheet metal) of the sandbox and by multiple reflections. Outside the sandbox, to the right and left, only the surface reflection of the laboratory floor appears. Since the sheet metal and floor are on the same level, the image shows well the lower propagation speed within the sand.

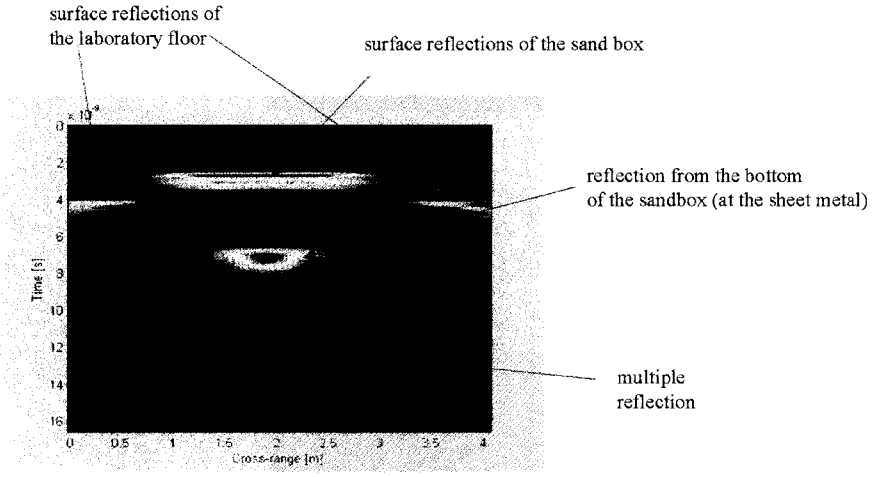

Fig. 17.24. Radargram of the sandbox (Courtesy of R. Zetik)

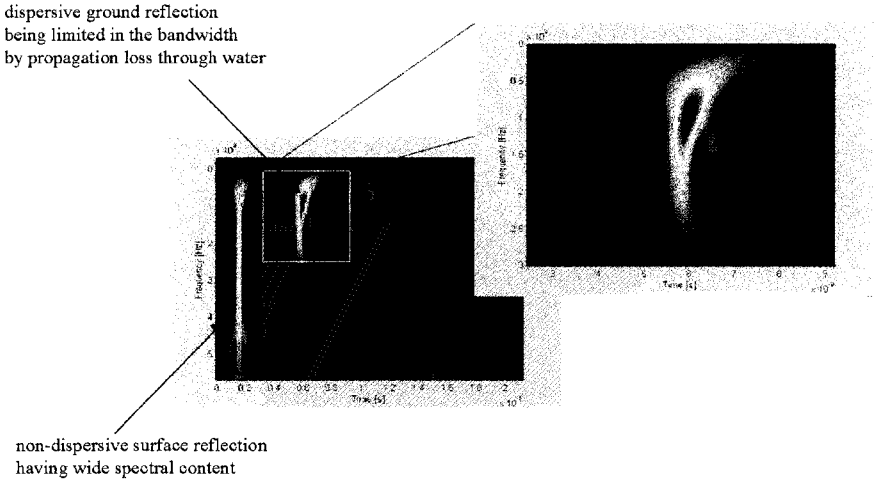

Fig. 17.25. Joint time – frequency representation of the central IRF of the radargram obtained by back scattering from a box containing moist sand. The ordinate of the images represents the propagation time of the sounding wave. The abscissa corresponds to the frequency and the color (gray scale) is a measure of energy. (Courtesy of R. Zetik)

In order to deal with the TRF, one IRF was isolated from the central part of the radargram in Fig. 17.24 and considered separately. Figure17.25 shows the joint time frequency representation of this IRF.

How do we interpret this image? The first impulse is caused by reflection at the surface of the sandbox. The sounding wave propagates only through air, thus no dispersion arises within the TFR. That means all spectral components of the pulse are delayed by the same time. This will give rise to a trace parallel to the frequency axis. The length of the trace reflects the bandwidth of the pulse.

The pulse which travels through the sandbox and back is subjected to a dispersion the strength of which depends upon the moisture content. That means the spectral components of the low-frequency part are delayed more strongly than the higher spectral components. Furthermore, higher frequencies will be attenuated. In the TFR, this results in a skewing of the impulse trace and a reduction of its length. For simplicity, we used only the skewing for the first trials to demonstrate the moisture effect. The results are presented in Fig. 17.26. But the joint time – frequency approach offers many more possibilities to separate different effects.

