RFID HANDBOOK

Fundamentals and Applications in Contactless Smart Cards and Identification

Second Edition

KLAUS FINKENZELLER

RFID Handbook

Fundamentals and Applications in Contactless Smart Cards and Identification

Second Edition

Klaus Finkenzeller
Giesecke & Devrient GmbH, Munich, Germany

Translated by
Rachel Waddington
Member of the Institute of Translation and Interpreting

WILEY

First published under the title *RFID-Handbuch, 2 Auflage* by Carl Hanser Verlag

Authorized translation from the 2nd edition in the original German language
published by Carl Hanser Verlag, Munich/FRG

Copyright © 2003 John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,
West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wileyeurope.com or www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 33 Park Road, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Finkenzeller, Klaus.
[RFID Handbuch. English]
RFID handbook : fundamentals and applications in contactless smart cards and identifcation/Klaus Finkenzeller; translated by Rachel Waddington. — 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-470-84402-7 (alk. paper)
1. Inventory control — Automation. 2. Radio frequency identification systems. 3. Smart cards. I. Title.

TS160.F5513 2003
658.7'87 — dc21

2002192439

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0-470-84402-7

Typeset in 10/12pt Times by Laserwords Private Limited, Chennai, India
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Contents

3

Fundamental Operating Principles

This chapter describes the basic interaction between transponder and reader, in particular the power supply to the transponder and the data transfer between transponder and reader (Figure 3.1). For a more in-depth description of the physical interactions and mathematical models relating to inductive coupling or backscatter systems please refer to Chapter 4.

3.1 1-Bit Transponder

A bit is the smallest unit of information that can be represented and has only two states: 1 and 0. This means that only two states can be represented by systems based upon a *1-bit transponder*: 'transponder in interrogation zone' and 'no transponder in interrogation zone'. Despite this limitation, 1-bit transponders are very widespread — their main field of application is in electronic *anti-theft devices* in shops (*EAS*, electronic article surveillance).

An EAS system is made up of the following components: the antenna of a 'reader' or interrogator, the *security element* or *tag*, and an optional *deactivation device* for deactivating the tag after payment. In modern systems deactivation takes place when the price code is registered at the till. Some systems also incorporate an *activator*, which is used to reactivate the security element after deactivation (Gillert, 1997). The main performance characteristic for all systems is the recognition or *detection rate* in relation to the gate width (maximum distance between transponder and interrogator antenna).

The procedure for the inspection and testing of installed article surveillance systems is specified in the guideline *VDI 4470* entitled 'Anti-theft systems for goods — detection gates. Inspection guidelines for customers'. This guideline contains definitions and testing procedures for the calculation of the detection rate and false alarm ratio. It can be used by the retail trade as the basis for sales contracts or for monitoring the performance of installed systems on an ongoing basis. For the product manufacturer, the Inspection Guidelines for Customers represents an effective benchmark in the development and optimisation of integrated solutions for security projects (in accordance with VDI 4470).

RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification. Klaus Finkenzeller
 ISBN: 0-470-84402-7

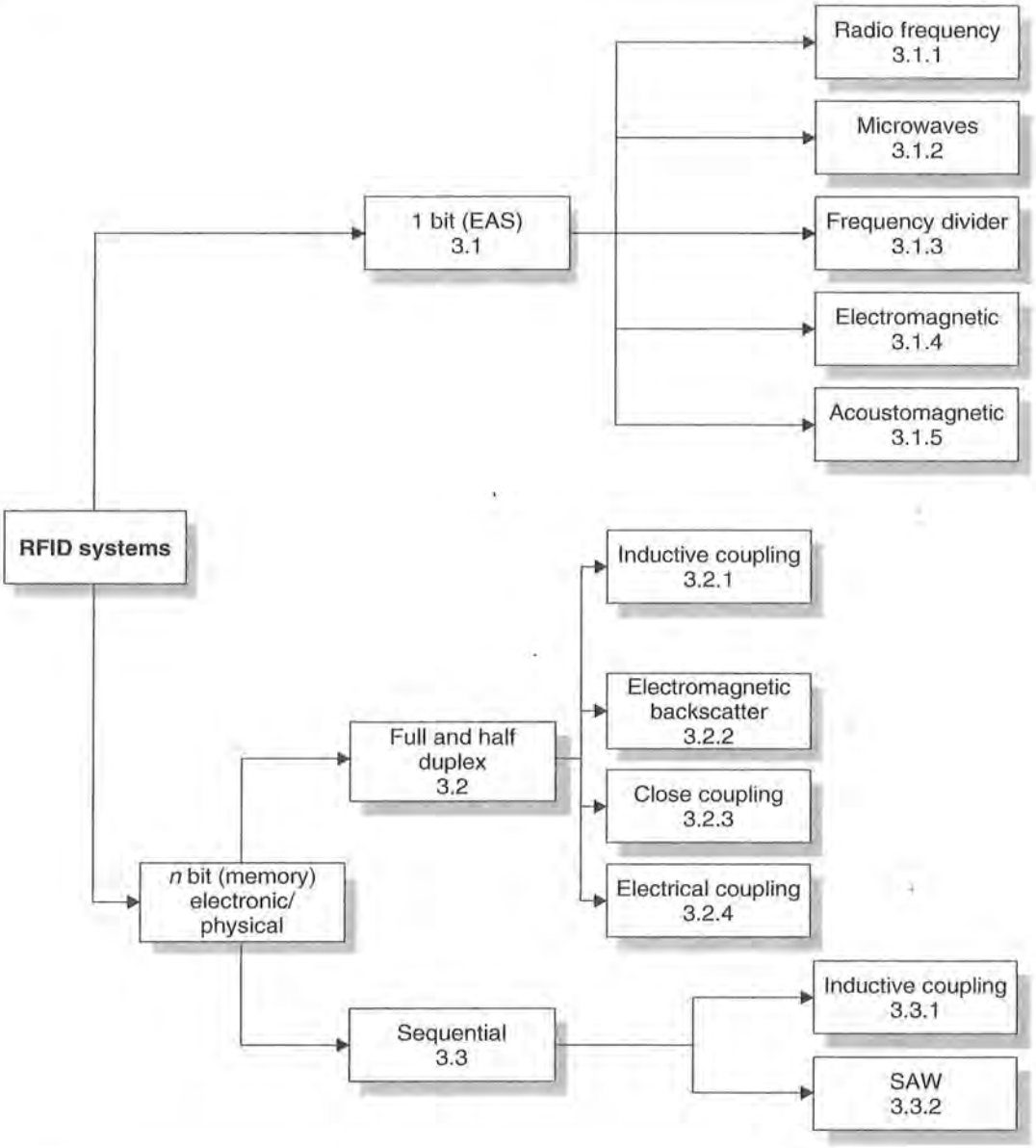


Figure 3.1 The allocation of the different operating principles of RFID systems into the sections of the chapter

3.1.1 Radio frequency

The *radio frequency (RF) procedure* is based upon LC resonant circuits adjusted to a defined resonant frequency f_R. Early versions employed inductive resistors made of wound enamelled copper wire with a soldered on capacitor in a plastic housing (*hard tag*). Modern systems employ coils etched between foils in the form of stick-on labels. To ensure that the damping resistance does not become too high and reduce the quality of the resonant circuit to an unacceptable level, the thickness of

the aluminium conduction tracks on the 25 μm thick *polyethylene foil* must be at least 50 μm (Jörn, 1994). Intermediate foils of 10 μm thickness are used to manufacture the capacitor plates.

The reader (detector) generates a magnetic alternating field in the radio frequency range (Figure 3.2). If the LC resonant circuit is moved into the vicinity of the magnetic alternating field, energy from the alternating field can be induced in the resonant circuit via its coils (Faraday's law). If the frequency f_G of the alternating field corresponds with the resonant frequency f_R of the LC resonant circuit the resonant circuit produces a *sympathetic oscillation*. The current that flows in the resonant circuit as a result of this acts against its cause, i.e. it acts against the external magnetic alternating field (see Section 4.1.10.1). This effect is noticeable as a result of a small change in the voltage drop across the transmitter's generator coil and ultimately leads to a weakening of the measurable magnetic field strength. A change to the induced voltage can also be detected in an optional sensor coil as soon as a resonant oscillating circuit is brought into the magnetic field of the generator coil.

The relative magnitude of this dip is dependent upon the gap between the two coils (generator coil — security element, security element — sensor coil) and the quality Q of the induced resonant circuit (in the security element).

The relative magnitude of the changes in voltage at the generator and sensor coils is generally very low and thus difficult to detect. However, the signal should be as clear as possible so that the security element can be reliably detected. This is achieved using a bit of a trick: the frequency of the magnetic field generated is not constant, it is 'swept'. This means that the generator frequency continuously crosses the range between minimum and maximum. The frequency range available to the swept systems is 8.2 MHz ±10% (Jörn, 1994).

Whenever the swept generator frequency exactly corresponds with the resonant frequency of the resonant circuit (in the transponder), the transponder begins to oscillate, producing a clear dip in the voltages at the generator and sensor coils (Figure 3.3). Frequency tolerances of the security element, which depend upon manufacturing tolerances

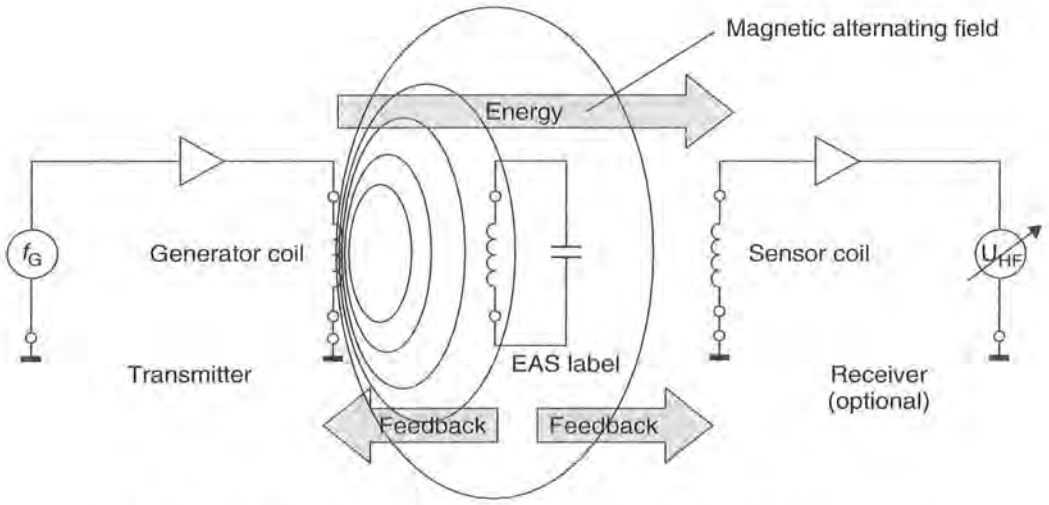


Figure 3.2 Operating principle of the EAS radio frequency procedure

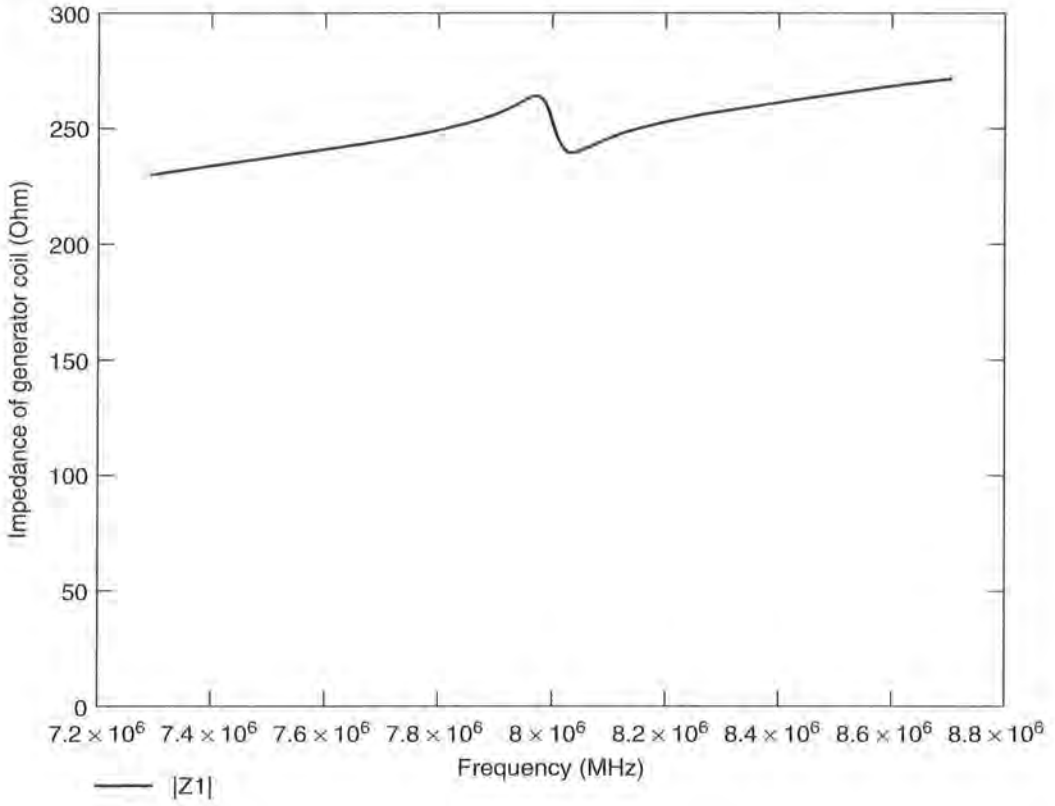


Figure 3.3 The occurrence of an impedance 'dip' at the generator coil at the resonant frequency of the security element ($Q = 90$, $k = 1\%$). The generator frequency f_G is continuously swept between two cut-off frequencies. An RF tag in the generator field generates a clear dip at its resonant frequency f_R

and vary in the presence of a metallic environment, no longer play a role as a result of the 'scanning' of the entire frequency range.

Because the tags are not removed at the till, they must be altered so that they do not activate the anti-theft system. To achieve this, the cashier places the protected product into a device — the deactivator — that generates a sufficiently high magnetic field that the induced voltage destroys the foil capacitor of the transponder. The capacitors are designed with intentional short-circuit points, so-called *dimples*. The breakdown of the capacitors is irreversible and detunes the resonant circuit to such a degree that this can no longer be excited by the *sweep signal*.

Large area *frame antennas* are used to generate the required magnetic alternating field in the detection area. The frame antennas are integrated into columns and combined to form gates. The classic design that can be seen in every large department store is illustrated in Figure 3.4. Gate widths of up to 2 m can be achieved using the RF procedure. The relatively low detection rate of 70% (Gillert, 1997) is disproportionately influenced by certain product materials. Metals in particular (e.g. food tins) affect the resonant frequency of the tags and the coupling to the detector coil and thus have a negative effect on the detection rate. Tags of 50 mm × 50 mm must be used to achieve the gate width and detection rate mentioned above.

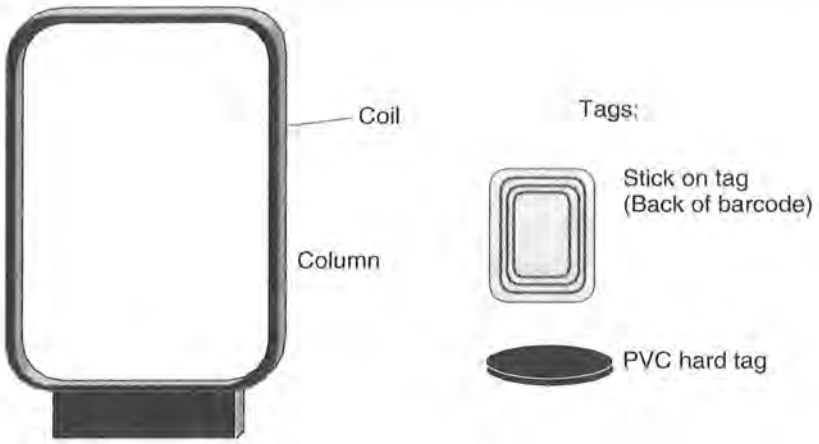


Figure 3.4 Left, typical frame antenna of an RF system (height 1.20–1.60 m); right, tag designs

Table 3.1 Typical system parameters for RF systems (VDI 4471)

Quality factor Q of the security element	>60–80
Minimum deactivation field strength H_D	1.5 A/m
Maximum field strength in the deactivation range	0.9 A/m

Table 3.2 Frequency range of different RF security systems (Plotzke *et al.*, 1994)

	System 1	System 2	System 3	System 4
Frequency (MHz)	1.86–2.18	7.44–8.73	7.30–8.70	7.40–8.60
Sweep frequency (Hz)	141	141	85	85

The range of products that have their own resonant frequencies (e.g. cable drums) presents a great challenge for system manufacturers. If these resonant frequencies lie within the sweep frequency 8.2 MHz ± 10% they will always trigger false alarms.

3.1.2 Microwaves

EAS systems in the *microwave range* exploit the generation of harmonics at components with nonlinear characteristic lines (e.g. diodes). The *harmonic* of a sinusoidal voltage A with a defined frequency f_A is a sinusoidal voltage B, whose frequency f_B is an integer multiple of the frequency f_A. The subharmonics of the frequency f_A are thus the frequencies $2f_A$, $3f_A$, $4f_A$ etc. The Nth multiple of the output frequency is termed the Nth harmonic (Nth harmonic wave) in radio-engineering; the output frequency itself is termed the carrier wave or first harmonic.

In principle, every two-terminal network with a nonlinear characteristic generates harmonics at the first harmonic. In the case of *nonlinear resistances*, however, energy is consumed, so that only a small part of the first harmonic power is converted into the harmonic oscillation. Under favourable conditions, the multiplication of f to $n \times f$

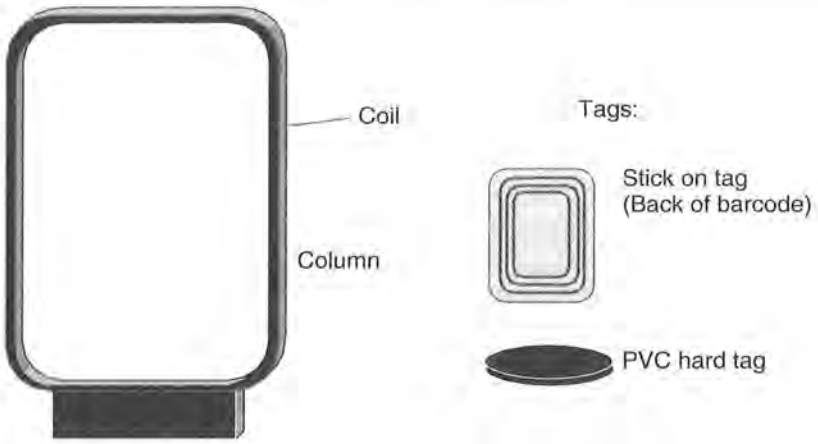


Figure 3.4 Left, typical frame antenna of an RF system (height 1.20–1.60 m); right, tag designs

Table 3.1 Typical system parameters for RF systems (VDI 4471)

Quality factor Q of the security element	>60–80
Minimum deactivation field strength H_D	1.5 A/m
Maximum field strength in the deactivation range	0.9 A/m

Table 3.2 Frequency range of different RF security systems (Plotzke *et al.*, 1994)

	System 1	System 2	System 3	System 4
Frequency (MHz)	1.86–2.18	7.44–8.73	7.30–8.70	7.40–8.60
Sweep frequency (Hz)	141	141	85	85

The range of products that have their own resonant frequencies (e.g. cable drums) presents a great challenge for system manufacturers. If these resonant frequencies lie within the sweep frequency 8.2 MHz ± 10% they will always trigger false alarms.

3.1.2 Microwaves

EAS systems in the *microwave range* exploit the generation of harmonics at components with nonlinear characteristic lines (e.g. diodes). The *harmonic* of a sinusoidal voltage A with a defined frequency f_A is a sinusoidal voltage B, whose frequency f_B is an integer multiple of the frequency f_A. The subharmonics of the frequency f_A are thus the frequencies $2f_A$, $3f_A$, $4f_A$ etc. The Nth multiple of the output frequency is termed the Nth harmonic (Nth harmonic wave) in radio-engineering; the output frequency itself is termed the carrier wave or first harmonic.

In principle, every two-terminal network with a nonlinear characteristic generates harmonics at the first harmonic. In the case of *nonlinear resistances*, however, energy is consumed, so that only a small part of the first harmonic power is converted into the harmonic oscillation. Under favourable conditions, the multiplication of f to $n \times f$

occurs with an efficiency of $\eta = 1/n^2$. However, if nonlinear energy storage is used for multiplication, then in the ideal case there are no losses (Fleckner, 1987).

Capacitance diodes are particularly suitable nonlinear energy stores for frequency multiplication. The number and intensity of the harmonics that are generated depend upon the capacitance diode's *dopant profile* and characteristic line gradient. The exponent n (also γ) is a measure for the gradient (=capacitance-voltage characteristic). For simple diffused diodes, this is 0.33 (e.g. BA110), for alloyed diodes it is 0.5 and for tuner diodes with a hyper-abrupt P-N junction it is around 0.75 (e.g. BB 141) (Intermetal Semiconductors ITT, 1996).

The capacitance-voltage characteristic of alloyed capacitance diodes has a quadratic path and is therefore best suited for the doubling of frequencies. Simple diffused diodes can be used to produce higher harmonics (Fleckner, 1987).

The layout of a 1-bit transponder for the generation of harmonics is extremely simple: a capacitance diode is connected to the base of a *dipole* adjusted to the carrier wave (Figure 3.5). Given a carrier wave frequency of 2.45 GHz the dipole has a total length of 6 cm. The carrier wave frequencies used are 915 MHz (outside Europe), 2.45 GHz or 5.6 GHz. If the transponder is located within the transmitter's range, then the flow of current within the diode generates and re-emits harmonics of the carrier wave. Particularly distinctive signals are obtained at two or three times the carrier wave, depending upon the type of diode used.

Transponders of this type cast in plastic (hard tags) are used mainly to protect textiles. The tags are removed at the till when the goods are paid for and they are subsequently reused.

Figure 3.6 shows a transponder being placed within the range of a microwave transmitter operating at 2.45 GHz. The second harmonic of 4.90 GHz generated in the diode characteristic of the transponder is re-transmitted and detected by a receiver, which is adjusted to this precise frequency. The reception of a signal at the frequency of the second harmonic can then trigger an alarm system.

If the amplitude or frequency of the carrier wave is modulated (ASK, FSK), then all harmonics incorporate the same modulation. This can be used to distinguish between 'interference' and 'useful' signals, preventing false alarms caused by external signals.

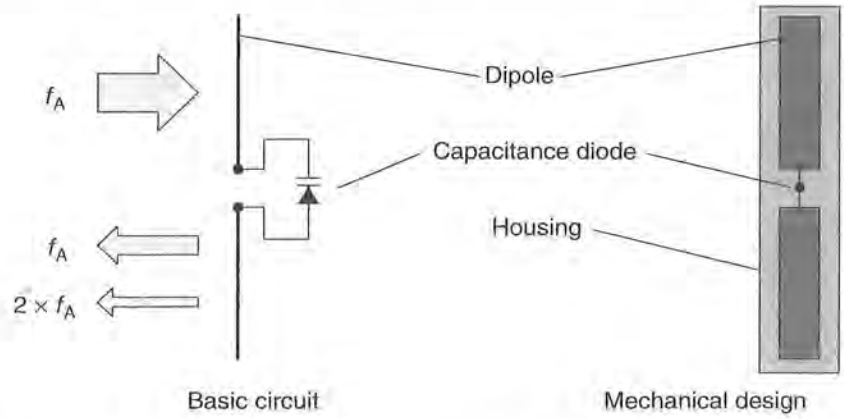


Figure 3.5 Basic circuit and typical construction format of a microwave tag

Figure 3.6 Microwave tag in the interrogation zone of a detector

In the example above, the amplitude of the carrier wave is modulated with a signal of 1 kHz (100% ASK). The second harmonic generated at the transponder is also modulated at 1 kHz ASK. The signal received at the receiver is demodulated and forwarded to a 1 kHz detector. Interference signals that happen to be at the reception frequency of 4.90 GHz cannot trigger false alarms because these are not normally modulated and, if they are, they will have a different modulation.

3.1.3 Frequency divider

This procedure operates in the long wave range at 100–135.5 kHz. The security tags contain a semiconductor circuit (microchip) and a resonant circuit coil made of wound enamelled copper. The resonant circuit is made to resonate at the operating frequency of the EAS system using a soldered capacitor. These transponders can be obtained in the form of hard tags (plastic) and are removed when goods are purchased.

The microchip in the transponder receives its power supply from the magnetic field of the security device (see Section 3.2.1.1). The frequency at the self-inductive coil is divided by two by the microchip and sent back to the security device. The signal at half the original frequency is fed by a tap into the resonant circuit coil (Figure 3.7).

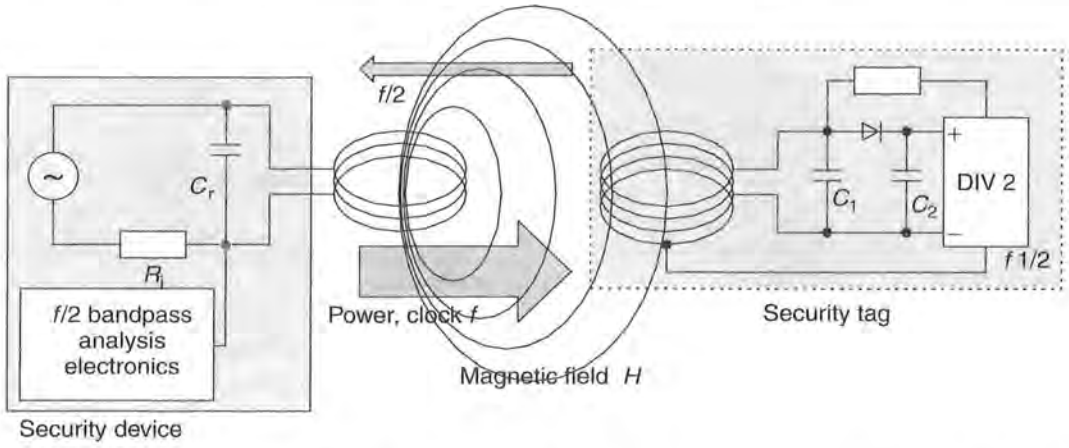


Figure 3.7 Basic circuit diagram of the EAS frequency division procedure: security tag (transponder) and detector (evaluation device)

Table 3.3 Typical system parameters (Plotzke *et al.*, 1994)

Frequency	130 kHz
Modulation type:	100% ASK
Modulation frequency/modulation signal:	12.5 Hz or 25 Hz, rectangle 50%

The magnetic field of the security device is pulsed at a lower frequency (ASK modulated) to improve the detection rate. Similarly to the procedure for the generation of harmonics, the modulation of the carrier wave (ASK or FSK) is maintained at half the frequency (*subharmonic*). This is used to differentiate between 'interference' and 'useful' signals. This system almost entirely rules out false alarms.

Frame antennas, described in Section 3.1.1, are used as sensor antennas.

3.1.4 Electromagnetic types

Electromagnetic types operate using strong magnetic fields in the *NF range* from 10 Hz to around 20 kHz. The security elements contain a soft magnetic *amorphous metal* strip with a steep flanked hysteresis curve (see also Section 4.1.12). The magnetisation of these strips is periodically reversed and the strips taken to magnetic saturation by a strong magnetic alternating field. The markedly nonlinear relationship between the applied field strength H and the magnetic flux density B near saturation (see also Figure 4.50), plus the sudden change of flux density B in the vicinity of the zero crossover of the applied field strength H, generates harmonics at the basic frequency of the security device, and these harmonics can be received and evaluated by the security device.

The electromagnetic type is optimised by superimposing additional signal sections with higher frequencies over the main signal. The marked nonlinearity of the strip's hysteresis curve generates not only harmonics but also signal sections with summation and differential frequencies of the supplied signals. Given a main signal of frequency $f_S = 20$ Hz and the additional signals $f_1 = 3.5$ and $f_2 = 5.3$ kHz, the following signals are generated (first order):

$$f_1 + f_2 = f_{1+2} = 8.80\,\text{kHz}$$
$$f_1 - f_2 = f_{1-2} = 1.80\,\text{kHz}$$
$$f_S + f_1 = f_{S+1} = 3.52\,\text{kHz} \text{ and so on}$$

The security device does not react to the harmonic of the basic frequency in this case, but rather to the summation or differential frequency of the extra signals.

The tags are available in the form of self-adhesive strips with lengths ranging from a few centimetres to 20 cm. Due to the extremely low operating frequency, electromagnetic systems are the only systems suitable for products containing metal. However, these systems have the disadvantage that the function of the tags is dependent upon position: for reliable detection the magnetic field lines of the security device must run vertically through the amorphous metal strip. Figure 3.8 shows a typical design for a security system.

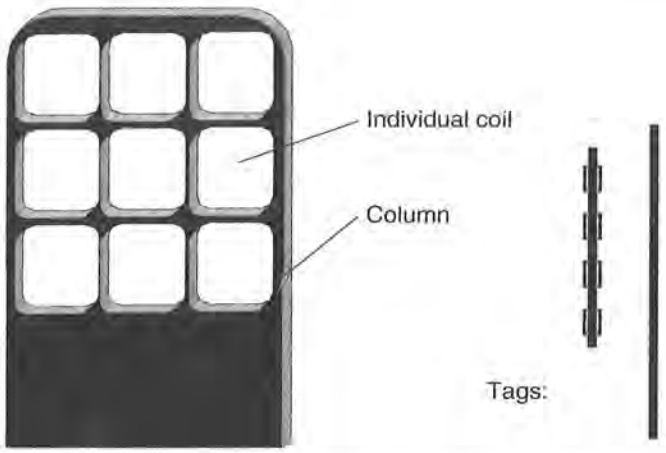


Figure 3.8 Left, typical antenna design for a security system (height approximately 1.40 m); right, possible tag designs

For deactivation, the tags are coated with a layer of hard magnetic metal or partially covered by hard magnetic plates. At the till the cashier runs a strong *permanent magnet* along the metal strip to deactivate the security elements (Plotzke *et al.*, 1994). This magnetises the hard magnetic metal plates. The metal strips are designed such that the remanence field strength (see Section 4.1.12) of the plate is sufficient to keep the amorphous metal strips at saturation point so that the magnetic alternating field of the security system can no longer be activated.

The tags can be reactivated at any time by demagnetisation. The process of deactivation and reactivation can be performed any number of times. For this reason, electromagnetic goods protection systems were originally used mainly in lending libraries. Because the tags are small (min. 32 mm short strips) and cheap, these systems are now being used increasingly in the grocery industry. See Figure 3.9.

In order to achieve the field strength necessary for demagnetisation of the permalloy strips, the field is generated by two coil systems in the columns at either side of a narrow passage. Several individual coils, typically 9 to 12, are located in the two pillars, and these generate weak magnetic fields in the centre and stronger magnetic fields on the outside (Plotzke *et al.*, 1994). Gate widths of up to 1.50 m can now be realised using this method, while still achieving detection rates of 70% (Gillert, 1997) (Figure 3.10).

3.1.5 Acoustomagnetic

Acoustomagnetic systems for security elements consist of extremely small plastic boxes around 40 mm long, 8 to 14 mm wide depending upon design, and just a millimetre

Table 3.4 Typical system parameters (Plotzke *et al.*, 1997)

Frequency	70 Hz
Optional combination frequencies of different systems	12 Hz, 215 Hz, 3.3 kHz, 5 kHz
Field strength H_{eff} in the detection zone	25–120 A/m
Minimum field strength for deactivation	16 000 A/m

Figure 3.9 Electromagnetic labels in use (reproduced by permission of Schreiner Codedruck, Munich)

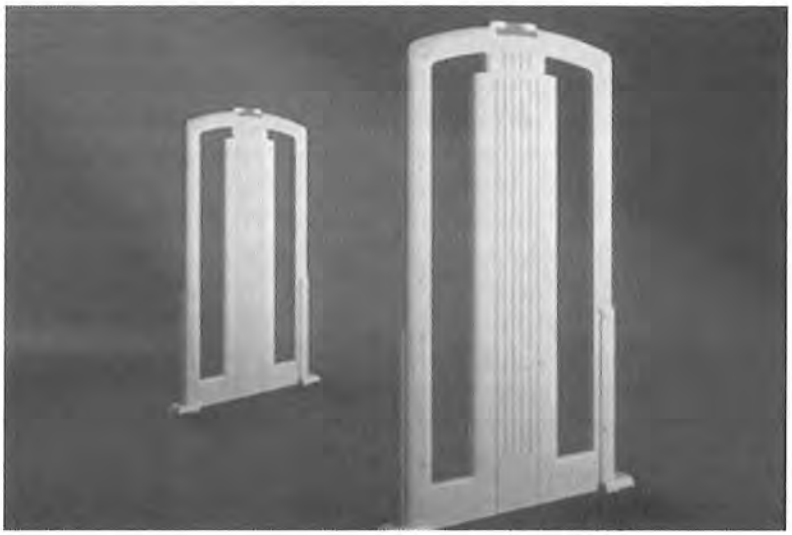


Figure 3.10 Practical design of an antenna for an article surveillance system (reproduced by permission of METO EAS System 2200, Esselte Meto, Hirschborn)

high. The boxes contain two metal strips, a *hard magnetic metal strip* permanently connected to the plastic box, plus a strip made of *amorphous metal*, positioned such that it is free to vibrate mechanically (Zechbauer, 1999).

Ferromagnetic metals (nickel, iron etc.) change slightly in length in a magnetic field under the influence of the field strength H. This effect is called *magnetostriction* and results from a small change in the interatomic distance as a result of magnetisation. In

a magnetic alternating field a magnetostrictive metal strip vibrates in the longitudinal direction at the frequency of the field. The amplitude of the vibration is especially high if the frequency of the magnetic alternating field corresponds with that of the (acoustic) resonant frequency of the metal strip. This effect is particularly marked in amorphous materials.

The decisive factor is that the magnetostrictive effect is also reversible. This means that an oscillating magnetostrictive metal strip emits a magnetic alternating field. *Acoustomagnetic security systems* are designed such that the frequency of the magnetic alternating field generated precisely coincides with the resonant frequencies of the metal strips in the security element. The amorphous metal strip begins to oscillate under the influence of the magnetic field. If the magnetic alternating field is switched off after some time, the excited magnetic strip continues to oscillate for a while like a tuning fork and thereby itself generates a magnetic alternating field that can easily be detected by the security system (Figure 3.11).

The great advantage of this procedure is that the security system is not itself transmitting while the security element is responding and the detection receiver can thus be designed with a corresponding degree of sensitivity.

In their activated state, acoustomagnetic security elements are magnetised, i.e. the above-mentioned hard magnetic metal strip has a high remanence field strength and thus forms a permanent magnet. To deactivate the security element the hard magnetic metal strip must be demagnetised. This detunes the resonant frequency of the amorphous

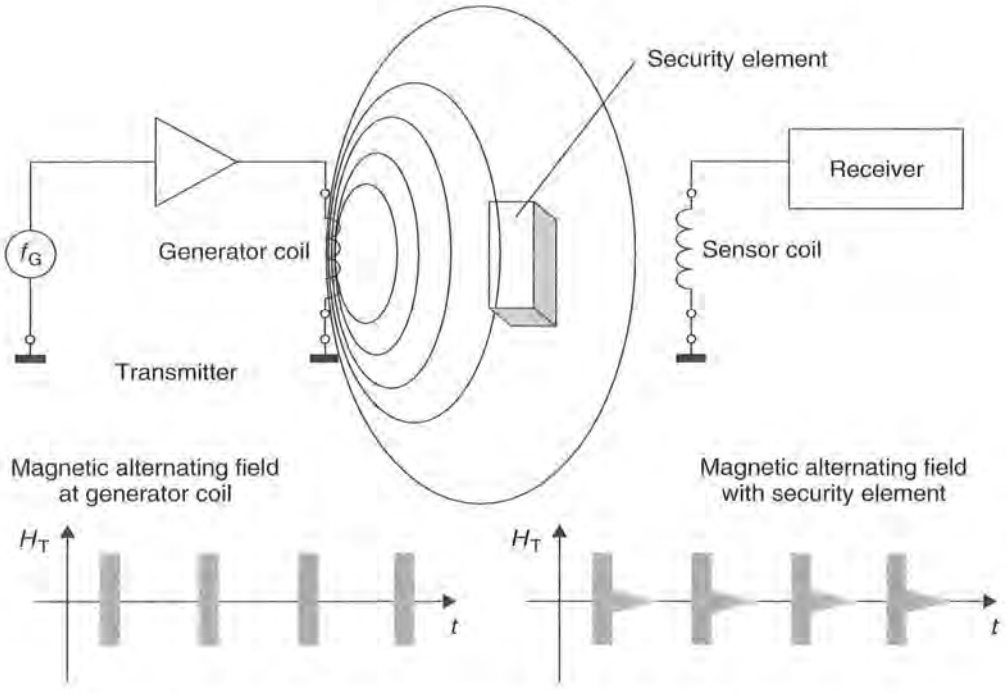


Figure 3.11 Acoustomagnetic system comprising transmitter and detection device (receiver). If a security element is within the field of the generator coil this oscillates like a tuning fork in time with the pulses of the generator coil. The transient characteristics can be detected by an analysing unit

Table 3.5 Typical operating parameters of acoustomagnetic systems (VDI 4471)

Parameter	Typical value
Resonant frequency f_0	58 kHz
Frequency tolerance	±0.52%
Quality factor Q	>150
Minimum field strength H_A for activation	>16 000 A/m
ON duration of the field	2 ms
Field pause (OFF duration)	20 ms
Decay process of the security element	5 ms

metal strip so it can no longer be excited by the operating frequency of the security system. The hard magnetic metal strip can only be demagnetised by a strong magnetic alternating field with a slowly decaying field strength. It is thus absolutely impossible for the security element to be manipulated by permanent magnets brought into the store by customers.

3.2 Full and Half Duplex Procedure

In contrast to 1-bit transponders, which normally exploit simple physical effects (oscillation stimulation procedures, stimulation of harmonics by diodes or the nonlinear hysteresis curve of metals), the transponders described in this and subsequent sections use an electronic microchip as the data-carrying device. This has a data storage capacity of up to a few kilobytes. To read from or write to the data-carrying device it must be possible to transfer data between the transponder and a reader. This transfer takes place according to one of two main procedures: full and half duplex procedures, which are described in this section, and sequential systems, which are described in the following section.

In the *half duplex procedure* (HDX) the data transfer from the transponder to the reader alternates with data transfer from the reader to the transponder. At frequencies below 30 MHz this is most often used with the load modulation procedure, either with or without a subcarrier, which involves very simple circuitry. Closely related to this is the modulated reflected cross-section procedure that is familiar from radar technology and is used at frequencies above 100 MHz. Load modulation and modulated reflected cross-section procedures directly influence the magnetic or electromagnetic field generated by the reader and are therefore known as *harmonic* procedures.

In the *full duplex procedure* (FDX) the data transfer from the transponder to the reader takes place at the same time as the data transfer from the reader to the transponder. This includes procedures in which data is transmitted from the transponder at a fraction of the frequency of the reader, i.e. a *subharmonic*, or at a completely independent, i.e. an *anharmonic*, frequency.

However, both procedures have in common the fact that the transfer of energy from the reader to the transponder is continuous, i.e. it is independent of the direction of data flow. In sequential systems (SEQ), on the other hand, the transfer of energy from the transponder to the reader takes place for a limited period of time only (pulse

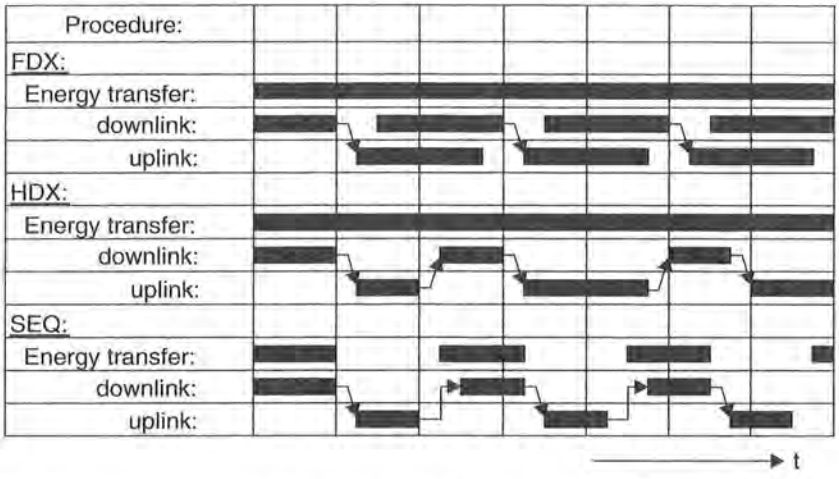


Figure 3.12 Representation of full duplex, half duplex and sequential systems over time. Data transfer from the reader to the transponder is termed downlink, while data transfer from the transponder to the reader is termed uplink

operation → *pulsed system*). Data transfer from the transponder to the reader occurs in the pauses between the power supply to the transponder. See Figure 3.12 for a representation of full duplex, half duplex and sequential systems.

Unfortunately, the literature relating to RFID has not yet been able to agree a consistent nomenclature for these system variants. Rather, there has been a confusing and inconsistent classification of individual systems into full and half duplex procedures. Thus pulsed systems are often termed half duplex systems — this is correct from the point of view of data transfer — and all unpulsed systems are falsely classified as full duplex systems. For this reason, in this book pulsed systems — for differentiation from other procedures, and unlike most RFID literature(!) — are termed sequential systems (SEQ).

3.2.1 Inductive coupling

3.2.1.1 Power supply to passive transponders

An inductively coupled transponder comprises an electronic data-carrying device, usually a single microchip, and a large area coil that functions as an antenna.

Inductively coupled transponders are almost always operated passively. This means that all the energy needed for the operation of the microchip has to be provided by the reader (Figure 3.13). For this purpose, the reader's antenna coil generates a strong, high frequency electromagnetic field, which penetrates the cross-section of the coil area and the area around the coil. Because the wavelength of the frequency range used (<135 kHz: 2400 m, 13.56 MHz: 22.1 m) is several times greater than the distance between the reader's antenna and the transponder, the electromagnetic field may be treated as a simple magnetic alternating field with regard to the distance between transponder and antenna (see Section 4.2.1.1 for further details).

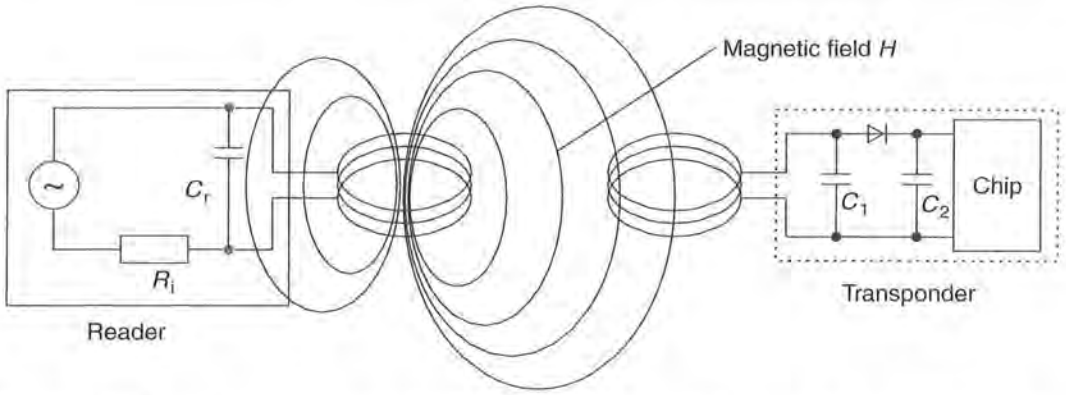


Figure 3.13 Power supply to an inductively coupled transponder from the energy of the magnetic alternating field generated by the reader

A small part of the emitted field penetrates the antenna coil of the transponder, which is some distance away from the coil of the reader. A voltage U_i is generated in the transponder's antenna coil by inductance. This voltage is rectified and serves as the power supply for the data-carrying device (microchip). A capacitor C_r is connected in parallel with the reader's antenna coil, the capacitance of this capacitor being selected such that it works with the coil inductance of the antenna coil to form a parallel resonant circuit with a resonant frequency that corresponds with the transmission frequency of the reader. Very high currents are generated in the antenna coil of the reader by resonance step-up in the parallel resonant circuit, which can be used to generate the required field strengths for the operation of the remote transponder.

The antenna coil of the transponder and the capacitor C_1 form a resonant circuit tuned to the transmission frequency of the reader. The voltage U at the transponder coil reaches a maximum due to resonance step-up in the parallel resonant circuit.

The layout of the two coils can also be interpreted as a transformer (*transformer coupling*), in which case there is only a very weak coupling between the two windings (Figure 3.14). The efficiency of power transfer between the antenna coil of the reader and the transponder is proportional to the operating frequency f, the number of windings n, the area A enclosed by the transponder coil, the angle of the two coils relative to each other and the distance between the two coils.

As frequency f increases, the required coil inductance of the transponder coil, and thus the number of windings n decreases (135 kHz: typical 100–1000 windings, 13.56 MHz: typical 3–10 windings). Because the voltage induced in the transponder is still proportional to frequency f (see Chapter 4), the reduced number of windings barely affects the efficiency of power transfer at higher frequencies. Figure 3.15 shows a reader for an inductively coupled transponder.

3.2.1.2 Data transfer transponder → reader

Load modulation As described above, inductively coupled systems are based upon a *transformer-type coupling* between the primary coil in the reader and the secondary coil in the transponder. This is true when the distance between the coils does not exceed

Figure 3.14 Different designs of inductively coupled transponders. The photo shows half finished transponders, i.e. transponders before injection into a plastic housing (reproduced by permission of AmaTech GmbH & Co. KG, D-Pfronten)

Figure 3.15 Reader for inductively coupled transponder in the frequency range <135 kHz with integral antenna (reproduced by permission of easy-key System, micron, Halbergmoos)

0.16 λ, so that the transponder is located in the *near field* of the transmitter antenna (for a more detailed definition of the near and far fields, please refer to Chapter 4).

If a resonant transponder (i.e. a transponder with a self-resonant frequency corresponding with the transmission frequency of the reader) is placed within the magnetic alternating field of the reader's antenna, the transponder draws energy from the magnetic field. The resulting feedback of the transponder on the reader's antenna can be

Table 3.6 Overview of the power consumption of various RFID-ASIC building blocks (Atmel, 1994). The minimum supply voltage required for the operation of the microchip is 1.8 V, the maximum permissible voltage is 10 V

	Memory (Bytes)	Write/read distance	Power consumption	Frequency	Application
ASIC#1	6	15 cm	10 µA	120 kHz	Animal ID
ASIC#2	32	13 cm	600 µA	120 kHz	Goods flow, access check
ASIC#3	256	2 cm	6 µA	128 kHz	Public transport
ASIC#4	256	0.5 cm	<1 mA	4 MHz*	Goods flow, public transport
ASIC#5	256	<2 cm	~1 mA	4/13.56 MHz	Goods flow
ASIC#6	256	100 cm	500 µA	125 kHz	Access check
ASIC#7	2048	0.3 cm	<10 mA	4.91 MHz*	Contactless chip cards
ASIC#8	1024	10 cm	~1 mA	13.56 MHz	Public transport
ASIC#9	8	100 cm	<1 mA	125 kHz	Goods flow
ASIC#10	128	100 cm	<1 mA	125 kHz	Access check

*Close coupling system.

represented as *transformed impedance* Z_T in the antenna coil of the reader. Switching a *load resistor* on and off at the transponder's antenna therefore brings about a change in the impedance Z_T, and thus voltage changes at the reader's antenna (see Section 4.1.10.3). This has the effect of an amplitude modulation of the voltage U_L at the reader's antenna coil by the remote transponder. If the timing with which the load resistor is switched on and off is controlled by data, this data can be transferred from the transponder to the reader. This type of data transfer is called *load modulation*.

To reclaim the data at the reader, the voltage tapped at the reader's antenna is rectified. This represents the demodulation of an amplitude modulated signal. An example circuit is shown in Section 11.3.

Load modulation with subcarrier Due to the weak coupling between the reader antenna and the transponder antenna, the voltage fluctuations at the antenna of the reader that represent the useful signal are smaller by orders of magnitude than the output voltage of the reader.

In practice, for a 13.56 MHz system, given an antenna voltage of approximately 100 V (voltage step-up by resonance) a useful signal of around 10 mV can be expected (=80 dB signal/noise ratio). Because detecting this slight voltage change requires highly complicated circuitry, the modulation sidebands created by the amplitude modulation of the antenna voltage are utilised (Figure 3.16).

If the additional load resistor in the transponder is switched on and off at a very high elementary frequency f_S, then two spectral lines are created at a distance of $\pm f_S$ around the transmission frequency of the reader f_{READER}, and these can be easily detected (however f_S must be less than f_{READER}). In the terminology of radio technology the new elementary frequency is called a *subcarrier*). Data transfer is by ASK, FSK or PSK modulation of the subcarrier in time with the data flow. This represents an amplitude modulation of the subcarrier.

Load modulation with a subcarrier creates two modulation sidebands at the reader's antenna at the distance of the subcarrier frequency around the operating frequency f_{READER} (Figure 3.17). These modulation sidebands can be separated from

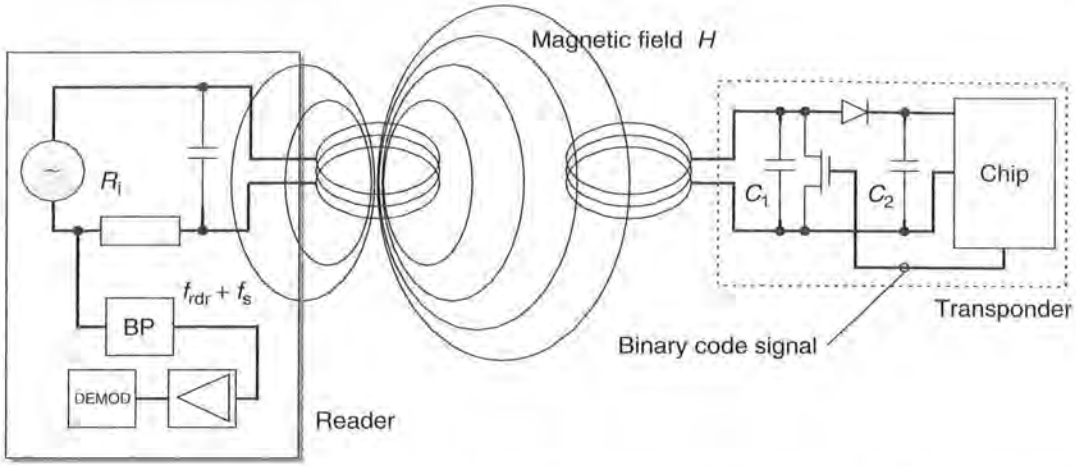


Figure 3.16 Generation of load modulation in the transponder by switching the drain-source resistance of an FET on the chip. The reader illustrated is designed for the detection of a subcarrier

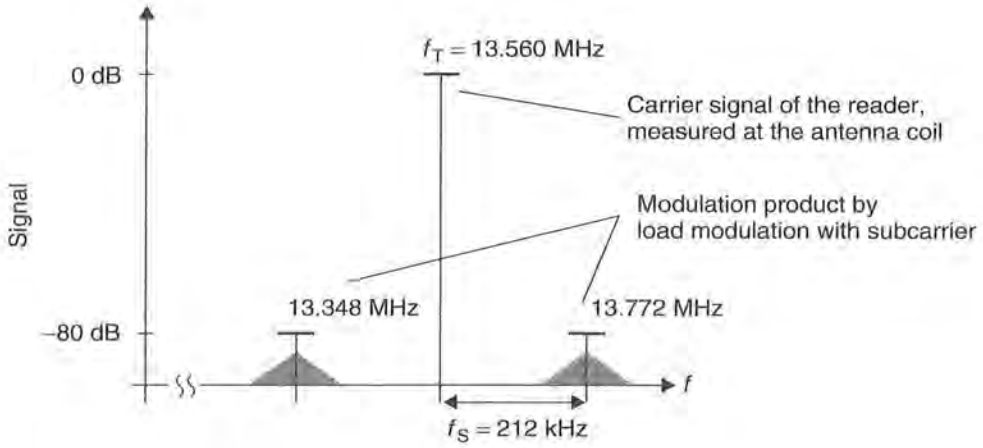


Figure 3.17 Load modulation creates two sidebands at a distance of the subcarrier frequency f_S around the transmission frequency of the reader. The actual information is carried in the sidebands of the two subcarrier sidebands, which are themselves created by the modulation of the subcarrier

the significantly stronger signal of the reader by bandpass (BP) filtering on one of the two frequencies $f_{READER} \pm f_S$. Once it has been amplified, the subcarrier signal is now very simple to demodulate.

Because of the large bandwidth required for the transmission of a subcarrier, this procedure can only be used in the ISM frequency ranges for which this is permitted, 6.78 MHz, 13.56 MHz and 27.125 MHz (see also Chapter 5).

Example circuit–load modulation with subcarrier Figure 3.18 shows an example circuit for a transponder using load modulation with a subcarrier. The circuit is designed for an operating frequency of 13.56 MHz and generates a subcarrier of 212 kHz.

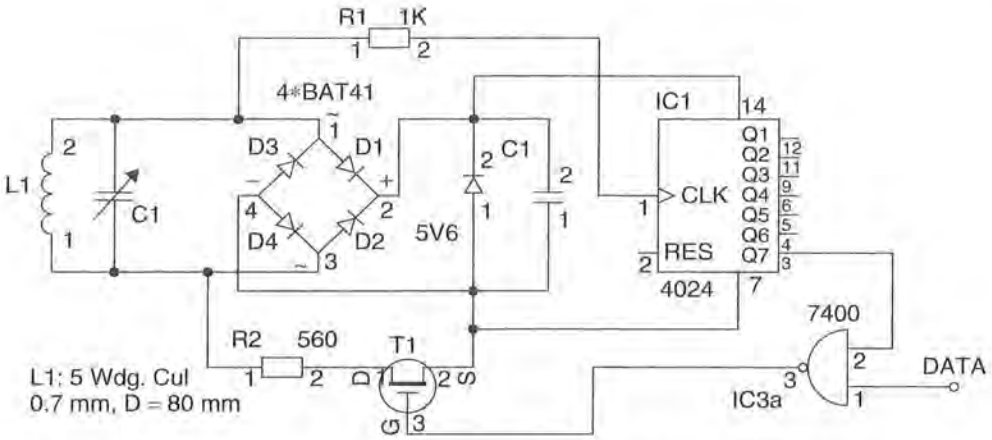


Figure 3.18 Example circuit for the generation of load modulation with subcarrier in an inductively coupled transponder

The voltage induced at the antenna coil L1 by the magnetic alternating field of the reader is rectified using the bridge rectifier (D1–D4) and after additional smoothing (C1) is available to the circuit as supply voltage. The parallel regulator (ZD 5V6) prevents the supply voltage from being subject to an uncontrolled increase when the transponder approaches the reader antenna.

Part of the high frequency antenna voltage (13.56 MHz) travels to the frequency divider's timing input (CLK) via the protective resistor (R1) and provides the transponder with the basis for the generation of an internal clocking signal. After division by $2^6 (= 64)$ a subcarrier clocking signal of 212 kHz is available at output Q7. The subcarrier clocking signal, controlled by a serial data flow at the data input (DATA), is passed to the switch (T1). If there is a logical HIGH signal at the data input (DATA), then the subcarrier clocking signal is passed to the switch (T1). The load resistor (R2) is then switched on and off in time with the subcarrier frequency.

Optionally in the depicted circuit the transponder resonant circuit can be brought into resonance with the capacitor C1 at 13.56 MHz. The range of this 'minimal transponder' can be significantly increased in this manner.

Subharmonic procedure The subharmonic of a sinusoidal voltage A with a defined frequency f_A is a sinusoidal voltage B, whose frequency f_B is derived from an integer division of the frequency f_A. The subharmonics of the frequency f_A are therefore the frequencies $f_A/2$, $f_A/3$, $f_A/4 \ldots$.

In the subharmonic transfer procedure, a second frequency f_B, which is usually lower by a factor of two, is derived by digital division by two of the reader's transmission frequency f_A. The output signal f_B of a binary divider can now be modulated with the data stream from the transponder. The modulated signal is then fed back into the transponder's antenna via an output driver.

One popular operating frequency for subharmonic systems is 128 kHz. This gives rise to a transponder response frequency of 64 kHz.

The transponder's antenna consists of a coil with a central tap, whereby the power supply is taken from one end. The transponder's return signal is fed into the coil's second connection (Figure 3.19).

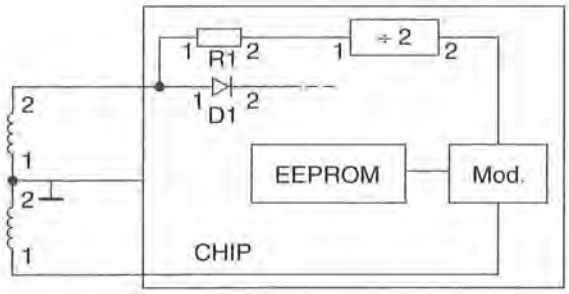


Figure 3.19 Basic circuit of a transponder with subharmonic back frequency. The received clocking signal is split into two, the data is modulated and fed into the transponder coil via a tap

3.2.2 Electromagnetic backscatter coupling

3.2.2.1 Power supply to the transponder

RFID systems in which the gap between reader and transponder is greater than 1 m are called *long-range systems*. These systems are operated at the *UHF frequencies* of 868 MHz (Europe) and 915 MHz (USA), and at the *microwave frequencies* 2.5 GHz and 5.8 GHz. The short wavelengths of these frequency ranges facilitate the construction of antennas with far smaller dimensions and greater efficiency than would be possible using frequency ranges below 30 MHz.

In order to be able to assess the energy available for the operation of a transponder we first calculate the *free space path loss* a_F in relation to the distance r between the transponder and the reader's antenna, the gain G_T and G_R of the transponder's and reader's antenna, plus the transmission frequency f of the reader:

$$a_F = -147.6 + 20\log(r) + 20\log(f) - 10\log(G_T) - 10\log(G_R) \qquad (3.1)$$

The free space path loss is a measure of the relationship between the HF power emitted by a reader into 'free space' and the HF power received by the transponder.

Using current low power semiconductor technology, transponder chips can be produced with a power consumption of no more than 5 μW (Friedrich and Annala, 2001). The efficiency of an integrated rectifier can be assumed to be 5–25% in the UHF and microwave range (Tanneberger, 1995). Given an efficiency of 10%, we thus require received power of $P_e = 50\,\mu W$ at the terminal of the transponder antenna for the operation of the transponder chip. This means that where the reader's transmission power is $P_s = 0.5$ W EIRP (effective isotropic radiated power) the free space path loss may not exceed 40 dB ($P_s/P_e = 10\,000/1$) if sufficiently high power is to be obtained at the transponder antenna for the operation of the transponder. A glance at Table 3.7 shows that at a transmission frequency of 868 MHz a *range* of a little over 3 m would be realisable; at 2.45 GHz a little over 1 m could be achieved. If the transponder's chip had a greater power consumption the achievable range would fall accordingly.

In order to achieve long ranges of up to 15 m or to be able to operate transponder chips with a greater power consumption at an acceptable range, backscatter transponders often have a backup battery to supply power to the transponder chip (Figure 3.20). To prevent this battery from being loaded unnecessarily, the microchips generally have

Table 3.7 Free space path loss a_F at different frequencies and distances. The gain of the transponder's antenna was assumed to be 1.64 (dipole), the gain of the reader's antenna was assumed to be 1 (isotropic emitter)

Distance r	868 MHz	915 MHz	2.45 GHz
0.3 m	18.6 dB	19.0 dB	27.6 dB
1 m	29.0 dB	29.5 dB	38.0 dB
3 m	38.6 dB	39.0 dB	47.6 dB
10 m	49.0 dB	49.5 dB	58.0 dB

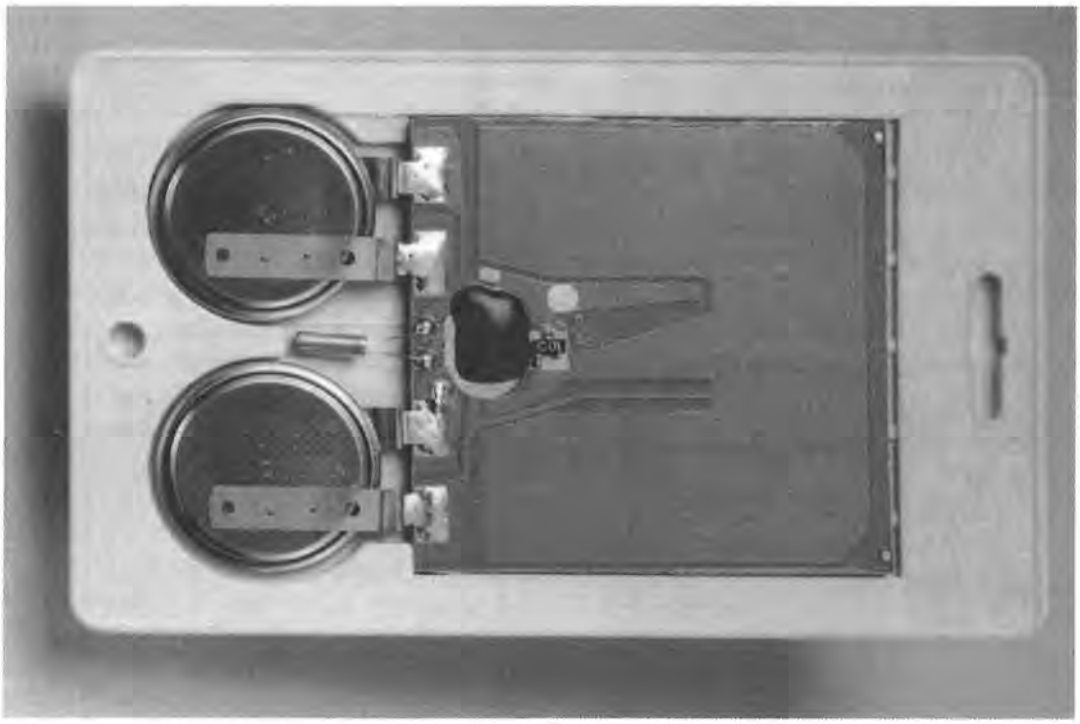


Figure 3.20 Active transponder for the frequency range 2.45 GHz. The data carrier is supplied with power by two *lithium batteries*. The transponder's microwave antenna is visible on the printed circuit board in the form of a u-shaped area (reproduced by permission of Pepperl & Fuchs, Mannheim)

a power saving 'power down' or 'stand-by' mode. If the transponder moves out of range of a reader, then the chip automatically switches over to the power saving 'power down' mode. In this state the power consumption is a few μA at most. The chip is not reactivated until a sufficiently strong signal is received in the read range of a reader, whereupon it switches back to normal operation. However, the battery of an active transponder never provides power for the transmission of data between transponder and reader, but serves exclusively for the supply of the microchip. Data transmission between transponder and reader relies exclusively upon the power of the electromagnetic field emitted by the reader.

3.2.2.2 Data transmission → reader

Modulated reflection cross-section We know from the field of *radar technology* that electromagnetic waves are reflected by objects with dimensions greater than around half the wavelength of the wave. The efficiency with which an object reflects electromagnetic waves is described by its *reflection cross-section*. Objects that are in resonance with the wave front that hits them, as is the case for antennas at the appropriate frequency, for example, have a particularly large reflection cross-section.

Power P_1 is emitted from the reader's antenna, a small proportion of which (free space attenuation) reaches the transponder's antenna (Figure 3.21). The power P_1' is supplied to the antenna connections as HF voltage and after rectification by the diodes D_1 and D_2 this can be used as turn-on voltage for the deactivation or activation of the power saving 'power down' mode. The diodes used here are *low barrier Schottky diodes*, which have a particularly low threshold voltage. The voltage obtained may also be sufficient to serve as a power supply for short ranges.

A proportion of the incoming power P_1' is reflected by the antenna and returned as power P_2. The *reflection characteristics* (=reflection cross-section) of the antenna can be influenced by altering the load connected to the antenna. In order to transmit data from the transponder to the reader, a load resistor R_L connected in parallel with the antenna is switched on and off in time with the data stream to be transmitted. The amplitude of the power P_2 reflected from the transponder can thus be modulated (→ modulated backscatter).

The power P_2 reflected from the transponder is radiated into free space. A small proportion of this (free space attenuation) is picked up by the reader's antenna. The reflected signal therefore travels into the antenna connection of the reader in the backwards direction and can be decoupled using a *directional coupler* and transferred to the receiver input of a reader. The forward signal of the transmitter, which is stronger by powers of ten, is to a large degree suppressed by the directional coupler.

The ratio of power transmitted by the reader and power returning from the transponder (P_1/P_2) can be estimated using the radar equation (for an explanation, refer to Chapter 4).

3.2.3 Close coupling

3.2.3.1 Power supply to the transponder

Close coupling systems are designed for ranges between 0.1 cm and a maximum of 1 cm. The transponder is therefore inserted into the reader or placed onto a marked surface ('*touch & go*') for operation.

Inserting the transponder into the reader, or placing it on the reader, allows the transponder coil to be precisely positioned in the *air gap* of a ring-shaped or U-shaped core. The functional layout of the transponder coil and reader coil corresponds with that of a transformer (Figure 3.22). The reader represents the primary winding and the transponder coil represents the secondary winding of a transformer. A high frequency alternating current in the primary winding generates a high frequency magnetic field in the core and air gap of the arrangement, which also flows through the transponder coil. This power is rectified to provide a power supply to the chip.

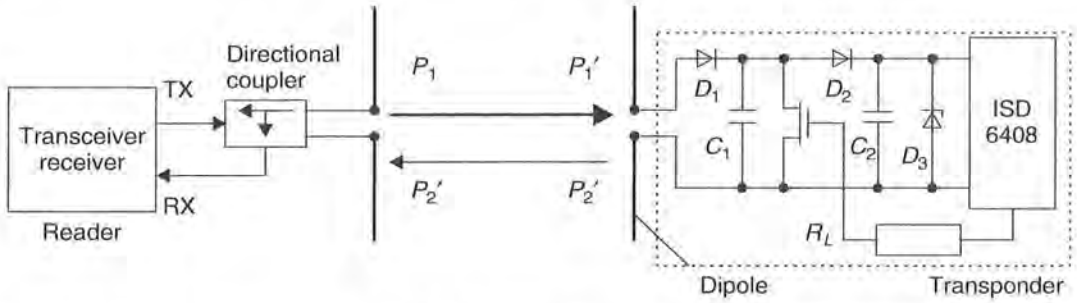


Figure 3.21 Operating principle of a backscatter transponder. The impedance of the chip is 'modulated' by switching the chip's FET (Integrated Silicon Design, 1996)

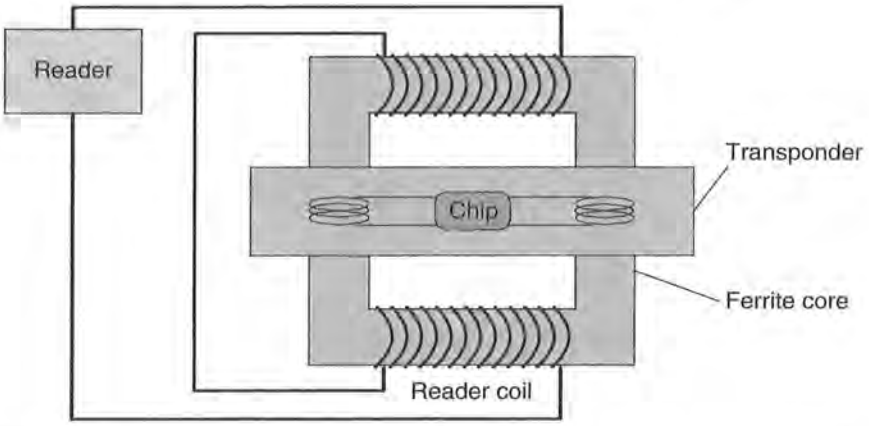


Figure 3.22 Close coupling transponder in an insertion reader with magnetic coupling coils

Because the voltage U induced in the transponder coil is proportional to the frequency f of the exciting current, the frequency selected for power transfer should be as high as possible. In practice, frequencies in the range 1–10 MHz are used. In order to keep the losses in the transformer core low, a ferrite material that is suitable for this frequency must be selected as the core material.

Because, in contrast to inductively coupled or microwave systems, the efficiency of power transfer from reader to transponder is very good, close coupling systems are excellently suited for the operation of chips with a high power consumption. This includes microprocessors, which still require some 10 mW power for operation (Sickert, 1994). For this reason, the close coupling chip card systems on the market all contain microprocessors.

The mechanical and electrical parameters of contactless close coupling chip cards are defined in their own standard, ISO 10536. For other designs the operating parameters can be freely defined.

3.2.3.2 Data transfer transponder → reader

Magnetic coupling Load modulation with subcarrier is also used for magnetically coupled data transfer from the transponder to the reader in close coupling systems.

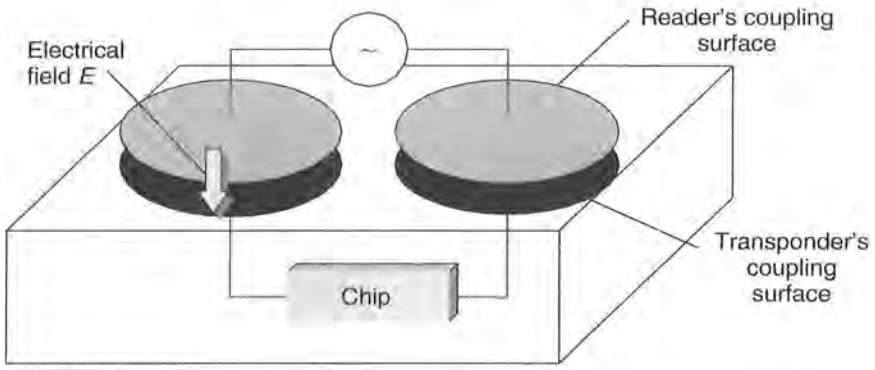


Figure 3.23 Capacitive coupling in close coupling systems occurs between two parallel metal surfaces positioned a short distance apart from each other

Subcarrier frequency and modulation is specified in ISO 10536 for close coupling chip cards.

Capacitive coupling Due to the short distance between the reader and transponder, close coupling systems may also employ *capacitive coupling* for data transmission. Plate capacitors are constructed from coupling surfaces isolated from one another, and these are arranged in the transponder and reader such that when a transponder is inserted they are exactly parallel to one another (Figure 3.23).

This procedure is also used in close coupling smart cards. The mechanical and electrical characteristics of these cards are defined in ISO 10536.

3.2.4 Electrical coupling

3.2.4.1 Power supply of passive transponders

In *electrically* (i.e. *capacitively*) coupled systems the reader generates a strong, high-frequency *electrical field*. The reader's antenna consists of a large, electrically conductive area (*electrode*), generally a metal foil or a metal plate. If a high-frequency voltage is applied to the electrode a high-frequency electric field forms between the electrode and the earth potential (ground). The voltages required for this, ranging between a few hundred volts and a few thousand volts, are generated in the reader by voltage rise in a resonant circuit made up of a coil L_1 in the reader, plus the parallel connection of an internal capacitor C_1 and the capacitance active between the electrode and the earth potential $C_{R\text{-}GND}$. The resonant frequency of the resonant circuit corresponds with the transmission frequency of the reader.

The antenna of the transponder is made up of two conductive surfaces lying in a plane (electrodes). If the transponder is placed within the electrical field of the reader, then an electric voltage arises between the two transponder electrodes, which is used to supply power to the transponder chips (Figure 3.24).

Since a capacitor is active both between the transponder and the transmission antenna ($C_{R\text{-}T}$) and between the transponder antenna and the earth potential ($C_{T\text{-}GND}$) the equivalent circuit diagram for an electrical coupling can be considered in a simplified form

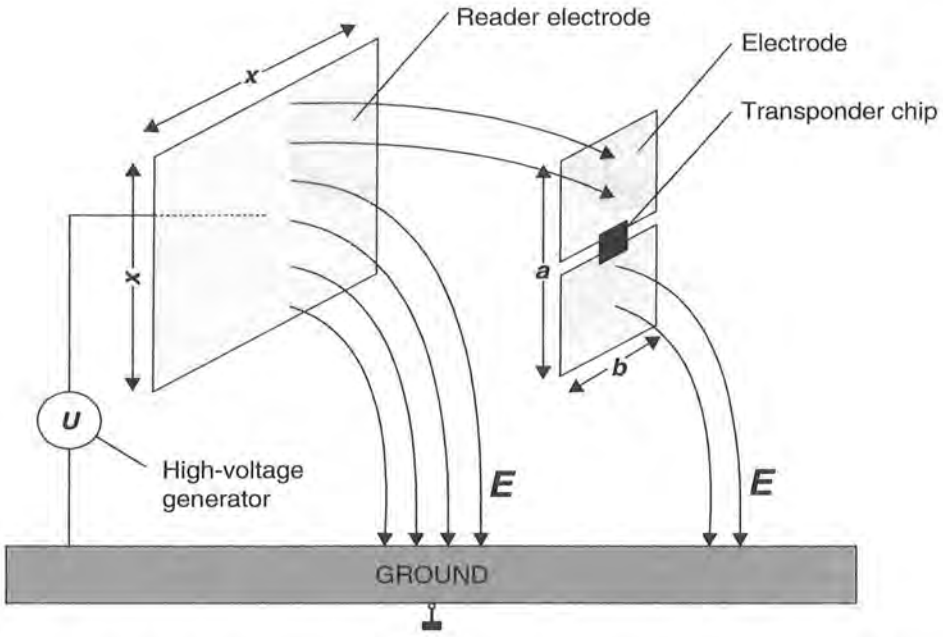


Figure 3.24 An electrically coupled system uses electrical (electrostatic) fields for the transmission of energy and data

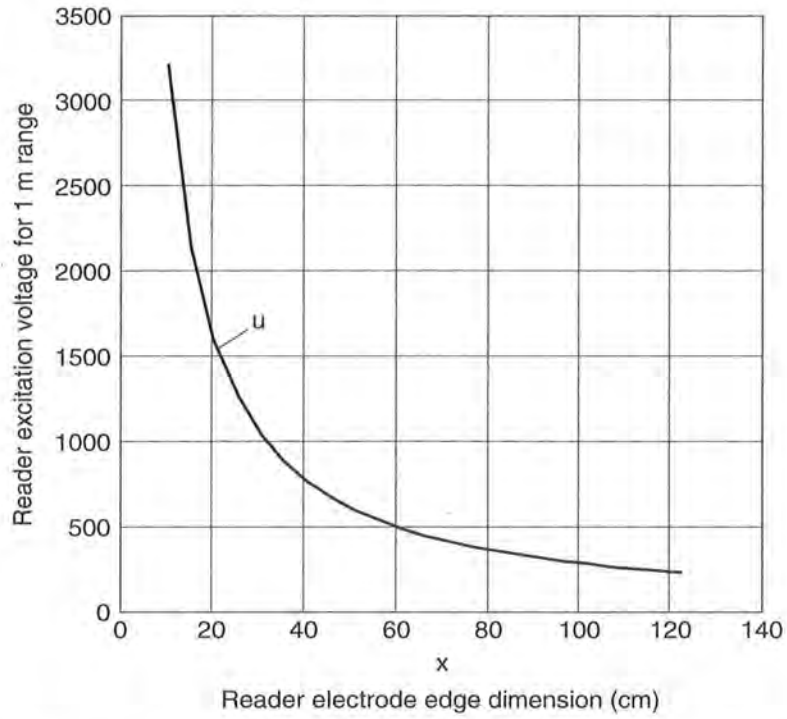


Figure 3.25 Necessary electrode voltage for the reading of a transponder with the electrode size a × b = 4.5 cm × 7 cm (format corresponds with a smart card), at a distance of 1 m ($f = 125$ kHz)

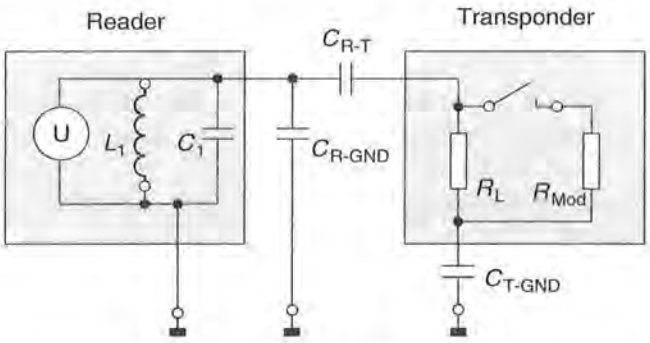


Figure 3.26 Equivalent circuit diagram of an electrically coupled RFID system

as a *voltage divider* with the elements $C_{R\text{-}T}$, R_L (input resistance of the transponder) and $C_{T\text{-}GND}$ (see Figure 3.26). Touching one of the transponder's electrodes results in the capacitance $C_{T\text{-}GND}$, and thus also the *read range*, becoming significantly greater.

The currents that flow in the electrode surfaces of the transponder are very small. Therefore, no particular requirements are imposed upon the conductivity of the electrode material. In addition to the normal metal surfaces (metal foil) the electrodes can thus also be made of conductive colours (e.g. a *silver conductive paste*) or a *graphite coating* (Motorola, Inc., 1999).

3.2.4.2 Data transfer transponder → reader

If an electrically coupled transponder is placed within the interrogation zone of a reader, the input resistance R_L of the transponder acts upon the resonant circuit of the reader via the coupling capacitance $C_{R\text{-}T}$ active between the reader and transponder electrodes, damping the resonant circuit slightly. This damping can be switched between two values by switching a modulation resistor R_{mod} in the transponder on and off. Switching the modulation resistor R_{mod} on and off thereby generates an amplitude modulation of the voltage present at L_1 and C_1 by the remote transponder. By switching the modulation resistor R_{mod} on and off in time with data, this data can be transmitted to the reader. This procedure is called *load modulation*.

3.2.5 Data transfer reader → transponder

All known digital modulation procedures are used in data transfer from the reader to the transponder in full and half duplex systems, irrespective of the operating frequency or the coupling procedure. There are three basic procedures:

- ASK: amplitude shift keying
- FSK: frequency shift keying
- PSK: phase shift keying

Because of the simplicity of demodulation, the majority of systems use ASK modulation.

3.3 Sequential Procedures

If the transmission of data and power from the reader to the data carrier alternates with data transfer from the transponder to the reader, then we speak of a *sequential procedure* (SEQ).

The characteristics used to differentiate between SEQ and other systems have already been described in Section 3.2.

3.3.1 Inductive coupling

3.3.1.1 Power supply to the transponder

Sequential systems using inductive coupling are operated exclusively at frequencies below 135 kHz. A transformer type coupling is created between the reader's coil and the transponder's coil. The induced voltage generated in the transponder coil by the effect of an alternating field from the reader is rectified and can be used as a power supply.

In order to achieve higher efficiency of data transfer, the transponder frequency must be precisely matched to that of the reader, and the quality of the transponder coil must be carefully specified. For this reason the transponder contains an *on-chip trimming capacitor* to compensate for resonant frequency manufacturing tolerances.

However, unlike full and half duplex systems, in sequential systems the reader's transmitter does not operate on a continuous basis. The energy transferred to the transmitter during the transmission operation charges up a *charging capacitor* to provide an energy store. The transponder chip is switched over to stand-by or power saving mode during the charging operation, so that almost all of the energy received is used to charge up the charging capacitor. After a fixed charging period the reader's transmitter is switched off again.

The energy stored in the transponder is used to send a reply to the reader. The minimum capacitance of the charging capacitor can be calculated from the necessary operating voltage and the chip's power consumption:

$$C = \frac{Q}{U} = \frac{It}{[V_{max} - V_{min}]} \tag{3.2}$$

where V_{max}, V_{min} are limit values for operating voltage that may not be exceeded, I is the power consumption of the chip during operation and t is the time required for the transmission of data from transponder to reader.

For example, the parameters $I = 5\,\mu A$, $t = 20\,ms$, $V_{max} = 4.5\,V$ and $V_{min} = 3.5\,V$ yield a charging capacitor of $C = 100\,nF$ (Schürmann, 1993).

3.3.1.2 A comparison between FDX/HDX and SEQ systems

Figure 3.27 illustrates the different conditions arising from full/half duplex (FDX/HDX) and sequential (SEQ) systems.

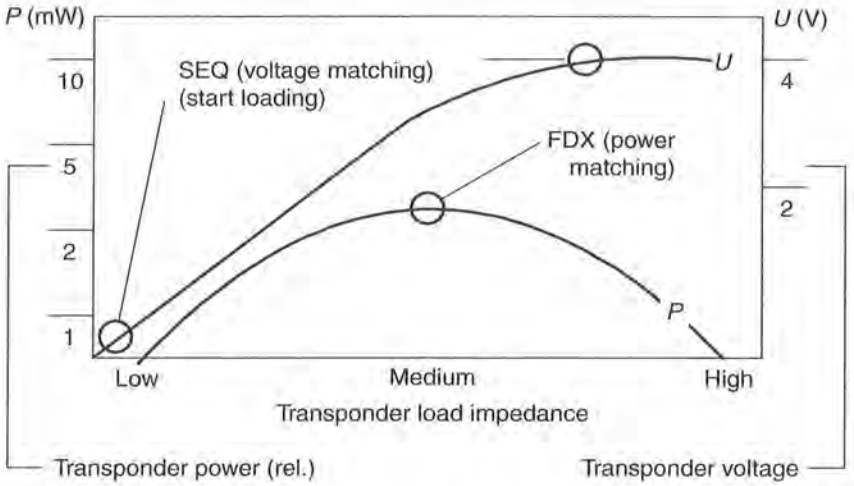


Figure 3.27 Comparison of induced transponder voltage in FDX/HDX and SEQ systems (Schürmann, 1993)

Because the power supply from the reader to the transponder in full duplex systems occurs at the same time as data transfer in both directions, the chip is permanently in operating mode. *Power matching* between the transponder antenna (current source) and the chip (current consumer) is desirable to utilise the transmitted energy optimally. However, if precise power matching is used only half of the source voltage (=open circuit voltage of the coil) is available. The only option for increasing the available operating voltage is to increase the impedance (=load resistance) of the chip. However, this is the same as decreasing the power consumption.

Therefore the design of full duplex systems is always a compromise between power matching (maximum power consumption P_{chip} at $U_{chip} = 1/2U_O$) and voltage matching (minimum power consumption P_{chip} at maximum voltage $U_{chip} = U_O$).

The situation is completely different in sequential systems: during the charging process the chip is in stand-by or power saving mode, which means that almost no power is drawn through the chip.

The charging capacitor is fully discharged at the beginning of the charging process and therefore represents a very low ohmic load for the voltage source (Figure 3.27: start loading). In this state, the maximum amount of current flows into the charging capacitor, whereas the voltage approaches zero (=*current matching*). As the charging capacitor is charged, the charging current starts to decrease according to an exponential function, and reaches zero when the capacitor is fully charged. The state of the charged capacitor corresponds with *voltage matching* at the transponder coil.

This achieves the following advantages for the chip power supply compared to a full/half duplex system:

- The full source voltage of the transponder coil is available for the operation of the chip. Thus the available operating voltage is up to twice that of a comparable full/half duplex system.
- The energy available to the chip is determined only by the capacitance of the charging capacitor and the charging period. Both values can in theory (!) be given any

required magnitude. In full/half duplex systems the maximum power consumption of the chip is fixed by the power matching point (i.e. by the coil geometry and field strength H).

3.3.1.3 Data transmission transponder → reader

In sequential systems (Figure 3.28) a full read cycle consists of two phases, the charging phase and the reading phase (Figure 3.29).

The end of the charging phase is detected by an *end of burst detector*, which monitors the path of voltage at the transponder coil and thus recognises the moment when the reader field is switched off. At the end of the charging phase an on-chip oscillator, which uses the resonant circuit formed by the transponder coil as a frequency determining component, is activated. A weak magnetic alternating field is generated by the transponder coil, and this can be received by the reader. This gives an improved signal-interference distance of typically 20 dB compared to full/half duplex systems, which has a positive effect upon the ranges that can be achieved using sequential systems.

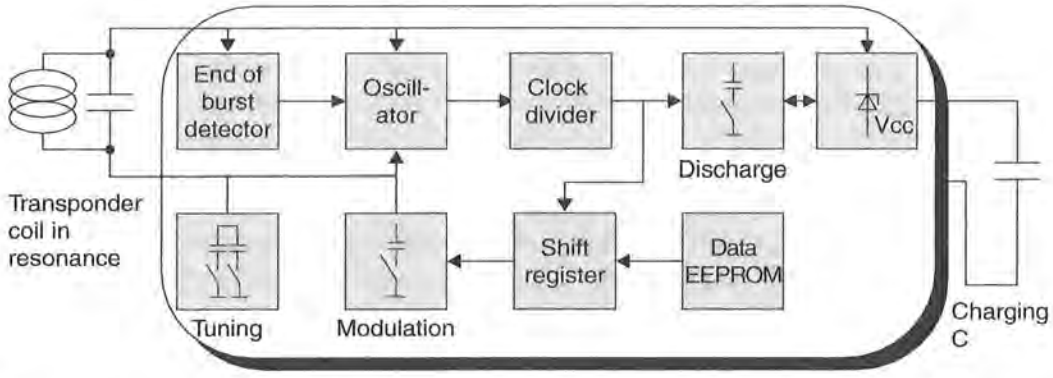


Figure 3.28 Block diagram of a sequential transponder by Texas Instruments TIRIS® Systems, using inductive coupling

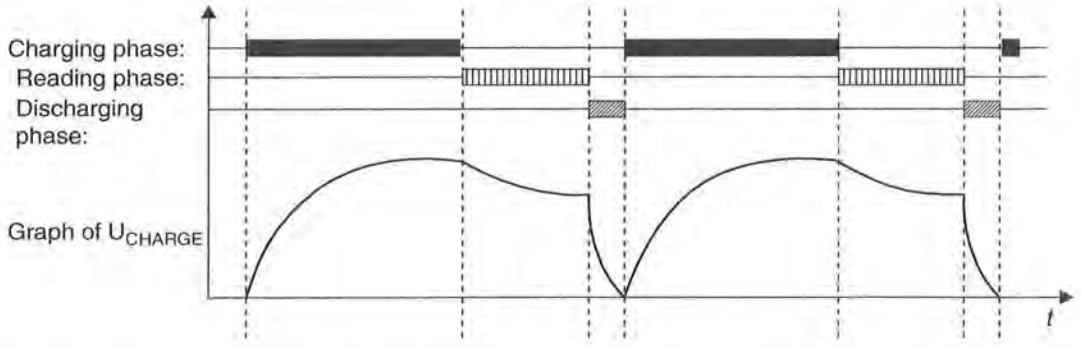


Figure 3.29 Voltage path of the charging capacitor of an inductively coupled SEQ transponder during operation

The transmission frequency of the transponder corresponds with the resonant frequency of the transponder coil, which was adjusted to the transmission frequency of the reader when it was generated.

In order to be able to modulate the HF signal generated in the absence of a power supply, an additional modulation capacitor is connected in parallel with the resonant circuit in time with the data flow. The resulting frequency shift keying provides a *2 FSK modulation*.

After all the data has been transmitted, the discharge mode is activated to fully discharge the charging capacitor. This guarantees a safe Power-On-Reset at the start of the next charging cycle.

3.3.2 Surface acoustic wave transponder

Surface acoustic wave (SAW) devices are based upon the piezoelectric effect and on the surface-related dispersion of elastic (=acoustic) waves at low speed. If an (ionic) crystal is elastically deformed in a certain direction, surface charges occur, giving rise to electric voltages in the crystal (application: piezo lighter). Conversely, the application of a surface charge to a crystal leads to an elastic deformation in the crystal grid (application: piezo buzzer). Surface acoustic wave devices are operated at microwave frequencies, normally in the ISM range 2.45 GHz.

Electroacoustic transducers (interdigital transducers) and *reflectors* can be created using planar electrode structures on piezoelectric substrates. The normal substrate used for this application is *lithium niobate* or *lithium tantalate*. The electrode structure is created by a photolithographic procedure, similar to the procedure used in microelectronics for the manufacture of integrated circuits.

Figure 3.30 illustrates the basic layout of a surface wave transponder. A finger-shaped electrode structure — the *interdigital transducer* — is positioned at the end of a long piezoelectrical substrate, and a suitable *dipole antenna* for the operating frequency is attached to its busbar. The interdigital transducer is used to convert between electrical signals and acoustic surface waves.

An electrical impulse applied to the busbar causes a mechanical deformation to the surface of the substrate due to the piezoelectrical effect between the electrodes (fingers), which disperses in both directions in the form of a surface wave (rayleigh wave). For a normal substrate the dispersion speed lies between 3000 and 4000 m/s. Similarly, a *surface wave* entering the converter creates an electrical impulse at the busbar of the interdigital transducer due to the piezoelectric effect.

Individual electrodes are positioned along the remaining length of the surface wave transponder. The edges of the electrodes form a reflective strip and reflect a small proportion of the incoming surface waves. Reflector strips are normally made of aluminium; however some reflector strips are also in the form of etched grooves (Meinke, 1992).

A high frequency *scanning pulse* generated by a reader is supplied from the dipole antenna of the transponder into the interdigital transducer and is thus converted into an acoustic surface wave, which flows through the substrate in the longitudinal direction. The frequency of the surface wave corresponds with the carrier frequency of the sampling pulse (e.g. 2.45 GHz) (Figure 3.31). The carrier frequency of the reflected

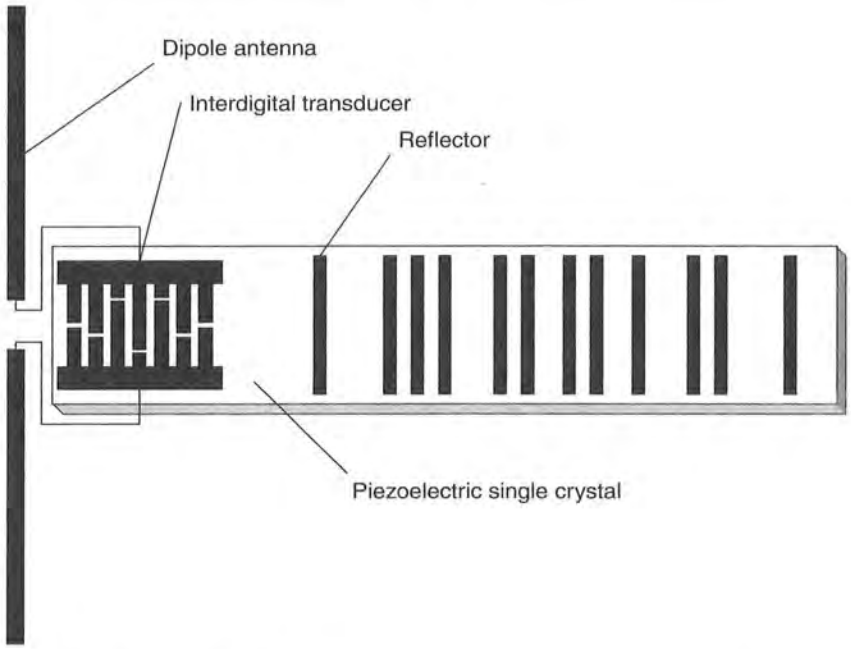


Figure 3.30 Basic layout of an SAW transponder. Interdigital transducers and reflectors are positioned on the piezoelectric crystal

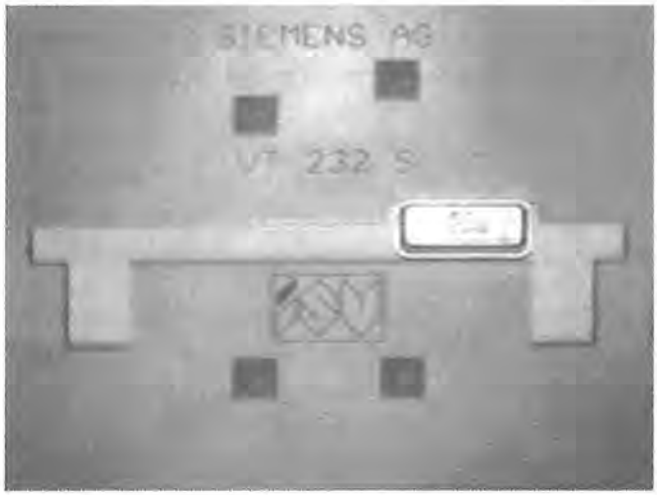


Figure 3.31 Surface acoustic wave transponder for the frequency range 2.45 GHz with antenna in the form of microstrip line. The piezocrystal itself is located in an additional metal housing to protect it against environmental influences (reproduced by permission of Siemens AG, ZT KM, Munich)

and returned pulse sequence thus corresponds with the transmission frequency of the sampling pulse.

Part of the surface wave is reflected off each of the reflective strips that are distributed across the substrate, while the remaining part of the surface wave continues to travel to the end of the substrate and is absorbed there.

The reflected parts of the wave travel back to the interdigital transducer, where they are converted into a high frequency pulse sequence and are emitted by the dipole antenna. This pulse sequence can be received by the reader. The number of pulses received corresponds with the number of reflective strips on the substrate. Likewise, the delay between the individual pulses is proportional to the spatial distance between the reflector strips on the substrate, and so the spatial layout of the reflector strips can represent a binary sequence of digits.

Due to the slow dispersion speed of the surface waves on the substrate the first response pulse is only received by the reader after a dead time of around 1.5 ms after the transmission of the scanning pulse. This gives decisive advantages for the reception of the pulse.

Reflections of the scanning pulse on the metal surfaces of the environment travel back to the antenna of the reader at the speed of light. A reflection over a distance of 100 m to the reader would arrive at the reader 0.6 ms after emission from the reader's antenna (travel time there and back, the signal is damped by >160 dB). Therefore, when the transponder signal returns after 1.5 ms all reflections from the environment of the reader have long since died away, so they cannot lead to errors in the pulse sequence (Dziggel, 1997).

The data storage capacity and data transfer speed of a surface wave transponder depend upon the size of the substrate and the realisable minimum distance between the reflector strips on the substrate. In practice, around 16–32 bits are transferred at a data transfer rate of 500 kbit/s (Siemens, n.d.).

The range of a surface wave system depends mainly upon the transmission power of the scanning pulse and can be estimated using the radar equation (Chapter 4). At the permissible transmission power in the 2.45 GHz ISM frequency range a range of 1–2 m can be expected.

4

Physical Principles of RFID Systems

The vast majority of RFID systems operate according to the principle of *inductive coupling*. Therefore, understanding of the procedures of power and data transfer requires a thorough grounding in the physical principles of magnetic phenomena. This chapter therefore contains a particularly intensive study of the theory of magnetic fields from the point of view of RFID.

Electromagnetic fields — radio waves in the classic sense — are used in RFID systems that operate at above 30 MHz. To aid understanding of these systems we will investigate the propagation of waves in the far field and the principles of radar technology.

Electric fields play a secondary role and are only exploited for capacitive data transmission in close coupling systems. Therefore, this type of field will not be discussed further.

4.1 Magnetic Field

4.1.1 Magnetic field strength *H*

Every moving charge (electrons in wires or in a vacuum), i.e. flow of current, is associated with a *magnetic field* (Figure 4.1). The intensity of the magnetic field can be demonstrated experimentally by the forces acting on a magnetic needle (compass) or a second electric current. The magnitude of the magnetic field is described by the *magnetic field strength* H regardless of the material properties of the space.

In the general form we can say that: 'the contour integral of magnetic field strength along a closed curve is equal to the sum of the current strengths of the currents within it' (Kuchling, 1985).

$$\sum I = \oint \vec{H} \cdot \vec{ds} \tag{4.1}$$

We can use this formula to calculate the field strength H for different types of conductor. See Figure 4.2.

RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification. Klaus Finkenzeller
 ISBN: 0-470-84402-7

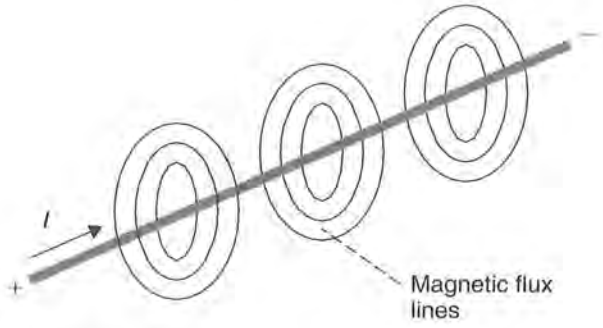


Figure 4.1 Lines of magnetic flux are generated around every current-carrying conductor

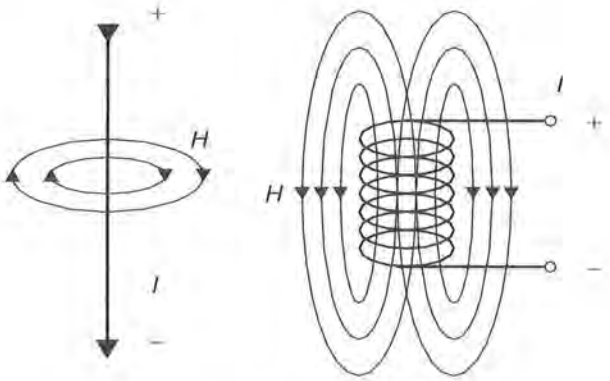


Figure 4.2 Lines of magnetic flux around a current-carrying conductor and a current-carrying cylindrical coil

Table 4.1 Constants used

Constant	Symbol	Value and unit
Electric field constant	ε_0	8.85×10^{-12} As/Vm
Magnetic field constant	μ_0	1.257×10^{-6} Vs/Am
Speed of light	c	299 792 km/s
Boltzmann constant	k	$1.380\,662 \times 10^{-23}$ J/K

In a straight conductor the field strength H along a circular *flux line* at a distance r is constant. The following is true (Kuchling, 1985):

$$H = \frac{1}{2\pi r} \tag{4.2}$$

4.1.1.1 Path of field strength H(x) in conductor loops

So-called 'short cylindrical coils' or conductor loops are used as magnetic antennas to generate the *magnetic alternating field* in the write/read devices of inductively coupled RFID systems (Figure 4.3).

Table 4.2 Units and abbreviations used

Variable	Symbol	Unit	Abbreviation
Magnetic field strength	H	Ampere per meter	A/m
Magnetic flux (n = number of windings)	Φ	Volt seconds	Vs
	$\Psi = n\Phi$		
Magnetic inductance	B	Volt seconds per meter squared	Vs/m^2
Inductance	L	Henry	H
Mutual inductance	M	Henry	H
Electric field strength	E	Volts per metre	V/m
Electric current	I	Ampere	A
Electric voltage	U	Volt	V
Capacitance	C	Farad	F
Frequency	f	Hertz	Hz
Angular frequency	$\omega = 2\pi f$	1/seconds	1/s
Length	l	Metre	m
Area	A	Metre squared	m^2
Speed	v	Metres per second	m/s
Impedance	Z	Ohm	Ω
Wavelength	λ	Metre	m
Power	P	Watt	W
Power density	S	Watts per metre squared	W/m^2

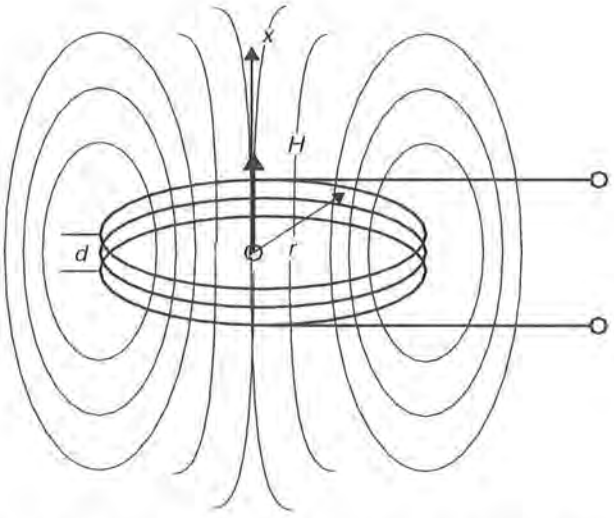


Figure 4.3 The path of the lines of magnetic flux around a short cylindrical coil, or conductor loop, similar to those employed in the transmitter antennas of inductively coupled RFID systems

If the measuring point is moved away from the centre of the coil along the coil axis (x axis), then the strength of the field H will decrease as the distance x is increased. A more in-depth investigation shows that the field strength in relation to the radius (or area) of the coil remains constant up to a certain distance and then falls rapidly (see Figure 4.4). In free space, the decay of field strength is approximately 60 dB per

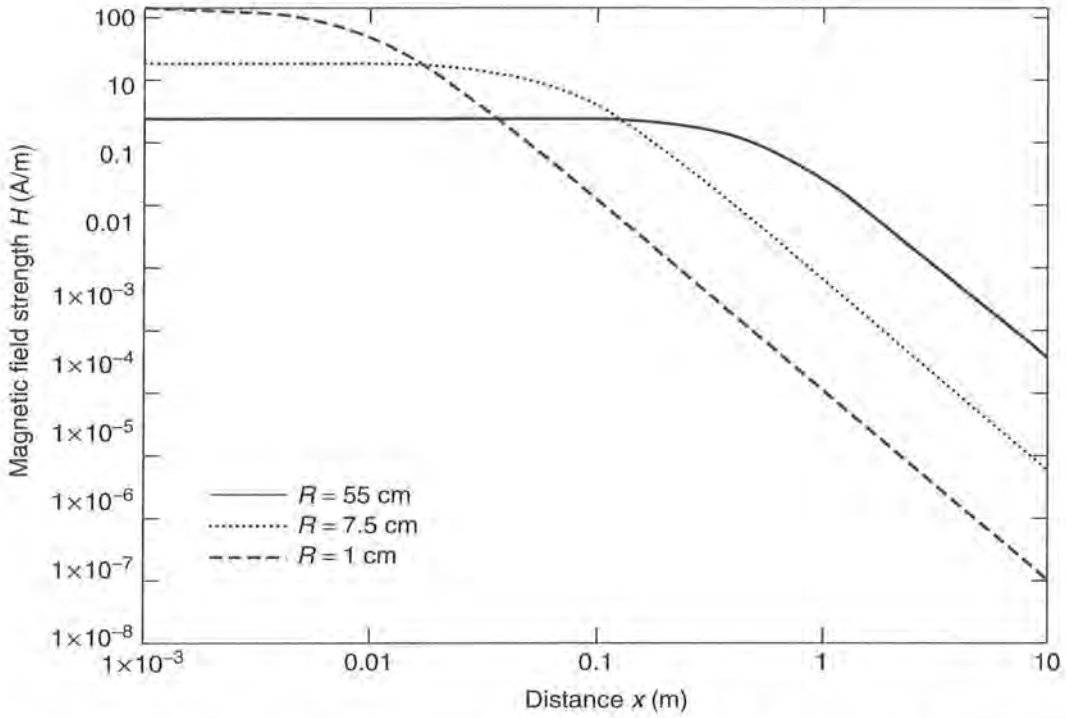


Figure 4.4 Path of magnetic field strength H in the near field of short cylinder coils, or conductor coils, as the distance in the x direction is increased

decade in the near field of the coil, and flattens out to 20 dB per decade in the far field of the electromagnetic wave that is generated (a more precise explanation of these effects can be found in Section 4.2.1).