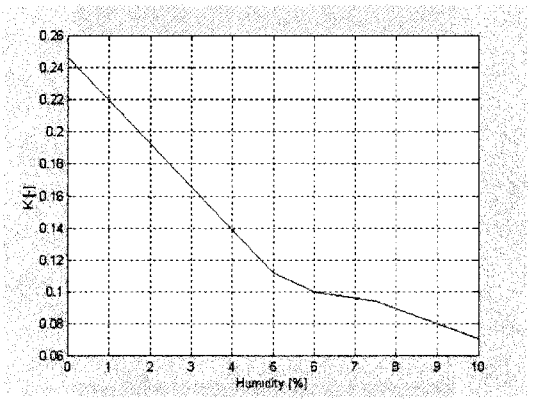

Fig. 17.26. Skew factor in dependence on the humidity of the sand. (Courtesy of R. Zetik)

17.7 Summary

UWB methods will open a promising perspective for further improvements in microwave moisture sensors and electromagnetic sensors in general. Measurements over a large frequency range are able to provide more information about the test object due to more and better data.

A successful application of UWB sensor principles requires:

- appropriate UWB applicators (antennas, electrode configurations, etc.),
- adequate UWB electronics, and
- a suitable model of the test arrangement in connection with (usually digital) data processing for extraction of the required information.

The greatest hindrance to the application of UWB sensor principles has been the lack of cost-effective, small-sized, power-saving, and robust UWB electronics in the past. But this situation is now changing. Different UWB electronics are available and under development. These electronics are supported by permanent improvements in size, power consumption, costs, and numeric power of the digital hardware which is inevitable in UWB sensors.

This chapter gave a short view of the philosophy of UWB measurements in order to enable the readers to estimate the potential of the UWB technique and its problems. Furthermore, some hints were given on how to choose the right measurement principle for the users purpose. But finally attention will be drawn the to the fact that the success of a UWB measurement strongly depends also on the quality of the signal processing.

References

1. Federal Communications Commission (2002) FCC 02-48, 14 Feb
2. Mirabbasi Sh, Martin K (2000) Classical and modern receiver architectures. IEEE Commun Mag Nov: 132-139
3. Stephan R, Loele H (1999) Ansätze zur technischen Realisierung einer Geschwindigkeitsmessung mit einem Breitband-Rausch-Radar. In: Proceedings of the workshop of the German IEEE/AP chapter on short range radars, Technische Universität Ilmenau, July 1999, pp 65-70
4. Narayanan RM, Xu Y, Hoffmeyer PD, Curtis JO (1995) Design and performance of a polarimetric random noise radar for detection of shallow buried targets. Proc SPIE 2496:20-30
5. Zollinger E (1993) Eigenschaften von Funkübertragungsstrecken in Gebäuden. Ph.D. thesis, ETH no 10064, Swiss Federal Institute of Technology, Zurich
6. Barrett TW (2001) History of ultra wideband communications and radar: Part II, UWB radars and sensors. Microwave J Feb
7. Musch T (2002) A high precision 24 GHz FMCW-radar using a phase-slope signal processing algorithm. In: Proceedings of the 32nd European microwave conference, Milan, vol 3, pp 945-948
8. Heide P, Vossiek M, Nalezinski M, Oréans L, Schubert R, Kunert M (1999) 24 GHz short-range microwave sensors for industrial and vehicular applications. In: Proceedings of the workshop of the German IEEE/AP chapter on short range radars, Technische Universität Ilmenau, July 1999, pp 4-9
9. Sachs J, Peyerl P (1999) A new principle for sensor-array-application. In: Proceedings of 16th IEEE instrumentation and measurement technology conference, IMTC/99, Venice, 24-26 May 1999, pp 1390-1395
10. Rossberg M, Sachs J, Rauschenbach P, Peyerl P, Pressel K, Winkler W, Knoll D (2000) 11 GHz SiGe circuits for ultra wideband radar. In: Bipolar/BiCMOS circuits and technology meeting, BCTM-2000, 25-26 Sept, Minneapolis
11. Sachs J, Peyerl P (2001) Integrated network analyser module for microwave moisture sensors. In: Proceedings of fourth international conference on "electromagnetic wave interaction with water and moist substances", 13-16 May 2001, Weimar, p 165ff

12. Sachs J, Peyerl P, Kmec M, Tkac F (2002) Digital ultra-wideband-sensor electronics integrated in SiGe-technology. Proceedings of 32nd European microwave conference, Milan, vol 2, pp 539-542, Milan
13. Zetik R, Sachs J (2002) Moisture determination of solid material by means of ultra-wideband radar and time-frequency signal representations. Acta Electron Inf 2(1):15
14. Zetik R (2000) Dual L-Wigner distribution and applications of time-frequency signal representations in ultra-wideband radar systems. Dissertation, University of Technology in Košice

Methods and Sensors for Quality Assesment to Products of Agriculture, Food, and Forestry