The following equation can be used to calculate the path of field strength along the x axis of a round coil (= conductor loop) similar to those employed in the transmitter antennas of inductively coupled RFID systems (Paul, 1993):

$$H = \frac{I \cdot N \cdot R^2}{2\sqrt{(R^2 + x^2)^3}} \tag{4.3}$$

where N is the number of windings, R is the circle radius r and x is the distance from the centre of the coil in the x direction. The following boundary condition applies to this equation: $d \ll R$ and $x < \lambda/2\pi$ (the transition into the electromagnetic far field begins at a distance $>2\pi$; see Section 4.2.1).

At distance 0 or, in other words, at the centre of the antenna, the formula can be simplified to (Kuchling, 1985):

$$H = \frac{I \cdot N}{2R} \tag{4.4}$$

We can calculate the *field strength path* of a rectangular conductor loop with edge length $a \times b$ at a distance of x using the following equation. This format is often used

as a transmitter antenna.

$$H = \frac{N \cdot I \cdot ab}{4\pi\sqrt{\left(\frac{a}{2}\right)^2 + \left(\frac{b}{2}\right)^2 + x^2}} \cdot \left(\frac{1}{\left(\frac{a}{2}\right)^2 + x^2} + \frac{1}{\left(\frac{b}{2}\right)^2 + x^2} \right) \tag{4.5}$$

Figure 4.4 shows the calculated field strength path $H(x)$ for three different antennas at a distance 0–20 m. The number of windings and the antenna current are constant in each case; the antennas differ only in radius R. The calculation is based upon the following values: $H1$: $R = 55$ cm, $H2$: $R = 7.5$ cm, $H3$: $R = 1$ cm.

The calculation results confirm that the increase in field strength flattens out at short distances ($x < R$) from the antenna coil. Interestingly, the smallest antenna exhibits a significantly higher field strength at the centre of the antenna (distance = 0), but at greater distances ($x > R$) the largest antenna generates a significantly higher field strength. It is vital that this effect is taken into account in the design of antennas for inductively coupled RFID systems.

4.1.1.2 Optimal antenna diameter

If the radius R of the transmitter antenna is varied at a constant distance x from the transmitter antenna under the simplifying assumption of constant coil current I in the transmitter antenna, then field strength H is found to be at its highest at a certain ratio of distance x to antenna radius R. This means that for every *read range* of an RFID system there is an optimal antenna radius R. This is quickly illustrated by a glance at Figure 4.4: if the selected antenna radius is too great, the field strength is too low even at a distance $x = 0$ from the transmission antenna. If, on the other hand, the selected antenna radius is too small, then we find ourselves within the range in which the field strength falls in proportion to x^3.

Figure 4.5 shows the graph of field strength H as the coil radius R is varied. The optimal coil radius for different read ranges is always the maximum point of the graph $H(R)$. To find the mathematical relationship between the maximum field strength H and the coil radius R we must first find the inflection point of the function $H(R)$ (see equation 4.3) (Lee, 1999). To do this we find the first derivative $H'(R)$ by differentiating $H(R)$ with respect to R:

$$H'(R) = \frac{\mathrm{d}}{\mathrm{d}R}H(R) = \frac{2 \cdot I \cdot N \cdot R}{\sqrt{(R^2 + x^2)^3}} - \frac{3 \cdot I \cdot N \cdot R^3}{(R^2 + x^2) \cdot \sqrt{(R^2 + x^2)^3}} \tag{4.6}$$

The inflection point, and thus the maximum value of the function $H(R)$, is found from the following zero points of the derivative $H'(R)$:

$$R_1 = x \cdot \sqrt{2}; \qquad R_2 = -x \cdot \sqrt{2} \tag{4.7}$$

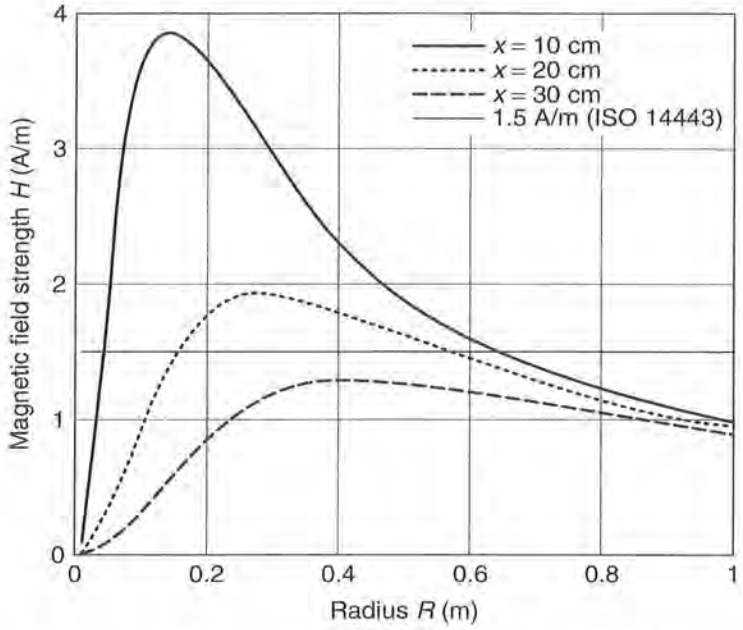


Figure 4.5 Field strength H of a transmission antenna given a constant distance x and variable radius R, where $I = 1$ A and $N = 1$

The optimal radius of a transmission antenna is thus twice the maximum desired read range. The second zero point is negative merely because the magnetic field H of a conductor loop propagates in both directions of the x axis (see also Figure 4.3).

However, an accurate assessment of a system's maximum read range requires knowledge of the *interrogation field strength* H_{min} of the transponder in question (see Section 4.1.9). If the selected antenna radius is too great, then there is the danger that the field strength H may be too low to supply the transponder with sufficient operating energy, even at a distance $x = 0$.

4.1.2 Magnetic flux and magnetic flux density

The magnetic field of a (cylindrical) coil will exert a force on a magnetic needle. If a soft iron core is inserted into a (cylindrical) coil — all other things remaining equal — then the force acting on the magnetic needle will increase. The quotient $I \times N$ (Section 4.1.1) remains constant and therefore so does field strength. However, the flux density — the total number of flux lines — which is decisive for the force generated (cf. Pauls, 1993), has increased.

The total number of lines of magnetic flux that pass through the inside of a cylindrical coil, for example, is denoted by *magnetic flux* Φ. Magnetic flux density B is a further variable related to area A (this variable is often referred to as 'magnetic inductance B in the literature') (Reichel, 1980). Magnetic flux is expressed as:

$$\Phi = B \cdot A \tag{4.8}$$

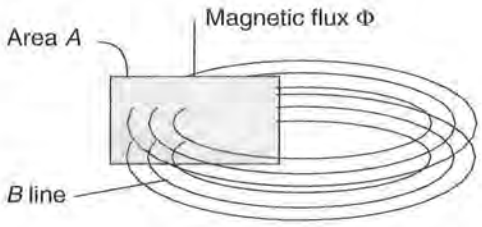


Figure 4.6 Relationship between magnetic flux Φ and flux density B

The material relationship between flux density B and field strength H (Figure 4.6) is expressed by the material equation:

$$B = \mu_0 \mu_r H = \mu H \tag{4.9}$$

The constant μ_0 is the magnetic field constant ($\mu_0 = 4\pi \times 10^{-6}$ Vs/Am) and describes the permeability (= magnetic conductivity) of a vacuum. The variable μ_r is called relative permeability and indicates how much greater than or less than μ_0 the permeability of a material is.

4.1.3 Inductance *L*

A magnetic field, and thus a magnetic flux Φ, will be generated around a conductor of any shape. This will be particularly intense if the conductor is in the form of a loop (coil). Normally, there is not one conduction loop, but N loops of the same area A, through which the same current I flows. Each of the conduction loops contributes the same proportion Φ to the total flux ψ (Paul, 1993).

$$\Psi = \sum_N \Phi_N = N \cdot \Phi = N \cdot \mu \cdot H \cdot A \tag{4.10}$$

The ratio of the interlinked flux ψ that arises in an area enclosed by current I, to the current in the conductor that encloses it (conductor loop) is denoted by *inductance* L (Figure 4.7):

$$L = \frac{\Psi}{I} = \frac{N \cdot \Phi}{I} = \frac{N \cdot \mu \cdot H \cdot A}{I} \tag{4.11}$$

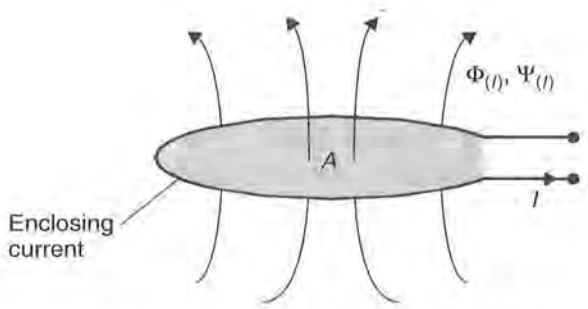


Figure 4.7 Definition of inductance L

Inductance is one of the characteristic variables of conductor loops (coils). The inductance of a conductor loop (coil) depends totally upon the material properties (permeability) of the space that the flux flows through and the geometry of the layout.

4.1.3.1 Inductance of a conductor loop

If we assume that the diameter d of the wire used is very small compared to the diameter D of the conductor coil ($d/D < 0.0001$) a very simple approximation can be used:

$$L = N^2 \mu_0 R \cdot \ln\left(\frac{2R}{d}\right) \tag{4.12}$$

where R is the radius of the conductor loop and d is the diameter of the wire used.

4.1.4 Mutual inductance *M*

If a second conductor loop 2 (area A_2) is located in the vicinity of conductor loop 1 (area A_1), through which a current is flowing, then this will be subject to a proportion of the total magnetic flux Φ flowing through A_1. The two circuits are connected together by this partial flux or coupling flux. The magnitude of the coupling flux ψ_{21} depends upon the geometric dimensions of both conductor loops, the position of the conductor loops in relation to one another, and the magnetic properties of the medium (e.g. permeability) in the layout.

Similarly to the definition of the (self) inductance L of a conductor loop, the *mutual inductance* M_{21} of conductor loop 2 in relation to conductor loop 1 is defined as the ratio of the partial flux ψ_{21} enclosed by conductor loop 2, to the current I_1 in conductor loop 1 (Paul, 1993):

$$M_{21} = \frac{\Psi_{21}(I_1)}{I_1} = \oint_{A2} \frac{B_2(I_1)}{I_1} \cdot dA_2 \tag{4.13}$$

Similarly, there is also a mutual inductance M_{12}. Here, current I_2 flows through the conductor loop 2, thereby determining the coupling flux ψ_{12} in loop 1. The following relationship applies:

$$M = M_{12} = M_{21} \tag{4.14}$$

Mutual inductance describes the coupling of two circuits via the medium of a magnetic field (Figure 4.8). Mutual inductance is always present between two electric circuits. Its dimension and unit are the same as for inductance.

The coupling of two electric circuits via the magnetic field is the physical principle upon which inductively coupled RFID systems are based. Figure 4.9 shows a calculation of the mutual inductance between a transponder antenna and three different reader antennas, which differ only in diameter. The calculation is based upon the following values: M_1: $R = 55\,\text{cm}$, M_2: $R = 7.5\,\text{cm}$, M_3: $R = 1\,\text{cm}$, transponder: $R = 3.5\,\text{cm}$. $N = 1$ for all reader antennas.

The graph of *mutual inductance* shows a strong similarity to the graph of magnetic field strength H along the x axis. Assuming a homogeneous magnetic field, the mutual

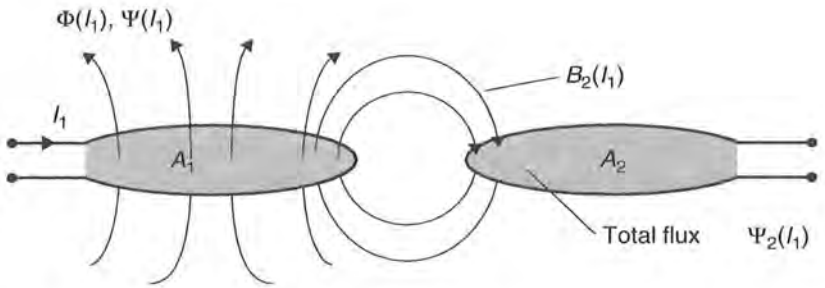


Figure 4.8 The definition of mutual inductance M_{21} by the coupling of two coils via a partial magnetic flow

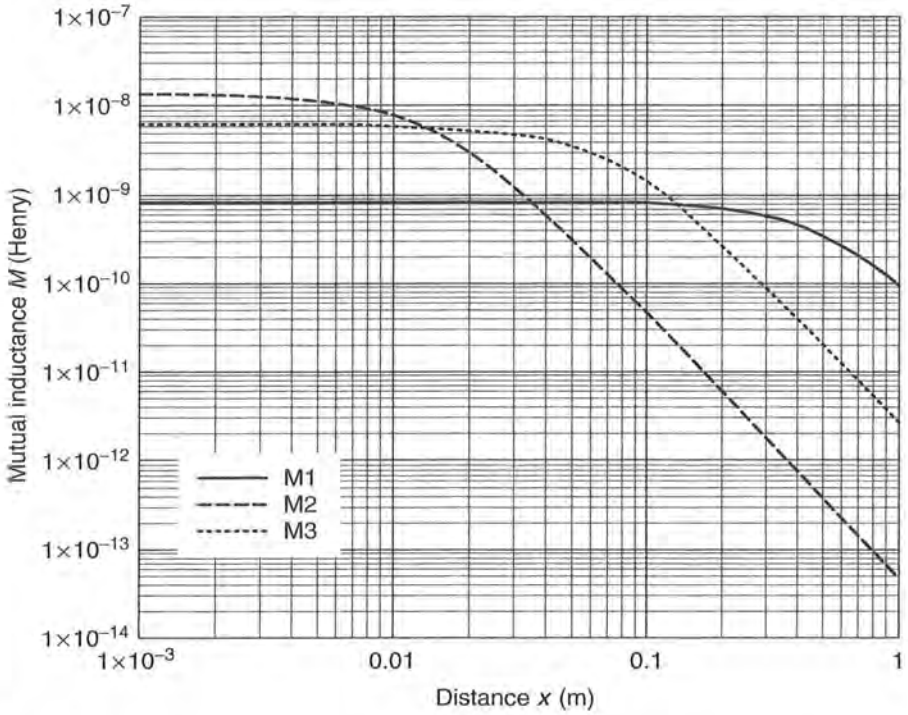


Figure 4.9 Graph of mutual inductance between reader and transponder antenna as the distance in the x direction increases

inductance M_{12} between two coils can be calculated using equation (4.13). It is found to be:

$$M_{12} = \frac{B_2(I_1) \cdot N_2 \cdot A_2}{I_1} = \frac{\mu_0 \cdot H(I_1) \cdot N_2 \cdot A_2}{I_1} \tag{4.15}$$

We first replace $H(I_1)$ with the expression in equation (4.4), and substitute $R^2\pi$ for A, thus obtaining:

$$M_{12} = \frac{\mu_0 \cdot N_1 \cdot R_1^2 \cdot N_2 \cdot R_2^2 \cdot \pi}{2\sqrt{(R_1^2 + x^2)^3}} \tag{4.16}$$

In order to guarantee the homogeneity of the magnetic field in the area A_2 the condition $A_2 \leq A_1$ should be fulfilled. Furthermore, this equation only applies to the case where the x axes of the two coils lie on the same plane. Due to the relationship $M = M_{12} = M_{21}$ the mutual inductance can be calculated as follows for the case $A_2 \geq A_1$:

$$M_{21} = \frac{\mu_0 \cdot N_1 \cdot R_1^2 \cdot N_2 \cdot R_2^2 \cdot \pi}{2\sqrt{(R_2^2 + x^2)^3}} \tag{4.17}$$

4.1.5 Coupling coefficient *k*

Mutual inductance is a quantitative description of the flux coupling of two conductor loops. The *coupling coefficient* k is introduced so that we can make a qualitative prediction about the coupling of the conductor loops independent of their geometric dimensions. The following applies:

$$k = \frac{M}{\sqrt{L_1 \cdot L_2}} \tag{4.18}$$

The coupling coefficient always varies between the two extreme cases $0 \leq k \leq 1$.

- $k = 0$: Full decoupling due to great distance or magnetic shielding.
- $k = 1$: Total coupling. Both coils are subject to the same magnetic flux Φ. The transformer is a technical application of total coupling, whereby two or more coils are wound onto a highly permeable iron core.

An analytic calculation is only possible for very simple antenna configurations. For two parallel conductor loops centred on a single x axis the coupling coefficient according to Roz and Fuentes (n.d.) can be approximated from the following equation. However, this only applies if the radii of the conductor loops fulfil the condition $r_{\text{Transp}} \leq r_{\text{Reader}}$. The distance between the conductor loops on the x axis is denoted by x.

$$k(x) \approx \frac{r_{\text{Transp}}^2 \cdot r_{\text{Reader}}^2}{\sqrt{r_{\text{Transp}} \cdot r_{\text{Reader}}} \cdot \left(\sqrt{x^2 + r_{\text{Reader}}^2}\right)^3} \tag{4.19}$$

Due to the fixed link between the coupling coefficient and mutual inductance M, and because of the relationship $M = M_{12} = M_{21}$, the formula is also applicable to transmitter antennas that are smaller than the transponder antenna. Where $r_{\text{Transp}} \geq r_{\text{Reader}}$, we write:

$$k(x) \approx \frac{r_{\text{Transp}}^2 \cdot r_{\text{Reader}}^2}{\sqrt{r_{\text{Transp}} \cdot r_{\text{Reader}}} \cdot \left(\sqrt{x^2 + r_{\text{Transp}}^2}\right)^3} \tag{4.20}$$

The coupling coefficient $k(x) = 1$ (= 100%) is achieved where the distance between the conductor loops is zero ($x = 0$) and the antenna radii are identical ($r_{\text{Transp}} = r_{\text{Reader}}$),

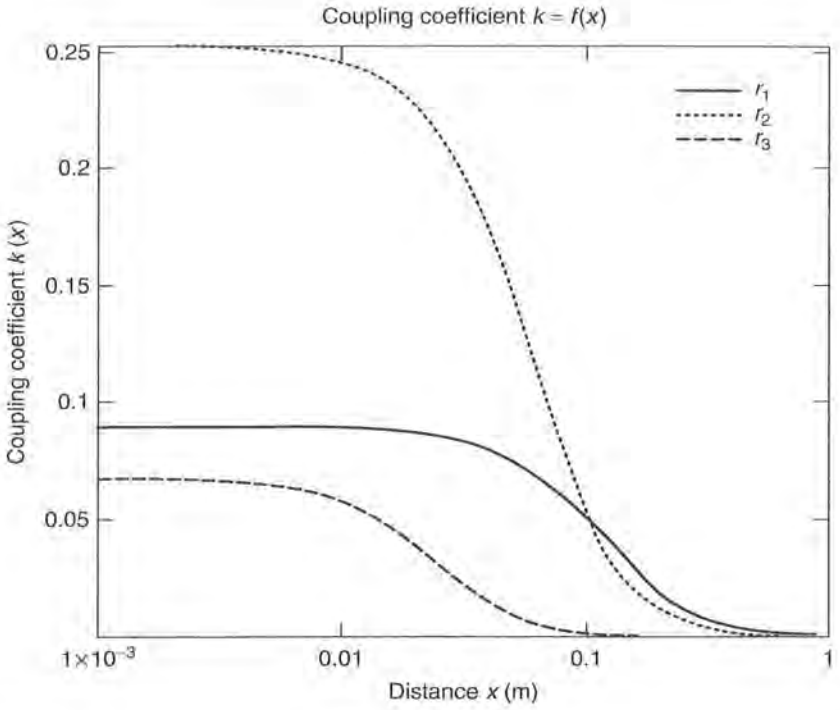


Figure 4.10 Graph of the coupling coefficient for different sized conductor loops. Transponder antenna: $r_{\text{Transp}} = 2\,\text{cm}$, reader antenna: $r_1 = 10\,\text{cm}$, $r_2 = 7.5\,\text{cm}$, $r_3 = 1\,\text{cm}$

because in this case the conductor loops are in the same place and are exposed to exactly the same magnetic flux ψ.

In practice, however, inductively coupled transponder systems operate with coupling coefficients that may be as low as 0.01 (<1%) (Figure 4.10).

4.1.6 Faraday's law

Any change to the magnetic flux Φ generates an electric field strength E_i. This characteristic of the magnetic field is described by *Faraday's law*.

The effect of the electric field generated in this manner depends upon the material properties of the surrounding area. Figure 4.11 shows some of the possible effects (Paul, 1993):

- Vacuum: in this case, the field strength E gives rise to an *electric rotational field*. Periodic changes in magnetic flux (high frequency current in an antenna coil) generate an electromagnetic field that propagates itself into the distance.
- Open conductor loop: an open circuit voltage builds up across the ends of an almost closed conductor loop, which is normally called *induced voltage*. This voltage corresponds with the line integral (path integral) of the field strength E that is generated along the path of the conductor loop in space.

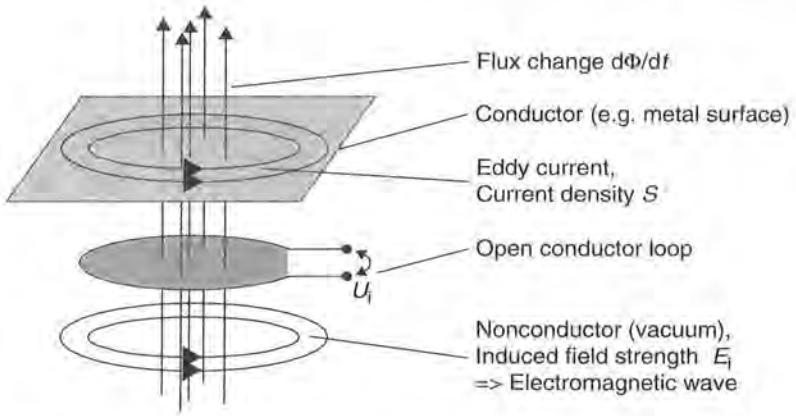


Figure 4.11 Induced electric field strength E in different materials. From top to bottom: metal surface, conductor loop and vacuum

- Metal surface: an electric field strength E is also induced in the metal surface. This causes free charge carriers to flow in the direction of the electric field strength. Currents flowing in circles are created, so-called *eddy currents*. This works against the exciting magnetic flux (Lenz's law), which may significantly damp the magnetic flux in the vicinity of *metal surfaces*. However, this effect is undesirable in inductively coupled RFID systems (installation of a transponder or reader antenna on a metal surface) and must therefore be prevented by suitable countermeasures (see Section 4.1.12.3).

In its general form Faraday's law is written as follows:

$$u_i = \oint E_i \cdot ds = -\frac{d\Psi(t)}{dt} \tag{4.21}$$

For a conductor loop configuration with N windings, we can also say that $u_i = N \cdot d\Psi/dt$. (The value of the contour integral $\int E_i \cdot ds$ can be increased N times if the closed integration path is carried out N times; Paul, 1993).

To improve our understanding of inductively coupled RFID systems we will now consider the effect of inductance on magnetically coupled conduction loops.

A time variant current i_1 (t) in conduction loop L_1 generates a time variant magnetic flux $d\Phi(i_1)/dt$. In accordance with the inductance law, a voltage is induced in the conductor loops L_1 and L_2 through which some degree of magnetic flux is flowing. We can differentiate between two cases:

- *Self-inductance*: the flux change generated by the current change di_n/dt induces a voltage u_n in the same conductor circuit.
- *Mutual inductance*: the flux change generated by the current change di_n/dt induces a voltage in the adjacent conductor circuit L_m. Both circuits are coupled by mutual inductance.

Figure 4.12 shows the equivalent circuit diagram for coupled conductor loops. In an inductively coupled RFID system L_1 would be the transmitter antenna of the reader.

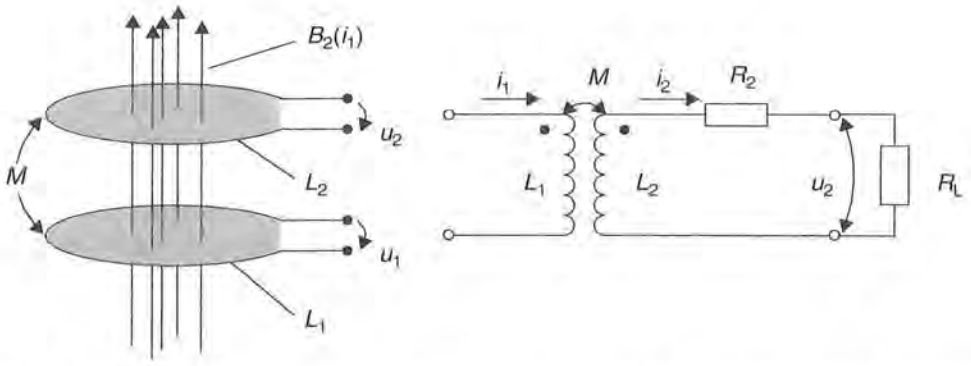


Figure 4.12 Left, magnetically coupled conductor loops; right, equivalent circuit diagram for magnetically coupled conductor loops

L_2 represents the antenna of the transponder, where R_2 is the *coil resistance* of the transponder antenna. The current consumption of the data memory is symbolised by the load resistor R_L.

A time varying flux in the conductor loop L_1 induces voltage u_{2i} in the conductor loop L_2 due to mutual inductance M. The flow of current creates an additional voltage drop across the coil resistance R_2, meaning that the voltage u_2 can be measured at the terminals. The current through the load resistor R_L is calculated from the expression u_2/R_L. The current through L_2 also generates an additional magnetic flux, which opposes the magnetic flux $\Psi_1(i_1)$. The above is summed up in the following equation:

$$u_2 = +\frac{d\Psi_2}{dt} = M\frac{di_1}{dt} - L_2\frac{di_2}{dt} - i_2R_2 \tag{4.22}$$

Because, in practice, i_1 and i_2 are sinusoidal (HF) alternating currents, we write equation (4.22) in the more appropriate complex notation (where $\omega = 2\pi f$):

$$u_2 = j\omega M \cdot i_1 - j\omega L_2 \cdot i_2 - i_2R_2 \tag{4.23}$$

If i_2 is replaced by u_2/R_L in equation (4.23), then we can solve the equation for u_2:

$$u_2 = \frac{j\omega M \cdot i_1}{1 + \dfrac{j\omega L_2 + R_2}{R_L}} \quad \begin{cases} \rightarrow R_L \rightarrow \infty\text{: } u_2 = j\omega M \cdot i_1 \\ \rightarrow R_L \rightarrow 0\text{: } u_2 \rightarrow 0 \end{cases} \tag{4.24}$$

4.1.7 Resonance

The voltage u_2 induced in the transponder coil is used to provide the power supply to the data memory (*microchip*) of a passive transponder (see Section 4.1.8.1). In order to significantly improve the efficiency of the equivalent circuit illustrated in Figure 4.12, an additional capacitor C_2 is connected in parallel with the transponder coil L_2 to form a *parallel resonant circuit* with a *resonant frequency* that corresponds with the

operating frequency of the RFID system in question.[1] The resonant frequency of the parallel resonant circuit can be calculated using the Thomson equation:

$$f = \frac{1}{2\pi\sqrt{L_2 \cdot C_2}} \tag{4.25}$$

In practice, C_2 is made up of a parallel capacitor C_2' and a parasitic capacitance C_p from the real circuit. $C_2 = (C_2' + C_p)$. The required capacitance for the parallel capacitor C_2' is found using the Thomson equation, taking into account the parasitic capacitance C_p:

$$C_2' = \frac{1}{(2\pi f)^2 L_2} - C_p \tag{4.26}$$

Figure 4.13 shows the equivalent circuit diagram of a real transponder. R_2 is the natural resistance of the transponder coil L_2 and the current consumption of the data carrier (chip) is represented by the load resistor R_L.

If a voltage $u_{Q2} = u_i$ is induced in the coil L_2, the following voltage u_2 can be measured at the data carrier load resistor R_L in the equivalent circuit diagram shown in Figure 4.13:

$$u_2 = \frac{u_{Q2}}{1 + (j\omega L_2 + R_2) \cdot \left(\frac{1}{R_L} + j\omega C_2\right)} \tag{4.27}$$

We now replace the induced voltage $u_{Q2} = u_i$ by the factor responsible for its generation, $u_{Q2} = u_i = j\omega M \cdot i_1 = \omega \cdot k \cdot \sqrt{L_1 \cdot L_2} \cdot i_1$, thus obtaining the relationship

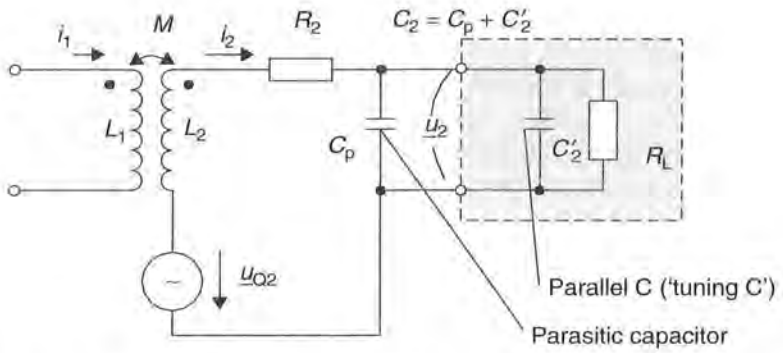


Figure 4.13 Equivalent circuit diagram for magnetically coupled conductor loops. Transponder coil L_2 and parallel capacitor C_2 form a parallel resonant circuit to improve the efficiency of voltage transfer. The transponder's data carrier is represented by the grey box

[1] However, in 13.56 MHz systems with anticollision procedures, the resonant frequency selected for the transponder is often 1–5 MHz higher to minimise the effect of the interaction between transponders on overall performance. This is because the overall resonant frequency of two transponders directly adjacent to one another is always lower than the resonant frequency of a single transponder.

between voltage u_2 and the magnetic coupling of transmitter coil and transponder coil:

$$u_2 = \frac{j\omega M \cdot i_1}{1 + (j\omega L_2 + R_2) \cdot \left(\frac{1}{R_L} + j\omega C_2\right)} \tag{4.28}$$

and:

$$u_2 = \frac{j\omega \cdot k \cdot \sqrt{L_1 \cdot L_2} \cdot i_1}{1 + (j\omega L_2 + R_2) \cdot \left(\frac{1}{R_L} + j\omega C_2\right)} \tag{4.29}$$

or in the non-complex form (Jurisch, 1994):

$$u_2 = \frac{\omega \cdot k \cdot \sqrt{L_1 L_2} \cdot i_1}{\sqrt{\left(\frac{\omega L_2}{R_L} + \omega R_2 C_2\right)^2 + \left(1 - \omega^2 L_2 C_2 + \frac{R_2}{R_L}\right)^2}} \tag{4.30}$$

where $C_2 = C_2' + C_p$.

Figure 4.14 shows the simulated graph of u_2 with and without resonance over a large frequency range for a possible transponder system. The current i_1 in the transmitter antenna (and thus also $\Phi(i_1)$), inductance L_2, mutual inductance M, R_2 and R_L are held constant over the entire frequency range.

We see that the graph of voltage u_2 for the circuit with the coil alone (circuit from Figure 4.12) is almost identical to that of the *parallel resonant circuit* (circuit from

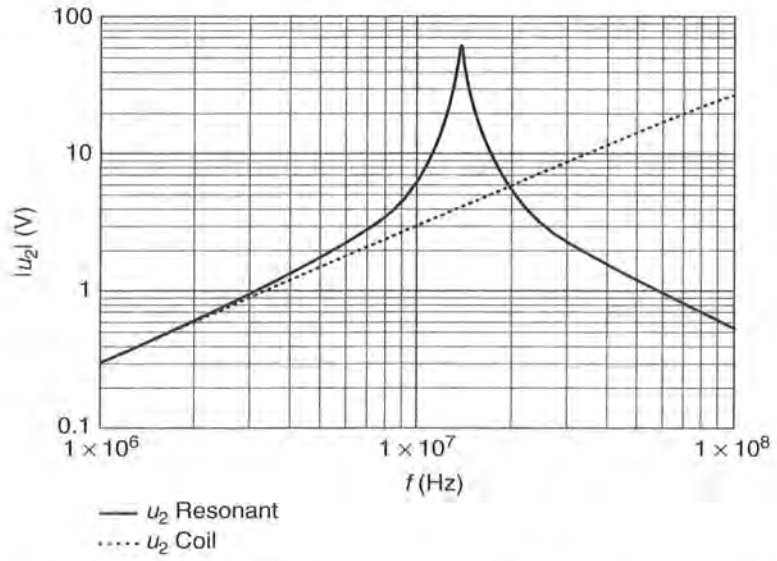


Figure 4.14 Plot of voltage at a transponder coil in the frequency range 1 to 100 MHz, given a constant magnetic field strength H or constant current i_1. A transponder coil with a parallel capacitor shows a clear voltage step-up when excited at its resonant frequency ($f_{RES} = 13.56$ MHz)

Figure 4.13) at frequencies well below the resonant frequencies of both circuits, but that when the resonant frequency is reached, voltage u_2 increases by more than a power of ten in the parallel resonant circuit compared to the voltage u_2 for the coil alone. Above the resonant frequency, however, voltage u_2 falls rapidly in the parallel resonant circuit, even falling below the value for the coil alone.

For transponders in the frequency range below 135 kHz, the transponder coil L_2 is generally connected in parallel with a chip capacitor ($C_2' = 20–220$ pF) to achieve the desired resonant frequency. At the higher frequencies of 13.56 MHz and 27.125 MHz, the required capacitance C_2 is usually so low that it is provided by the input capacitance of the data carrier together with the parasitic capacitance of the transponder coil.

Let us now investigate the influence of the circuit elements R_2, R_L and L_2 on voltage u_2. To gain a better understanding of the interactions between the individual parameters we will now introduce the Q factor (the Q factor crops up again when we investigate the connection of transmitter antennas in Section 11.4.1.3). We will refrain from deriving formulas because the electric resonant circuit is dealt with in detail in the background reading.

The *Q factor* is a measure of the voltage and current step-up in the resonant circuit at its resonant frequency. Its reciprocal 1/Q denotes the expressively named *circuit damping d*. The Q factor is very simple to calculate for the equivalent circuit in Figure 4.13. In this case ω is the angular frequency ($\omega = 2\pi f$) of the transponder resonant circuit:

$$Q = \frac{1}{R_2 \cdot \sqrt{\frac{C_2}{L_2}} + \frac{1}{R_L} \cdot \sqrt{\frac{L_2}{C_2}}} = \frac{1}{\frac{R_2}{\omega L_2} + \frac{\omega L_2}{R_L}} \tag{4.31}$$

A glance at equation (4.31) shows that when $R_2 \to \infty$ and $R_L \to 0$, the Q factor also tends towards zero. On the other hand, when the transponder coil has a very low coil resistance $R_2 \to 0$ and there is a high load resistor $R_L \gg 0$ (corresponding with very low transponder chip power consumption), very high Q factors can be achieved. The voltage u_2 is now proportional to the quality of the resonant circuit, which means that the dependency of voltage u_2 upon R_2 and R_L is clearly defined.

Voltage u_2 thus tends towards zero where $R_2 \to \infty$ and $R_L \to 0$. At a very low transponder coil resistance $R_2 \to 0$ and a high value load resistor $R_L \gg 0$, on the other hand, a very high voltage u_2 can be achieved (compare equation (4.30)).

It is interesting to note the path taken by the graph of voltage u_2 when the inductance of the transponder coil L_2 is changed, thus maintaining the resonance condition (i.e. $C_2 = 1/\omega^2 L_2$ for all values of L_2). We see that for certain values of L_2, voltage u_2 reaches a clear peak (Figure 4.15).

If we now consider the graph of the Q factor as a function of L_2 (Figure 4.16), then we observe a maximum at the same value of transponder inductance L_2. The maximum voltage $u_2 = f(L_2)$ is therefore derived from the maximum Q factor, $Q = f(L_2)$, at this point.

This indicates that for every pair of parameters (R_2, R_L), there is an inductance value L_2 at which the Q factor, and thus also the supply voltage u_2 to the data carrier, is at a maximum. This should always be taken into consideration when designing a transponder, because this effect can be exploited to optimise the energy range of

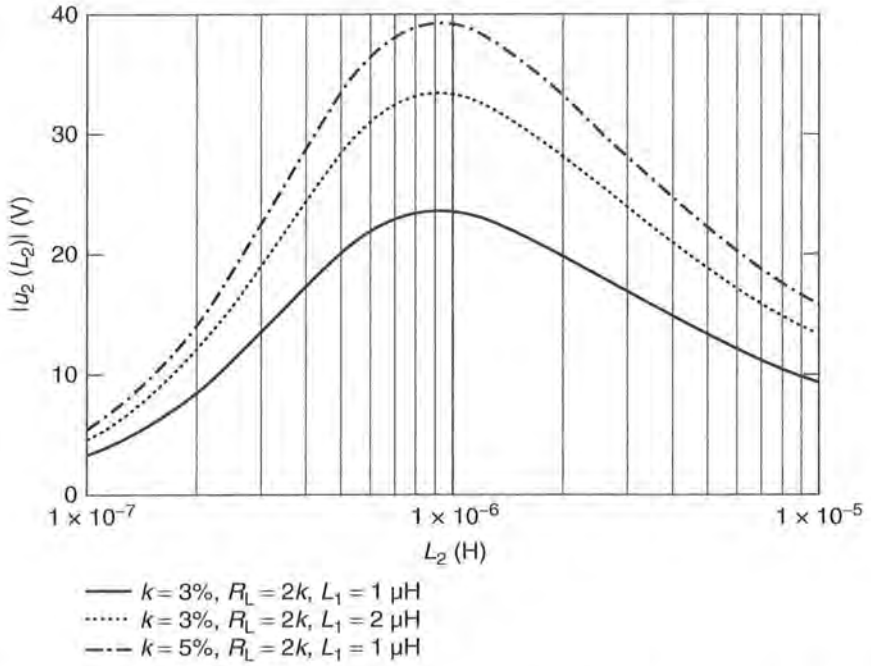


Figure 4.15 Plot of voltage u_2 for different values of transponder inductance L_2. The resonant frequency of the transponder is equal to the transmission frequency of the reader for all values of L_2 ($i_1 = 0.5\,A$, $f = 13.56\,MHz$, $R_2 = 1\,\Omega$)

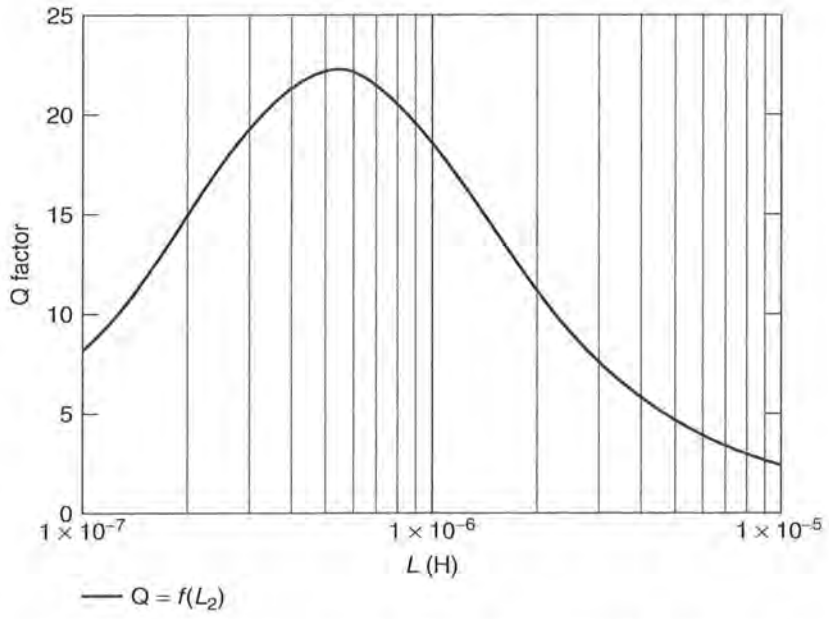


Figure 4.16 Graph of the Q factor as a function of transponder inductance L_2, where the resonant frequency of the transponder is constant ($f = 13.56\,MHz$, $R_2 = 1\,\Omega$)

an inductively coupled RFID system. However, we must also bear in mind that the influence of component tolerances in the system also reaches a maximum in the Q_{max} range. This is particularly important in systems designed for mass production. Such systems should be designed so that reliable operation is still possible in the range $Q \ll Q_{max}$ at the maximum distance between transponder and reader.

R_L should be set at the same value as the input resistance of the data carrier after setting the 'power on' reset, i.e. before the activation of the voltage regulator, as is the case for the maximum energy range of the system.

4.1.8 Practical operation of the transponder

4.1.8.1 Power supply to the transponder

Transponders are classified as active or passive depending upon the type of power supply they use.

Active transponders incorporate their own battery to provide the power supply to the data carrier. In these transponders, the voltage u_2 is generally only required to generate a 'wake up' signal. As soon as the voltage u_2 exceeds a certain limit this signal is activated and puts the data carrier into operating mode. The transponder returns to the power saving 'sleep' or 'stand-by mode' after the completion of a transaction with the reader, or when the voltage u_2 falls below a minimum value.

In *passive transponders* the data carrier has to obtain its power supply from the voltage u_2. To achieve this, the voltage u_2 is converted into direct current using a low loss bridge rectifier and then smoothed. A simple basic circuit for this application is shown in Figure 3.18.

4.1.8.2 Voltage regulation

The induced voltage u_2 in the transponder coil very rapidly reaches high values due to resonance step-up in the resonant circuit. Considering the example in Figure 4.14, if we increase the coupling coefficient k — possibly by reducing the gap between reader and transponder — or the value of the load resistor R_L, then voltage u_2 will reach a level much greater than 100 V. However, the operation of a data carrier requires a constant *operating voltage* of 3–5 V (after rectification).

In order to regulate voltage u_2 independently of the coupling coefficient k or other parameters, and to hold it constant in practice, a voltage-dependent shunt resistor R_S is connected in parallel with the load resistor R_L. The equivalent circuit diagram for this is shown in Figure 4.17.

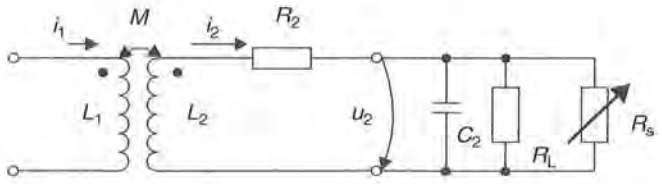


Figure 4.17 Operating principle for voltage regulation in the transponder using a shunt regulator

As induced voltage $u_{Q2} = u_i$ increases, the value of the shunt resistor R_S falls, thus reducing the quality of the transponder resonant circuit to such a degree that the voltage u_2 remains constant. To calculate the value of the shunt resistor for different variables, we refer back to equation (4.29) and introduce the parallel connection of R_L and R_S in place of the constant load resistor R_L. The equation can now be solved with respect to R_S. The variable voltage u_2 is replaced by the constant voltage u_{Transp} — the desired input voltage of the data carrier — giving the following equation for R_S:

$$R_S = \left| \frac{1}{\frac{\left(\frac{j\omega \cdot k \cdot \sqrt{L_1 L_2} \cdot i_1}{u_{Transp}}\right) - 1}{j\omega L_2 + R_2} - j\omega C_2 - \frac{1}{R_L}} \right| \quad | u_{2\text{-unreg}} > u_{Transp} \qquad (4.32)$$

Figure 4.18 shows the graph of voltage u_2 when such an 'ideal' shunt regulator is used. Voltage u_2 initially increases in proportion with the coupling coefficient k. When u_2 reaches its desired value, the value of the shunt resistor begins to fall in inverse proportion to k, thus maintaining an almost constant value for voltage u_2.

Figure 4.19 shows the variable value of the shunt resistor R_S as a function of the coupling coefficient. In this example the value range for the shunt resistor covers several powers of ten. This can only be achieved using a semiconductor circuit, therefore so-called *shunt* or *parallel regulators* are used in inductively coupled transponders. These terms describe an electronic regulator circuit, the internal resistance of which

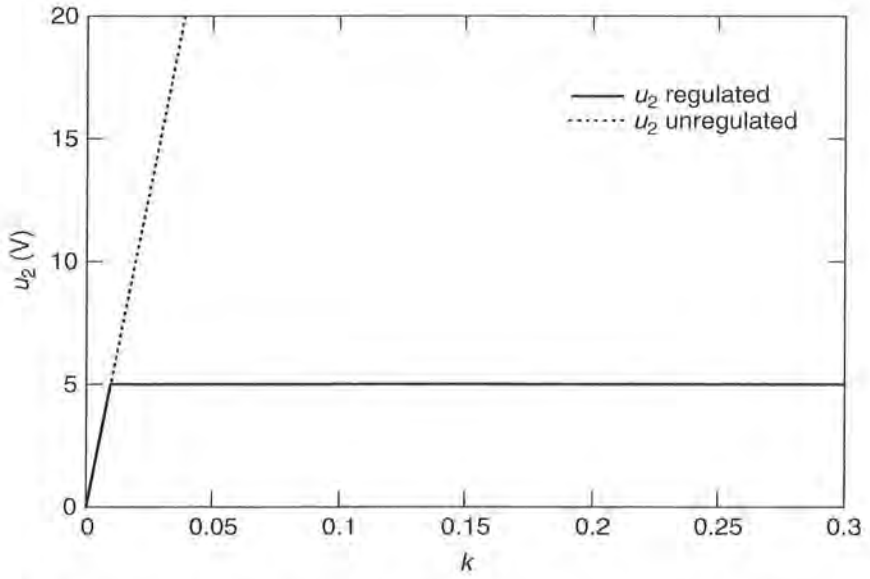


Figure 4.18 Example of the path of voltage u_2 with and without shunt regulation in the transponder, where the coupling coefficient k is varied by altering the distance between transponder and reader antenna. (The calculation is based upon the following parameters: $i_1 = 0.5$ A, $L_1 = 1\,\mu$H, $L_2 = 3.5\,\mu$H, $R_L = 2\,k\Omega$, $C_2 = 1/\omega_2 L_2$)

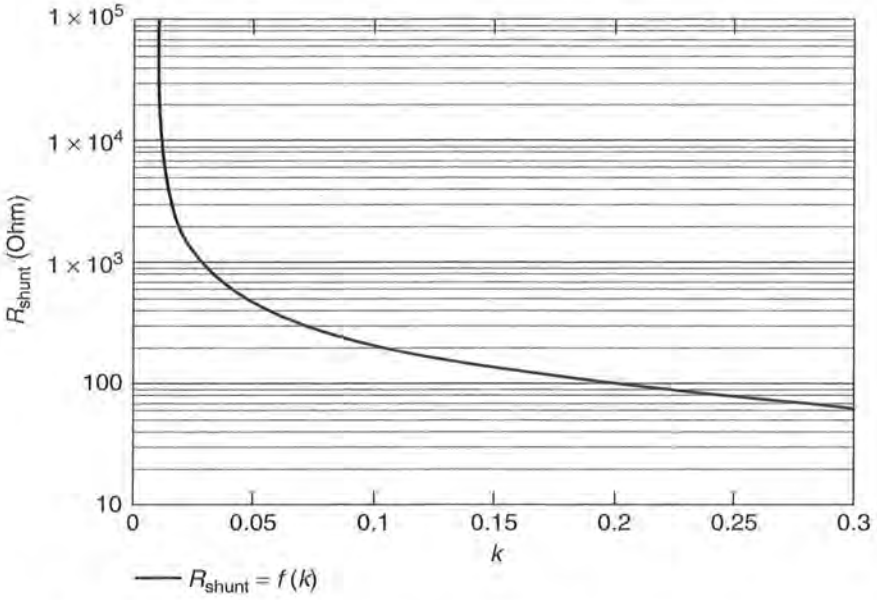


Figure 4.19 The value of the shunt resistor R_S must be adjustable over a wide range to keep voltage u_2 constant regardless of the coupling coefficient k (parameters as Figure 4.18)

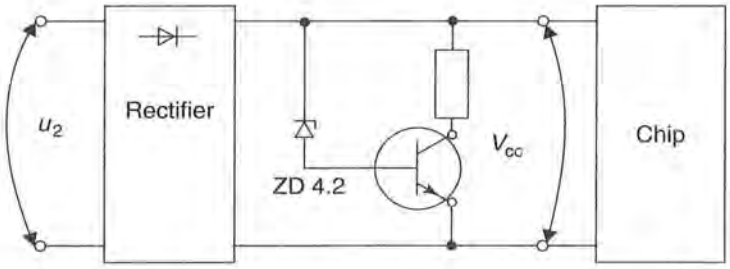


Figure 4.20 Example circuit for a simple shunt regulator

falls disproportionately sharply when a threshold voltage is exceeded. A simple shunt regulator based upon a zener diode (Nührmann, 1994) is shown in Figure 4.20.

4.1.9 Interrogation field strength H_{min}

We can now use the results obtained in Section 4.1.7 to calculate the *interrogation field strength* of a transponder. This is the minimum field strength H_{min} (at a maximum distance x between transponder and reader) at which the supply voltage u_2 is just high enough for the operation of the data carrier.

However, u_2 is not the internal operating voltage of the data carrier (3V or 5V) here; it is the *HF input voltage* at the terminal of the transponder coil L_2 on the data carrier, i.e. prior to rectification. The voltage regulator (shunt regulator) should not yet be active at this supply voltage. R_L corresponds with the input resistance of the data carrier after the 'power on reset', C_2 is made up of the input capacitance

C_p of the data carrier (chip) and the parasitic capacitance of the transponder layout $C_2' : C_2 = (C_2' + C_p)$.

The inductive voltage (source voltage $u_{Q2} = u_i$) of a transponder coil can be calculated using equation (4.21) for the general case. If we assume a homogeneous, sinusoidal magnetic field in air (permeability constant $= \mu_0$) we can derive the following, more appropriate, formula:

$$u_i = \mu_0 \cdot A \cdot N \cdot \omega \cdot H_{eff} \tag{4.33}$$

where H_{eff} is the effective field strength of a sinusoidal magnetic field, ω is the angular frequency of the magnetic field, N is the number of windings of the transponder coil L_2, and A is the cross-sectional area of the transponder coil.

We now replace $u_{Q2} = u_i = j\omega M \cdot i_1$ from equation (4.29) with equation (4.33) and thus obtain the following equation for the circuit in Figure 4.13:

$$u_2 = \frac{j\omega \cdot \mu_0 \cdot H_{eff} \cdot A \cdot N}{1 + (j\omega L_2 + R_2)\left(\frac{1}{R_L} + j\omega C_2\right)} \tag{4.34}$$

Multiplying out the denominator:

$$u_2 = \frac{j\omega \cdot \mu_0 \cdot H_{eff} \cdot A \cdot N}{j\omega\left(\frac{L_2}{R_L} + R_2 C_2\right) + \left(1 - \omega^2 L_2 C_2 + \frac{R_2}{R_L}\right)} \tag{4.35}$$

We now solve this equation for H_{eff} and obtain the value of the complex form. This yields the following relationship for the interrogation field H_{min} in the general case:

$$H_{min} = \frac{u_2 \cdot \sqrt{\left(\frac{\omega L_2}{R_L} + \omega R_2 C_2\right)^2 + \left(1 - \omega^2 L_2 C_2 + \frac{R_2}{R_L}\right)^2}}{\omega \cdot \mu_0 \cdot A \cdot N} \tag{4.36}$$

A more detailed analysis of equation (4.36) shows that the interrogation field strength is dependent upon the frequency $\omega = 2\pi f$ in addition to the antenna area A, the number of windings N (of the transponder coil), the minimum voltage u_2 and the input resistance R_2. This is not surprising, because we have determined a resonance step-up of u_2 at the resonant frequency of the transponder resonant circuit. Therefore, when the transmission frequency of the reader corresponds with the resonant frequency of the transponder, the interrogation field strength H_{min} is at its minimum value.

To optimise the interrogation sensitivity of an inductively coupled RFID system, the resonant frequency of the transponder should be matched precisely to the transmission frequency of the reader. Unfortunately, this is not always possible in practice. First, tolerances occur during the manufacture of a transponder, which lead to a deviation in the transponder resonant frequency. Second, there are also technical reasons for setting the resonant frequency of the transponder a few percentage points higher than the transmission frequency of the reader (for example in systems using anticollision procedures to keep the interaction of nearby transponders low).

Some semiconductor manufacturers incorporate additional smoothing capacitors into the transponder chip to smooth out frequency deviations in the transponder caused by manufacturing tolerances (see Figure 3.28, 'tuning C'). During manufacture the transponder is adjusted to the desired frequency by switching individual smoothing capacitors on and off (Schürmann, 1993).

In equation (4.36) the resonant frequency of the transponder is expressed as the product L_2C_2. This is not recognisable at first glance. In order to make a direct prediction regarding the frequency dependency of interrogation sensitivity, we rearrange equation (4.25) to obtain:

$$L_2C_2 = \frac{1}{(2\pi\ f_0)^2} = \frac{1}{\omega_0^2} \tag{4.37}$$

By substituting this expression into the right-hand term under the root of equation (4.36) we obtain a function in which the dependence of the interrogation field strength $H_{\min}$ on the relationship between the transmission frequency of the reader (ω) and the resonant frequency of the transponder (ω_0) is clearly expressed. This is based upon the assumption that the change in the resonant frequency of the transponder is caused by a change in the capacitance of C_2 (e.g. due to temperature dependence or manufacturing tolerances of this capacitance), whereas the inductance L_2 of the coil remains constant. To express this, the capacitor C_2 in the left-hand term under the root of equation (4.36) is replaced by $C_2 = (\omega_0^2 \cdot L_2)^{-1}$:

$$H_{\min} = \frac{u_2 \cdot \sqrt{\omega^2 \left(\frac{L_2}{R_L} + \frac{R_2}{\omega_0^2 L_2}\right)^2 + \left(\frac{\omega_0^2 - \omega^2}{\omega_0^2} + \frac{R_2}{R_L}\right)^2}}{\omega\mu_0 \cdot A \cdot N} \tag{4.38}$$

Therefore a deviation of the transponder resonant frequency from the transmission frequency of the reader will lead to a higher transponder interrogation field strength and thus to a lower *read range* (Figure 4.21).

4.1.9.1 Energy range of transponder systems

If the interrogation field strength of a transponder is known, then we can also assess the energy range associated with a certain reader. The *energy range* of a transponder is the distance from the reader antenna at which there is just enough energy to operate the transponder (defined by $u_{2\,\min}$ and R_L). However, the question of whether the energy range obtained corresponds with the maximum functional range of the system also depends upon whether the data transmitted from the transponder can be detected by the reader at the distance in question.

Given a known antenna current[2] I, radius R, and number of windings of the transmitter antenna N_1, the path of the field strength in the x direction can be calculated using equation (4.3) (see Section 4.1.1.1). If we solve the equation with respect to x

[2] If the antenna current of the transmitter antenna is not known it can be calculated from the measured field strength $H(x)$ at a distance x, where the antenna radius R and the number of windings N_1 are known (see Section 4.1.1.1).

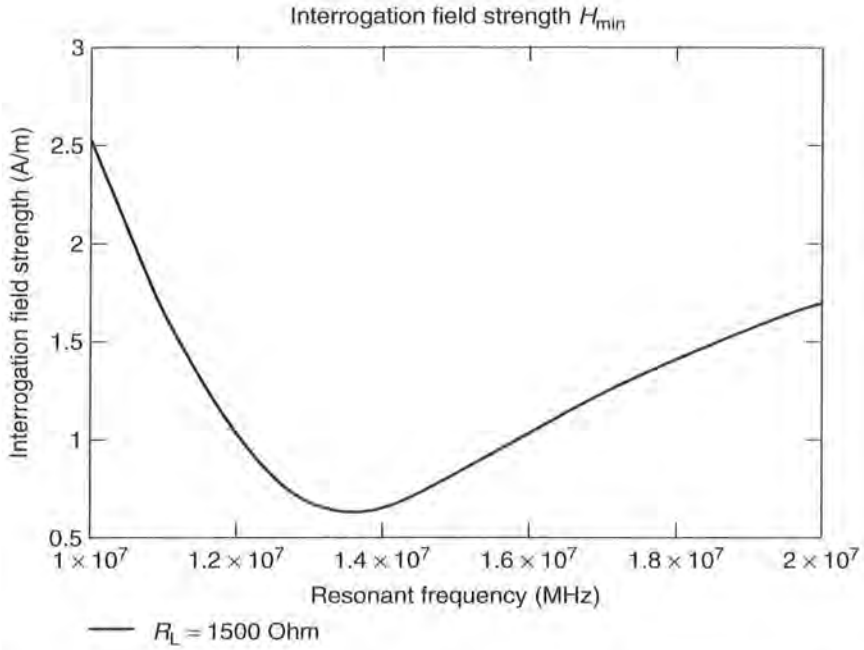


Figure 4.21 Interrogation sensitivity of a contactless smart card where the transponder resonant frequency is detuned in the range 10–20 MHz ($N = 4$, $A = 0.05 \times 0.08\,\text{m}^2$, $u_2 = 5\text{V}$, $L_2 = 3.5\,\mu\text{H}$, $R_2 = 5\,\Omega$, $R_L = 1.5\,\text{k}\Omega$). If the transponder resonant frequency deviates from the transmission frequency (13.56 MHz) of the reader an increasingly high field strength is required to address the transponder. In practical operation this results in a reduction of the read range

we obtain the following relationship between the energy range and interrogation field H_{min} of a transponder for a given reader:

$$x = \sqrt{\sqrt[3]{\left(\frac{I \cdot N_1 \cdot R^2}{2 \cdot H_{min}}\right)^2} - R^2} \tag{4.39}$$

As an example (see Figure 4.22), let us now consider the energy range of a transponder as a function of the power consumption of the data carrier ($R_L = u_2 / i_2$). The reader in this example generates a field strength of 0.115 A/m at a distance of 80 cm from the transmitter antenna (radius R of transmitter antenna: 40 cm). This is a typical value for RFID systems in accordance with ISO 15693.

As the current consumption of the transponder (lower R_L) increases, the interrogation sensitivity of the transponder also increases and the energy range falls.

The maximum energy range of the transponder is determined by the distance between transponder and reader antenna at which the minimum power supply $u_{2\,min}$ required for the operation of the data carrier exists even with an unloaded transponder resonant circuit (i.e. $i_2 \to 0$, $R_L \to \infty$). Where distance $x = 0$ the maximum current i_2 represents a limit, above which the supply voltage for the data carrier falls below $u_{2\,min}$, which means that the reliable operation of the data carrier can no longer be guaranteed in this operating state.

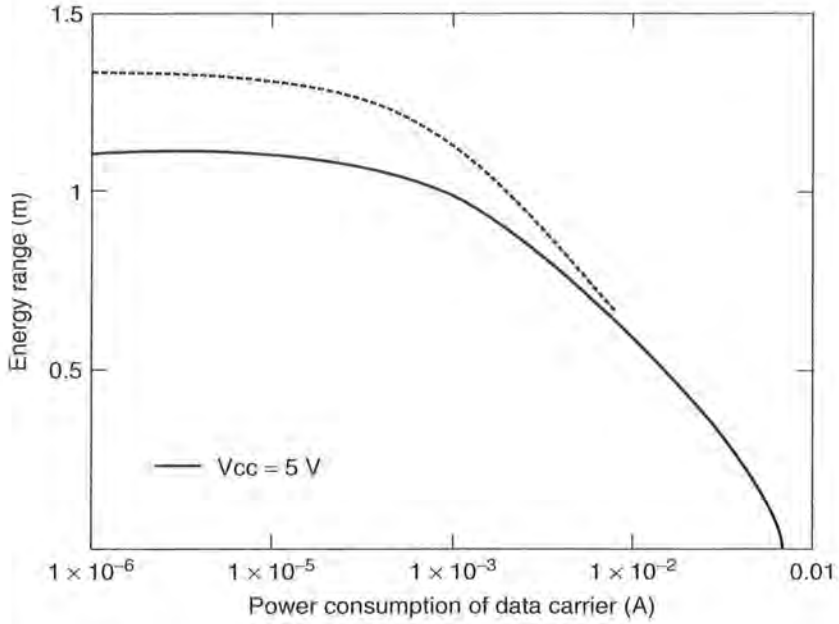


Figure 4.22 The energy range of a transponder also depends upon the power consumption of the data carrier (R_L). The transmitter antenna of the simulated system generates a field strength of 0.115 A/m at a distance of 80 cm, a value typical for RFID systems in accordance with ISO 15693 (transmitter: $I = 1$A, $N_1 = 1$, $R = 0.4$ m. Transponder: $A = 0.048 \times 0.076\,\text{m}^2$ (smart card), $N = 4$, $L_2 = 3.6\,\mu\text{H}$, $u_{2\,\text{min}} = 5\text{V}/3\text{V}$)

4.1.9.2 Interrogation zone of readers

In the calculations above the implicit assumption was made of a homogeneous magnetic field H parallel to the coil axis x. A glance at Figure 4.23 shows that this only applies for an arrangement of reader coil and transponder coil with a common central axis x. If the transponder is tilted away from this central axis or displaced in the direction of the y or z axis this condition is no longer fulfilled.

If a coil is magnetised by a magnetic field H, which is tilted by the angle ϑ in relation to the central axis of the coil, then in very general terms the following applies:

$$u_{0\vartheta} = u_0 \cdot \cos(\vartheta) \tag{4.40}$$

where u_0 is the voltage that is induced when the coil is perpendicular to the magnetic field. At an angle $\vartheta = 90°$, in which case the field lines run in the plane of the coil radius R, no voltage is induced in the coil.

As a result of the bending of the *magnetic field lines* in the entire area around the reader coil, here too there are different angles ϑ of the magnetic field H in relation to the transponder coil. This leads to a characteristic *interrogation zone* (Figure 4.24, grey area) around the reader antenna. Areas with an angle $\vartheta = 0°$ in relation to the transponder antenna — for example along the coil axis x, but also to the side of the antenna windings (returning field lines) — give rise to an optimal read range. Areas in which the magnetic field lines run parallel to the plane of the transponder coil radius

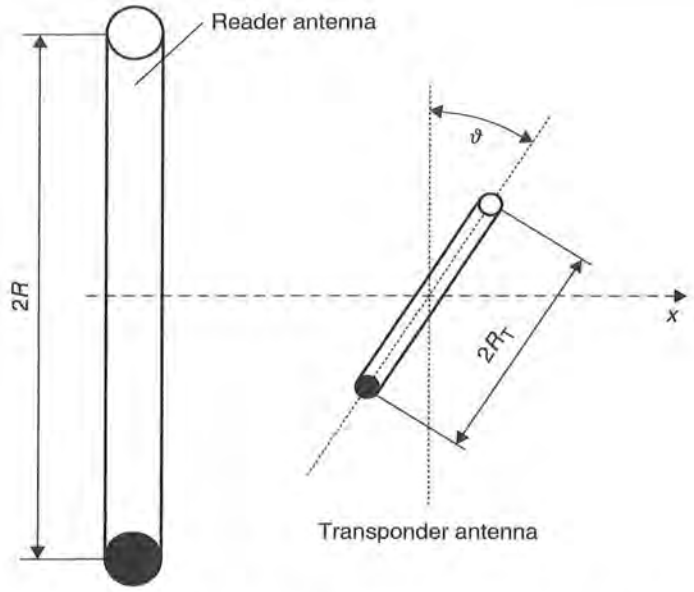


Figure 4.23 Cross-section through reader and transponder antennas. The transponder antenna is tilted at an angle ϑ in relation to the reader antenna

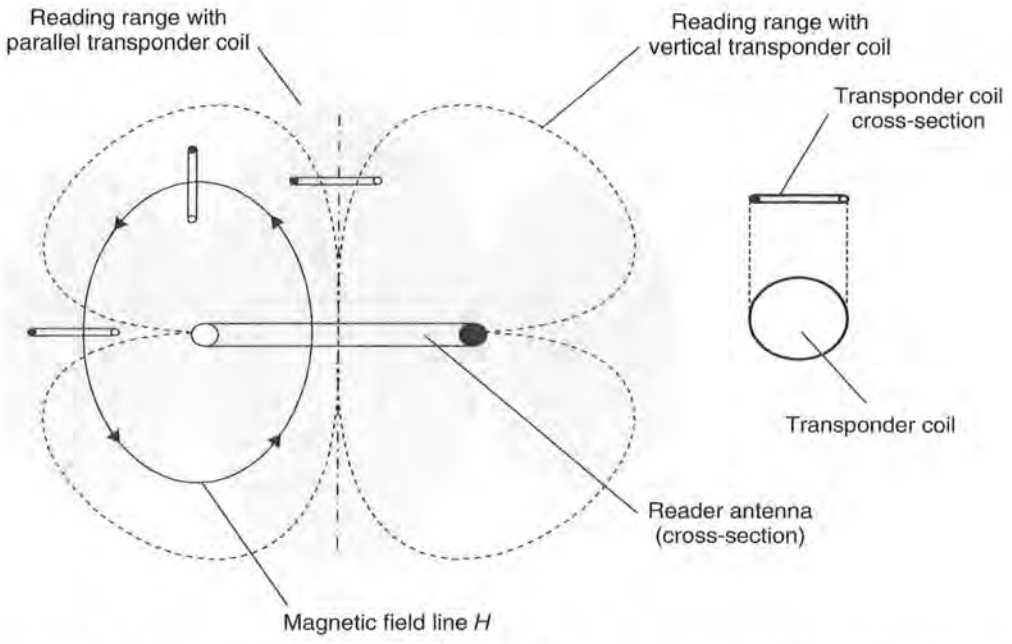


Figure 4.24 Interrogation zone of a reader at different alignments of the transponder coil

R — for example, exactly above and below the coil windings — exhibit a significantly reduced read range. If the transponder itself is tilted through 90° a completely different picture of the interrogation zone emerges (Figure 4.24, dotted line). Field lines that run parallel to the R-plane of the reader coil now penetrate the transponder coil at an angle $\vartheta = 0°$ and thus lead to an optimal *range* in this area.

4.1.10 Total transponder – reader system

Up to this point we have considered the characteristics of inductively coupled systems primarily from the point of view of the transponder. In order to analyse in more detail the interaction between transponder and *reader* in the system, we need to take a slightly different view and first examine the electrical properties of the reader so that we can then go on to study the system as a whole.

Figure 4.25 shows the equivalent circuit diagram for a reader (the practical realisation of this circuit configuration can be found in Section 11.4). The *conductor loop* necessary to generate the magnetic alternating field is represented by the coil L_1. The series resistor R_1 corresponds with the ohmic losses of the wire resistance in the conductor loop L_1. In order to obtain maximum current in the conductor coil L_1 at the reader *operating frequency* f_{TX}, a *series resonant circuit* with the resonant frequency $f_{RES} = f_{TX}$ is created by the serial connection of the capacitor C_1. The resonant frequency of the series resonant circuit can be calculated very easily using the Thomson equation (4.25). The operating state of the reader can be described by:

$$f_{TX} = f_{RES} = \frac{1}{2\pi\sqrt{L_1 \cdot C_1}} \tag{4.41}$$

Because of the series configuration, the total impedance Z_1 of the series resonant circuit is the sum of individual impedances, i.e.:

$$Z_1 = R_1 + j\omega L_1 + \frac{1}{j\omega C_1} \tag{4.42}$$

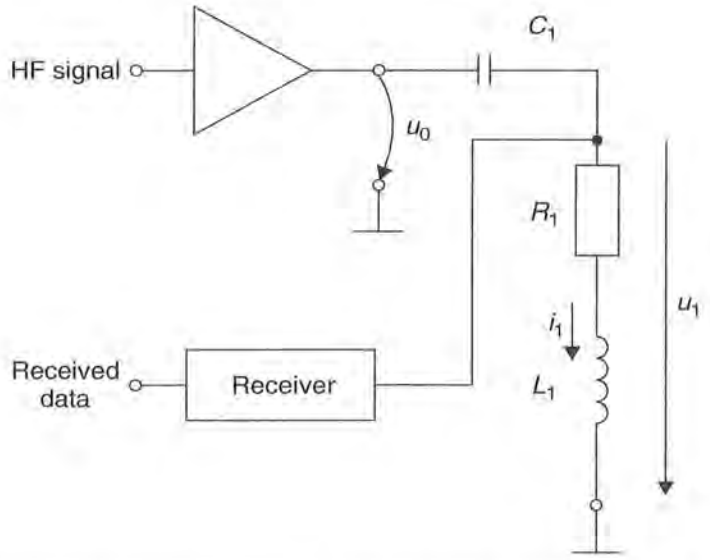


Figure 4.25 Equivalent circuit diagram of a reader with antenna L_1. The transmitter output branch of the reader generates the HF voltage u_0. The receiver of the reader is directly connected to the antenna coil L_1

At the resonant frequency f_{RES}, however, the impedances of L_1 and C_1 cancel each other out. In this case the total impedance Z_1 is determined by R_1 only and thus reaches a minimum.

$$j\omega L_1 + \frac{1}{j\omega C_1} = 0 \Big|_{\omega=2\pi \cdot f_{RES}} \Rightarrow Z_1(f_{RES}) = R_1 \tag{4.43}$$

The *antenna current* i_1 reaches a maximum at the resonant frequency and is calculated (based upon the assumption of an ideal voltage source where $R_i = 0$) from the source voltage u_0 of the transmitter high level stage, and the ohmic coil resistance R_1.

$$i_1(f_{res}) = \frac{u_0}{Z_1(f_{RES})} = \frac{u_0}{R_1} \tag{4.44}$$

The two voltages, u_1 at the conductor loop L_1, and u_{C1} at the capacitor C_1, are in antiphase and cancel each other out at the resonant frequency because current i_1 is the same. However, the individual values may be very high. Despite the low source voltage u_0, which is usually just a few volts, figures of a few hundred volts can easily be reached at L_1 and C_1. Designs for *conductor loop antennas* for high currents must therefore incorporate sufficient voltage resistance in the components used, in particular the capacitors, because otherwise these would easily be destroyed by arcing. Figure 4.26 shows an example of voltage step-up at resonance.

Despite the fact that the voltage may reach very high levels, it is completely safe to touch the voltage-carrying components of the reader antenna. Because of the additional

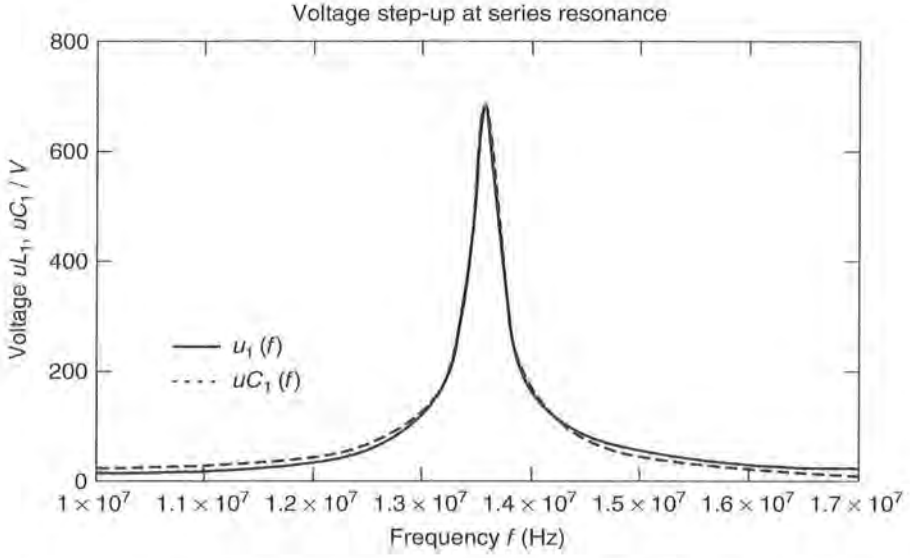


Figure 4.26 Voltage step-up at the coil and capacitor in a series resonant circuit in the frequency range 10–17 MHz ($f_{RES} = 13.56$ MHz, $u_0 = 10$V(!), $R_1 = 2.5\,\Omega$, $L_1 = 2\,\mu$H, $C_1 = 68.8$ pF). The voltage at the conductor coil and series capacitor reaches a maximum of above 700 V at the resonant frequency. Because the resonant frequency of the reader antenna of an inductively coupled system always corresponds with the transmission frequency of the reader, components should be sufficiently voltage resistant

capacitance of the hand, the series resonant circuit is rapidly detuned, thus reducing the resonance step-up of voltage.

4.1.10.1 Transformed transponder impedance Z'_T

If a transponder enters the magnetic alternating field of the conductor coil L_1 a change can be detected in the current i_1. The current i_2 induced in the transponder coil thus acts upon current i_1 responsible for its generation via the magnetic *mutual inductance* M.[3]

In order to simplify the mathematical description of the mutual inductance on the current i_1, let us now introduce an imaginary impedance, the *complex transformed transponder impedance* Z'_T. The electrical behaviour of the reader's series resonant circuit in the presence of mutual inductance is as if the imaginary impedance Z'_T were actually present as a discrete component: Z'_T takes on a finite value $|Z'_T| > 0$. If the mutual inductance is removed, e.g. by withdrawing the transponder from the field of the conductor loop, then $|Z'_T| = 0$. We will now derive the calculation of this transformed impedance step by step.

The source voltage u_0 of the reader can be divided into the individual voltages u_{C1}, u_{R1}, u_{L1} and u_{ZT} in the series resonant circuit, as illustrated in Figure 4.27. Figure 4.28 shows the vector diagram for the individual voltages in this circuit at resonance.

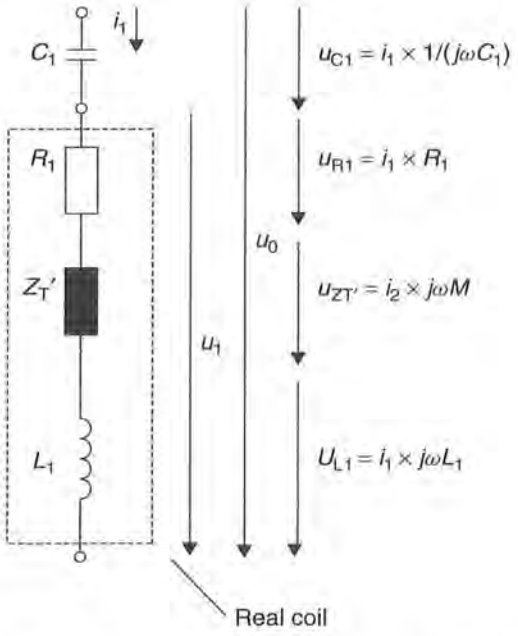


Figure 4.27 Equivalent circuit diagram of the series resonant circuit — the change in current i_1 in the conductor loop of the transmitter due to the influence of a magnetically coupled transponder is represented by the impedance Z'_T

[3] This is in accordance with Lenz's law, which states that 'the induced voltage always attempts to set up a current in the conductor circuit, the direction of which opposes that of the voltage that induced it' (Paul, 1993).

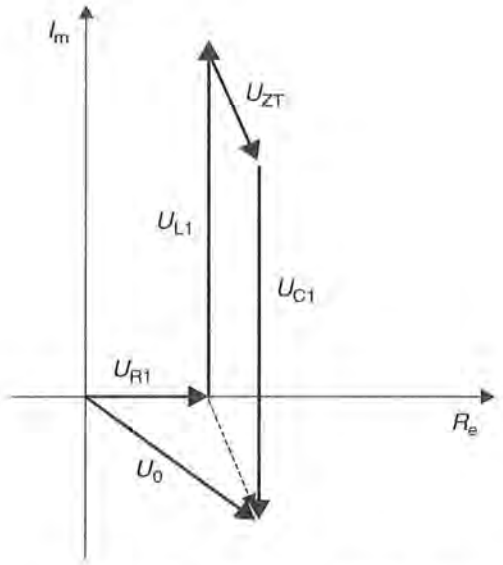


Figure 4.28 The vector diagram for voltages in the series resonance circuit of the reader antenna at resonant frequency. The figures for individual voltages u_{L1} and u_{C1} can reach much higher levels than the total voltage u_0

Due to the constant current i_1 in the series circuit, the source voltage u_0 can be represented as the sum of the products of the individual impedances and the current i_1. The transformed impedance Z'_T is expressed by the product $j\omega M \cdot i_2$:

$$u_0 = \frac{1}{j\omega C_1} \cdot i_1 + j\omega L_1 \cdot i_1 + R_1 \cdot i_1 - j\omega M \cdot i_2 \tag{4.45}$$

Since the series resonant circuit is operated at its *resonant frequency*, the individual impedances $(j\omega C_1)^{-1}$ and $j\omega L_1$ cancel each other out. The voltage u_0 is therefore only divided between the resistance R_1 and the transformed transponder impedance Z'_T, as we can see from the vector diagram (Figure 4.28). Equation 4.45 can therefore be further simplified to:

$$u_0 = R_1 \cdot i_1 - j\omega M \cdot i_2 \tag{4.46}$$

We now require an expression for the current i_2 in the coil of the transponder, so that we can calculate the value of the transformed transponder impedance. Figure 4.29 gives an overview of the currents and voltages in the transponder in the form of an equivalent circuit diagram:

The source voltage u_{Q2} is induced in the transponder coil L_2 by mutual inductance M. The current i_2 in the transponder is calculated from the quotient of the voltage u_2 divided by the sum of the individual impedances $j\omega L_2$, R_2 and Z_2 (here Z_2 represents the total input impedance of the data carrier and the parallel capacitor C_2). In the next step, we replace the voltage u_{Q2} by the voltage responsible for its generation $u_{Q2} = j\omega M \cdot i_1$, yielding the following expression for u_0:

$$u_0 = R_1 \cdot i_1 - j\omega M \cdot \frac{u_{Q2}}{R_2 + j\omega L_2 + Z_2} = R_1 \cdot i_1 - j\omega M \cdot \frac{j\omega M \cdot i_1}{R_2 + j\omega L_2 + Z_2} \tag{4.47}$$

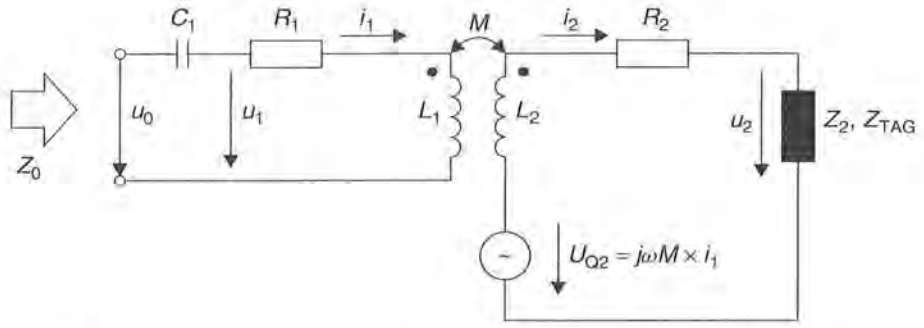


Figure 4.29 Simple equivalent circuit diagram of a transponder in the vicinity of a reader. The impedance Z_2 of the transponder is made up of the load resistor R_L (data carrier) and the capacitor C_2

As it is generally impractical to work with the mutual inductance M, in a final step we replace M with $M = k\sqrt{L_1 \cdot L_2}$ because the values k, L_1 and L_2 of a transponder are generally known. We write:

$$u_0 = R_1 \cdot i_1 + \frac{\omega^2 k^2 \cdot L_1 \cdot L_2}{R_2 + j\omega L_2 + Z_2} \cdot i_1 \tag{4.48}$$

Dividing both sides of equation (4.48) by i_1 yields the total impedance $Z_0 = u_0/i_1$ of the series resonant circuit in the reader as the sum of R_1 and the transformed transponder impedance Z'_T. Thus Z'_T is found to be:

$$Z'_T = \frac{\omega^2 k^2 \cdot L_1 \cdot L_2}{R_2 + j\omega L_2 + Z_2} \tag{4.49}$$

Impedance Z_2 represents the parallel connection of C_2 and R_L in the transponder. We replace Z_2 with the full expression containing C_2 and R_L and thus finally obtain an expression for Z'_T that incorporates all components of the transponder and is thus applicable in practice:

$$Z'_T = \frac{\omega^2 k^2 \cdot L_1 \cdot L_2}{R_2 + j\omega L_2 + \dfrac{R_L}{1 + j\omega R_L C_2}} \tag{4.50}$$

4.1.10.2 Influencing variables of Z'_T

Let us now investigate the influence of individual parameters on the *transformed transponder impedance* Z'_T. In addition to line diagrams, locus curves are also suitable for this investigation: there is precisely one vector in the complex Z plane for every parameter value x in the function $Z'_T = f(x)$ and thus exactly one point on the curve.

Table 4.3 Parameters for line diagrams and locus curves, if not stated otherwise

$L_1 = 1\,\mu H$	$L_2 = 3.5\,\mu H$
$C_1 = 1/(\omega_{TX})^2 \cdot L_1$ (resonance)	$R_2 = 5\,\Omega$
$C_2 = 1/(\omega_{RX})^2 \cdot L_2$ (resonance)	$R_L = 5\,k\Omega$
$f_{RES} = f_{TX} = 13.56\,MHz$	$k = 15\%$

All line diagrams and locus curves from Section 4.1.10 are — unless stated otherwise — calculated using the constant parameter values listed in Table 4.3.

Transmission frequency f_{TX} Let us first change the *transmission frequency* f_{TX} of the reader, while the transponder resonant frequency f_{RES} is kept constant. Although this case does not occur in practice it is very useful as a theoretical experiment to help us to understand the principles behind the transformed transponder impedance Z'_T.

Figure 4.30 shows the locus curve $Z'_T = f(f_{TX})$ for this case. The impedance vector Z'_T traces a circle in the clockwise direction in the complex Z plane as transmission frequency f_{TX} increases.

In the frequency range below the transponder resonant frequency ($f_{TX} < f_{RES}$) the impedance vector Z'_T is initially found in quadrant I of the complex Z plane. The transformed transponder impedance Z'_T is inductive in this frequency range.

If the transmission frequency precisely corresponds with the transponder resonant frequency ($f_{TX} = f_{RES}$) then the reactive impedances for L_2 and C_2 in the transponder

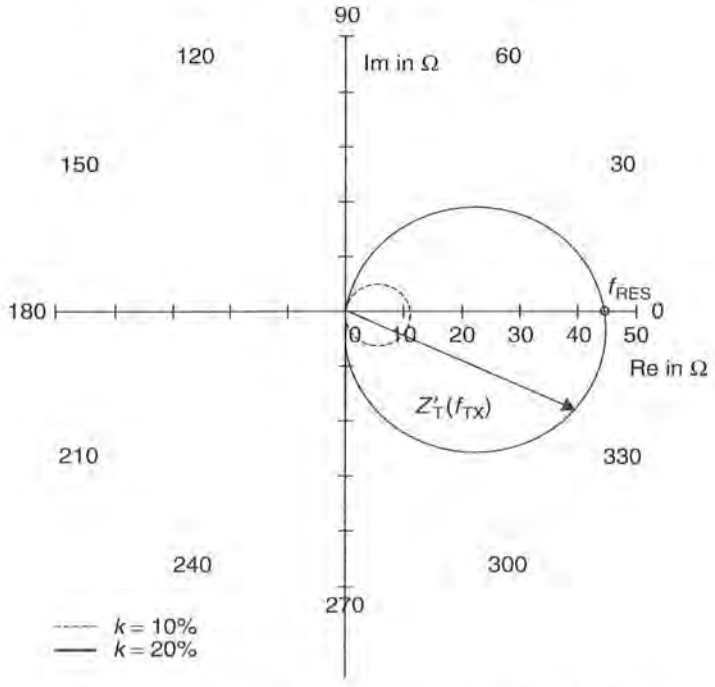


Figure 4.30 The impedance locus curve of the complex transformed transponder impedance Z'_T as a function of transmission frequency ($f_{TX} = 1$–$30\,MHz$) of the reader corresponds with the impedance locus curve of a parallel resonant circuit

cancel each other out. Z'_T acts as an ohmic (real) resistor — the locus curve thus intersects the real x axis of the complex Z plane at this point.

In the frequency range above the transponder resonant frequency ($f_{TX} > f_{RES}$), the locus curve finally passes through quadrant IV of the complex Z plane — Z'_T has a capacitive effect in this range.

The impedance locus curve of the complex transformed transponder impedance Z'_T corresponds with the impedance locus curve of a damped parallel resonant circuit with a parallel resonant frequency equal to the resonant frequency of the transponder. Figure 4.31 shows an equivalent circuit diagram for this. The complex current i_2 in the coil L_2 of the transponder resonant circuit is transformed by the magnetic mutual inductance M in the antenna coil L_1 of the reader and acts there as a parallel resonant circuit with the (frequency dependent) impedance Z'_T. The value of the real resistor R' in the equivalent circuit diagram corresponds with the point of intersection of the locus curve Z'_T with the real axis in the Z plane.

Coupling coefficient k Given constant geometry of the transponder and reader antenna, the *coupling coefficient* is defined by the distance and angle of the two coils in relation to each other (see Section 4.1.5). The influence of metals in the vicinity of the transmitter or transponder coil on the coupling coefficient should not be disregarded (e.g. shielding effect caused by eddy current losses). In practice, therefore, the coupling coefficient is the parameter that varies the most. Figure 4.32 shows the locus curve of the complex transformed transponder impedance for the range $0 \leq k \leq 1$. We differentiate between three ranges:

- $k = 0$: If the transponder coil L_2 is removed from the field of the reader antenna L_1 entirely, then no mutual inductance occurs. For this limit case, the transformed transponder impedance is no longer effective, that is $Z'_T(k = 0) = 0$.
- $0 < k < 1$: If the transponder coil L_2 is slowly moved towards the reader antenna L_1, then the coupling coefficient, and thus also the mutual inductance M between the two coils, increases continuously. The value of complex transformed transponder impedance increases proportionately, whereby $Z'_T \sim k_2$. When f_{TX} exactly

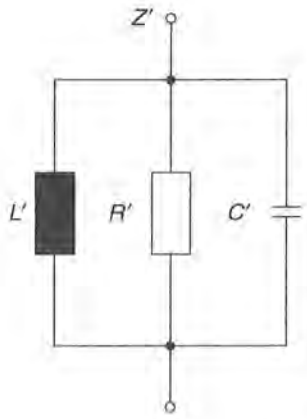


Figure 4.31 The equivalent circuit diagram of complex transformed transponder impedance Z'_T is a damped parallel resonant circuit

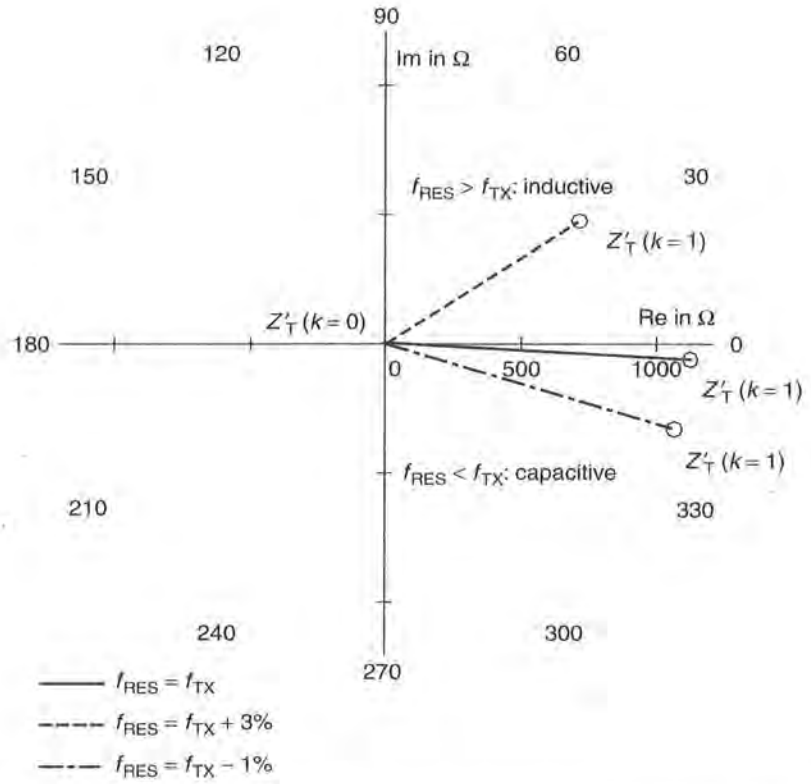


Figure 4.32 The locus curve of $Z'_T(k = 0{-}1)$ in the complex impedance plane as a function of the coupling coefficient k is a straight line

corresponds with f_{RES}, $Z'_T(k)$ remains real for all values of k.[4] Given a detuning of the transponder resonant frequency ($f_{RES} \neq f_{TX}$), on the other hand, Z'_T also has an inductive or capacitive component.

- $k = 1$: This case only occurs if both coils are identical in format, so that the windings of the two coils L_1 and L_2 lie directly on top of each other at distance $d = 0$. $Z'_T(k)$ reaches a maximum in this case. In general the following applies: $|Z'_T(k)_{max}| = |Z'_T(K_{max})|$.

Transponder capacitance C_2 We will now change the value of transponder capacitance C_2, while keeping all other parameters constant. This naturally detunes the resonant frequency f_{RES} of the transponder in relation to the transmission frequency f_{TX} of the reader. In practice, different factors may be responsible for a change in C_2:

- manufacturing tolerances, leading to a static deviation from the target value;
- a dependence of the data carrier's input capacitance on the input voltage u_2 due to effects in the semiconductor: $C_2 = f(u_2)$;

[4] The low angular deviation in the locus curve in Figure 4.32 where $f_{RES} = f_{TX}$ is therefore due to the fact that the resonant frequency calculated according to equation (4.34) is only valid without limitations for the undamped parallel resonant circuit. Given damping by R_L and R_2, on the other hand, there is a slight detuning of the resonant frequency. However, this effect can be largely disregarded in practice and thus will not be considered further here.

- intentional variation of the capacitance of C_2 for the purpose of data transmission (we will deal with so-called 'capacitive load modulation' in more detail in Section 4.1.10.3).
- detuning due to environmental influences such as metal, temperature, moisture, and 'hand capacitance' when the smart card is touched.

Figure 4.33 shows the locus curve for $Z'_T(C_2)$ in the complex impedance plane. As expected, the locus curve obtained is the circle in the complex Z plane that is typical of a parallel resonant circuit. Let us now consider the extreme values for C_2:

- $C_2 = 1/\omega_{TX}^2 L_2$: The resonant frequency of the transponder in this case precisely corresponds with the transmission frequency of the reader (see equation (4.25)). The current i_2 in the transponder coil reaches a maximum at this value due to resonance step-up and is real. Because $Z'_T \sim j\omega M \cdot i_2$ the value for impedance Z'_T also reaches a maximum — the locus curve intersects the real axis in the complex Z plane. The following applies: $|Z'_T(C_2)|_{max} = |Z'_T(C_2 = 1/\omega_{TX})^2 \cdot L_2)|$.
- $C_2 \neq 1/\omega^2 L_2$: If the capacitance C_2 is less than or greater than $C_2 = 1/\omega_{TX}^2 L_2$ then the resonant frequency of the transponder will be detuned and will vary significantly from the transmission frequency of the reader. The polarity of the current i_2 in the resonant circuit of the transponder varies when the resonant frequency is exceeded,

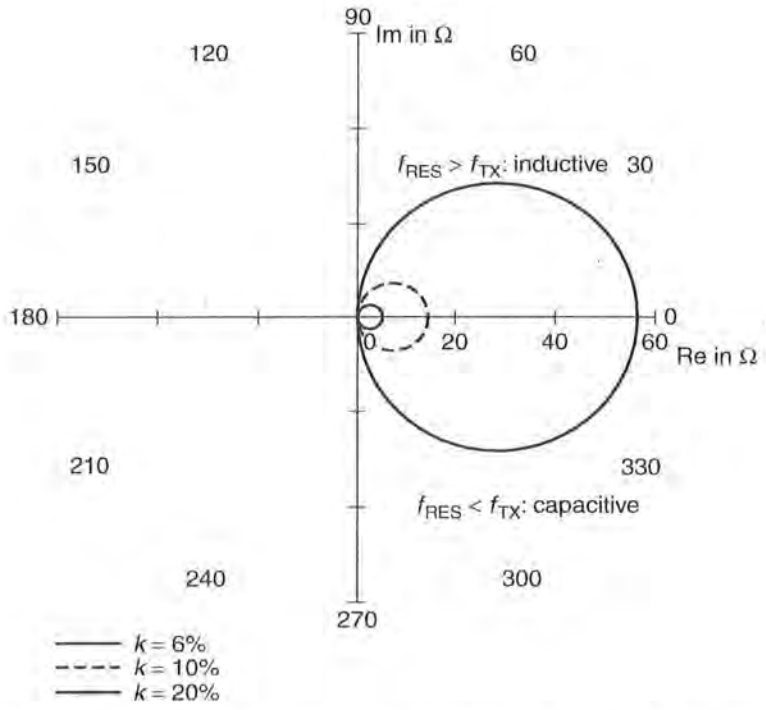


Figure 4.33 The locus curve of Z'_T ($C_2 = 10$–110 pF) in the complex impedance plane as a function of the capacitance C_2 in the transponder is a circle in the complex Z plane. The diameter of the circle is proportional to k_2

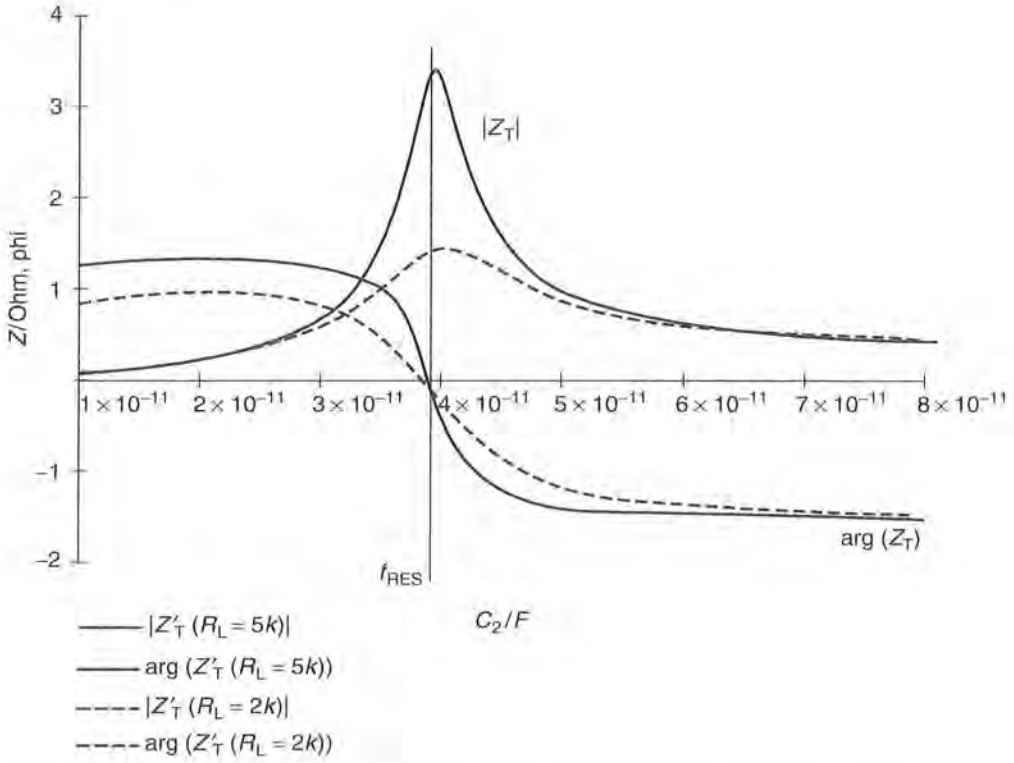


Figure 4.34 Value and phase of the transformed transponder impedance Z'_T as a function of C_2. The maximum value of Z'_T is reached when the transponder resonant frequency matches the transmission frequency of the reader. The polarity of the phase angle of Z'_T varies

as we can see from Figure 4.34. Similarly, the locus curve of Z'_T describes the familiar circular path in the complex Z plane. For both extreme values:

$$Z'_T(C_2 \to 0) = \frac{\omega^2 k^2 \cdot L_1 \cdot L_2}{j\omega L_2 + R_2 + R_L} \tag{4.51}$$

(no resonance step-up)

$$Z'_T(C_2 \to \infty) = \frac{\omega^2 k^2 \cdot L_1 \cdot L_2}{j\omega L_2 + R_2} \tag{4.52}$$

('short-circuited' transponder coil).

Load resistance R_L The *load resistance* R_L is an expression for the *power consumption* of the data carrier (microchip) in the transponder. Unfortunately, the load resistance is generally not constant, but falls as the coupling coefficient increases due to the influence of the shunt regulator (voltage regulator). The power consumption of the data carrier also varies, for example during the read or write operation. Furthermore, the value of the load resistance is often intentionally altered in order to transmit data to the reader (see Section 4.1.10.3).

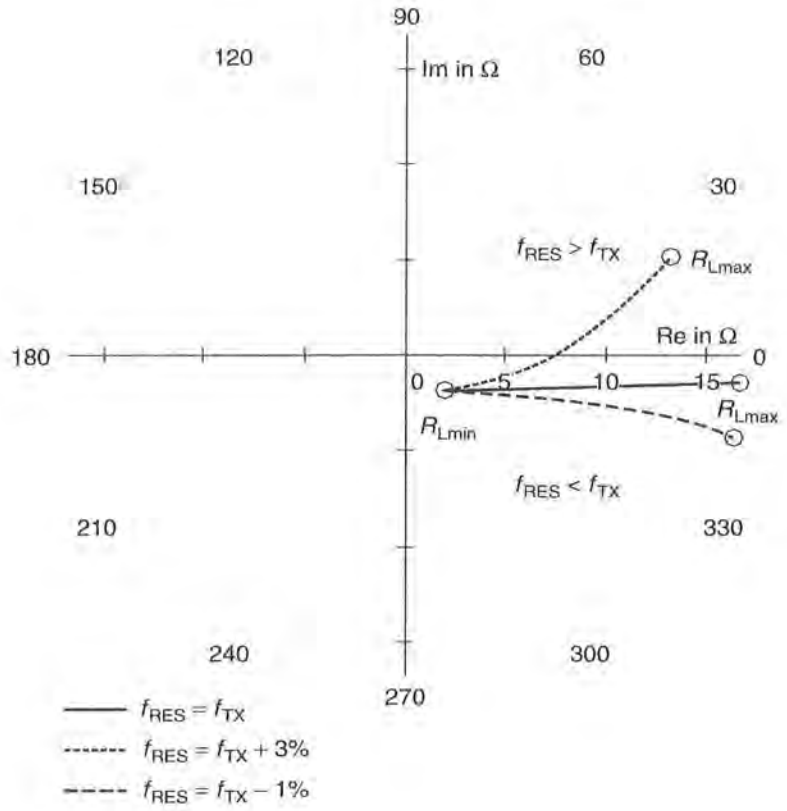


Figure 4.35 Locus curve of Z'_T ($R_L = 0.3–3\,k\Omega$) in the impedance plane as a function of the load resistance R_L in the transponder at different transponder resonant frequencies

Figure 4.35 shows the corresponding locus curve for $Z'_T = f(R_L)$. This shows that the transformed transponder impedance is proportional to R_L. Increasing load resistance R_L, which corresponds with a lower(!) current in the data carrier, thus also leads to a greater value for the transformed transponder impedance Z'_T. This can be explained by the influence of the load resistance R_L on the Q factor: a high-ohmic load resistance R_L leads to a high Q factor in the resonant circuit and thus to a greater current step-up in the transponder resonant circuit. Due to the proportionality $Z'_T \sim j\omega M \cdot i_2$ — and not to i_{RL} — we obtain a correspondingly high value for the transformed transponder impedance.

If the transponder resonant frequency is detuned we obtain a curved locus curve for the transformed transponder impedance Z'_T. This can also be traced back to the influence of the Q factor, because the phase angle of a detuned parallel resonant circuit also increases as the Q factor increases ($R_L \uparrow$), as we can see from a glance at Figure 4.34.

Let us reconsider the two extreme values of R_L:

$$Z'_T(R_L \to 0) = \frac{\omega^2 k^2 \cdot L_1 \cdot L_2}{R_2 + j\omega L_2} \tag{4.53}$$

('short-circuited' transponder coil)

$$Z'_T(R_L \to \infty) = \frac{\omega^2 k^2 \cdot L_1 \cdot L_2}{j\omega L_2 + R_2 + \dfrac{1}{j\omega C_2}} \tag{4.54}$$

(unloaded transponder resonant circuit).

Transponder inductance L_2 Let us now investigate the influence of inductance L_2 on the transformed transponder impedance, whereby the resonant frequency of the transponder is again held constant, so that $C_2 = 1/\omega_{TX}^2 L_2$.

Transformed transponder impedance reaches a clear peak at a given inductance value, as a glance at the line diagram shows (Figure 4.36). This behaviour is reminiscent of the graph of voltage $u_2 = f(L_2)$ (see also Figure 4.15). Here too the peak transformed transponder impedance occurs where the Q factor, and thus the current i_2 in the transponder, is at a maximum ($Z'_T \sim j\omega M \cdot i_2$). Please refer to Section 4.1.7 for an explanation of the mathematical relationship between load resistance and the Q factor.

4.1.10.3 Load modulation

Apart from a few other methods (see Chapter 3), so-called *load modulation* is the most common procedure for *data transmission* from transponder to reader by some margin.

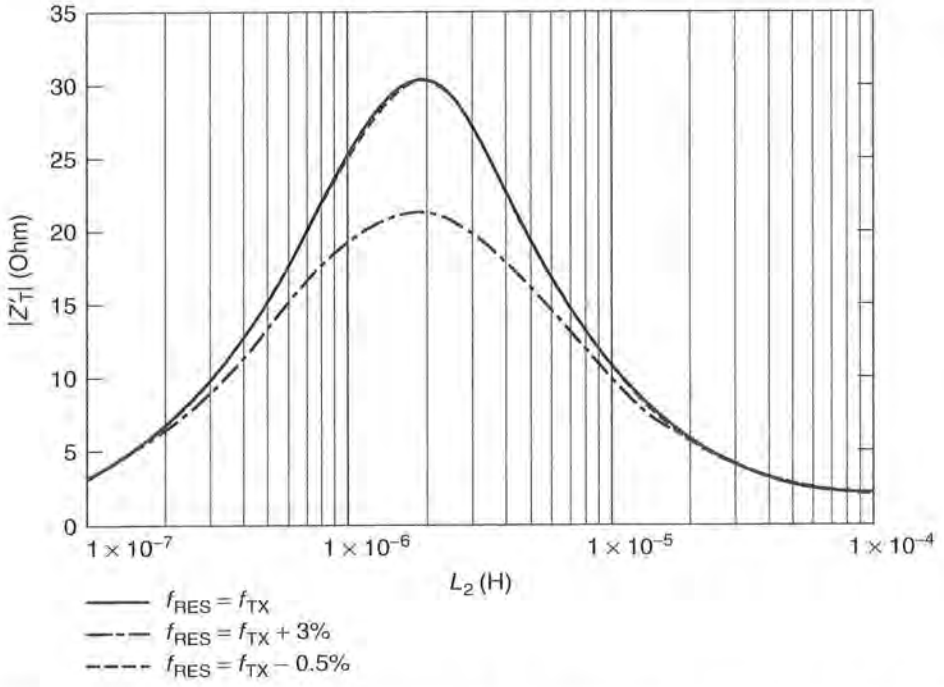


Figure 4.36 The value of Z'_T as a function of the transponder inductance L_2 at a constant resonant frequency f_{RES} of the transponder. The maximum value of Z'_T coincides with the maximum value of the Q factor in the transponder

By varying the circuit parameters of the *transponder resonant circuit* in time with the data stream, the magnitude and phase of the *transformed transponder impedance* can be influenced (modulation) such that the data from the transponder can be reconstructed by an appropriate evaluation procedure in the reader (demodulation).

However, of all the circuit parameters in the transponder resonant circuit, only two can be altered by the data carrier: the load resistance R_L and the parallel capacitance C_2. Therefore RFID literature distinguishes between ohmic (or real) and capacitive load modulation.

Ohmic load modulation In this type of load modulation a parallel resistor R_{mod} is switched on and off within the data carrier of the transponder in time with the data stream (or in time with a modulated subcarrier) (Figure 4.37). We already know from the previous section that the parallel connection of R_{mod} ($\rightarrow$ reduced total resistance) will reduce the Q factor and thus also the transformed transponder impedance Z'_T. This is also evident from the locus curve for the ohmic load modulator: Z'_T is switched between the values Z'_T (R_L) and Z'_T ($R_L \| R_{mod}$) by the load modulator in the transponder (Figure 4.38). The phase of Z'_T remains almost constant during this process (assuming $f_{TX} = f_{RES}$).

In order to be able to reconstruct (i.e. demodulate) the transmitted data, the falling voltage u_{ZT} at Z'_T must be sent to the receiver (RX) of the reader. Unfortunately, Z'_T is not accessible in the reader as a discrete component because the voltage u_{ZT} is induced in the real antenna coil L_1. However, the voltages u_{L1} and u_{R1} also occur at the antenna coil L_1, and they can only be measured at the terminals of the antenna coil as the total voltage u_{RX}. This total voltage is available to the receiver branch of the reader (see also Figure 4.25).

The vector diagram in Figure 4.39 shows the magnitude and phase of the voltage components u_{ZT}, u_{L1} and u_{R1} which make up the total voltage u_{RX}. The magnitude and phase of u_{RX} is varied by the modulation of the voltage component u_{ZT} by the load modulator in the transponder. *Load modulation* in the transponder thus brings about the *amplitude modulation* of the reader antenna voltage u_{RX}. The transmitted data is therefore not available in the baseband at L_1; instead it is found in the modulation products (= modulation sidebands) of the (load) modulated voltage u_1 (see Chapter 6).

Capacitive load modulation In *capacitive load modulation* it is an additional capacitor C_{mod}, rather than a modulation resistance, that is switched on and off in time with the data stream (or in time with a modulated subcarrier) (Figure 4.40). This causes the resonant frequency of the transponder to be switched between two frequencies.

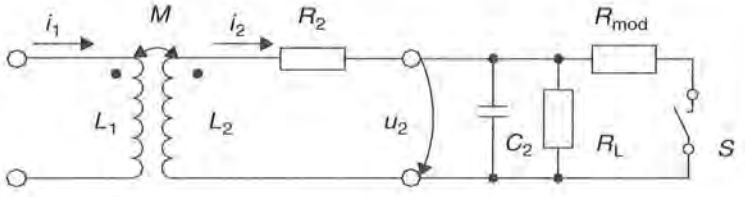


Figure 4.37 Equivalent circuit diagram for a transponder with load modulator. Switch S is closed in time with the data stream — or a modulated subcarrier signal — for the transmission of data

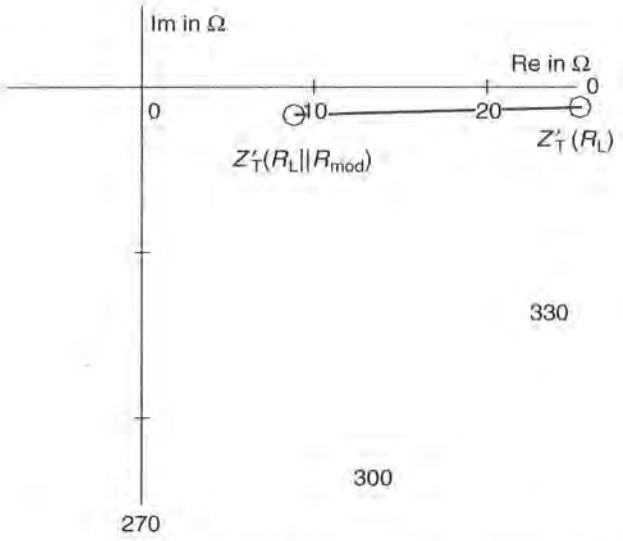


Figure 4.38 Locus curve of the transformed transponder impedance with ohmic load modulation ($R_L \| R_{mod} = 1.5\text{–}5\,k\Omega$) of an inductively coupled transponder. The parallel connection of the *modulation resistor* R_{mod} results in a lower value of Z'_T

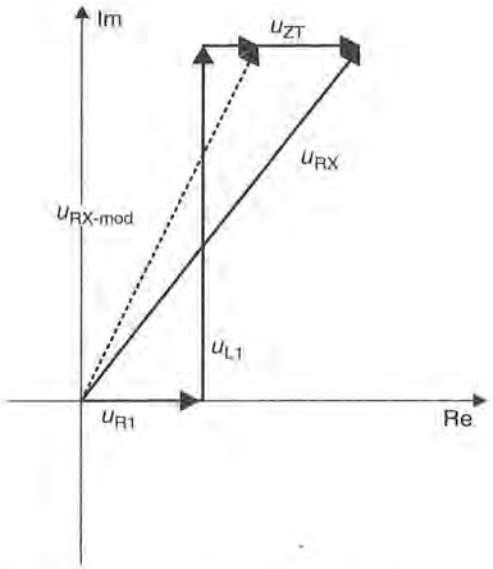


Figure 4.39 Vector diagram for the total voltage u_{RX} that is available to the receiver of a reader. The magnitude and phase of u_{RX} are modulated at the antenna coil of the reader (L_1) by an ohmic load modulator

We know from the previous section that the detuning of the transponder resonant frequency markedly influences the magnitude and phase of the transformed transponder impedance Z'_T. This is also clearly visible from the locus curve for the capacitive load modulator (Figure 4.41): Z'_T is switched between the values $Z'_T(\omega_{RES1})$ and $Z'_T(\omega_{RES2})$ by the load modulator in the transponder. The locus curve for Z'_T thereby passes

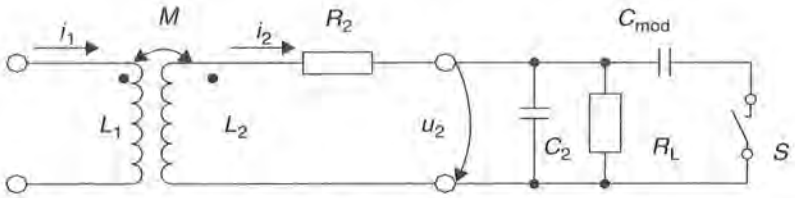


Figure 4.40 Equivalent circuit diagram for a transponder with capacitive load modulator. To transmit data the switch S is closed in time with the data stream — or a modulated subcarrier signal

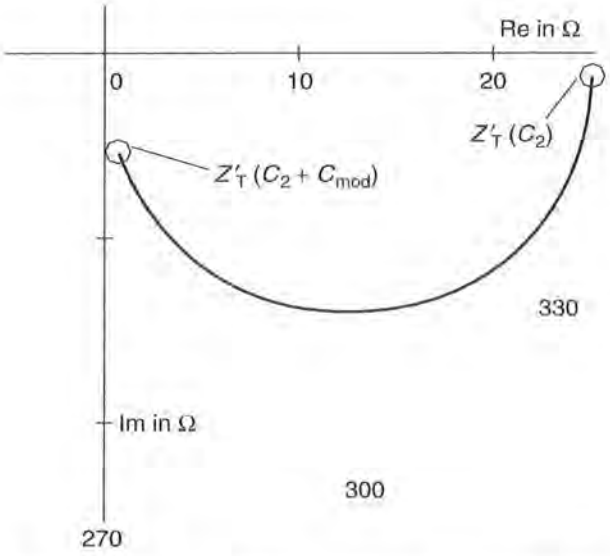


Figure 4.41 Locus curve of transformed transponder impedance for the capacitive load modulation ($C_2 \| C_{mod} = 40–60$ pF) of an inductively coupled transponder. The parallel connection of a *modulation capacitor* C_{mod} results in a modulation of the magnitude and phase of the transformed transponder impedance Z'_T

through a segment of the circle in the complex Z plane that is typical of the parallel resonant circuit.

Demodulation of the data signal is similar to the procedure used with ohmic load modulation. Capacitive *load modulation* generates a combination of *amplitude and phase modulation* of the reader antenna voltage u_{RX} and should therefore be processed in an appropriate manner in the receiver branch of the reader. The relevant vector diagram is shown in Figure 4.42.

Demodulation in the reader For transponders in the frequency range <135 kHz the load modulator is generally controlled directly by a serial data stream encoded in the baseband, e.g. a Manchester encoded bit sequence. The modulation signal from the transponder can be recreated by the rectification of the amplitude modulated voltage at the antenna coil of the reader (see Section 11.3).

In higher frequency systems operating at 6.78 MHz or 13.56 MHz, on the other hand, the transponder's load modulator is controlled by a modulated subcarrier signal

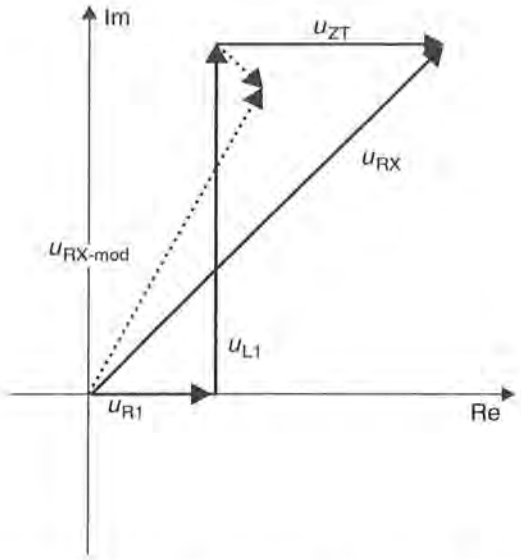


Figure 4.42 Vector diagram of the total voltage u_{RX} available to the receiver of the reader. The magnitude and phase of this voltage are modulated at the antenna coil of the reader (L_1) by a capacitive load modulator

(see Section 6.2.4). The subcarrier frequency f_H is normally 847 kHz (ISO 14443-2), 423 kHz (ISO 15693) or 212 kHz.

Load modulation with a subcarrier generates two sidebands at a distance of $\pm f_H$ to either side of the transmission frequency (see Section 6.2.4). The information to be transmitted is held in the two sidebands, with each sideband containing the same information. One of the two sidebands is filtered in the reader and finally demodulated to reclaim the baseband signal of the modulated data stream.

The influence of the Q factor As we know from the preceding section, we attempt to maximise the *Q factor* in order to maximise the energy range and the retroactive transformed transponder impedance. From the point of view of the energy range, a high Q factor in the transponder resonant circuit is definitely desirable. If we want to transmit data from or to the transponder a certain minimum bandwidth of the transmission path from the data carrier in the transponder to the receiver in the reader will be required. However, the bandwidth B of the *transponder resonant circuit* is inversely proportional to the Q factor.

$$B = \frac{f_{RES}}{Q} \tag{4.55}$$

Each load modulation operation in the transponder causes a corresponding amplitude modulation of the current i_2 in the transponder coil. The modulation sidebands of the current i_2 that this generates are damped to some degree by the bandwidth of the transponder resonant circuit, which is limited in practice. The bandwidth B determines a frequency range around the resonant frequency f_{RES}, at the limits of which the modulation sidebands of the current i_2 in the transponder reach a damping of 3 dB relative to the resonant frequency (Figure 4.43). If the Q factor of the transponder is

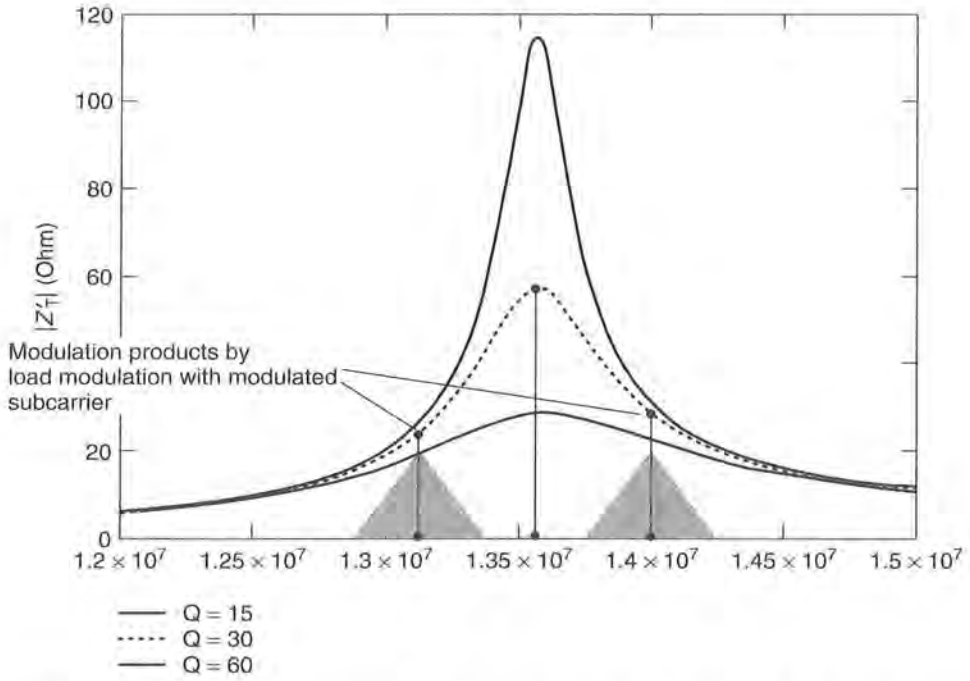


Figure 4.43 The transformed transponder impedance reaches a peak at the resonant frequency of the transponder. The amplitude of the modulation sidebands of the current i_2 is damped due to the influence of the bandwidth B of the transponder resonant circuit (where $f_H = 440\,\text{kHz}$, $Q = 30$)

too high, then the modulation sidebands of the current i_2 are damped to such a degree due to the low bandwidth that the range is reduced (transponder signal range).

Transponders used in 13.56 MHz systems that support an *anticollision algorithm* are adjusted to a resonant frequency of 15–18 MHz to minimise the mutual influence of several transponders. Due to the marked detuning of the transponder resonant frequency relative to the transmission frequency of the reader the two modulation sidebands of a load modulation system with subcarrier are transmitted at a different level (see Figure 4.44).

The term bandwidth is problematic here (the frequencies of the reader and the modulation sidebands may even lie outside the bandwidth of the transponder resonant circuit). However, the selection of the correct Q factor for the transponder resonant circuit is still important, because the Q factor can influence the transient effects during load modulation.

Ideally, the 'mean Q factor' of the transponder will be selected such that the energy range and transponder signal range of the system are identical. However, the calculation of an ideal Q factor is non-trivial and should not be underestimated because the Q factor is also strongly influenced by the *shunt regulator* (in connection with the distance d between transponder and reader antenna) and by the *load modulator* itself. Furthermore, the influence of the bandwidth of the transmitter antenna (series resonant circuit) on the level of the load modulation sidebands should not be underestimated.

Therefore, the development of an inductively coupled RFID system is always a compromise between the system's range and its data transmission speed (baud

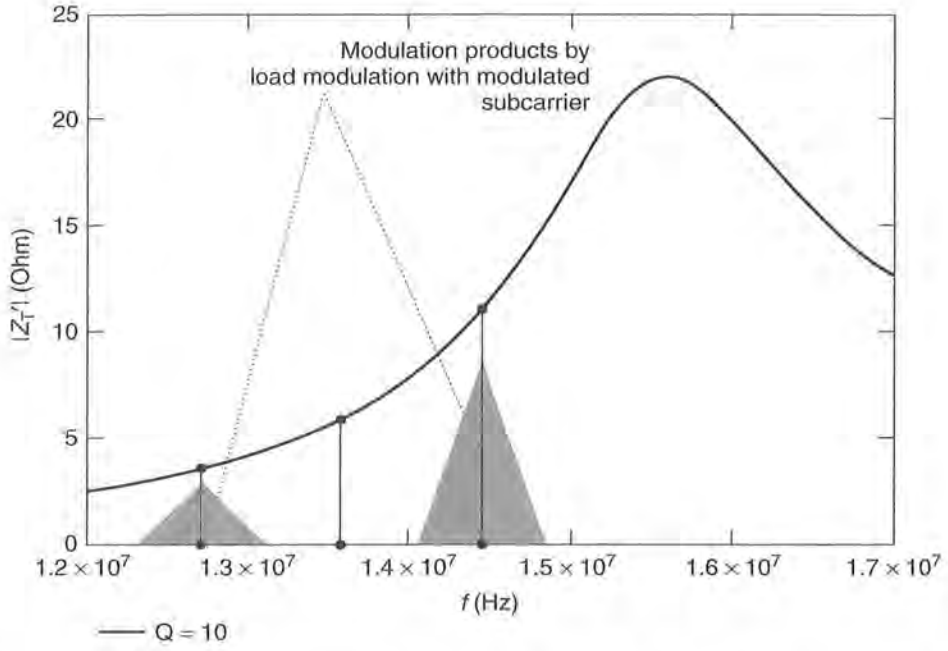


Figure 4.44 If the transponder resonant frequency is markedly detuned compared to the transmission frequency of the reader the two modulation sidebands will be transmitted at different levels. (Example based upon subcarrier frequency $f_H = 847\,\text{kHz}$)

rate/subcarrier frequency). Systems that require a short transaction time (that is, rapid data transmission and large bandwidth) often only have a range of a few centimetres, whereas systems with relatively long transaction times (that is, slow data transmission and low bandwidth) can be designed to achieve a greater range. A good example of the former case is provided by contactless smart cards for local public transport applications, which carry out authentication with the reader within a few 100 ms and must also transmit booking data. Contactless smart cards for 'hands free' access systems that transmit just a few bytes — usually the serial number of the data carrier — within 1–2 seconds are an example of the latter case. A further consideration is that in systems with a 'large' transmission antenna the data rate of the reader is restricted by the fact that only small sidebands may be generated because of the need to comply with the radio licensing regulations (ETS, FCC). Table 4.4 gives a brief overview of the relationship between range and bandwidth in inductively coupled RFID systems.

4.1.11 Measurement of system parameters

4.1.11.1 Measuring the coupling coefficient k

The coupling coefficient k and the associated mutual inductance M are the most important parameters for the design of an inductively coupled RFID system. It is precisely these parameters that are most difficult to determine analytically as a result of the — often complicated — field pattern. Mathematics may be fun, but has its limits.

Table 4.4 Typical relationship between range and bandwidth in 13.56 MHz systems. An increasing Q factor in the transponder permits a greater range in the transponder system. However, this is at the expense of the bandwidth and thus also the data transmission speed (baud rate) between transponder and reader

System	Baud rate	$f_{Subcarrier}$	f_{TX}	Range
ISO 14443	106 kBd	847 kHz	13.56 MHz	0–10 cm
ISO 15693 short	26.48 kBd	484 kHz	13.56 MHz	0–30 cm
ISO 15693 long	6.62 kBd	484 kHz	13.56 MHz	0–70 cm
Long-range system	9.0 kBd	212 kHz	13.56 MHz	0–1 m
LF system	–0–10 kBd	No subcarrier	<125 kHz	0–1.5 m

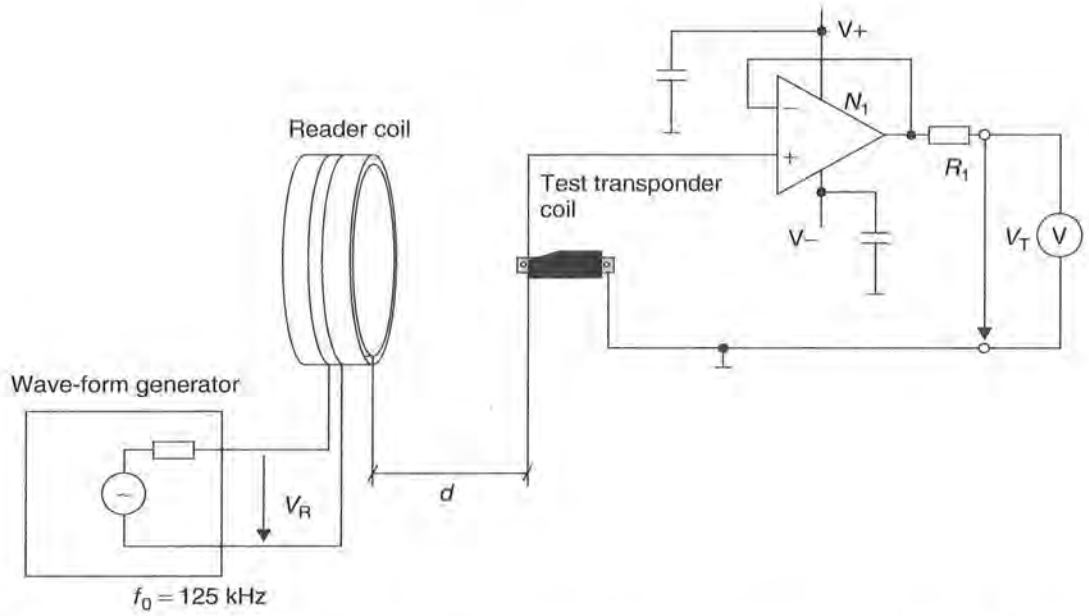


Figure 4.45 Measurement circuit for the measurement of the magnetic coupling coefficient k. N1: TL081 or LF 356N, R1: 100–500 Ω (reproduced by permission of TEMIC Semiconductor GmbH, Heilbronn)

Furthermore, the software necessary to calculate a numeric simulation is often unavailable — or it may simply be that the time or patience is lacking.

However, the coupling coefficient k for an existing system can be quickly determined by means of a simple measurement. This requires a test transponder coil with electrical and mechanical parameters that correspond with those of the 'real' transponder. The coupling coefficient can be simply calculated from the measured voltages U_R at the reader coil and U_T at the transponder coil (in Figure 4.45 these are denoted as V_R and V_T):

$$k = A_k \cdot \frac{U_T}{U_R} \cdot \sqrt{\frac{L_R}{L_T}} \tag{4.56}$$

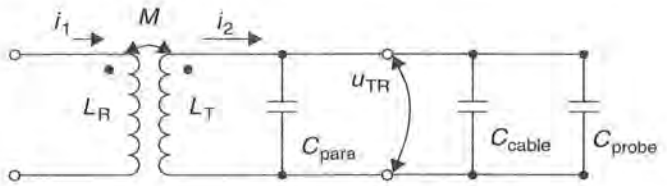


Figure 4.46 Equivalent circuit diagram of the test transponder coil with the parasitic capacitances of the measuring circuit

where U_T is the voltage at the transponder coil, U_R is the voltage at the reader coil, L_T and L_R are the inductance of the coils and A_K is the correction factor (<1).

The parallel, parasitic capacitances of the measuring circuit and the test transponder coil itself influence the result of the measurement because of the undesired current i_2. To compensate for this effect, equation (4.56) includes a correction factor A_K. Where $C_{TOT} = C_{para} + C_{cable} + C_{probe}$ (see Figure 4.46) the correction factor is defined as:

$$A_k = 2 - \frac{1}{1 - (\omega^2 \cdot C_{TOT} \cdot L_T)} \tag{4.57}$$

In practice, the correction factor in the low capacitance layout of the measuring circuit is $A_K \sim 0.99 - 0.8$ (TEMIC, 1977).

4.1.11.2 Measuring the transponder resonant frequency

The precise measurement of the transponder resonant frequency so that deviations from the desired value can be detected is particularly important in the manufacture of inductively coupled transponders. However, since transponders are usually packed in a glass or plastic housing, which renders them inaccessible, the measurement of the resonant frequency can only be realised by means of an inductive coupling.

The measurement circuit for this is shown in Figure 4.47. A coupling coil (conductor loop with several windings) is used to achieve the inductive coupling between transponder and measuring device. The self-resonant frequency of this coupling coil should be significantly higher (by a factor of at least 2) than the self-resonant frequency of the transponder in order to minimise measuring errors.

A phase and impedance analyser (or a network analyser) is now used to measure the impedance Z_1 of the coupling coil as a function of frequency. If Z_1 is represented

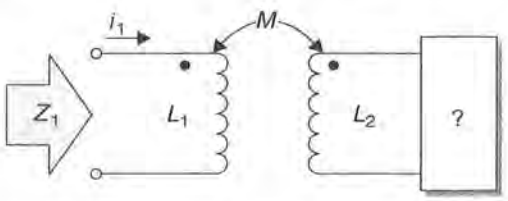


Figure 4.47 The circuit for the measurement of the *transponder resonant frequency* consists of a coupling coil L_1 and a measuring device that can precisely measure the complex impedance of Z_1 over a certain frequency range

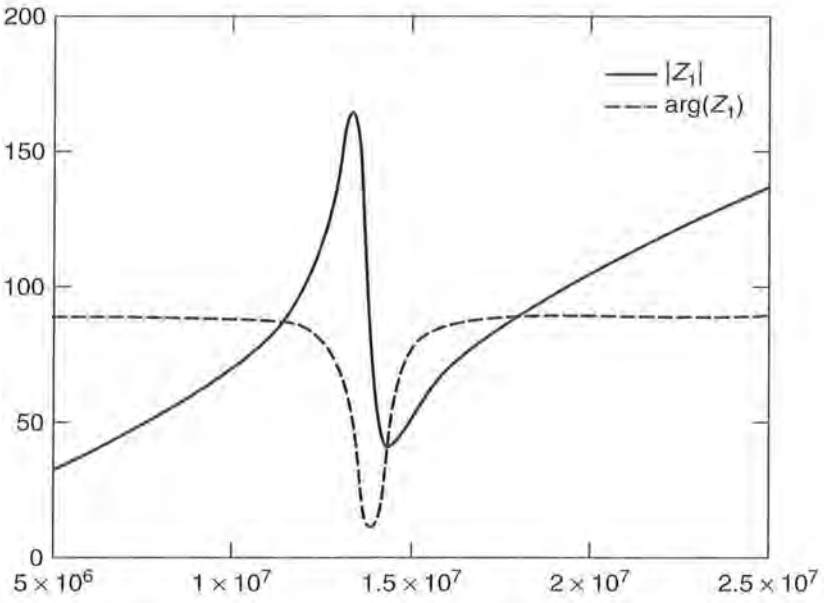


Figure 4.48 The measurement of impedance and phase at the measuring coil permits no conclusion to be drawn regarding the frequency of the transponder

in the form of a line diagram it has a curved path, as shown in Figure 4.48. As the measuring frequency rises the line diagram passes through various local maxima and minima for the magnitude and phase of Z_1. The sequence of the individual maxima and minima is always the same.

In the event of mutual inductance with a transponder the impedance Z_1 of the coupling coil L_1 is made up of several individual impedances:

$$Z_1 = R_1 + j\omega L_1 + Z'_T \tag{4.58}$$

Apart from at the transponder resonant frequency f_{RES}, Z'_T tends towards zero, so $Z_1 = R_L + j\omega L_1$. The locus curve in this range is a line parallel to the imaginary y axis of the complex Z plane at a distance of R_1 from it. If the measuring frequency approaches the transponder resonant frequency this straight line becomes a circle as a result of the influence of Z'_T. The locus curve for this is shown in Figure 4.49. The transponder resonant frequency corresponds with the maximum value of the real component of Z_1 (however this is not visible in the line diagram shown in Figure 4.48). The appearance of the individual maxima and minima of the line diagram can also be seen in the locus curve. A precise measurement of the transponder resonant frequency is therefore only possible using measuring devices that permit a separate measurement of R and X or can display a locus curve or line diagram.

4.1.12 Magnetic materials

Materials with a relative *permeability* >1 are termed ferromagnetic materials. These materials are iron, cobalt, nickel, various alloys and ferrite.

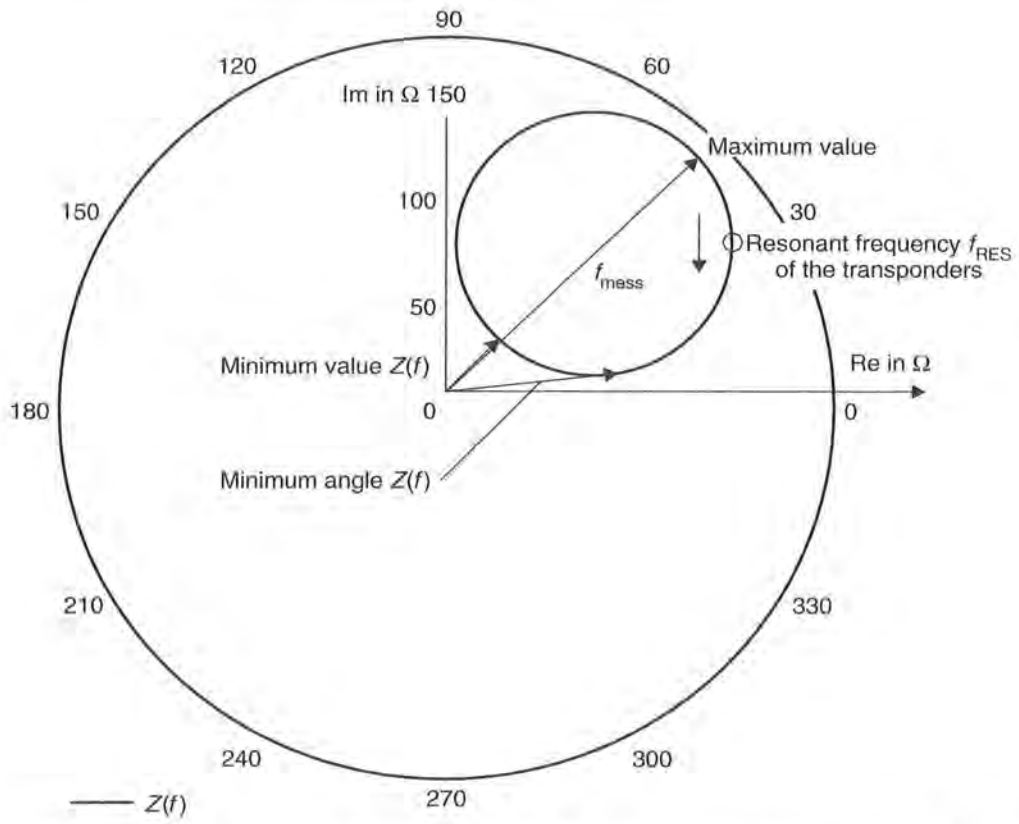


Figure 4.49 The locus curve of impedance Z_1 in the frequency range 1–30 MHz

4.1.12.1 Properties of magnetic materials and ferrite

One important characteristic of a magnetic material is the *magnetisation characteristic* or *hysteresis curve*. This describes $B = f(\mathrm{H})$, which displays a typical path for all ferromagnetic materials.

Starting from the unmagnetized state of the ferromagnetic material, the virgin curve $A \rightarrow B$ is obtained as the magnetic field strength H increases. During this process, the molecular magnets in the material align themselves in the B direction. (Ferromagnetism is based upon the presence of molecular magnetic dipoles. In these, the electron circling the atomic core represents a current and generates a magnetic field. In addition to the movement of the electron along its path, the rotation of the electron around itself, the spin, also generates a magnetic moment, which is of even greater importance for the material's magnetic behaviour.) Because there is a finite number of these molecular magnets, the number that remain to be aligned falls as the magnetic field increases, thus the gradient of the hysteresis curve falls. When all molecular magnets have been aligned, B rises in proportion to H only to the same degree as in a vacuum (Figure 4.50).

When the field strength H falls to $H = 0$, the flux density B falls to the positive residual value B_R, the remanence. Only after the application of an opposing field

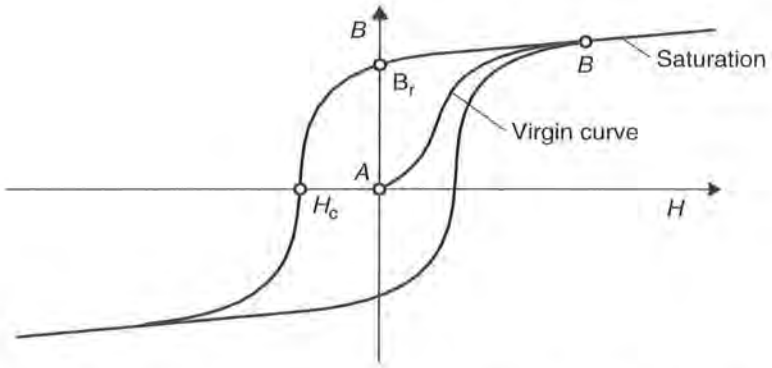


Figure 4.50 Typical magnetisation or hysteresis curve for a ferromagnetic material

$(-H)$ does the flux density B fall further and finally return to zero. The field strength necessary for this is termed the coercive field strength H_C.

Ferrite is the main material used in high frequency technology. This is used in the form of soft magnetic ceramic materials (low B_r), composed mainly of mixed crystals or compounds of iron oxide (Fe_2O_3) with one or more oxides of bivalent metals (NiO, ZnO, MnO etc.) (Vogt. Elektronik, 1990). The manufacturing process is similar to that for ceramic technologies (sintering).

The main characteristic of ferrite is its high specific electrical resistance, which varies between 1 and 10^6 Ωm depending upon the material type, compared to the range for metals, which vary between 10^{-5} and 10^{-4} Ωm. Because of this, eddy current losses are low and can be disregarded over a wide frequency range.

The relative permeability of ferrites can reach the order of magnitude of $\mu_r = 2000$.

An important characteristic of ferrite materials is their material-dependent limit frequency, which is listed in the datasheets provided by the ferrite manufacturer. Above the limit frequency increased losses occur in the ferrite material, and therefore ferrite should not be used outside the specified frequency range.

4.1.12.2 Ferrite antennas in LF transponders

Some applications require extremely small transponder coils (Figure 4.51). In transponders for animal identification, typical dimensions for cylinder coils are $d \times l = 5\,\text{mm} \times 0.75\,\text{mm}$. The mutual inductance that is decisive for the power supply of the transponder falls sharply due to its proportionality with the cross-sectional area of the coil ($M \sim A$; equation (4.13)). By inserting a ferrite material with a high permeability μ into the coil ($M \sim \Psi \rightarrow M \sim \mu \cdot H \cdot A$; equation (4.13)), the mutual inductance can be significantly increased, thus compensating for the small cross-sectional area of the coil.

The inductance of a *ferrite antenna* can be calculated according to the following equation (Philips Components, 1994):

$$L = \frac{\mu_0 \mu_{\text{Ferrite}} \cdot n^2 \cdot A}{l} \tag{4.59}$$

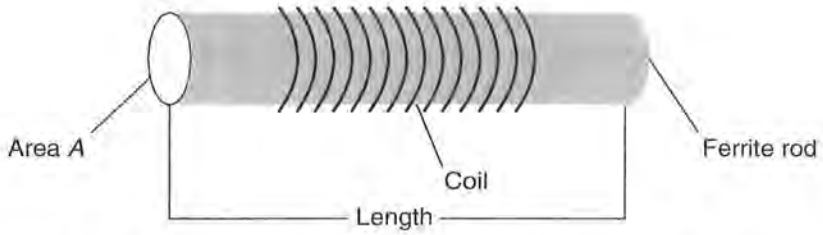


Figure 4.51 Configuration of a ferrite antenna in a 135 kHz glass transponder

4.1.12.3 Ferrite shielding in a metallic environment

The use of (inductively coupled) RFID systems often requires that the reader or transponder antenna be mounted directly upon a metallic surface. This might be the reader antenna of an automatic ticket dispenser or a transponder for mounting on gas bottles (see Figure 4.52).

However, it is not possible to fit a magnetic antenna directly onto a metallic surface. The magnetic flux through the metal surface induces eddy currents within the metal, which oppose the field responsible for their creation, i.e. the reader's field (Lenz's law), thus damping the magnetic field in the surface of the metal to such a degree that communication between reader and transponder is no longer possible. It makes no difference here whether the magnetic field is generated by the coil mounted upon the metal surface (reader antenna) or the field approaches the metal surface from 'outside' (transponder on metal surface).

By inserting highly permeable ferrite between the coil and metal surface it is possible to largely prevent the occurrence of eddy currents. This makes it possible to mount the antenna on metal surfaces.

When fitting antennas onto ferrite surfaces it is necessary to take into account the fact that the inductance of the conductor loop or coils may be significantly increased by

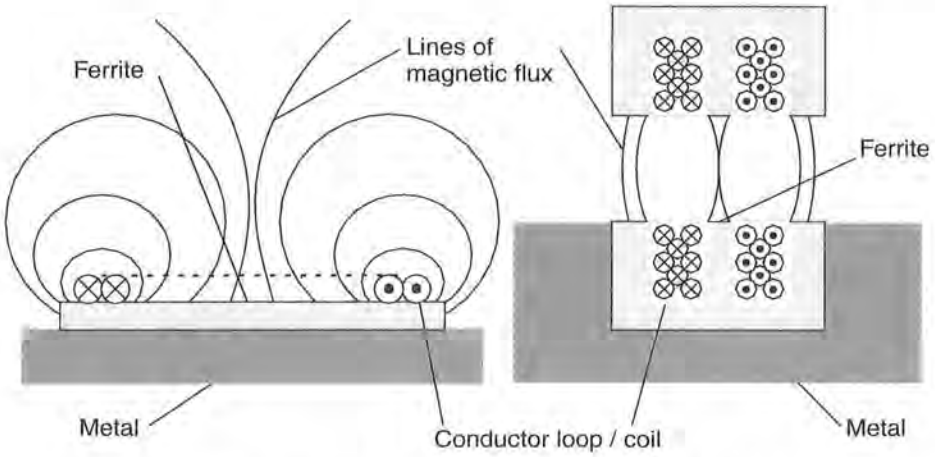


Figure 4.52 Reader antenna (left) and gas bottle transponder in a u-shaped core with read head (right) can be mounted directly upon or within metal surfaces using ferrite shielding

the permeability of the ferrite material, and it may therefore be necessary to readjust the resonant frequency or even redimension the matching network (in readers) altogether (see Section 11.4).

4.1.12.4 Fitting transponders in metal

Under certain circumstances it is possible to fit transponders directly into a metallic environment (Figure 4.53). *Glass transponders* are used for this because they contain a coil on a highly permeable *ferrite rod*. If such a transponder is inserted horizontally into a long groove on the metal surface somewhat larger than the transponder itself, then the transponder can be read without any problems. When the transponder is fitted horizontally the field lines through the transponder's ferrite rod run in parallel to the *metal surface* and therefore the eddy current losses remain low. The insertion of the transponder into a vertical bore would be unsuccessful in this situation, since the field lines through the transponder's ferrite rod in this arrangement would end at the top of the bore at right angles to the metal surface. The *eddy current losses* that occur in this case hinder the interrogation of a transponder.

It is even possible to cover such an arrangement with a *metal lid*. However, a narrow gap of dielectric material (e.g. paint, plastic, air) is required between the two metal surfaces in order to interrogate the transponder. The field lines running parallel

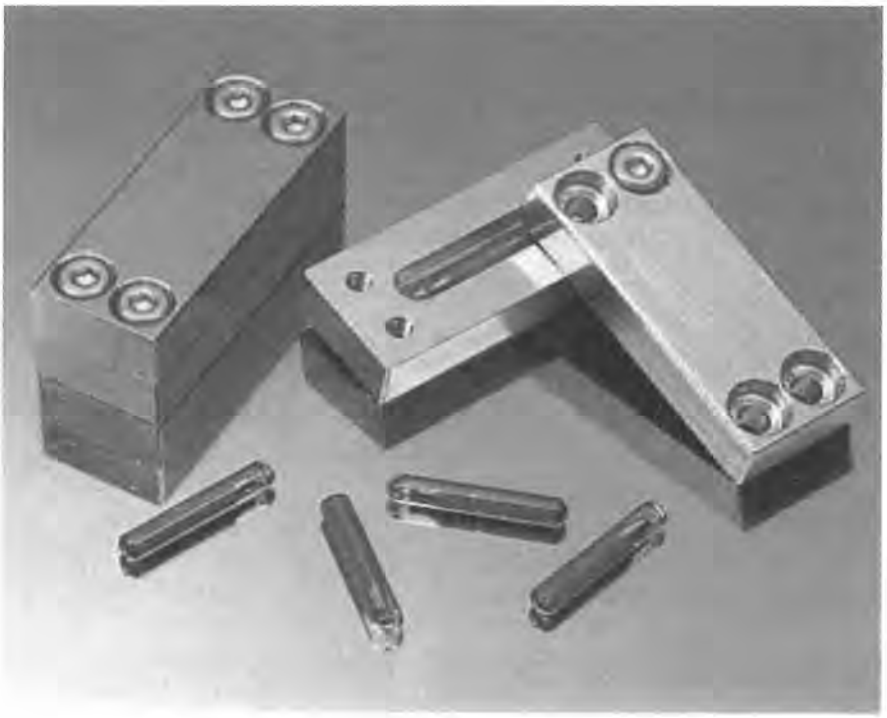


Figure 4.53 Right, fitting a glass transponder into a metal surface; left, the use of a thin dielectric gap allows the transponders to be read even through a metal casing (Photo: HANEX HXID system with Sokymat glass transponder in metal, reproduced by permission of HANEX Co. Ltd, Japan)

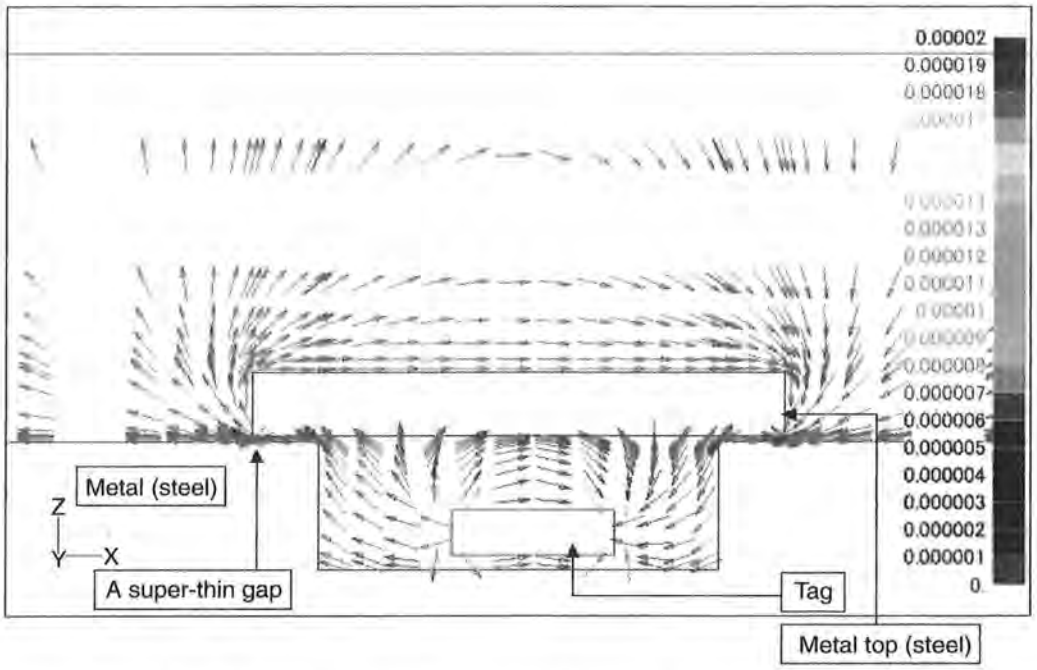


Figure 4.54 Path of field lines around a transponder encapsulated in metal. As a result of the dielectric gap the field lines run in parallel to the metal surface, so that eddy current losses are kept low (reproduced by permission of HANEX Co. Ltd, Japan)

to the metal surface enter the cavity through the *dielectric gap* (see Figure 4.54), so that the transponder can be read. Fitting transponders in metal allows them to be used in particularly hostile environments. They can even be run over by vehicles weighing several tonnes without suffering any damage.

Disk tags and contactless smart cards can also be embedded between metal plates. In order to prevent the magnetic field lines from penetrating into the metal cover, metal foils made of a highly permeable *amorphous metal* are placed above and below the tag (Hanex, n.d.). It is of crucial importance for the functionality of the system that the amorphous foils each cover only one half of the tag.

The magnetic field lines enter the amorphous material in parallel to the surface of the metal plates and are carried through it as in a conductor (Figure 4.55). At the gap between the two part foils a magnetic flux is generated through the transponder coil, so that this can be read.

4.2 Electromagnetic Waves

4.2.1 The generation of electromagnetic waves

Earlier in the book we described how a time varying *magnetic field* in space induces an *electric field* with closed field lines (rotational field) (see also Figure 4.11). The electric field surrounds the magnetic field and itself varies over time. Due to the variation of the electric rotational field over time, a magnetic field with closed field lines occurs in

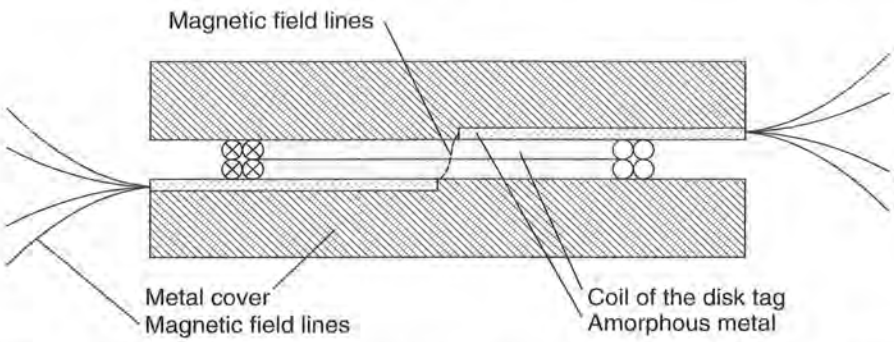


Figure 4.55 Cross-section through a sandwich made of disk transponder and metal plates. Foils made of amorphous metal cause the magnetic field lines to be directed outwards

space (rotational field). It surrounds the electric field and itself varies over time, thus generating another electric field. Due to the mutual dependence of the time varying fields there is a chain effect of electric and magnetic fields in space (Fricke *et al.*, 1979).

Radiation can only occur given a finite propagation speed ($c \approx 300\,000$ km/s; *speed of light*) for the electromagnetic field, which prevents a change in the voltage at the antenna from being followed immediately by the field in the vicinity of the change. Figure 4.56 shows the creation of an *electromagnetic wave* at a *dipole antenna*. Even at the alternating voltage's zero crossover (Figure 4.56c), the field lines remaining in space from the previous half wave cannot end at the antenna, but close into themselves, forming eddies. The eddies in the opposite direction that occur in the next half wave propel the existing eddies, and thus the energy stored in this field, away from the emitter at the speed of light c. The magnetic field is interlinked with the varying electrical field that propagates at the same time. When a certain distance is reached, the fields are released from the emitter, and this point represents the beginning of electromagnetic radiation ($\rightarrow$ far field). At high frequencies, that is small wavelengths, the radiation generated is particularly effective, because in this case the separation takes place in the direct vicinity of the emitter, where high field strengths still exist (Fricke *et al.*, 1979).

The distance between two field eddies rotating in the same direction is called the *wavelength* λ of the electromagnetic wave, and is calculated from the quotient of the speed of light c and the frequency of the radiation:

$$\lambda = \frac{c}{f} \tag{4.60}$$

4.2.1.1 Transition from near field to far field in conductor loops

The primary magnetic field generated by a *conductor loop* begins at the antenna (see also Section 4.1.1.1). As the magnetic field propagates an electric field increasingly also develops by induction (compare Figure 4.11). The field, which was originally purely magnetic, is thus continuously transformed into an electromagnetic field. Moreover,

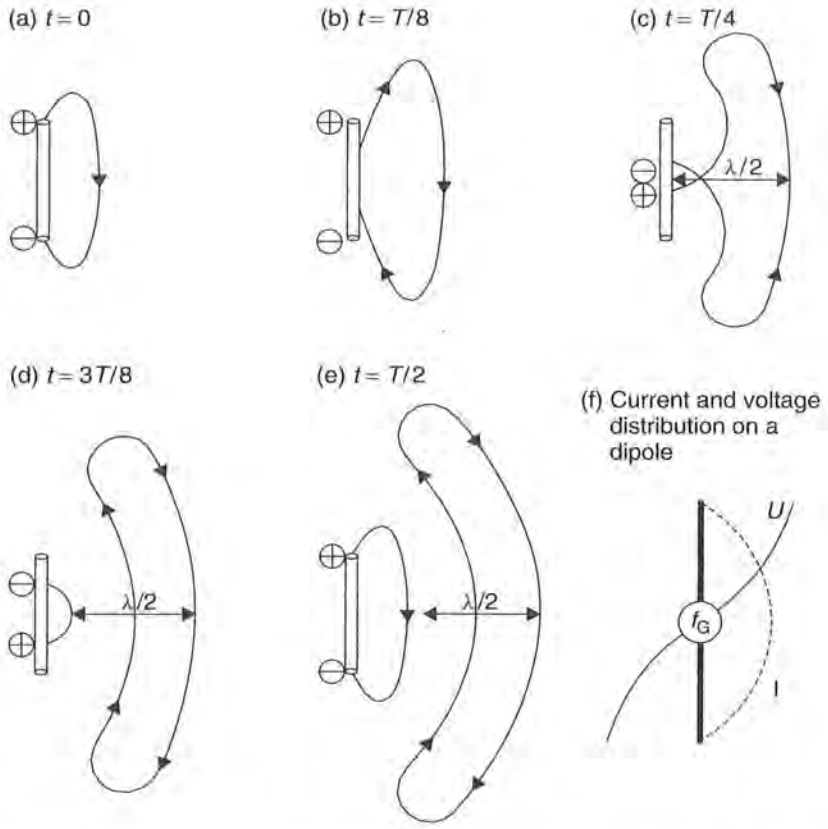


Figure 4.56 The creation of an electromagnetic wave at a dipole antenna. The electric field E is shown. The magnetic field H forms as a ring around the antenna and thus lies at right angles to the electric field

at a distance of $\lambda/2\pi$ the electromagnetic field begins to separate from the antenna and wanders into space in the form of an electromagnetic wave. The area from the antenna to the point where the electromagnetic field forms is called the *near field* of the antenna. The area after the point at which the electromagnetic wave has fully formed and separated from the antenna is called the *far field*.

A separated electromagnetic wave can no longer retroact upon the antenna that generated it by inductive or capacitive coupling. For inductively coupled RFID systems this means that once the far field has begun a *transformer (inductive) coupling* is no

Table 4.5 Frequency and wavelengths of different VHF–UHF frequencies

Frequency	Wavelength (cm)
433 MHz	69 (70 cm band)
868 MHz	34
915 MHz	33
2.45 GHz	12
5.8 GHz	5.2

Table 4.6 r_F and λ for different frequency ranges

Frequency	Wavelength λ (m)	$\lambda/2\pi$ (m)
<135 kHz	>2222	>353
6.78 MHz	44.7	7.1
13.56 MHz	22.1	3.5
27.125 MHz	11.0	1.7

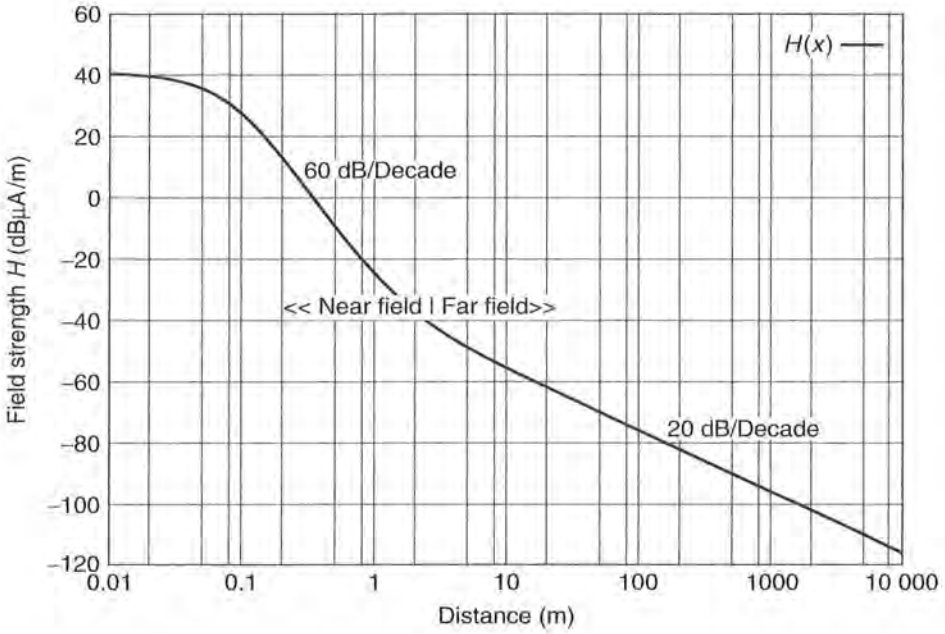


Figure 4.57 Graph of the magnetic field strength H in the transition from near to far field at a frequency of 13.56 MHz

longer possible. The beginning of the far field (the radius $r_F = \lambda/2\pi$ can be used as a rule of thumb) around the antenna thus represents an insurmountable *range limit* for inductively coupled systems.

The field strength path of a magnetic antenna along the coil x axis follows the relationship $1/d^3$ in the near field, as demonstrated above. This corresponds with a damping of 60 dB per decade (of distance). Upon the transition to the far field, on the other hand, the damping path flattens out, because after the separation of the field from the antenna only the *free space attenuation* of the electromagnetic waves is relevant to the field strength path (Figure 4.57). The field strength then decreases only according to the relationship $1/d$ as distance increases (see equation (4.65)). This corresponds with a damping of just 20 dB per decade (of distance).

4.2.2 Radiation density *S*

An *electromagnetic wave* propagates into space spherically from the point of its creation. At the same time, the electromagnetic wave transports energy in the surrounding

space. As the distance from the radiation source increases, this energy is divided over an increasing sphere surface area. In this connection we talk of the *radiation power* per unit area, also called *radiation density* S.

In a *spherical emitter*, the so-called *isotropic emitter*, the energy is radiated uniformly in all directions. At distance r the radiation density S can be calculated very easily as the quotient of the energy supplied by the emitter (thus the transmission power P_{EIRP}) and the surface area of the sphere.

$$S = \frac{P_{\mathrm{EIRP}}}{4\pi r^2} \tag{4.61}$$

4.2.3 Characteristic wave impedance and field strength *E*

The energy transported by the electromagnetic wave is stored in the electric and magnetic field of the wave. There is therefore a fixed relationship between the radiation density S and the field strengths E and H of the interconnected electric and magnetic fields. The electric field with electric field strength E is at right angles to the magnetic field H. The area between the vectors E and H forms the wave front and is at right angles to the direction of propagation. The radiation density S is found from the Poynting radiation vector S as a vector product of E and H (Figure 4.58).

$$S = E \times H \tag{4.62}$$

The relationship between the field strengths E and H is defined by the permittivity and the dielectric constant of the propagation medium of the electromagnetic wave. In a vacuum and also in air as an approximation:

$$E = H \cdot \sqrt{\mu_0 \varepsilon_0} = H \cdot Z_{\mathrm{F}} \tag{4.63}$$

Z_{F} is termed the *characteristic wave impedance* ($Z_{\mathrm{F}} = 120\pi\,\Omega = 377\,\Omega$). Furthermore, the following relationship holds:

$$E = \sqrt{S \cdot Z_{\mathrm{F}}} \tag{4.64}$$

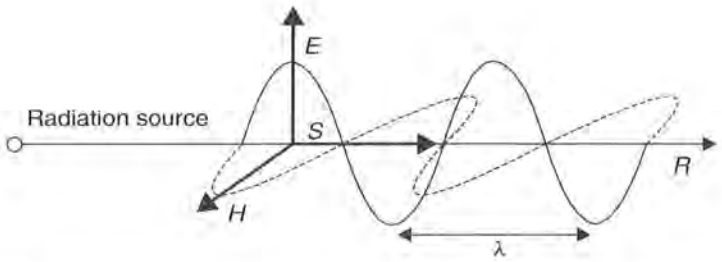


Figure 4.58 The Poynting radiation vector S as the vector product of E and H

Therefore, the field strength E at a certain distance r from the radiation source can be calculated using equation (4.61). P_{EIRP} is the transmission power emitted from the isotropic emitter:

$$E = \sqrt{\frac{P_{EIRP} \cdot Z_F}{4\pi r^2}} \tag{4.65}$$

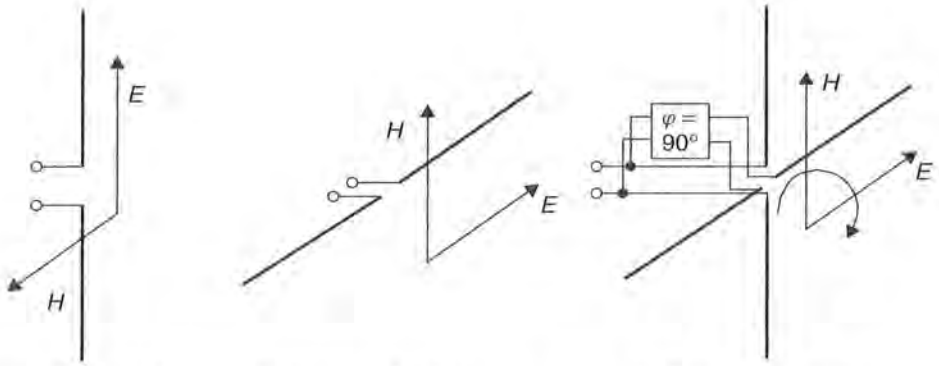
4.2.4 Polarisation of electromagnetic waves

The *polarisation* of an electromagnetic wave is determined by the direction of the electric field of the wave. We differentiate between *linear polarisation* and *circular polarisation*. In linear polarisation the direction of the field lines of the electric field E in relation to the surface of the earth provide the distinction between *horizontal* (the electric field lines run parallel to the surface of the earth) and *vertical* (the electric field lines run at right angles to the surface of the earth) *polarisation*.

So, for example, the dipole antenna is a linear polarised antenna in which the electric field lines run parallel to the dipole axis. A dipole antenna mounted at right angles to the earth's surface thus generates a vertically polarised electromagnetic field.

The transmission of energy between two linear polarised antennas is optimal if the two antennas have the same polarisation direction. Energy transmission is at its lowest point, on the other hand, when the polarisation directions of transmission and receiving antennas are arranged at exactly 90° or 270° in relation to one another (e.g. a horizontal antenna and a vertical antenna). In this situation an additional damping of 20 dB has to be taken into account in the power transmission due to *polarisation losses* (Rothammel, 1981), i.e. the receiving antenna draws just 1/100 of the maximum possible power from the emitted electromagnetic field.

In RFID systems, there is generally no fixed relationship between the position of the portable transponder antenna and the reader antenna. This can lead to fluctuations in the read range that are both high and unpredictable. This problem is aided by the use of circular polarisation in the reader antenna. The principle generation of circular polarisation is shown in Figure 4.59: two dipoles are fitted in the form of a cross. One of the two dipoles is fed via a 90° ($\lambda/4$) delay line. The polarisation direction of



(a) Vertical polarisation (b) Horizontal polarisation (c) Circular polarisation

Figure 4.59 Definition of the polarisation of electromagnetic waves

the electromagnetic field generated in this manner rotates through 360° every time the wave front moves forward by a wavelength. The rotation direction of the field can be determined by the arrangement of the delay line. We differentiate between left-handed and right-handed circular polarisation.

A polarisation loss of 3 dB should be taken into account between a linear and a circular polarised antenna; however, this is independent of the polarisation direction of the receiving antenna (e.g. the transponder).

4.2.4.1 Reflection of electromagnetic waves

An *electromagnetic wave* emitted into the surrounding space by an antenna encounters various objects. Part of the high frequency energy that reaches the object is absorbed by the object and converted into heat; the rest is scattered in many directions with varying intensity.

A small part of the reflected energy finds its way back to the transmitter antenna. *Radar technology* uses this reflection to measure the distance and position of distant objects (Figure 4.60).

In RFID systems the reflection of electromagnetic waves (*backscatter system, modulated radar cross-section*) is used for the transmission of data from a transponder to a reader. Because the *reflective properties* of objects generally increase with increasing frequency, these systems are used mainly in the frequency ranges of 868 MHz (Europe), 915 MHz (USA), 2.45 GHz and above.

Let us now consider the relationships in an RFID system. The antenna of a reader emits an electromagnetic wave in all directions of space at the transmission power P_{EIRP}. The *radiation density* S that reaches the location of the transponder can easily be calculated using equation (4.61). The transponder's antenna reflects a power P_S that is proportional to the power density S and the so-called *radar cross-section* σ is:

$$P_S = \sigma \cdot S \tag{4.66}$$

The reflected electromagnetic wave also propagates into space spherically from the point of reflection. Thus the radiation power of the reflected wave also decreases in proportion to the square of the distance (r^2) from the radiation source (i.e. the reflection). The following power density finally returns to the reader's antenna:

$$S_{Back} = \frac{P_S}{4\pi r^2} = S \cdot \frac{\sigma}{4\pi r^2} = \frac{P_{EIRP}}{4\pi r^2} \cdot \frac{\sigma}{4\pi r^2} = \frac{P_{EIRP} \cdot \sigma}{(4\pi)^2 \cdot r^4} \tag{4.67}$$

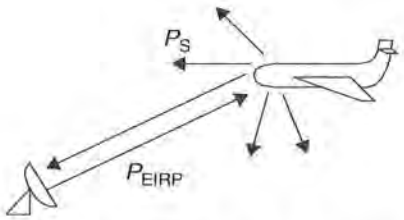


Figure 4.60 Reflection off a distant object is also used in radar technology

The radar cross-section σ (RCS, scatter aperture) is a measure of how well an object reflects electromagnetic waves. The radar cross-section depends upon a range of parameters, such as object size, shape, material, surface structure, but also wavelength and polarisation.

The radar cross-section can only be calculated precisely for simple surfaces such as spheres, flat surfaces and the like (for example see Baur, 1985). The material also has a significant influence. For example, *metal surfaces* reflect much better than plastic or composite materials. Because the dependence of the radar cross-section σ on wavelength plays such an important role, objects are divided into three categories:

- Rayleigh range: the wavelength is large compared with the object dimensions. For objects smaller than around half the wavelength, σ exhibits a λ^{-4} dependency and so the reflective properties of objects smaller than 0.1 λ can be completely disregarded in practice.
- Resonance range: the wavelength is comparable with the object dimensions. Varying the wavelength causes σ to fluctuate by a few decibels around the geometric value. Objects with sharp resonance, such as sharp edges, slits and points may, at certain wavelengths, exhibit resonance step-up of σ. Under certain circumstances this is particularly true for antennas that are being irradiated at their resonant wavelengths (resonant frequency).
- Optical range: the wavelength is small compared to the object dimensions. In this case, only the geometry and position (angle of incidence of the electromagnetic wave) of the object influence the radar cross-section.

Backscatter RFID systems employ antennas with different construction formats as reflection areas. Reflections at transponders therefore occur exclusively in the resonance range. In order to understand and make calculations about these systems we need to know the radar cross-section σ of a resonant antenna. A detailed introduction to the calculation of the radar cross-section can therefore be found in the following sections.

It also follows from equation (4.67) that the power reflected back from the transponder is proportional to the fourth root of the power transmitted by the reader (Figure 4.61). In other words: if we wish to double the power density S of the reflected

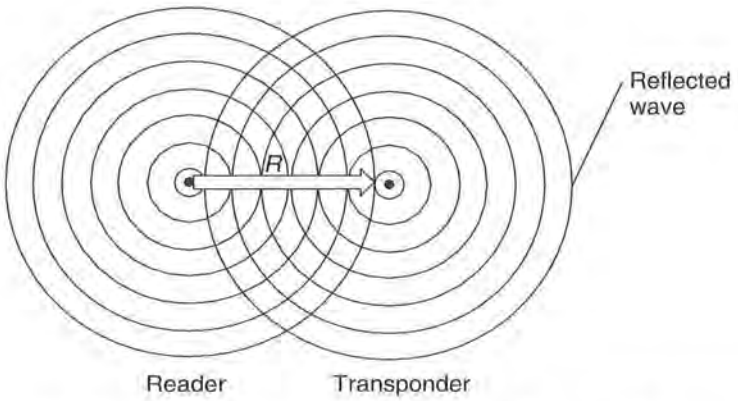


Figure 4.61 Propagation of waves emitted and reflected at the transponder

signal from the transponder that arrives at the reader, then, all other things being equal, the transmission power must be multiplied by sixteen!

4.2.5 Antennas

The creation of electromagnetic waves has already been described in detail in the previous section (see also Sections 4.1.6 and 4.2.1). The laws of physics tell us that the radiation of electromagnetic waves can be observed in all conductors that carry voltage and/or current. In contrast to these effects, which tend to be parasitic, an *antenna* is a component in which the radiation or reception of electromagnetic waves has been to a large degree optimised for certain frequency ranges by the fine-tuning of design properties. In this connection, the behaviour of an antenna can be precisely predicted and is exactly defined mathematically.

4.2.5.1 Gain and directional effect

Section 4.2.2 demonstrated how the power P_{EIRP} emitted from an *isotropic emitter* at a distance r is distributed in a fully uniform manner over a spherical surface area. If we integrate the power density S of the electromagnetic wave over the entire surface area of the sphere the result we obtain is, once again, the power P_{EIRP} emitted by the isotropic emitter.

$$P_{\mathrm{EIRP}} = \int_{A_{\mathrm{sphere}}} S \cdot \mathrm{d}A \tag{4.68}$$

However, a real antenna, for example a dipole, does not radiate the supplied power uniformly in all directions. For example, no power at all is radiated by a *dipole antenna* in the axial direction in relation to the antenna.

Equation (4.68) applies for all types of antennas. If the antenna emits the supplied power with varying intensity in different directions, then equation (4.68) can only be fulfilled if the radiation density S is greater in the preferred direction of the antenna than would be the case for an isotropic emitter. Figure 4.62 shows the *radiation pattern* of a dipole antenna in comparison to that of an isotropic emitter. The length of the vector $G(\Theta)$ indicates the relative radiation density in the direction of the vector. In the *main radiation direction* (G_{i}) the radiation density can be calculated as follows:

$$S = \frac{P_{\mathrm{i}} \cdot G_{\mathrm{i}}}{4\pi \cdot r^2} \tag{4.69}$$

P_{i} is the power supplied to the antenna. G_{i} is termed the gain of the antenna and indicates the factor by which the radiation density S is greater than that of an isotropic emitter at the same transmission power.

An important radio technology term in this connection is the *EIRP* (effective isotropic radiated power).

$$P_{\mathrm{EIRP}} = P_{\mathrm{i}} \cdot G_{\mathrm{i}} \tag{4.70}$$

This figure can often be found in radio licensing regulations (e.g. Section 5.2.4) and indicates the transmission power at which an isotropic emitter (i.e. $G_{\mathrm{i}} = 1$) would

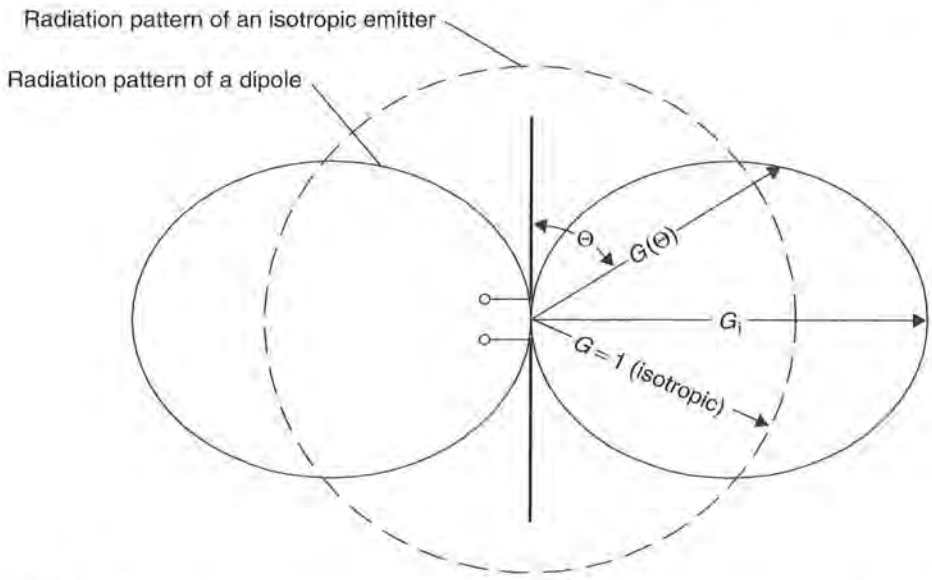


Figure 4.62 Radiation pattern of a dipole antenna in comparison to the radiation pattern of an isotropic emitter

have to be supplied in order to generate a defined radiation power at distance r. An antenna with a gain G_i may therefore only be supplied with a transmission power P_1 that is lower by this factor so that the specified limit value is not exceeded:

$$P_1 = \frac{P_{EIRP}}{G_i} \tag{4.71}$$

4.2.5.2 EIRP and ERP

In addition to power figures in EIRP we frequently come across the power figure ERP (equivalent radiated power) in radio regulations and technical literature. The ERP is also a reference power figure. However, in contrast to the EIRP, ERP relates to a dipole antenna rather than a spherical emitter. An ERP power figure thus expresses the transmission power at which a dipole antenna must be supplied in order to generate a defined emitted power at a distance of r. Since the gain of the dipole antenna

Table 4.7 In order to emit a constant EIRP in the main radiation direction less transmission power must be supplied to the antenna as the antenna gain G increases

EIRP = 4 W	**Power P_1 fed to the antenna**
Isotropic emitter $G_i = 1$	4 W
Dipole antenna	2.44 W
Antenna $G_i = 3$	1.33 W

($G_i = 1.64$) in relation to an isotropic emitter is known, it is easy to convert between the two figures:

$$P_{EIRP} = P_{ERP} \cdot 1.64 \tag{4.72}$$

4.2.5.3 Input impedance

A particularly important property of the antenna is the complex *input impedance* Z_A. This is made up of a complex resistance X_A, a loss resistance R_V and the so-called *radiation resistance* R_r:

$$Z_A = R_r + R_V + jX_A \tag{4.73}$$

The loss resistance R_V is an effective resistance and describes all losses resulting from the ohmic resistance of all current-carrying line sections of the antenna (Figure 4.63). The power converted by this resistance is converted into heat.

The radiation resistance R_r also takes the units of an effective resistance but the power converted within it corresponds with the power emitted from the antenna into space in the form of electromagnetic waves.

At the operating frequency (i.e. the resonant frequency of the antenna) the complex resistance X_A of the antenna tends towards zero. For a loss-free antenna (i.e. $R_V = 0$):

$$Z_A(f_{RES}) = R_r \tag{4.74}$$

The input impedance of an ideal antenna in the resonant case is thus a real resistance with the value of the radiation resistance R_r. For a $\lambda/2$ dipole the radiation resistance $R_r = 73\,\Omega$.

4.2.5.4 Effective aperture and scatter aperture

The maximum received power that can be drawn from an antenna, given optimal alignment and correct polarisation, is proportional to the power density S of an incoming

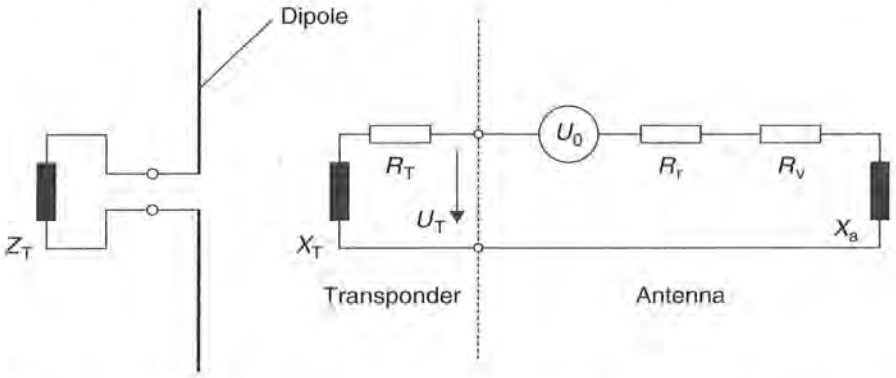


Figure 4.63 Equivalent circuit of an antenna with a connected transponder

plane wave and a proportionality factor. The proportionality factor has the dimension of an area and is thus called the *effective aperture* A_e. The following applies:

$$P_e = A_e \cdot S \tag{4.75}$$

We can envisage A_e as an area at right angles to the direction of propagation, through which, at a given radiation density S, the power P_e passes (Meinke and Gundlach, 1992). The power that passes through the effective aperture is absorbed and transferred to the connected terminating impedance Z_T (Figure 4.64).

In addition to the effective aperture A_e, an antenna also possesses a *scatter aperture* $\sigma = A_s$ at which the electromagnetic waves are reflected.

In order to improve our understanding of this, let us once again consider Figure 4.63. When an electromagnetic field with radiation density S is received a voltage U_0 is induced in the antenna, which represents the cause of a current I through the antenna impedance Z_A and the terminating impedance Z_T. The current I is found from the quotient of the induced voltage U_0 and the series connection of the individual impedances (Kraus, 1988):

$$\mathrm{I} = \frac{\mathrm{U}_0}{\mathrm{Z}_\mathrm{T} + \mathrm{Z}_\mathrm{A}} = \frac{\mathrm{U}_0}{\sqrt{(\mathrm{R}_\mathrm{r} + \mathrm{R}_\mathrm{V} + \mathrm{R}_\mathrm{T})^2 + (\mathrm{X}_\mathrm{A} + \mathrm{X}_\mathrm{T})^2}} \tag{4.76}$$

Furthermore for the received power P_e transferred to Z_T:

$$P_e = I^2 \cdot R_T \tag{4.77}$$

Let us now substitute I^2 in equation (4.77) for the expression in equation (4.76), obtaining:

$$P_e = \frac{U_0^2 \cdot R_T}{(R_r + R_V + R_T)^2 + (X_A + X_T)^2} \tag{4.78}$$

According to equation (4.75) the effective aperture A_e is the quotient of the received power P_e and the radiation density S. This finally yields:

$$A_e = \frac{P_e}{S} = \frac{U_0^2 \cdot R_T}{S \cdot [(R_r + R_V + R_T)^2 + (X_A + X_T)^2]} \tag{4.79}$$

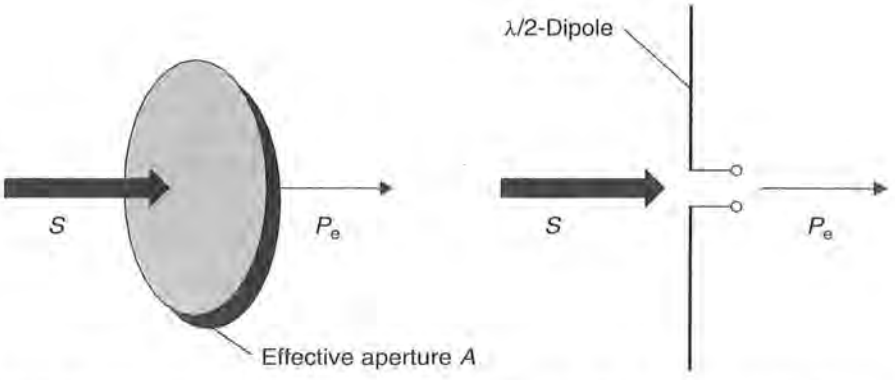


Figure 4.64 Relationship between the radiation density S and the received power P of an antenna

If the antenna is operated using power matching, i.e. $R_T = R_V$ and $X_T = -X_A$, then the following simplification can be used:

$$A_e = \frac{U_0^2}{4SR_r} \tag{4.80}$$

As can be seen from Figure 4.63 the current I also flows through the radiation resistance R_r of the antenna. The converted power P_S is emitted from the antenna and it makes no difference whether the current I was caused by an incoming electromagnetic field or by supply from a transmitter. The power P_S emitted from the antenna, i.e. the reflected power in the received case, can be calculated from:

$$P_S = I^2 \cdot R_r \tag{4.81}$$

Like the derivation for equation (4.79), for the scatter aperture A_s we find:

$$\sigma = A_S = \frac{P_S}{S} = \frac{I^2 \cdot R_r}{S} = \frac{U_0^2 \cdot R_r}{S \cdot [(R_r + R_V + R_T)^2 + (X_A + X_T)^2]} \tag{4.82}$$

If the antenna is again operated using power matching and is also loss-free, i.e. $R_V = 0$, $R_T = R_r$ and $X_T = -X_A$, then as a simplification:

$$\sigma = A_S = \frac{U_0^2}{4SR_r} \tag{4.83}$$

Therefore, in the case of the power matched antenna $\sigma = A_s = A_e$. This means that only half of the total power drawn from the electromagnetic field is supplied to the terminating resistor R_T; the other half is reflected back into space by the antenna.

The behaviour of the scatter aperture A_s at different values of the terminating impedance Z_T is interesting. Of particular significance for RFID technology is the limit case $Z_T = 0$. This represents a short-circuit at the terminals of the antennas. From equation (4.82) this is found to be:

$$\sigma_{max} = A_{S\text{-}max} = \frac{U_0^2}{SR_r} = 4A_e|_{Z_T=0} \tag{4.84}$$

The opposite limit case consists of the connection of an infinitely high-ohmic terminating resistor to the antenna, i.e. $Z_T \to \infty$. From equation (4.82) it is easy to see that the scatter aperture A_s, just like the current I, tends towards zero.

$$\sigma_{min} = A_{S\text{-}min} = 0|_{Z_T \to \infty} \tag{4.85}$$

The scatter aperture can thus take on any desired value in the range 0–4 A_e at various values of the terminating impedance Z_T (Figure 4.65). This property of antennas is utilised for the data transmission from transponder to reader in backscatter RFID systems (see Section 4.2.6.6).

Equation (4.82) shows only the relationship between the scatter aperture A_S and the individual resistors of the equivalent circuit from Figure 4.63. However, if we are

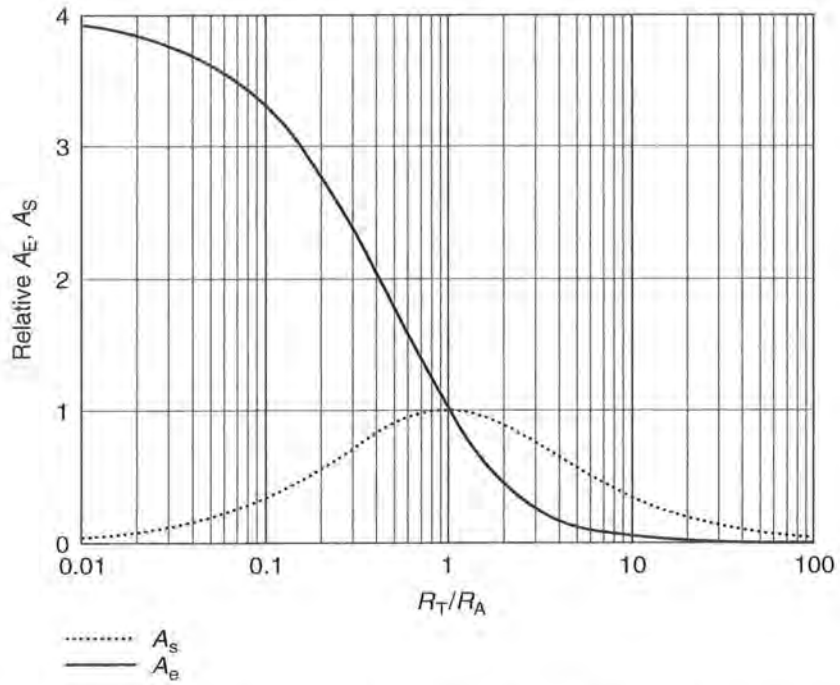


Figure 4.65 Graph of the relative effective aperture A_e and the relative scatter aperture σ in relation to the ratio of the resistances R_A and R_r. Where $R_T/R_A = 1$ the antenna is operated using power matching ($R_T = R_r$). The case $R_T/R_A = 0$ represents a short-circuit at the terminals of the antenna

to calculate the reflected power P_S of an antenna (see Section 4.2.4.1) we need the absolute value for A_S. The *effective aperture* A_e of an antenna is proportional to its gain G (Kraus, 1988; Meinke and Gundlach, 1992). Since the gain is known for most antenna designs, the effective aperture A_e, and thus also the scatter aperture A_S, is simple to calculate for the case of matching ($Z_A = Z_T$). The following is true[5]:

$$\sigma = A_e = \frac{\lambda_0^2}{4\pi} \cdot G \tag{4.86}$$

From equation (4.75) it thus follows that:

$$P_e = A_e \cdot S = \frac{\lambda_0^2}{4\pi} \cdot G \cdot S \tag{4.87}$$

4.2.5.5 Effective length

As we have seen, a voltage U_0 is induced in the antenna by an electromagnetic field. The voltage U_0 is proportional to the electric field strength E of the incoming wave.

[5] The derivation of this relationship is not important for the understanding of RFID systems, but can be found in Kraus (1988, chapter 2–22) if required.

The proportionality factor has the dimension of a length and is therefore called the *effective length* l_0 (also *effective height* h) (Meinke and Gundlach, 1992). The following is true:

$$U_0 = l_0 \cdot E = l_0 \cdot \sqrt{S \cdot Z_F} \tag{4.88}$$

For the case of the matched antenna (i.e. $R_r = R_T$) the effective length can be calculated from the effective aperture A_e (Kraus, 1988):

$$l_0 = 2\sqrt{\frac{A_e \cdot R_r}{Z_F}} \tag{4.89}$$

If we substitute the expression in equation (4.86) for A_e, then the effective length of a matched antenna can be calculated from the gain G, which is normally known (or easy to find by measuring):

$$l_0 = \lambda_0\sqrt{\frac{G \cdot R_r}{\pi \cdot Z_F}} \tag{4.90}$$

4.2.5.6 Dipole antennas

In its simplest form the *dipole antenna* consists solely of a straight piece of line (e.g. a copper wire) of a defined length (Figure 4.66). By suitable shaping the characteristic properties, in particular the *radiation resistance* and bandwidth, can be influenced.

A simple, extended *half-wave dipole* ($\lambda/2$ dipole) consists of a piece of line of length $l = \lambda/2$, which is interrupted half way along. The dipole is supplied at this break-point (Figure 4.67).

The parallel connection of two $\lambda/2$ pieces of line a small distance apart ($d < 0.05\lambda$) creates the *2-wire folded dipole*. This has around four times the radiation resistance of

Figure 4.66 915 MHz transponder with a simple, extended dipole antenna. The transponder can be seen half way along (reproduced by permission of Trolleyscan, South Africa)

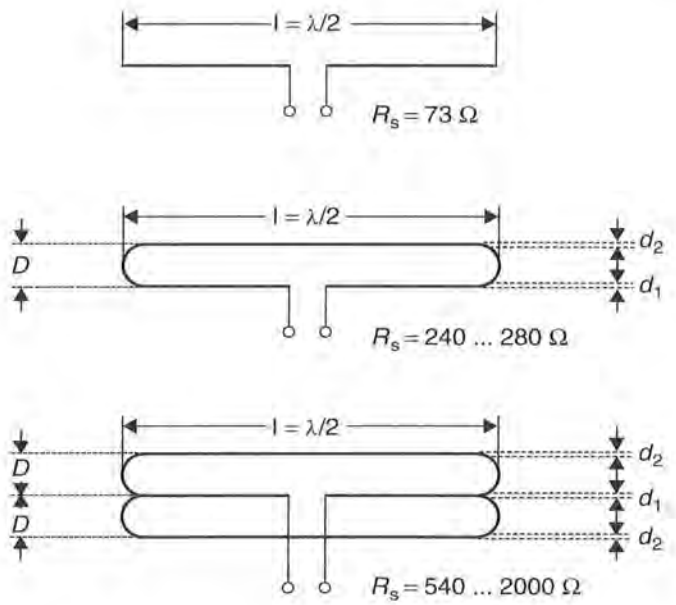


Figure 4.67 Different dipole antenna designs — from top to bottom: simple extended dipole, 2-wire folded dipole, 3-wire folded dipole

the single $\lambda/2$ dipole ($R_r = 240–280\,\Omega$). According to Rothammel (1981) the following relationship applies:

$$R_r = 73.2\,\Omega \cdot \left(\frac{\lg\left(\frac{4D^2}{d_1 \cdot d_2}\right)}{\lg\left(\frac{2D}{d_2}\right)} \right)^2 \tag{4.91}$$

A special variant of the loop dipole is the 3-wire folded dipole. The radiation resistance of the 3-wire folded dipole is greatly dependent upon the conductor diameter and the distance between the $\lambda/2$ line sections. In practice, the radiation resistance of the 3-wire folded dipole takes on values of 540–2000 Ω. According to Rothammel (1981) the following relationship applies:

$$R_r = 73.2\,\Omega \cdot \left(\frac{\lg\left(\frac{4D^3}{d_1^2 \cdot d_2}\right)}{\lg\left(\frac{D}{d_2}\right)} \right)^2 \tag{4.92}$$

The bandwidth of a dipole can be influenced by the ratio of the diameter of the $\lambda/2$ line section to its length, increasing as the diameter increases. However, the dipole must then be shortened somewhat in order to allow it to resonate at the desired frequency. In practice, the *shortening factor* is around 0.90–0.99. For a more precise calculation

Table 4.8 Electrical properties of the dipole and 2-wire folded dipole

Parameter	Gain G	Effective aperture	Effective length	Apex angle
$\lambda/2$ dipole	1.64	$0.13\,\lambda^2$	$0.32\,\lambda$	78°
$\lambda/2$ 2-wire folded dipole	1.64	$0.13\,\lambda^2$	$0.64\,\lambda$	78°

of this topic, the reader is referred to the antenna literature, e.g. Rothammel (1981), Kraus (1988).

4.2.5.7 Yagi–Uda antenna

The *Yagi–Uda antenna*, named after its inventors, could well be the most important variant of a *directional antenna* in radio technology.

The antenna is an alignment array, made up of a driven emitter and a series of parasitic elements. A typical Yagi-Uda antenna is shown in Figure 4.68. Parasitic dipoles are arranged in front of the driven emitter (usually a dipole or 2-wire folded dipole) in the desired *direction of maximum radiation*. These parasitic dipoles function as *directors*, while a rod, usually a single rod, behind the exciter acts as a *reflector*. To create the directional transmission, the rods acting as directors must be shorter, and the rod acting as a reflector must be longer, than the exciter operating at resonance (Meinke and Gundlach, 1992). Compared to an isotropic emitter, gains of 9 dBi (based upon three elements) to 12 dB (based upon seven elements) can be achieved with a Yagi–Uda antenna. So-called 'long' Yagi antennas (10, 15 or more elements) can even achieve gains of up to 15 dBi in the main radiation direction.

Due to their size, Yagi–Uda antennas are used exclusively as antennas for readers. Like a torch, the Yagi–Uda antenna transmits in only one direction of maximum radiation, at a precisely known apex angle. Interference from adjacent devices or readers to the side can thus be suppressed and tuned out.

Figure 4.68 Typical design of a Yagi–Uda directional antenna (six elements), comprising a driven emitter (second transverse rod from left), a reflector (first transverse rod from left) and four directors (third to sixth transverse rods from left) (reproduced by permission of Trolleyscan, South Africa)

Due to the popularity of the Yagi–Uda antenna both as an antenna for radio and television reception and also in commercial radio technology, there is a huge amount of literature on the operation and construction of this antenna design. Therefore, we will not deal with this antenna in more detail at this point.

4.2.5.8 Patch or microstrip antenna

Patch antennas (also known as *microstrip* or *planar antennas*) can be found in many modern communication devices. For example, they are used in the latest generations of GPS receivers and mobile telephones, which are becoming smaller all the time. Thanks to their special construction format, patch antennas also offer some advantages for RFID systems.

In its simplest form, a patch antenna comprises a printed circuit board (e.g. Teflon or PTFE for higher frequencies) coated (i.e. metallised) on both sides, the underside of which forms a continuous ground (Kraus, 2000). On the top there is a small rectangle, which is supplied via a microstrip feed on one side, feeders through the base plate (see Figure 4.71) or capacitive coupling via an intermediate layer (aperture coupled patch antenna; see Kossel (n.d.), Fries and Kossel (n.d.)). Planar antennas can therefore be manufactured cheaply and with high levels of reproducibility using PCB etching technology (see Figures 4.69 and 4.70).

The length L_p of the patch determines the resonant frequency of the antenna. Under the condition $h_D \leq \lambda$:

$$L_P = \frac{\lambda}{2} - h_D \tag{4.93}$$

Normally the substrate thickness h_D is 1–2% of the wavelength.

The width w_p influences the resonant frequency of the antenna only slightly, but determines the *radiation resistance* R_r of the antenna (Krug, 1985). Where $w_p < \lambda/2$:

$$R_r = \frac{90}{\frac{\varepsilon_r + 1}{2} + (\varepsilon_r - 1)\sqrt{4 + \frac{48 \cdot h_p}{w_p}}} \cdot \left(\frac{\lambda}{w_p}\right)^2 \tag{4.94}$$

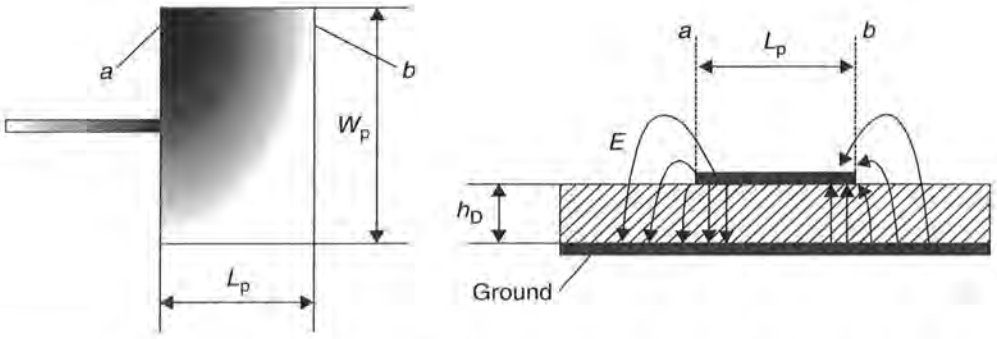


Figure 4.69 Fundamental layout of a patch antenna. The ratio of L_p to h_D is not shown to scale

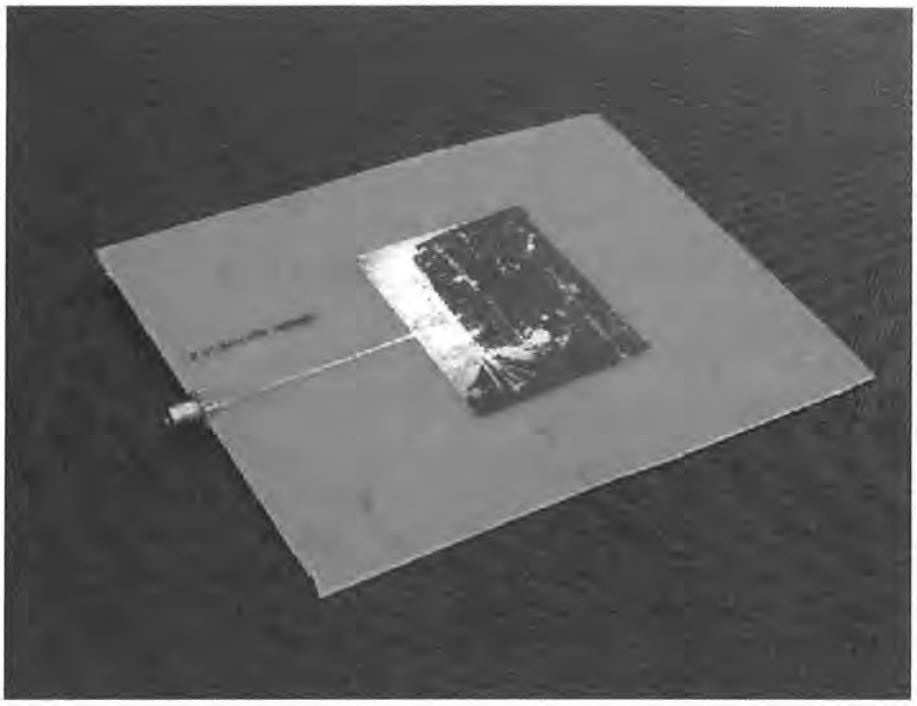


Figure 4.70 Practical layout of a patch antenna for 915 MHz on a printed circuit board made of epoxy resin (reproduced by permission of Trolleyscan, South Africa)

where $w_p > 3\lambda/2$:

$$R_r = \frac{120}{\frac{\varepsilon_r + 1}{2} + (\varepsilon_r - 1)\sqrt{4 + \frac{48 \cdot h_p}{w_p}}} \cdot \frac{\lambda}{w_p} \tag{4.95}$$

If the patch antenna is operated at its resonant frequency the phase difference between the patch edges a and b is precisely 180°. Figure 4.69 shows the path of the electrical field lines. At the entry and exit edges of the patch the field lines run in phase. The patch edges a and b thus behave like two in-phase fed slot antennas. The polarisation of the antenna is linear and parallel to the longitudinal edge L_p. See Figure 4.71.

Due to the type of power supply, patch antennas can also be used with *circular polarisation*. To generate circular polarisation, an emitter element must be supplied with signals with a phase angle of 90° at only two edges that are geometrically offset by 90°.

It is a relatively simple matter to amalgamate patch antennas to form *group antennas* (Figure 4.72). As a result, the gain increases in relation to that of an individual element. The layout shown in the figure comprises in-phase fed emitter elements. The approximately $\lambda/2$ long patch elements are fed via almost non-radiative line sections of around $\lambda/2$ in length connected in series, so that the transverse edges $a-a$ or $b-b$ of the patch element lie precisely wavelength λ apart. Thus the in-phase feed to the

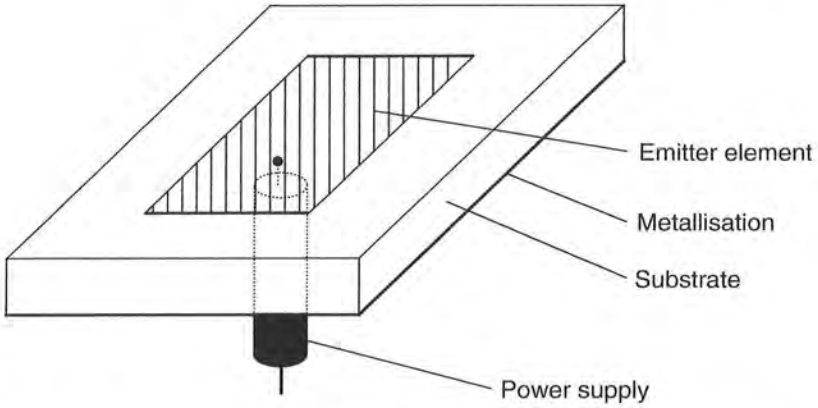


Figure 4.71 Supply of a $\lambda/2$ emitter quad of a patch antenna via the supply line on the reverse

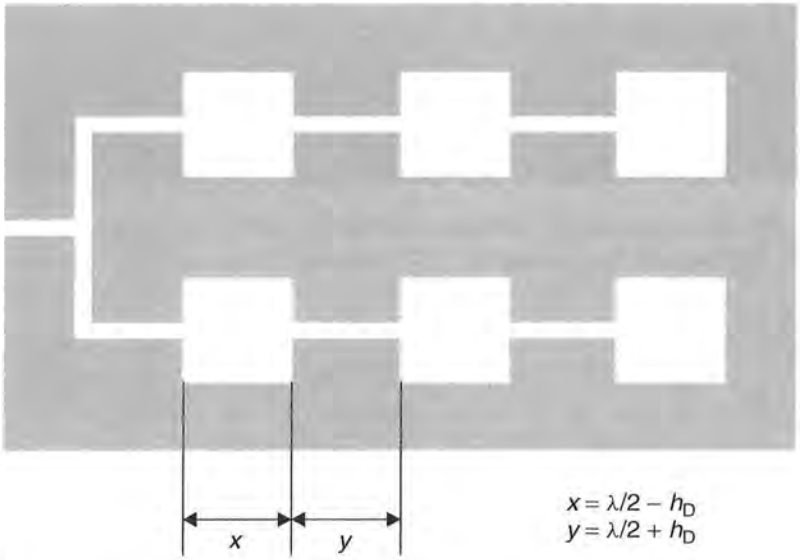


Figure 4.72 The interconnection of patch elements to form a group increases the directional effect and gain of the antenna

individual elements is guaranteed. The arrangement is polarised in the direction of the line sections.

4.2.5.9 Slot antennas

If we cut a strip of length $\lambda/2$ out of the centre of a large metal surface the slot can be used as an emitter (Rothammel, 1981). The width of the slot must be small in relation to its length. The base point of the emitter is located at the mid-point of its longitudinal side (Figure 4.73).

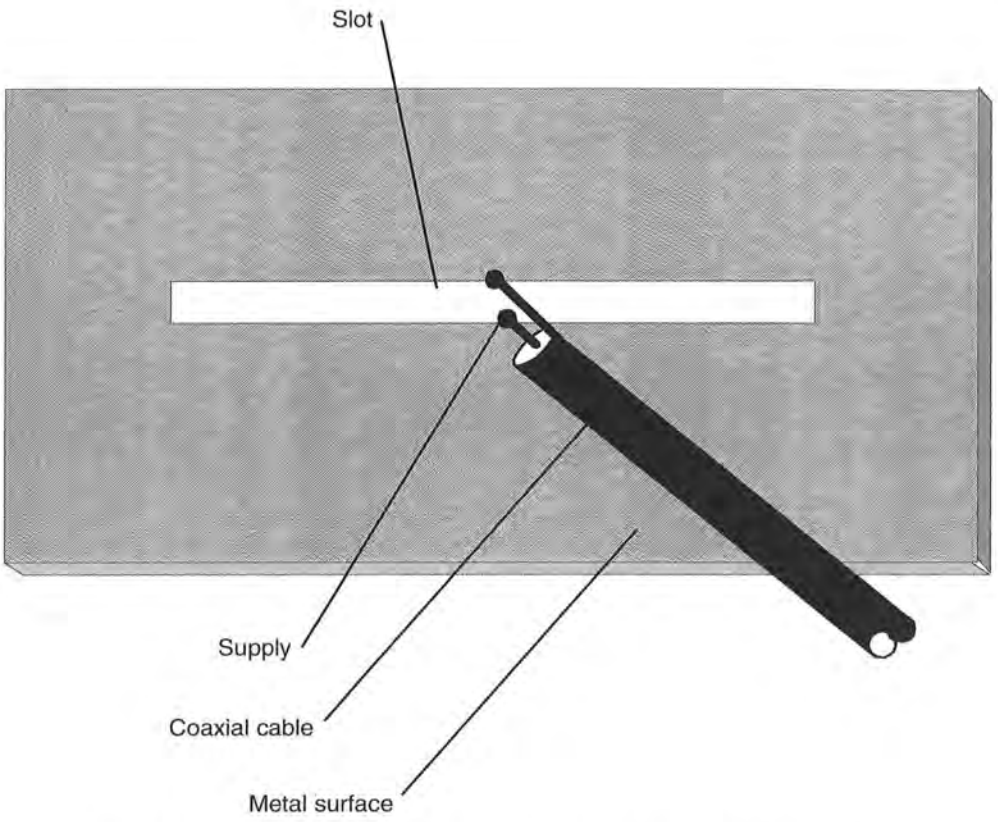


Figure 4.73 Layout of a slot antenna for the UHF and microwave range

4.2.6 Practical operation of microwave transponders

Let us now turn our attention to practical operation when a transponder is located in the *interrogation zone* of a reader. Figure 4.74 shows the simplified model of such a *backscatter system*. The reader emits an electromagnetic wave with the effective radiated power $P_1 \cdot G_1$ into the surrounding space. Of this, a transponder receives power $P_2 = P_e$, proportional to the field strength E, at distance r.

Power P_S is also reflected by the transponder's antenna, of which power P_3 is again received by the reader at distance r.

4.2.6.1 Equivalent circuits of the transponder

In the previous sections we have quoted the simplified equation for the impedance of the transformer $Z_T = R_T + jX_T$ (simplified equivalent circuit). In practice, however, the *input impedance* of a transponder can be represented more clearly in the form of the parallel circuit consisting of a *load resistor* R_L, an *input capacitor* C_2, and possibly a modulation impedance Z_{mod} (see also Section 4.2.6.6) (functional equivalent circuit).

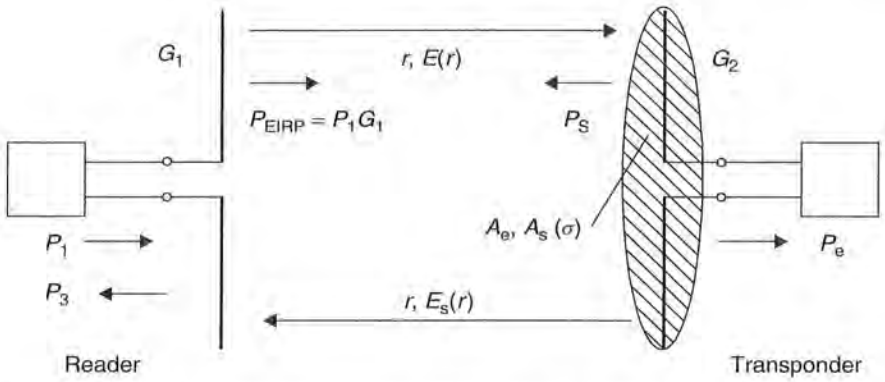


Figure 4.74 Model of a microwave RFID system when a transponder is located in the interrogation zone of a reader. The figure shows the flow of HF power throughout the entire system

It is relatively simple to make the conversion between the components of the two equivalent circuits. For example, the transponder impedance Z_T can be determined from the functional or the simplified equivalent circuit, as desired (Figure 4.75).

$$Z_T = jX_T + R_T = \frac{1}{j\omega C_2 + \frac{1}{R_L} + \frac{1}{Z_{mod}}} \tag{4.96}$$

The individual components R_T and X_T of the simplified equivalent circuit can also be simply determined from the components of the functional equivalent circuit. The following is true:

$$R_T = \mathrm{Re}\left(\frac{1}{j\omega C_2 + \frac{1}{R_L} + \frac{1}{Z_{mod}}} \right) \tag{4.97}$$

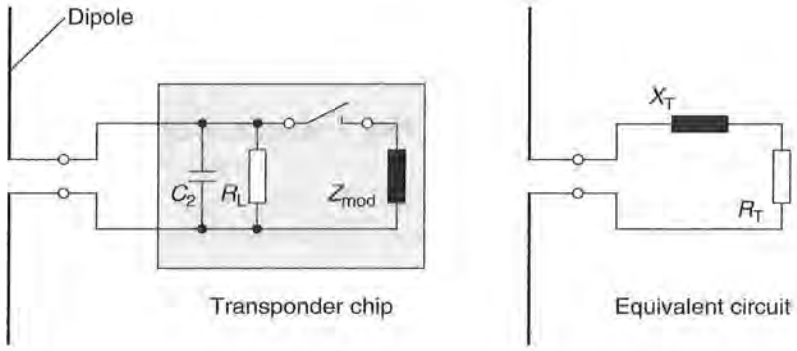


Figure 4.75 Functional equivalent circuit of the main circuit components of a microwave transponder (left) and the simplified equivalent circuit (right)

$$X_T = \mathrm{Im}\left(\frac{1}{j\omega C_2 + \frac{1}{R_L} + \frac{1}{Z_{mod}}}\right) \tag{4.98}$$

4.2.6.2 Power supply of passive transponders

A passive transponder does not have its own *power supply* from an internal voltage source, such as a battery or solar cell. If the transponder is within range of the reader a voltage U_0 is induced in the transponder antenna by the field strength E that occurs at distance r. Part of this voltage is available at the terminals of the antenna as voltage U_T. Only this voltage U_T is rectified and is available to the transponder as supply voltage (rectenna) (Jurianto and Chia, n.d. a, b).

In the case of power matching between the radiation resistor R_r and the input impedance Z_T of the transponder, power $P_2 = P_e$ can be derived from equation (4.87). Figure 4.76 shows the power available in RFID systems at different distances at the

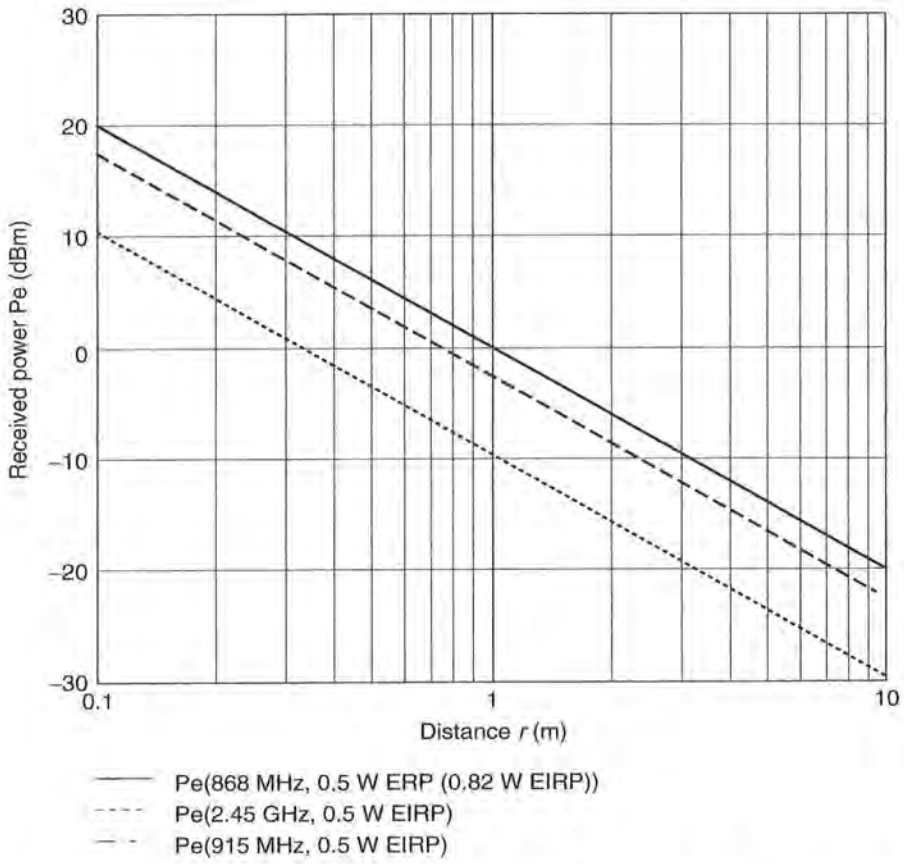


Figure 4.76 The maximum power $P_e(0\,\mathrm{dBm} = 1\,\mathrm{mW})$ available for the operation of the transponder, in the case of power matching at the distance r, using a dipole antenna at the transponder

reader's normal transmission power. In order to use this low power as effectively as possible a *Schottky detector* with *impedance matching* is typically used as a rectifier.

A *Schottky diode* consists of a metal–semiconductor sequence of layers. At the boundary layer there is, as in the p-n junction, a charge-free space-charge zone and a potential barrier that hinders charge transport. The current–voltage characteristic of the metal–semiconductor transition has a diode characteristic. Schottky diodes function as a rectifier at wavelengths below the microwave range since, unlike the pn diode, there are no inertia effects caused by minority carrier injection. Further advantages in comparison to pn diodes are the low voltage drop in the direction of flow and the low noise. A possible layout of a Schottky diode is shown in Figure 4.77 (Agilent Technologies, n.d.).

A Schottky diode can be represented by a linear *equivalent circuit* (Figure 4.77b). C_j represents the parasitic *junction capacitance* of the chip and R_s is the loss resistance in the terminals of the diode. R_j is the *junction resistor* of the diode, which can be calculated as follows (Agilent Technologies, n.d.):

$$R_j = \frac{8.33 \cdot 10^{-5} \cdot n \cdot T}{I_s + I_b} \tag{4.99}$$

where n is the ideality factor, T the temperature in Kelvin, I_s the saturation current and I_b the bias current through the Schottky diode.

By a suitable combination of the p- or n-doped semiconductor with the various metals the properties of the Schottky diode can be varied across a wide range. In RFID transponders primarily p-doped Schottky diodes are used, since these are particularly suitable for detectors with no zero bias in small signal operation, i.e. for the conditions that occur in every transponder (Hewlett Packard, 988).

The circuit of a Schottky detector for voltage rectification is shown in Figure 4.78. Such a Schottky detector has different operating ranges. If it is driven at power above −10 dBm (0.1 mW) the Schottky detector lies in the range of *linear detection* (Hewlett Packard, 986). Here there is *peak value rectification*, as is familiar from the field of power electronics. The following holds:

$$u_{chip} \sim \hat{u}_{in} \Rightarrow u_{chip} \sim \sqrt{P_{in}} \tag{4.100}$$

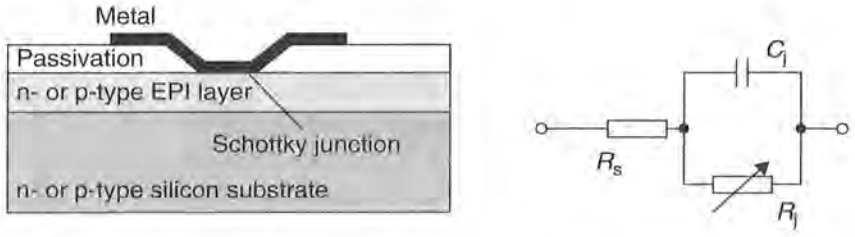
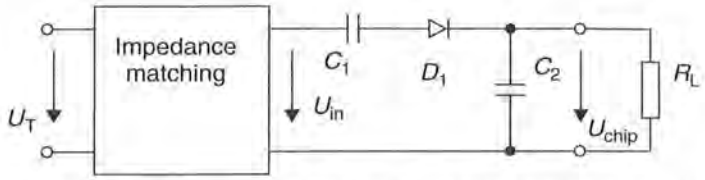
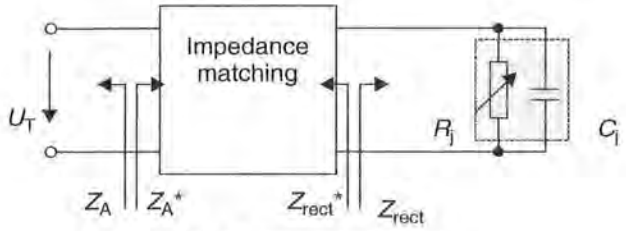


Figure 4.77 A Schottky diode is created by a metal–semiconductor junction. In small signal operation a Schottky diode can be represented by a linear equivalent circuit



(a) Practical circuit: voltage doubler with impedance matching



(b) AC equivalent circuit

Figure 4.78 (a) Circuit of a Schottky detector with impedance transformation for power matching at the voltage source and (b) the HF equivalent circuit of the Schottky detector

In the case of operation at powers below −20 dBm (10 μW) the detector is in the range of square law detection. The following holds (Hewlett Packard, 986):

$$u_{chip} \sim \hat{u}_{in}^2 \Rightarrow u_{chip} \sim P_{in} \tag{4.101}$$

Schottky detectors in RFID transponders operate in the range of *square law detection* at greater distances from the reader, but also in the transition range to linear detection at smaller distances (Figure 4.79).

The relationship between the input power and output voltage of a Schottky detector can be expressed using a Bessel function of the zeroth order (Hewlett Packard, 1088):

$$I0\left(\frac{\Lambda}{n}\sqrt{8R_g \cdot P_{in}}\right) = \left(1 + \frac{I_b}{I_s} + \frac{u_{chip}}{R_L \cdot I_s}\right) \cdot e^{\left[\left(1+\frac{R_g+R_s}{R_L}\right)\cdot\frac{\Lambda\cdot u_{chip}}{n}+\frac{\Lambda\cdot R_s\cdot I_b}{n}\right]} \tag{4.102}$$

Where $\Lambda = q/(k \cdot T)$, q is the elementary charge, k is the Boltzmann constant, T is the temperature of the diode in Kelvin, R_g is the internal resistance of the voltage source (in transponders this is the *radiation resistance* R_r of the antenna), P_{in} is the supplied power, R_L is the connected load resistor (transponder chip) and u_{chip} is the output voltage (supply voltage of the transponder chip).

By numerical iteration using a program such as Mathcad (n.d.) this equation can easily be solved, yielding a diagram $u_{chip}(P_{in})$ (see Figure 4.81). The transition from square law detection to linear law detection at around −20 (10 μW) to −10 dBM (0.1 mW) input power is clearly visible in this figure.

Evaluating equation (4.102), we see that a higher saturation current I_s leads to good sensitivity in the square law detection range. However, in the range that is of interest for RFID transponders, with output voltages u_{chip} of 0.8–3 V, this effect is unfortunately no longer marked.

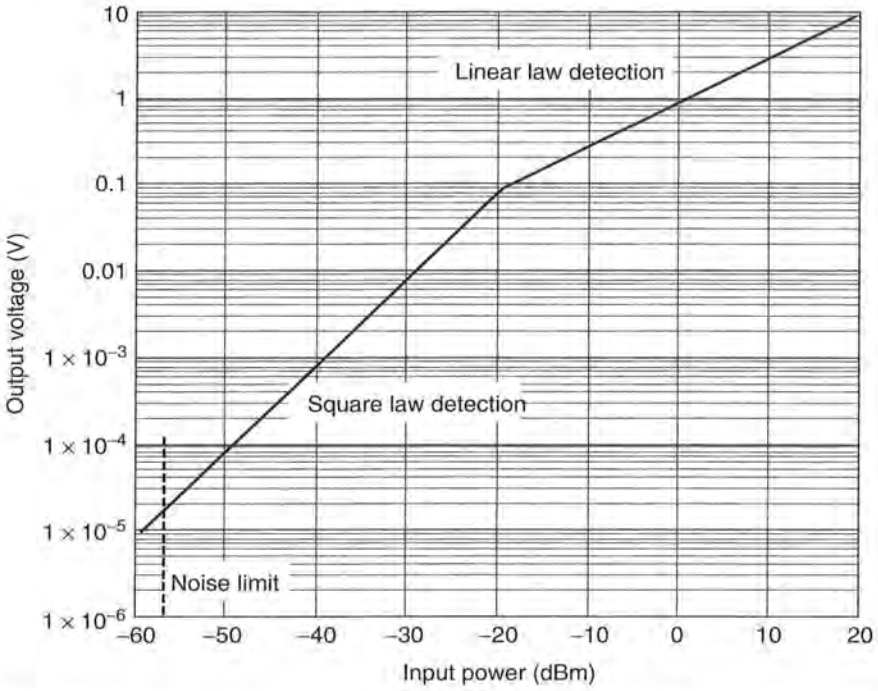


Figure 4.79 When operated at powers below −20 dBm (10 μW) the Schottky diode is in the square law range

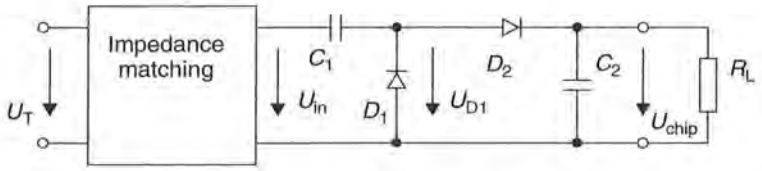


Figure 4.80 Circuit of a Schottky detector in a voltage doubler circuit (villard-rectifier)

In order to further increase the output voltage, *voltage doublers* (Hewlett Packard, 956-4) are used. The circuit of a voltage doubler is shown in Figure 4.80. The output voltage u_{chip} at constant input power P_{in} is almost doubled in comparison to the single Schottky detector (Figure 4.81). The Bessel function (equation (4.102)) can also be used for the calculation of the relationship of P_{in} to u_{chip} in voltage doublers. However, the value used for R_g should be doubled, the value R_L should be halved, and the calculated values for the output voltage u_{chip} should also be doubled (Figure 4.81).

The influence of various operating frequencies on the output voltage is not taken into account in equation (4.102). In practice, however, a frequency-dependent current flows through the parasitic capacitor C_j, which has a detrimental effect upon the efficiency of the Schottky detector. The influence of the junction capacitance on the output voltage can be expressed by a factor M (Hewlett Packard, 1088). The following holds:

$$M = \frac{1}{1 + \omega^2 C_j^2 R_s R_j} \tag{4.103}$$

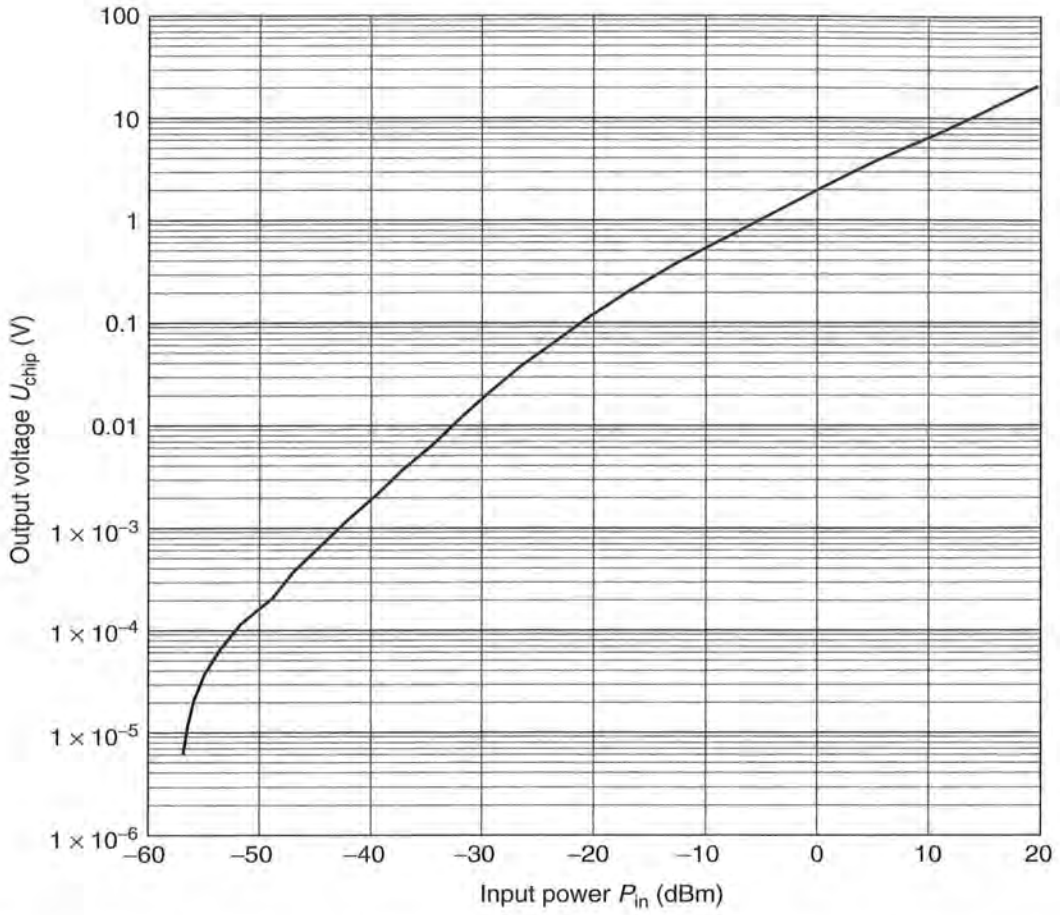


Figure 4.81 Output voltage of a Schottky detector in a voltage doubler circuit. In the input power range −20 to −10 dBm the transition from square law to linear law detection can be clearly seen ($R_L = 500\,k\Omega$, $I_s = 2\,\mu A$, $n = 1.12$)

However, in the range that is of interest for RFID transponders at output voltages u_{chip} 0.8–3 V and the resulting junction resistances R_j in the range <250 Ω (Hewlett Packard, 1088) the influence of the junction capacitance can largely be disregarded (Figure 4.82; see also Figure 4.81).

In order to utilise the received power P_e as effectively as possible, the input impedance Z_{rect} of the Schottky detector would have to represent the complex conjugate of the antenna impedance Z_A (voltage source), i.e. $Z_{rect} = Z_A^*$. If this condition is not fulfilled, then only part of the power is available to the Schottky detector, as a glance at Figure 4.65 makes unmistakably clear.

The HF equivalent circuit of a Schottky detector is shown in Figure 4.78. It is the job of the capacitor C_2 to filter out all HF components of the generated direct voltage and it is therefore dimensioned such that X_{C2} tends towards zero at the transmission frequency of the reader. In this frequency range the diode (or the equivalent circuit of the diode) thus appears to lie directly parallel to the input of the circuit. The load resistor R_L is short-circuited by the capacitor C_2 for the HF voltages and is thus not present in the HF equivalent circuit. R_L, however, determines the current I_b through

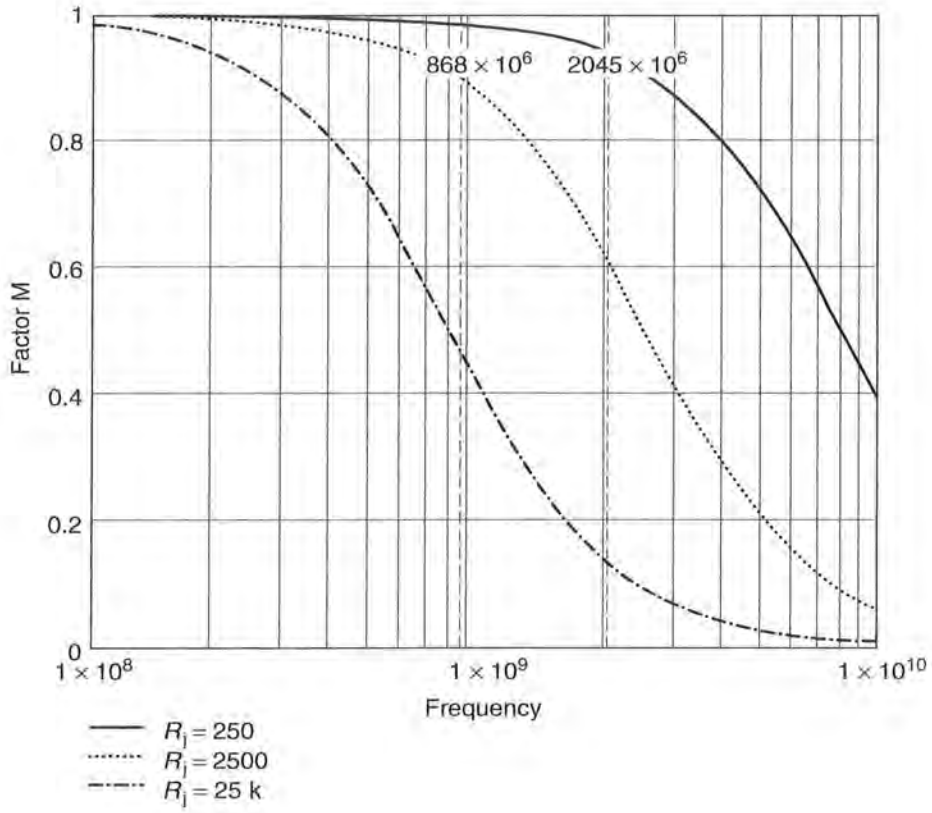


Figure 4.82 The factor M describes the influence of the parasitic junction capacitance C_j upon the output voltage u_{chip} at different frequencies. As the junction resistance R_j falls, the influence of the junction capacitance C_j also declines markedly. Markers at 868 MHz and 2.45 GHz

the Schottky detector and thus also the junction resistance R_j of the Schottky diode. The HF equivalent circuit of a voltage doubler correspondingly consists of the parallel connection of two Schottky diodes.

In order to now achieve the required power matching between the antenna and the Schottky detector, the input impedance Z_{rect} of the Schottky detector must be matched by means of a circuit for the impedance matching at the antenna impedance Z_A. In HF technology, discrete components, i.e. L and C, but also line sections of differing impedances (line transformation), can be used for this.

At ideal matching, the voltage sensitivity $\gamma 2xs$ (in mV/μW) of a Schottky detector can be simply calculated (Figure 4.83; Hewlett Packard, 963, 1089):

$$\gamma^2 = \frac{0.52}{(I_s + I_b) \cdot (1 + \varpi^2 C_j^2 R_s R_j) \cdot \left(1 + \frac{R_j}{R_L}\right)} \tag{4.104}$$

The theoretical maximum of $\gamma 2$ lies at 200 mV/μW (868 MHz) for a Schottky diode of type HSM 2801, and occurs at a total diode current $I_T = I_s + I_b$ of 0.65 μA. The saturation current I_s of the selected Schottky diode is, however, as low as 2 μA, which means that in theory this voltage sensitivity is completely out of reach even

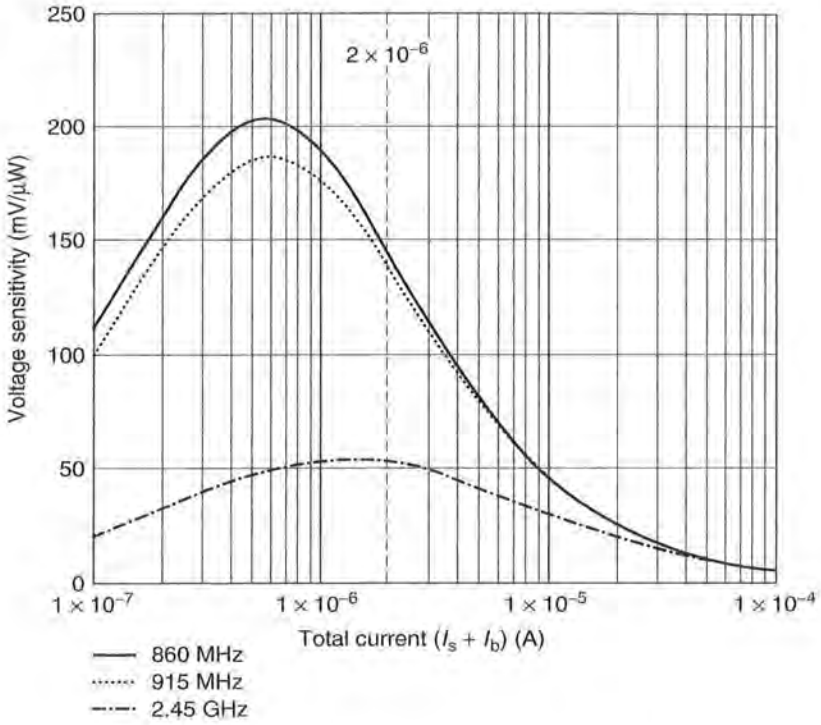


Figure 4.83 Voltage sensitivity $\gamma 2$ of a Schottky detector in relation to the total current $I_T \cdot C_j = 0.25\,\text{pF}$, $R_s = 25\,\Omega$, $R_L = 100\,\text{k}\Omega$

at an operating current $I_b = 0$. Additionally, the junction resistance $R_j = 40\,\text{k}\Omega$ that results at $I_T = 0.65\,\mu\text{A}$ can hardly be approximately transformed without loss using real components at the low-ohmic source impedance of the antenna $Z_A = 73\,\Omega + j0\,\Omega$. Finally, the influence of the parasitic junction capacitance C_j at such high-ohmic junction resistances is clearly visible, and shows itself in a further reduction in the voltage sensitivity, particularly at 2.45 GHz.

In practice, Schottky detectors are operated at currents of 2.5–25 μA, which leads to a significantly lower junction resistance. In practice, values around 50 mV/μW can be assumed for voltage sensitivity (Hewlett Packard, 1089).

Due to the influencing parameters described above it is a great challenge for the designer designing a Schottky detector for an RFID transponder to select a suitable Schottky diode for the operating case in question and to set all operating parameters such that the voltage sensitivity of the Schottky detector is as high as possible.

Let us finally consider another example of the *matching* of a Schottky voltage doubler to a dipole antenna. Based upon two Schottky diodes connected in parallel in the HF equivalent circuit ($L_p = 2\,\text{nH}$, $C_p = 0.08\,\text{pF}$, $R_s = 20\,\Omega$, $C_j = 0.16\,\text{pF}$, $I_T = 3\,\mu\text{A}$, $R_j = 8.6\,\text{k}\Omega$) we obtain an impedance $Z_{rect} = 37 - \text{j}374\,\Omega\,(|Z_{rect}| = 375\,\Omega)$. The Smith diagram in Figure 4.84 shows a possible transformation route, plus the values and sequence of the components used in this example that would be necessary to perform a matching to 72 Ω (dipole in resonance).

It is not always sensible or desirable to perform *impedance matching* between transponder chip and antenna by means of discrete components. Particularly in the

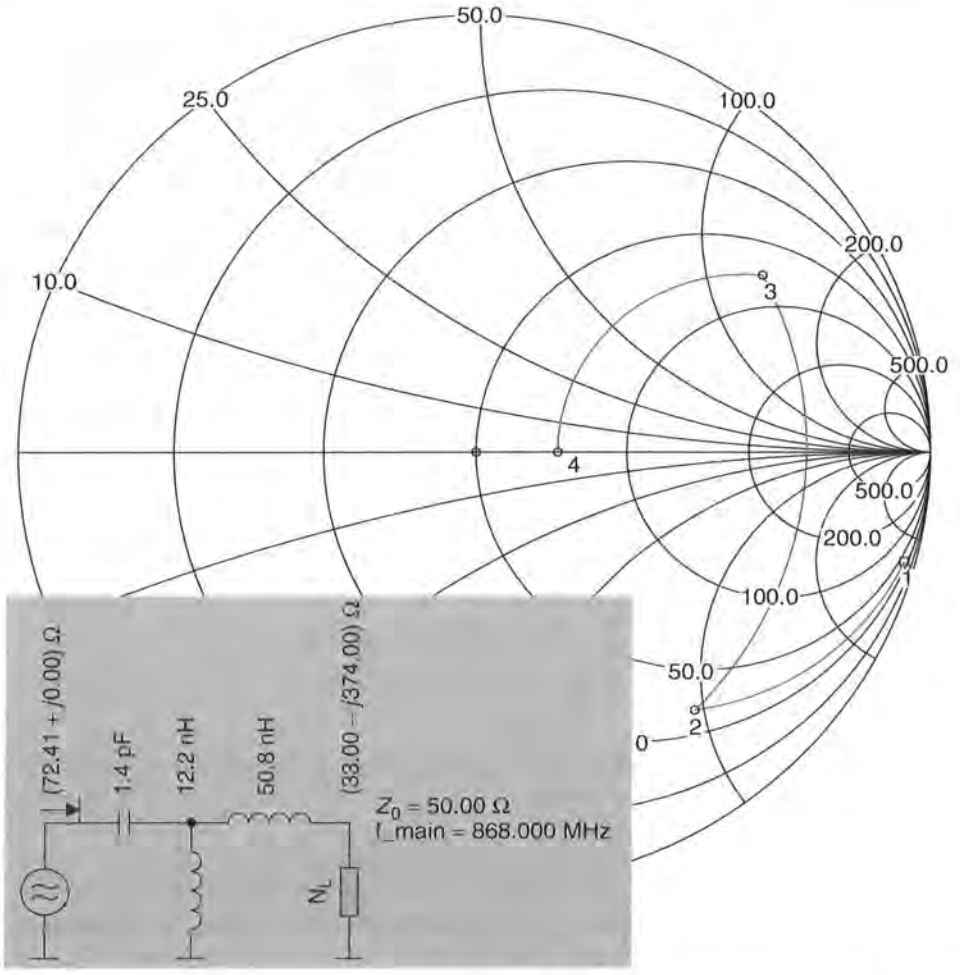


Figure 4.84 Matching of a Schottky detector (point 1) to a dipole antenna (point 4) by means of the series connection of a coil (point 1–2), the parallel connection of a second coil (point 2–3), and finally the series connection of a capacitor (point 3–4)

case of labels, in which the transponder chip is mounted directly upon foil, additional components are avoided where possible.

If the elements of a dipole are shortened or lengthened (i.e. operated at above or below their resonant frequency), then the impedance Z_A of the antenna contains an inductive or capacitive component $X_T \neq 0$. Furthermore, the radiation resistance R_s can be altered by the construction format. By a suitable antenna design it is thus possible to set the input impedance of the antenna to be the complex conjugate of the input impedance of the transponder, i.e. $Z_T = Z_A^*$ (Figure 4.85). The power matching between transponder chip and antenna is thereby only realised by the antenna.

4.2.6.3 Power supply of active transponders

In active transponders the power supply of the semiconductor chip is provided by a *battery*. Regardless of the distance between transponder and reader the voltage is

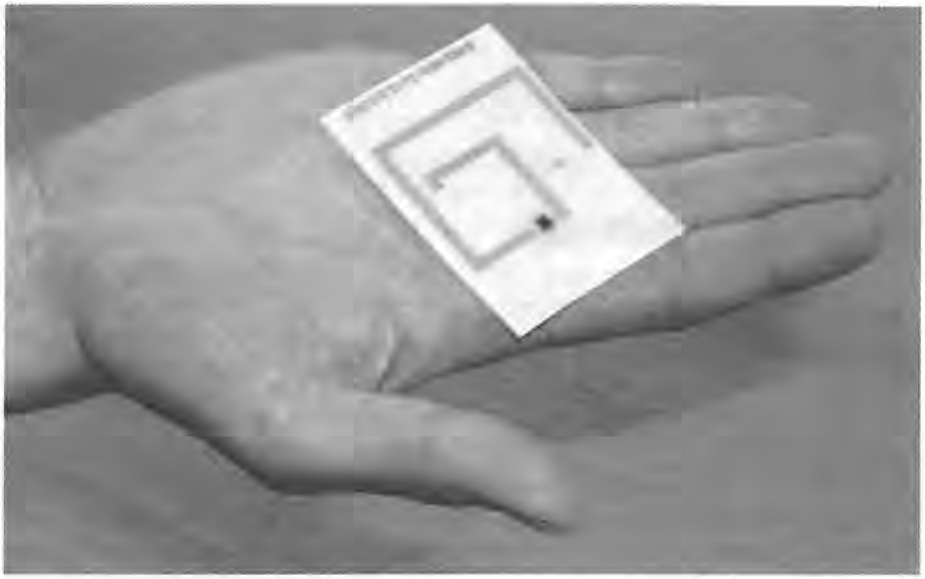


Figure 4.85 By suitable design of the transponder antenna the impedance of the antenna can be designed to be the complex conjugate of the input impedance of the transponder chip (reproduced by permission of Rafsec, Palomar-Konsortium, PALOMAR-Transponder)

always high enough to operate the circuit. The voltage supplied by the antenna is used to activate the transponder by means of a detection circuit. In the absence of external activation the transponder is switched into a power saving mode in order to save the battery from unnecessary discharge.

Depending upon the type of evaluation circuit a much lower received power P_e is needed to activate the transponder than is the case for a comparable passive transponder. Thus the read range is greater compared to a passive transponder. In practice, ranges of over 10 m are normal.

4.2.6.4 Reflection and cancellation

The electromagnetic field emitted by the reader is not only reflected by a transponder, but also by all objects in the vicinity, the spatial dimension of which is greater than the wavelength λ_o of the field (see also Section 4.2.4.1). The reflected fields are superimposed upon the primary field emitted by the reader. This leads alternately to a local damping or even so-called *cancellation* (antiphase superposition) and an amplification (in-phase superposition) of the field at intervals of $\lambda_0/2$ between the individual minima. The simultaneous occurrence of many individual reflections of varying intensity at different distances from the reader leads to a very erratic path of field strength E around the reader, with many local zones of cancellation of the field. Such effects should be expected particularly in an environment containing large metal objects, e.g. in an industrial operation (machines, metal pipes etc.).

We are all familiar with the effect of *reflection* and cancellation in our daily lives. In built-up areas it is not unusual to find that when you stop your car at traffic lights you are in a 'radio gap' (i.e. a local cancellation) and instead of your favourite radio station

all you can hear from the radio is noise. Experience shows that it is generally sufficient to roll the car forward just a short way, thus leaving the area of local cancellation, in order to restore the reception.

In RFID systems these effects are much more disruptive, since a transponder at a local field strength minimum may not have enough power at its disposal for operation. Figure 4.86 shows the results of the measurement of the reader's field strength E at an increasing distance from the transmission antenna when reflections occur in close proximity to the reader.

4.2.6.5 Sensitivity of the transponder

Regardless of the type of power supply of the transponder a minimum *field strength* E is necessary to activate the transponder or supply it with sufficient energy for the operation of the circuit. The minimum field strength is called the *interrogation field strength* $E_{\min}$ and is simple to calculate. Based upon the minimum required HF input power $P_{\text{e-min}}$ of the *Schottky detector* and of the transponder antenna gain G we find:

$$E_{\min} = \sqrt{\frac{4\pi \cdot Z_F \cdot P_{\text{e-min}}}{\lambda_0^2 \cdot G}} \tag{4.105}$$

This is based upon the prerequisite that the *polarisation directions* of the reader and transponder antennas precisely correspond. If the transponder is irradiated with a field that has a different polarisation direction, then $E_{\min}$ increases accordingly.

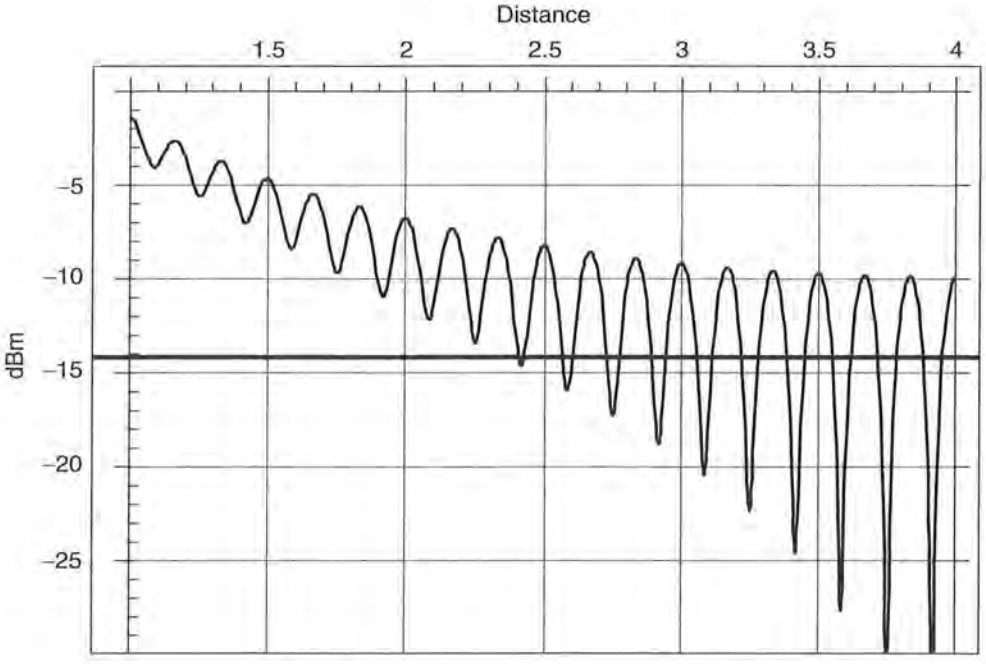


Figure 4.86 The superposition of the field originally emitted with reflections from the environment leads to local cancellations. x axis, distance from reader antenna; y axis, path attenuation in decibels (reproduced by permission of Rafsec, Palomar-Konsortium)

4.2.6.6 Modulated backscatter

As we have already seen, the transponder antenna reflects part of the irradiated power at the *scatter aperture* σ (A_s) of the *transponder antenna*. In this manner, a small part of the power P_1 that was originally emitted by the reader returns to the reader via the transponder as received power P_3.

The dependence of the scatter aperture σ on the relationship between Z_T and Z_A established in Section 4.2.5.4 is used in RFID transponders to send data from the transponder to the reader. To achieve this, the input impedance Z_T of the transponder is altered in time with the data stream to be transmitted by the switching on and off of an additional impedance Z_{mod} in time with the data stream to be transmitted. As a result, the scatter aperture σ, and thus the power P_S reflected by the transponder, is changed in time with the data, i.e. it is modulated. This procedure is therefore also known as *modulated backscatter* or *σ-modulation* (Figure 4.87).

In order to investigate the relationships in a RFID transponder more precisely, let us now refer back to equation (4.82), since this equation expresses the influence of the transponder impedance $Z_T = R_T + X_T$ on the scatter aperture σ. In order to replace U_0^2 by the general properties of the transponder antenna we first substitute equation (4.90) into equation (4.88) and obtain:

$$U_0 = \lambda_0 \cdot \sqrt{\frac{G \cdot R_r}{\pi \cdot Z_F}} \cdot \sqrt{S \cdot Z_F} = \lambda_0 \cdot \sqrt{\frac{G \cdot R_r \cdot S}{\pi}} \tag{4.106}$$

We now replace U_0 in equation (4.82) by the right-hand expression in equation (4.106) and finally obtain (PALOMAR, 18000):

$$\sigma = \frac{\lambda_0{}^2 \cdot R_r^2 \cdot G}{\pi \cdot [(R_r + R_V + R_T)^2 + (X_A + X_T)^2]} \tag{4.107}$$

where G is the gain of the transponder antenna.

However, a drawback of this equation is that it only expresses the value of the scatter aperture σ (PALOMAR, 18000). If, for the clarification of the resulting problems, we imagine a transponder, for which the imaginary component of the input impedance Z_T in unmodulated state takes the value $X_{Toff} = -X_A + \Delta X_{mod}$, but in the modulated state (modulation impedance Z_{mod} connected in parallel) it is $X_{Ton} = -X_A - \Delta X_{mod}$. We further assume that the real component R_T of the input impedance Z_T is not influenced by the modulation. For this special case the imaginary part of the impedance during modulation between the values $(+\Delta X_{mod})^2$ and $(-\Delta X_{mod})^2$ is switched. As can be clearly seen, the value of the scatter aperture σ remains constant. Equation (4.81), on the other hand, shows that the reflected power P_S is proportional to the square of the current I in the antenna. However, since by switching the imaginary part of the impedances between $-\Delta X_{mod}$ and $+\Delta X_{mod}$ we also change the phase θ of the current I, we can conclude that the phase θ of the reflected power P_S also changes to the same degree.

To sum up, therefore, we can say that modulating the input impedance Z_T of the transponder results in the modulation of the value and/or phase of the reflected power P_S and thus also of the scatter aperture σ. P_S and σ should thus not be considered

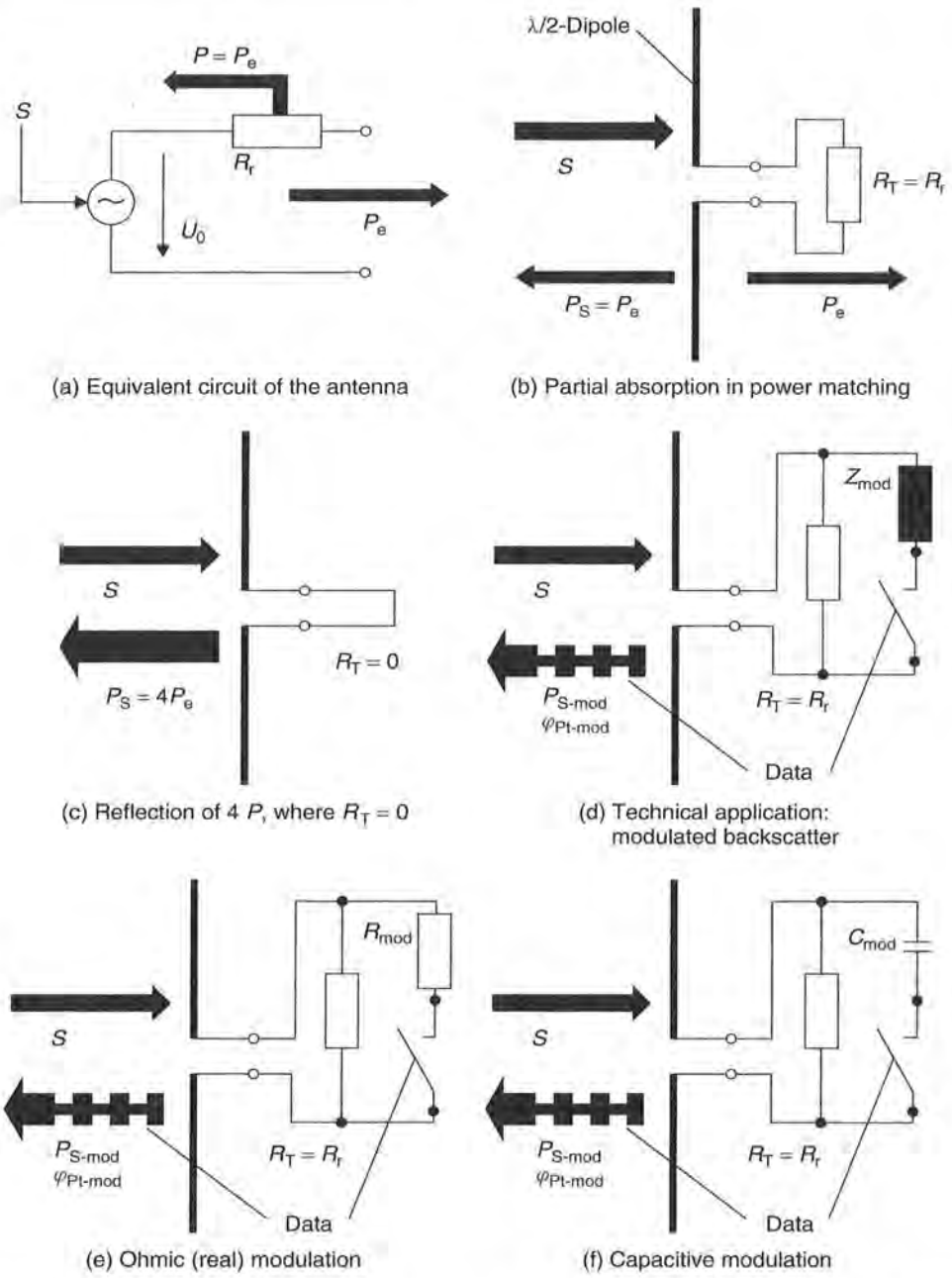


Figure 4.87 Generation of the modulated backscatter by the modulation of the transponder impedance $Z_T(= R_T)$

as real quantities in RFID systems, but as complex quantities. The relative change in value and phase of the scatter aperture σ can be expressed using the following equation (PALOMAR, 18000):

$$\Delta\sigma = \frac{\lambda_0^2 \cdot G \cdot \Delta Z_{\text{mod}}}{4 \cdot \pi \cdot R_r} \tag{4.108}$$

RFID transponders' property of generating mixed phase and amplitude modulation must also be taken into account in the development of readers. Modern readers thus often operate using I/Q demodulators in order to ensure that the transponder's signal can always be demodulated. See Figure 4.88.

4.2.6.7 Read range

Two conditions must be fulfilled for a reader to be able to communicate with a transponder.

First, the transponder must be supplied with sufficient power for its activation. We have already discussed the conditions for this in Section 4.2.6.2. Furthermore, the signal reflected by the transponder must still be sufficiently strong when it reaches the reader for it to be able to be detected without errors. The *sensitivity of a receiver* indicates how great the field strength or the induced voltage U must be at the receiver input for a signal to be received without errors. The level of *noise* that travels through the antenna and the primary stage of the receiver input, interfering with signals that are too weak or suppressing them altogether, is decisive for the sensitivity of a receiver.

In backscatter readers the permanently switched on transmitter, which is required for the activation of the transponder, induces a significant amount of additional noise, thereby drastically reducing the sensitivity of the receiver in the reader. This noise arises largely as a result of *phase noise* of the *oscillator* in the transmitter. As a rule of thumb in practice we can assume that for the transponder to be detected, the transponder's signal may lie no more than 100 dB below the level of the transmitter's

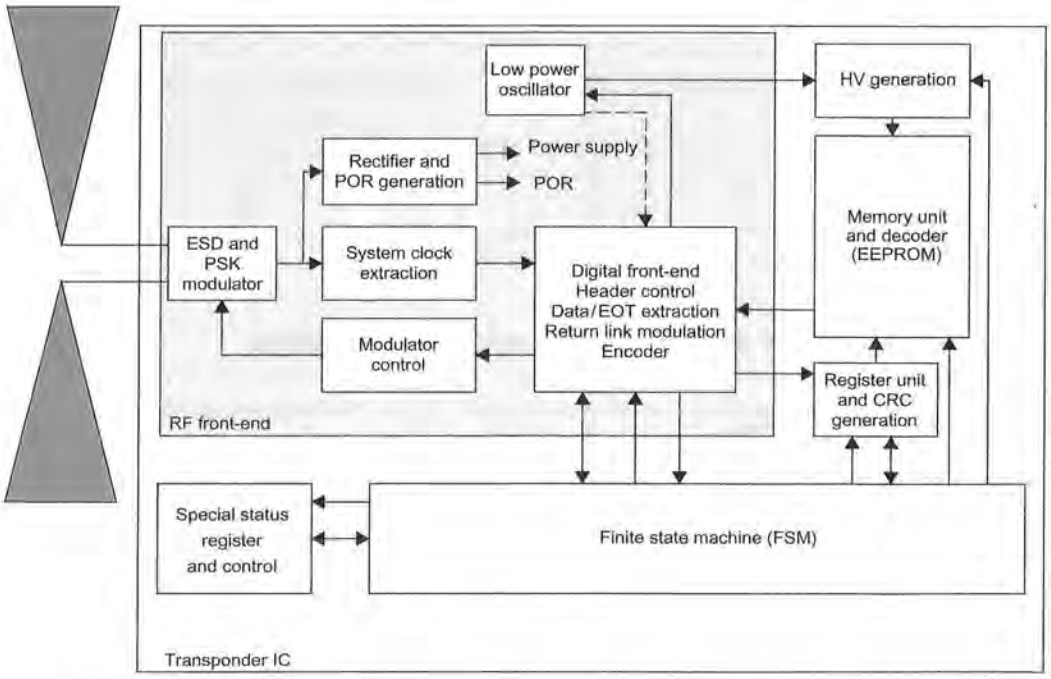


Figure 4.88 Block diagram of a passive UHF transponder (reproduced by permission of Rafsec, Palomar-Konsortium, PALOMAR Transponder)

carrier signal (GTAG, 2001). In order to make a precise prediction regarding the sensitivity of a reader, however, this value must be established for the individual case by measurement.

For the transmission of data the signal reflected by the transponder is modulated. It should be noted in this connection that as part of the process of *modulation* the reflected power P_S is broken down into a reflected 'carrier signal' and two sidebands. In pure ASK modulation with a theoretical modulation index of 100%[6] the two sidebands would each contain 25% of the total reflected power P_S (i.e. P_3 — 6 dB), and at a lower modulation index correspondingly less. Since the information is transmitted exclusively in the *sidebands*, a lower wanted signal should be specified according to the modulation index. The reflected carrier contains no information, but cannot be received by the reader since it is completely masked by a transmission signal of the same frequency, as Figure 4.89 shows.

Let us now consider the magnitude of the power P_3 arriving at the reader that is reflected by the transponder.

As in equation (4.75) the received power P_3 at the receiver of the reader is:

$$P_3 = A_{e\cdot Reader} \cdot S_{Back} \tag{4.109}$$

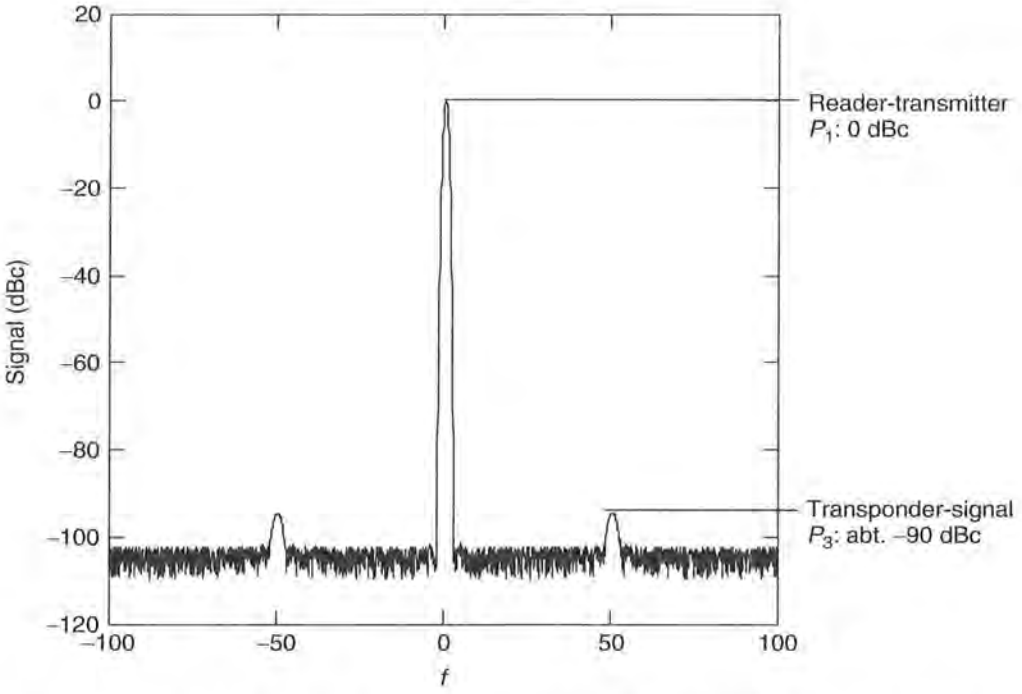


Figure 4.89 Example of the level relationships in a reader. The noise level at the receiver of the reader lies around 100 dB below the signal of the carrier. The modulation sidebands of the transponder can clearly be seen. The reflected carrier signal cannot be seen, since the level of the carrier signal of the reader's transmitter, which is the same frequency, is higher by orders of magnitude

[6] In practice 100% ASK modulation of the reflected signal cannot be achieved, since Z_T would have to have an infinite value in the modulated state.

The radiation density S_{Back} is found from equation (4.67). We thus obtain:

$$P_3 = A_{\text{e-Reader}} \cdot \frac{P_1 \cdot G_{\text{Reader}} \cdot \sigma}{(4\pi)^2 \cdot r^4} \tag{4.110}$$

We now replace $A_{\text{e-Reader}}$ by the expression in equation (4.86), since we have already used the gain G_{Reader} of the reader's antenna in the radar equation:

$$P_3 = \frac{P_1 \cdot G_{\text{Reader}}^2 \cdot \lambda_0^2 \cdot \sigma}{(4\pi)^3 \cdot r^4} \tag{4.111}$$

In the same way we replace $A_s = \sigma$ by equation (4.86), and thus finally obtain:

$$P_3 = \frac{P_1 \cdot G_{\text{Reader}}^2 \cdot \lambda_0^4 \cdot G_T^2}{(4\pi r)^4} \tag{4.112}$$

This equation naturally only applies in the case of power matching between the transponder's antenna and the connected consumer Z_T. In practical operation the scatter aperture σ can take on values between 0 and $4A_e$, as was shown in Section 4.2.5.4. In generalised form the following applies:

$$P_3 = \left. \frac{k \cdot P_1 \cdot G_{\text{Reader}}^2 \cdot \lambda_0^4 \cdot G_T^2}{(4\pi r)^4} \right|_{k=0..4} \tag{4.113}$$

The precise value for k is found from the relationship between the radiation resistance of the antenna R_r and the input impedance Z_T of the transponder chip and can be derived from Figure 4.65.

We solve equation (4.113) with respect to r, obtaining:

$$r = \frac{\lambda_0}{4\pi} \cdot 4 \left. \sqrt{\frac{k \cdot P_1 \cdot G_{\text{Reader}}^2 \cdot G_T^2}{P_3}} \right|_{k=0..4} \tag{4.114}$$

At known sensitivity of the reader's receiver $P_{3\,\text{min}}$ the maximum distance between transponder and reader at which the transponder's signal can just be received by the reader can thus be calculated (Figure 4.90). The fact that P_3 represents the total power reflected by the transponder must be taken into account. The splitting of power P_3 into a carrier signal and the two sidebands (i.e. $P_3 = P_{\text{carrier}} + P_{\text{USB}} + P_{\text{LSB}}$)[7] has not yet been taken into account here. In order to be able to detect a single sideband of the reflected, modulated signal, P_3 must be correspondingly greater.

[7] USB = upper sideband, i.e. the modulation sideband at the higher frequency position; LSB = lower sideband, i.e. the modulation sideband at the lower frequency position.

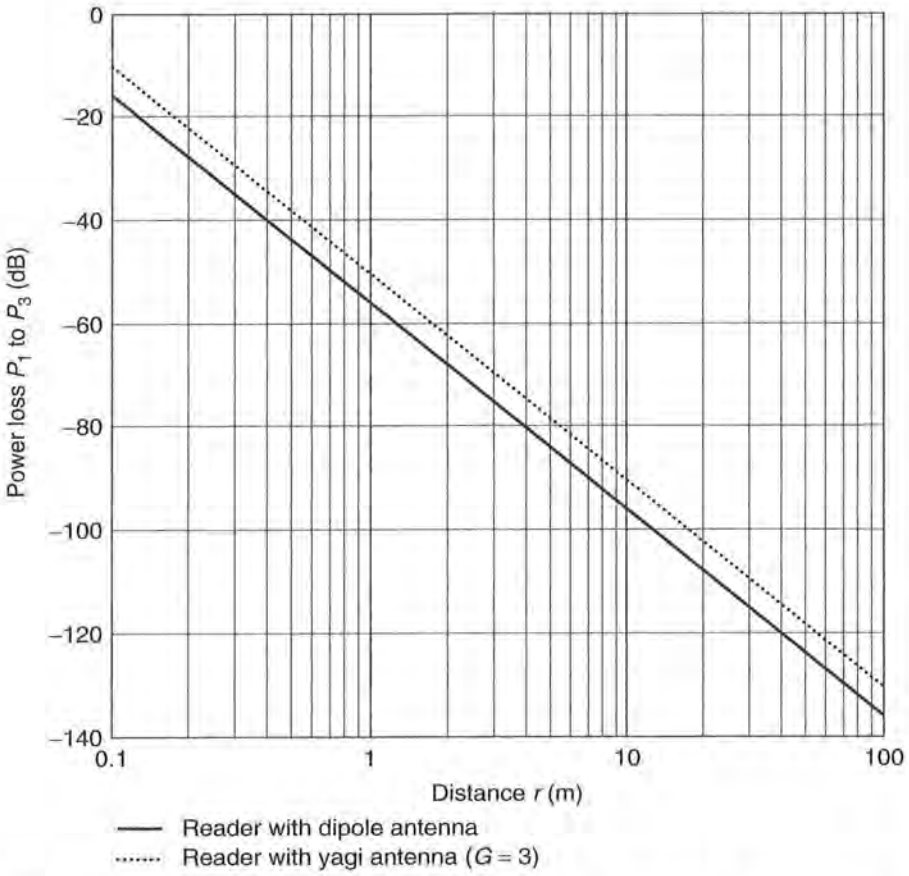


Figure 4.90 Damping of a signal on the way to and from the transponder

4.3 Surface Waves

4.3.1 The creation of a surface wave

If a voltage is applied to the electrodes of a *piezoelectric crystal* such as *quartz* (SiO_2), *lithium niobate* ($LiNbO_3$) or *lithium tantalate* ($LiTaO_3$) mechanical distortions arise in the *crystal lattice* as a result of the *piezoeffect*. This effect is used to generate *surface acoustic waves* on the crystal. To achieve this, electrode structures made of approximately 0.1 μm thick aluminium are applied to the polished surface of a piezoelectric single crystal in the form of an electroacoustic converter. When an alternating voltage is applied to the electroacoustic converter, surface acoustic waves — so-called *Rayleigh waves* — propagate on the surface of the crystal (Meinke and Gundlach, 1992). The deflections in the crystal lattice decrease exponentially as the depth increases.

The majority of the induced acoustic power is thus concentrated within a thin layer with a depth of approximately one wavelength λ on the surface of the crystal. The propagation of a surface acoustic wave on the highly polished surface of a substrate is almost undamped and dispersion free. The propagation speed v is approximately 3000 to 4000 m/s, i.e. only around 1/100 000 of the speed of light c.

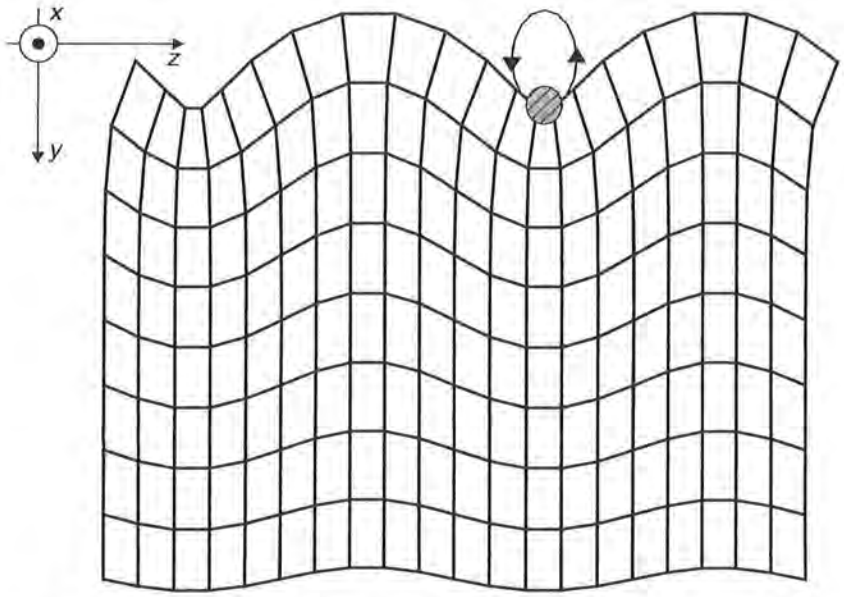


Figure 4.91 The section through a crystal shows the surface distortions of a surface wave propagating in the z-direction (reproduced by permission of Siemens AG, ZT KM, Munich)

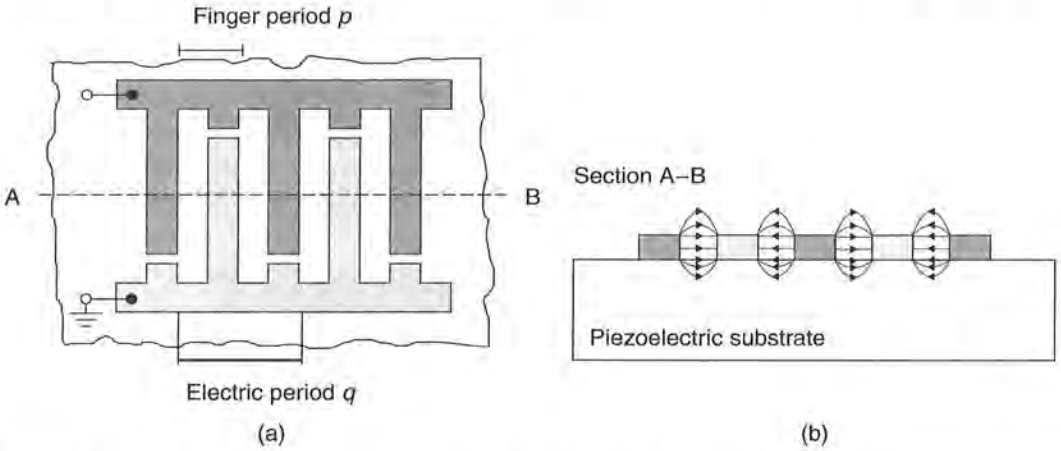


Figure 4.92 Principal structure of an interdigital transducer. Left, arrangement of the finger-shaped electrodes of an interdigital transducer; right, the creation of an electric field between electrodes of different polarity (reproduced by permission of Siemens AG, ZT KM, Munich)

Interdigital electrode structures in the form of interleaved fingers make effective electroacoustic transducers. Each pair of such interleaved fingers (Figure 4.91) form a so-called *interdigital transducer* (Latin *digitus* = finger, *inter* = between). A δ-shaped electrical pulse applied to the busbar of an interdigital transducer results in a mechanical deformation at the surface of the substrate between fingers of different polarity due to the piezoelectric effect. This deformation is proportional to the electrical field and propagates as a surface wave in both directions at velocity v (see Figure 4.92). Conversely, a surface wave entering the converter generates a signal proportional to

the finger structure at the busbar as a result of the piezoelectric effect (Meinke and Gundlach, 1992).

The distance between two fingers of the same polarity is termed the electrical period q of the interdigital transducer. The maximum electroacoustic interaction is obtained at the frequency f_0, the mid-frequency of the transducer. At this frequency the wavelength λ_0 of the surface acoustic wave precisely corresponds with the electrical period q of the interdigital transducer, so that all wave trains are superimposed in-phase and transmission is maximized (Reindl and Mágori; 1995).

$$\frac{v}{f_0} = \lambda_0 = q \tag{4.115}$$

The relationship between the electrical and mechanical power density of a surface wave is described by the material-dependent piezoelectric coupling coefficient k^2. Around k^{-2} overlaps of the transducer are required to convert the entire electrical power applied to the interdigital transducer into the acoustic power of a surface wave.

The bandwidth B of a transducer can be influenced by the length of the converter and is:

$$B = 2f_0/N (N = \text{number of fingers}) \tag{4.116}$$

4.3.2 Reflection of a surface wave

If a surface wave meets a mechanical or electrical discontinuity on the surface a small part of the surface wave is reflected. The transition between free and metallised surface represents such a discontinuity, therefore a periodic arrangement of N reflector strips can be used as a reflector. If the reflector period p (see Figure 4.93) is equal to half a wavelength λ_0, then all reflections are superimposed in-phase. The degree of reflection

Figure 4.93 Scanning electron microscope photograph of several surface wave packets on a piezoelectric crystal. The interdigital transducer itself can be seen to the bottom left of the picture. An electric alternating voltage at the electrodes of the interdigital transducer generates a surface wave in the crystal lattice as a result of the piezoelectric effect. Conversely, an incoming surface wave generates an electric alternating voltage of the same frequency at the electrodes of the transducer (reproduced by permission of Siemens AG, ZT KM, Munich)

thus reaches its maximum value for the associated frequency, the so-called Bragg frequency f_B. See Figure 4.94.

$$f_B = \frac{v}{2p} \tag{4.117}$$

4.3.3 Functional diagram of SAW transponders (Figure 4.95)

A surface wave transponder is created by the combination of an interdigital transducer and several reflectors on a piezoelectric monocrystal, with the two busbars of the interdigital transducer being connected by a (dipole) antenna.

A high-frequency *interrogation pulse* is emitted by the antenna of a reader at periodic intervals. If a surface wave transponder is located in the interrogation zone of the reader part of the power emitted is received by the transponder's antenna and travels to the terminals of the interdigital converter in the form of a high-frequency voltage pulse. The interdigital transducer converts part of this received power into a surface acoustic wave, which propagates in the crystal at right angles to the fingers of the transducer.[8]

Figure 4.94 Geometry of a simple reflector for surface waves (reproduced by permission of Siemens AG, ZT KM, Munich)

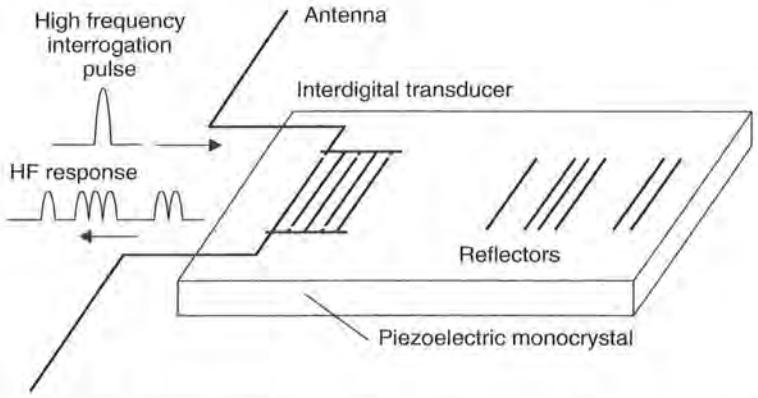


Figure 4.95 Functional diagram of a surface wave transponder (reproduced by permission of Siemens AG, ZT KM, Munich)

[8] To convert as much of the received power as possible into acoustic power, firstly the transmission frequency f_0 of the reader should correspond with the mid-frequency of the interdigital converter. Secondly, however, the number of transducer fingers should be matched to the coupling coefficient k_2.

Reflectors are now applied to the crystal in a characteristic sequence along the propagation path of the surface wave. At each of the reflectors a small part of the surface wave is reflected and runs back along the crystal in the direction of the interdigital transducer. Thus a number of pulses are generated from a single interrogation pulse. In the interdigital transducer the incoming acoustic pulses are converted back into high-frequency voltage pulses and are emitted from the antenna of the transponder as the transponder's response signal. Due to the low propagation speed of the surface wave the first response pulses arrive at the reader after a delay of a few microseconds. After this time delay the *interference reflections* from the vicinity of the reader have long since decayed and can no longer interfere with the transponder's response pulse. Interference reflections from a radius of 100 m around the reader have decayed after around 0.66 µs (propagation time for 2 × 100 m). A surface wave on a quartz substrate ($v = 3158$ m/s) covers 2 mm in this time and thus just reaches the first reflectors on the substrate. This type of surface wave transponder is therefore also known as 'reflective delay lines' (Figure 4.96).

Surface wave transponders are completely linear and thus respond with a defined phase in relation to the interrogation pulse (see Figure 4.97). Furthermore, the phase angle $\phi_{2\text{-}1}$ and the differential propagation time $\tau_{2\text{-}1}$ between the reflected individual signals is constant. This gives rise to the possibility of improving the *range* of a surface wave transponder by taking the mean of weak transponder response signals from many interrogation pulses. Since a read operation requires only a few microseconds, several hundreds of thousands of read cycles can be performed per second.

The range of a surface wave transponder system can be determined using the radar equation (see Section 4.2.4.1). The influence of coherent averaging is taken into account as 'integration time' t_I (Reindl *et al.*, 1998a).

$$d = \sqrt[4]{\frac{P_\text{T} \cdot G_\text{T}^2 \cdot G_\text{R}^2 \cdot t_\text{i} \cdot \lambda^4}{k \cdot \text{T}_0 \cdot \text{F} \cdot \frac{\text{S}}{\text{N}} \cdot \text{IL}}} \tag{4.118}$$

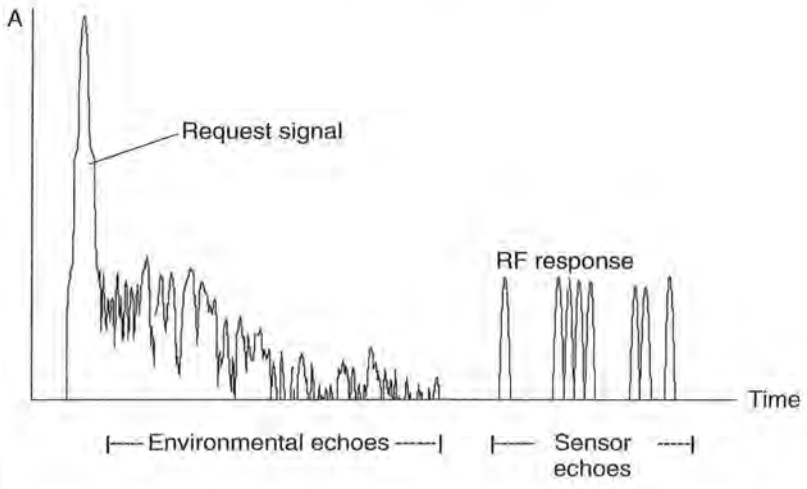


Figure 4.96 Sensor echoes from the surface wave transponder do not arrive until environmental echoes have decayed (reproduced by permission of Siemens AG, ZT KM, Munich)

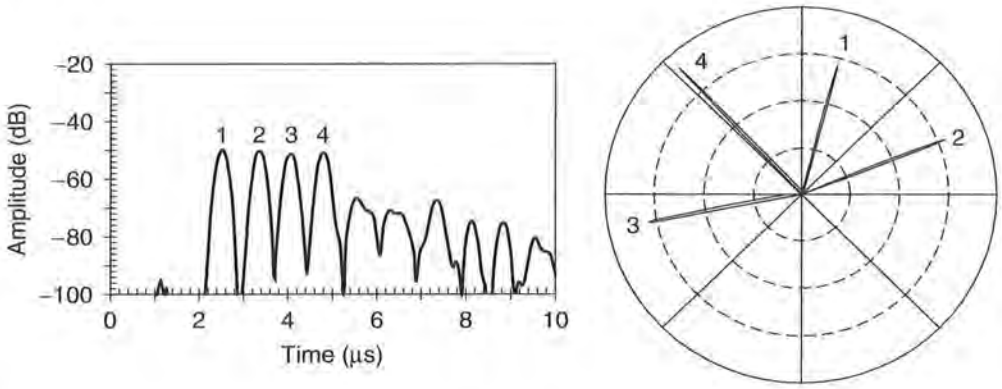


Figure 4.97 Surface wave transponders operate at a defined phase in relation to the interrogation pulse. Left, interrogation pulse, consisting of four individual pulses; right, the phase position of the response pulse, shown in a clockface diagram, is precisely defined (reproduced by permission of Siemens AG, ZT KM, Munich)

The relationship between the number of read cycles and the range of the system is shown in Figure 4.98 for two different frequency ranges. The calculation is based upon the system parameters listed in Table 4.9, which are typical of surface wave systems.

4.3.4 The sensor effect

The velocity v of a surface wave on the substrate, and thus also the propagation time τ and the mid-frequency f_0 of a surface wave component, can be influenced by a range of physical variables (Reindl and Mágori, 1995). In addition to temperature, mechanical forces such as static elongation, compression, shear, bending and acceleration have a particular influence upon the surface wave velocity v. This facilitates the remote interrogation of mechanical forces by surface wave sensors (Reindl and Mágori, 1995).

In general, the sensitivity S of the quantity x to a variation of the influence quantity y can be defined as:

$$S_y^x = \frac{1}{x} \cdot \frac{\partial x}{\partial y} \tag{4.119}$$

Table 4.9 System parameters for the range calculation shown in Figure 4.97

Value	At 433 MHz	At 2.45 GHz
P_S: transmission power	+14 dBm	
G_T: gain of transmission antenna	0 dB	
G_R: gain of transponder antenna	−3 dBl	0 dBl
Wavelength λ	70 cm	12 cm
F: Noise number of the receiver (reader)	12 dB	
S/N: Required signal/noise distance for error-free data detection	20 dB	
IL: Insertion loss: This is the additional damping of the electromagnetic response signal on the return path in the form of a surface wave	35 dB	40 dB
T_0: Noise temperature of the receiving antenna	300 K	

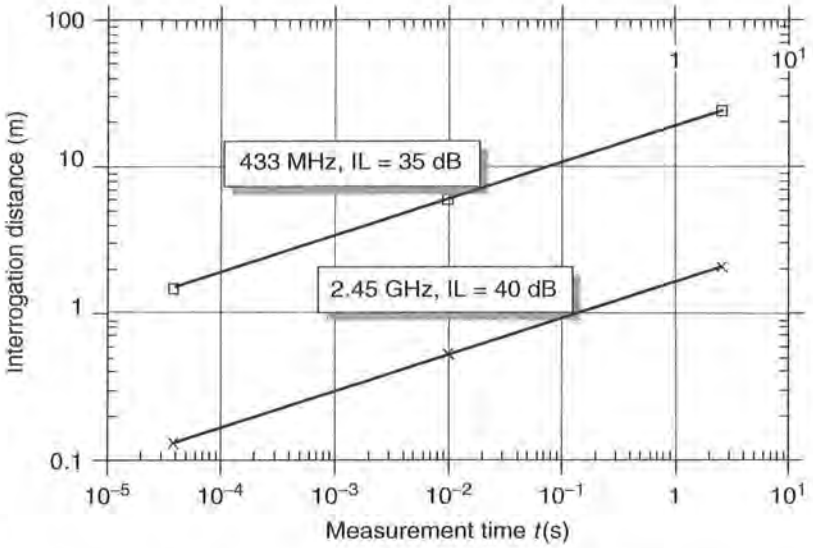


Figure 4.98 Calculation of the system range of a surface wave transponder system in relation to the integration time t_i at different frequencies (reproduced by permission of Siemens AG, ZT KM, Munich)

The sensitivity S to a certain influence quantity y is dependent here upon substrate material and crystal section. For example, the influence of temperature T upon propagation speed v for a surface wave on quartz is zero. Surface wave transponders are therefore particularly temperature stable on this material. On other substrate materials the propagation speed v varies with the temperature T.

The temperature dependency is described by the sensitivity S_T^v (also called the temperature coefficient Tk). The influence of temperature on the propagation speed v, the mid-frequency f_0 and the propagation time τ can be calculated as follows (Reindl and Mágori, 1995):

$$v(T) = v(T_0) \cdot [1 - S_T^v \cdot (T - T_0)] \tag{4.120}$$

$$f_0(T) = f_0(T_0) \cdot [1 - S_T^v \cdot (T - T_0)] \tag{4.121}$$

$$\tau(T) = \tau(T_0) \cdot [1 + S_T^v \cdot (T - T_0)] \tag{4.122}$$

4.3.4.1 Reflective delay lines

If only the differential propagation times or the differential phases between the individual reflected pulses are evaluated, the sensor signal is independent of the distance between the reader and the transponder. The differential propagation time $\tau_{2\text{-}1}$, and the differential phase $\theta_{2\text{-}1}$ between two received response pulses is obtained from the distance $L_{2\text{-}1}$ between the two reflectors, the velocity v of the surface wave and the frequency f of the interrogation pulse.

$$\tau_{2-1} = \frac{2 \cdot L_{2-1}}{v} \tag{4.123}$$

Table 4.10 The properties of some common surface wave substrate materials

Material	Crystal direction		V	k^2	S_T^v (Tk)	Damping (dB/μs)	
	Section ①	Prop ②	(m/s)	(%)	(ppm/°C)	433 MHz	2.45 GHz
Quartz	ST	X	3158	0.1	0	0.75	18.6
Quartz	37° rot-Y	90° rot-X	5092	=0.1	00	③	③
$LiNbO_3$	Y	Z	3488	4.1	94	0.25	5.8
$LiNbO_3$	128° rot-Y	X	3980	5.5	75	0.27	5.2
$LiTaO_3$	36° rot-Y	X	4112	=6.6	30	1.35 ③	20.9 ③
$LiTaO_3$	X	112° rot-Y	3301	0.88	18	—	—

① Section — surface normal to crystal axis.
② Crystal axis of the wave propagation.
③ Strong dependency of the value on the layer thickness.

$$\varphi_{2-1} = 2\pi f \cdot \tau_{2-1} = \frac{4\pi f \cdot L_{2-1}}{v} \tag{4.124}$$

The measurable change $\Delta\tau_{2\text{-}1}$ or $\Delta\theta_{2\text{-}1}$ when a physical quantity y is changed by the amount Δy is thus:

$$\Delta\tau_{2-1} = \tau_{2-1} \cdot S_y^\tau \cdot \Delta y \tag{4.125}$$

$$\Delta\varphi_{2-1} = 2\pi f \cdot \tau_{2-1} \cdot S_y^\tau \cdot \Delta y \tag{4.126}$$

The influence of the physical quantity y on the surface wave transponder can thus be determined only by the evaluation of the phase difference between the different pulses of the response signal. The measurement result is therefore also independent of the distance between reader and transponder.

For lithium niobate ($LiNbO_3$, YZ section), the linear temperature coefficient $T_k = S_T^v$ is approximately 90 ppm/°C. A reflective delay line on this crystal is thus a sensitive *temperature sensor* that can be interrogated by radio. Figure 4.99 shows the example of the pulse response of a temperature sensor and the temperature dependency of the associated phase values (Reindl *et al.*, 1998b). The precision of a temperature measurement based upon the evaluation of the associated phase value $\theta_{2\text{-}1}$ is approximately ±0.1°C and this precision can even be increased by special measures such as the use of longer propagation paths $L_{2\text{-}1}$ (see equation (4.124)) in the crystal. The unambiguity of the phase measurement can be assured over the entire measuring range by three to four correctly positioned reflectors.

4.3.4.2 Resonant sensors

In a reflective delay line the available path is used twice. However, if the interdigital transducer is positioned between two fully reflective structures, then the acoustic path can be used a much greater number of times due to multiple reflection. Such an arrangement (see Figure 4.99) is called a surface wave *one-port resonator*. The distance between the two reflectors must be an integer multiple of the half wavelength λ_0 at the resonant frequency f_1.

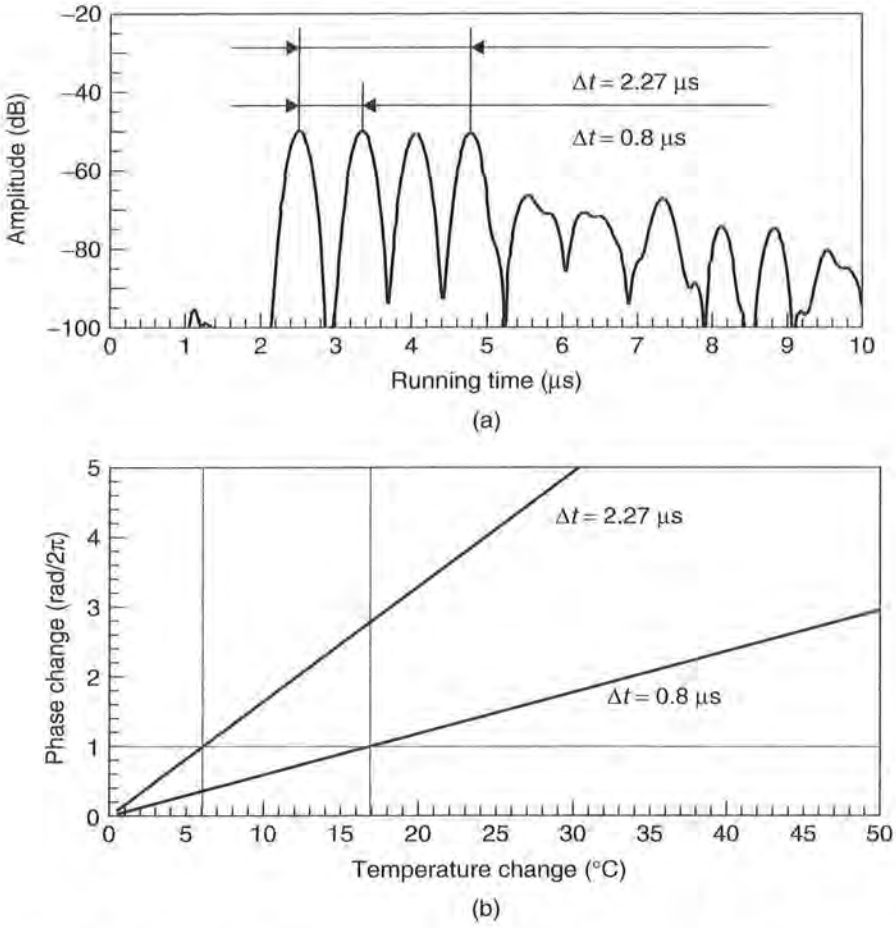


Figure 4.99 Impulse response of a temperature sensor and variation of the associated phase values between two pulses ($\Delta\tau = 0.8\,\mu s$) or four pulses ($\Delta\tau = 2.27\,\mu s$). The high degree of linearity of the measurement is striking (reproduced by permission of Siemens AG, ZT KM, Munich)

The number of wave trains stored in such a *resonator* will be determined by its loaded Q factor. Normally a *Q factor* of 10 000 is achieved at 434 MHz and at 2.45 GHz a Q factor of between 1500 and 3000 is reached (Reindl *et al.*, 1998b). The displacement of the mid-frequency Δf_1 and the displacement of the associated phase $\Delta\theta_1$ of a resonator due to a change of the physical quantity y with the loaded Q factor are (Reindl *et al.*, 1998a):

$$\Delta f_1 = -f_1(y_0) \cdot S_{y,1} \cdot \Delta y \tag{4.127}$$

and

$$\Delta\varphi = 2Q \cdot \frac{\Delta f}{f} \tag{4.128}$$

where f_1 is the unaffected resonant frequency of the resonator.

In practice, the same sensitivity is obtained as for a reflective delay line, but with a significant reduction in chip size (Reindl *et al.*, 1998b) (Figure 4.100).

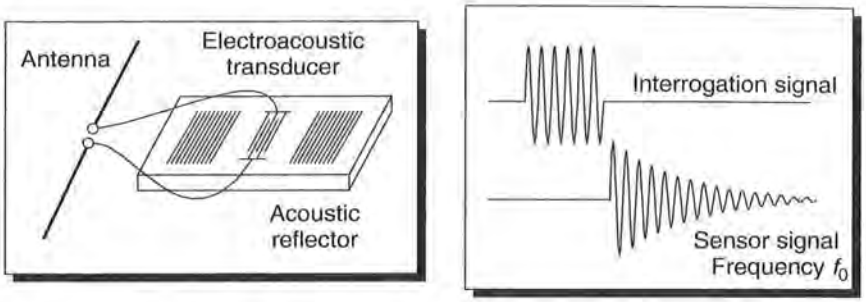


Figure 4.100 Principal layout of a resonant surface wave transponder and the associated pulse response (reproduced by permission of Siemens AG, ZT KM, Munich)

If, instead of one resonator, several resonators with different frequencies are placed on a crystal (Figure 4.101), then the situation is different: instead of a pulse sequence in the time domain, such an arrangement emits a characteristic line spectrum back to the interrogation device (Reindl *et al.*, 1998b,c), which can be obtained from the received sensor signal by a Fourier transformation (Figure 4.102).

The difference $\Delta f_{2\text{-}1}$ between the resonant frequencies of the two resonators is determined to measure a physical quantity y in a surface wave transponder with two resonators. Similarly to equation (4.127), this yields the following relationship (Reindl *et al.*, 1998c).

$$\Delta f_{2-1} = -[f_2(y_0) \cdot S_{y,2} - f_1(y_0) \cdot S_{y,1}] \cdot \Delta y \tag{4.129}$$

4.3.4.3 Impedance sensors

Using surface wave transponders, even conventional sensors can be passively interrogated by radio if the impedance of the sensor changes as a result of the change of a physical quantity y (e.g. photoresistor, Hall sensor, NTC or PTC resistor). To achieve this a second interdigital transducer is used as a reflector and connected to the external sensor (Figure 4.103). A measured quantity Δy thus changes the terminating impedance of the additional interdigital transducer. This changes the acoustic transmission and reflection ρ of the converter that is connected to this load, and thus also

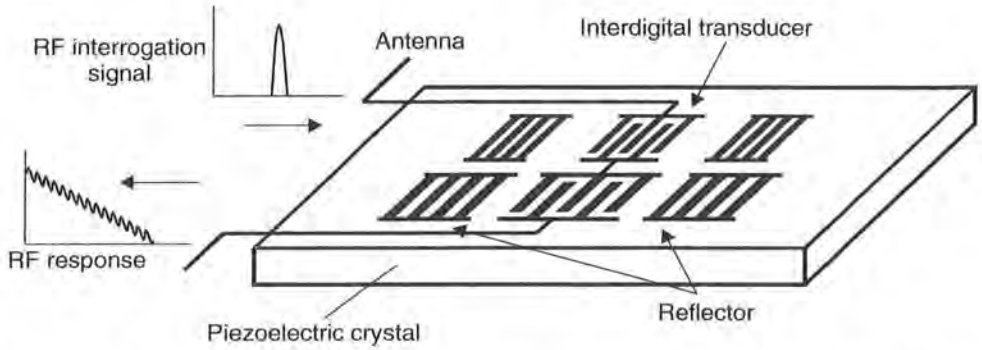


Figure 4.101 Principal layout of a surface wave transponder with two resonators of different frequency (f_1, f_2) (reproduced by permission of Siemens AG, ZT KM, Munich)

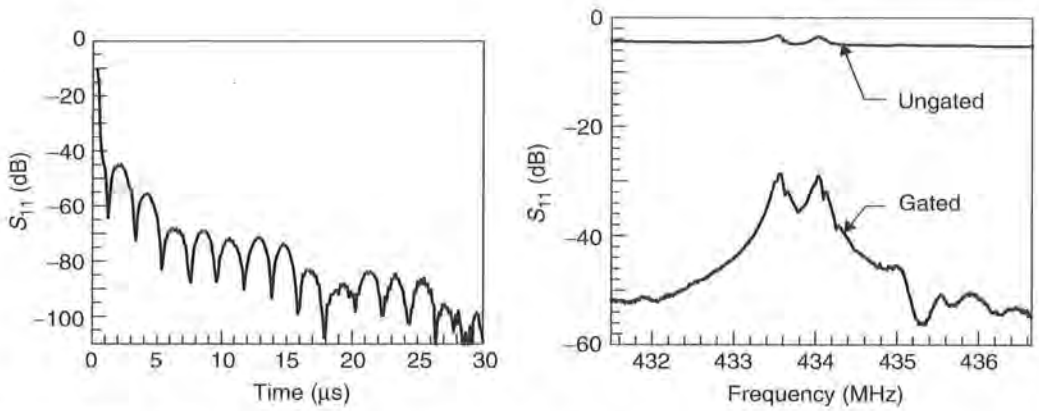


Figure 4.102 Left, measured impulse response of a surface wave transponder with two resonators of different frequency; right, after the Fourier transformation of the impulse response the different resonant frequencies of the two resonators are visible in the line spectrum (here: approx. 433.5 MHz and 434 MHz) (reproduced by permission of Siemens AG, ZT KM, Munich)

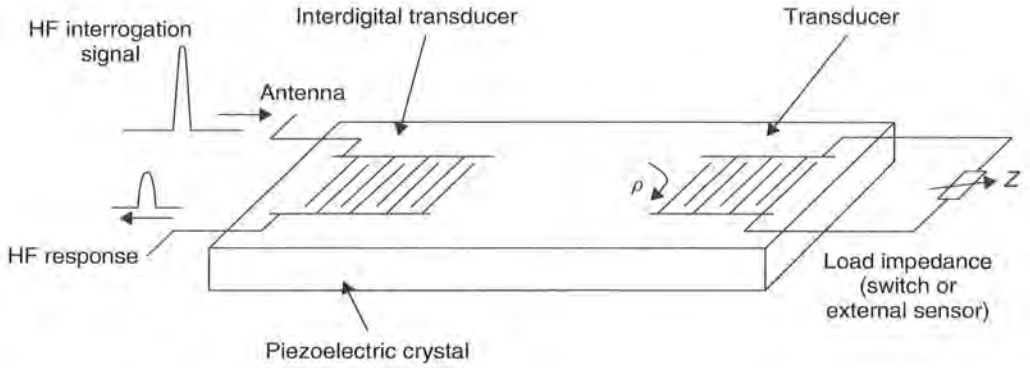


Figure 4.103 Principal layout of a passive surface wave transponder connected to an external sensor (reproduced by permission of Siemens AG, ZT KM, Munich)

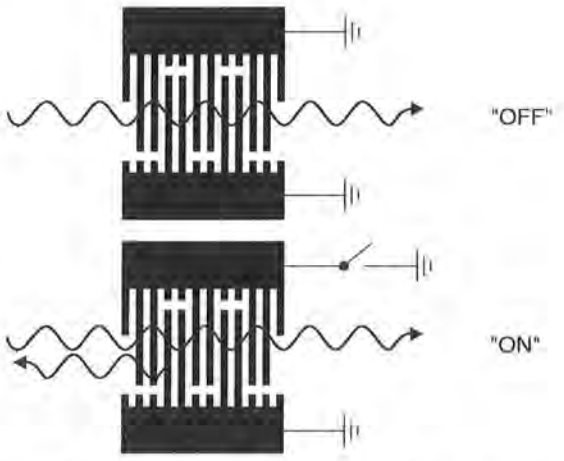


Figure 4.104 Passive recoding of a surface wave transponder by a switched interdigital transducer (reproduced by permission of Siemens AG, ZT KM, Munich)

changes the magnitude and phase of the reflected HF pulse, which can be detected by the reader.

4.3.5 Switched sensors

Surface wave transponders can also be passively recoded (Figure 4.104). As is the case for an impedance sensor, a second interdigital transducer is used as a reflector. External circuit elements of the interdigital transducer's busbar make it possible to switch between the states 'short-circuited' and 'open'. This significantly changes the acoustic transmission and reflection ρ of the transducer and thus also the magnitude and phase of the reflected HF impulse that can be detected by the reader.

11
Readers

11.1 Data Flow in an Application

A *software application* that is designed to read data from a contactless data carrier (transponder) or write data to a contactless data carrier, requires a contactless *reader* as an interface. From the point of view of the application software, access to the data carrier should be as transparent as possible. In other words, the read and write operations should differ as little as possible from the process of accessing comparable data carriers (smart card with contacts, serial EEPROM).

Write and read operations involving a contactless data carrier are performed on the basis of the *master–slave principle* (Figure 11.1). This means that all reader and transponder activities are initiated by the application software. In a hierarchical system structure the application software represents the master, while the reader, as the slave, is only activated when write/read commands are received from the application software.

To execute a command from the application software, the reader first enters into communication with a transponder. The reader now plays the role of the master in relation to the transponder. The transponder therefore only responds to commands from the reader and is never active independently (except for the simplest read-only transponders. See Chapter 10).

A simple read command from the application software to the reader can initiate a series of communication steps between the reader and a transponder. In the example in Table 11.1, a read command first leads to the activation of a transponder, followed by the execution of the authentication sequence and finally the transmission of the requested data.

The reader's main functions are therefore to activate the data carrier (transponder), structure the communication sequence with the data carrier, and transfer data between the application software and a contactless data carrier. All features of the contactless communication, i.e. making the connection, and performing anticollision and authentication procedures, are handled entirely by the reader.

11.2 Components of a Reader

A number of contactless transmission procedures have already been described in the preceding chapters. Despite the fundamental differences in the type of coupling

RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification. Klaus Finkenzeller
 ISBN: 0-470-84402-7

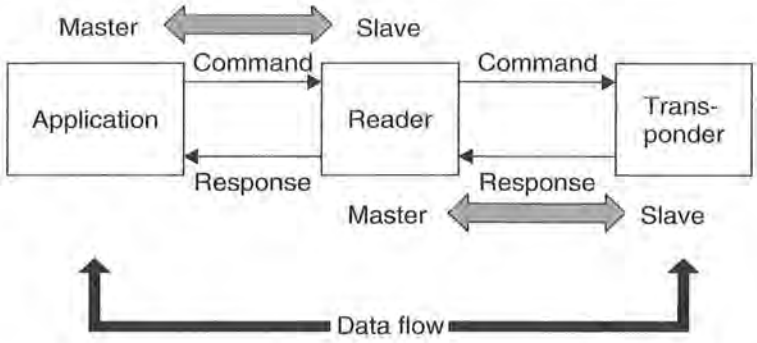


Figure 11.1 Master–slave principle between application software (application), reader and transponder

Table 11.1 Example of the execution of a read command by the application software, reader and transponder

Application ↔ reader	Reader ↔ transponder	Comment
→ Blockread_Address[00]		Read transponder memory [address]
	→ Request	Transponder in the field?
	← ATR_SNR[4712]	Transponder operates with serial number
	→ GET_Random	Initiate authentication
	← Random[081514]	
	→ SEND_Token1	
	← GET_Token2	Authentication successfully completed
	→ Read_@[00]	Read command [address]
	← Data[9876543210]	Data from transponder
← Data[9876543210]		Data to application

(inductive — electromagnetic), the communication sequence (FDX, HDX, SEQ), the data transmission procedure from the transponder to the reader (load modulation, backscatter, subharmonic) and, last but not least, the frequency range, all readers are similar in their basic operating principle and thus in their design.

Readers in all systems can be reduced to two fundamental functional blocks: the control system and the *HF interface*, consisting of a transmitter and receiver (Figure 11.2). Figure 11.3 shows a reader for an inductively coupled RFID system. On the right-hand side we can see the HF interface, which is shielded against undesired spurious emissions by a tinplate housing. The control system is located on the left-hand side of the reader and, in this case, it comprises an ASIC module and microcontroller. In order that it can be integrated into a software application, this reader has an RS232 interface to perform the data exchange between the reader (slave) and the external application software (master).

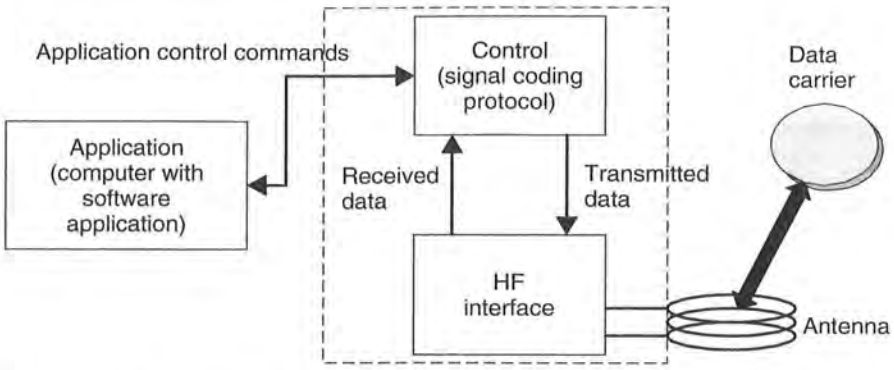


Figure 11.2 Block diagram of a reader consisting of control system and HF interface. The entire system is controlled by an external application via control commands

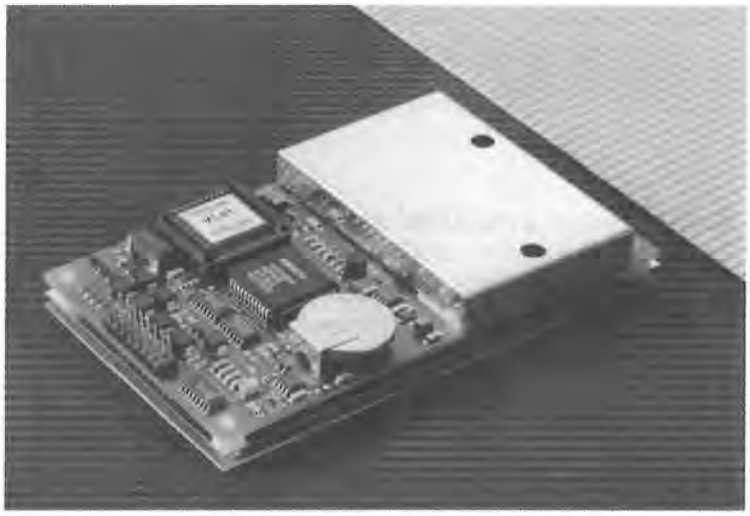


Figure 11.3 Example of a reader. The two functional blocks, HF interface and control system, can be clearly differentiated (MIFARE® reader, reproduced by permission of Philips Electronics N.V.)

11.2.1 HF interface

The reader's HF interface performs the following functions:

- generation of high frequency transmission power to activate the transponder and supply it with power;
- modulation of the transmission signal to send data to the transponder;
- reception and demodulation of HF signals transmitted by a transponder.

The HF interface contains two separate signal paths to correspond with the two directions of data flow from and to the transponder (Figure 11.4). Data transmitted to

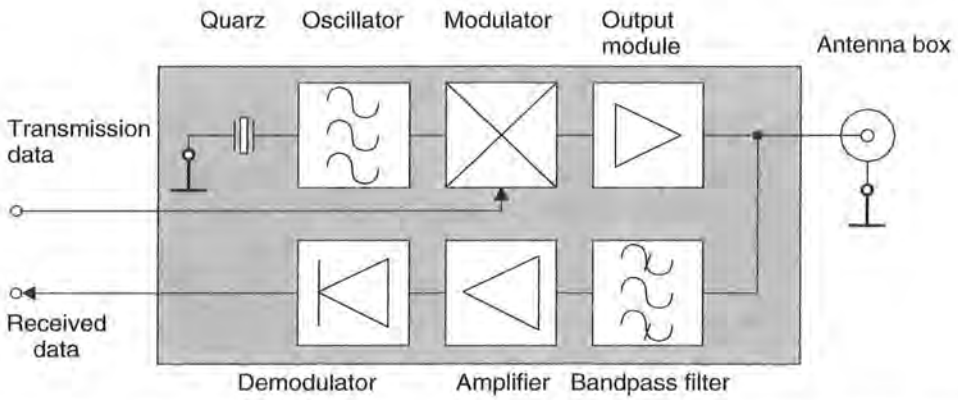


Figure 11.4 Block diagram of an HF interface for an inductively coupled RFID system

the transponder travels through the *transmitter arm*. Conversely, data received from the transponder is processed in the *receiver arm*. We will now analyse the two signal channels in more detail, giving consideration to the differences between the different systems.

11.2.1.1 Inductively coupled system, FDX/HDX

First, a signal of the required operating frequency, i.e. 135 kHz or 13.56 MHz, is generated in the transmitter arm by a stable (frequency) quartz oscillator. To avoid worsening the noise ratio in relation to the extremely weak received signal from the transponder, the *oscillator* is subject to high demands regarding phase stability and sideband noise.

The oscillator signal is fed into a modulation module controlled by the baseband signal of the signal coding system. This *baseband signal* is a keyed direct voltage signal (TTL level), in which the binary data is represented using a serial code (Manchester, Miller, NRZ). Depending upon the modulator type, *ASK* or *PSK modulation* is performed on the oscillator signal.

FSK modulation is also possible, in which case the baseband signal is fed directly into the frequency synthesiser.

The modulated signal is then brought to the required level by a power output module and can then be decoupled to the antenna box.

The *receiver arm* begins at the antenna box, with the first component being a steep edge bandpass filter or a notch filter. In FDX/HDX systems this filter has the task of largely blocking the strong signal from the transmission output module and filtering out just the response signal from the transponder. In subharmonic systems, this is a simple process, because transmission and reception frequencies are usually a whole octave apart. In systems with load modulation using a *subcarrier* the task of developing a suitable filter should not be underestimated because, in this case, the transmitted and received signals are only separated by the subcarrier frequency. Typical subcarrier frequencies in 13.56 MHz systems are 847 kHz or 212 kHz.

Some LF systems with load modulation and no subcarrier use a notch filter to increase the modulation depth (duty factor) — the ratio of the level to the load modulation sidebands — and thus the duty factor by reducing their own carrier signal.

A different procedure is the rectification and thus demodulation of the (load) amplitude modulated voltage directly at the reader antenna. A sample circuit for this can be found in Section 11.3.

11.2.1.2 Microwave systems – half duplex

The main difference between *microwave systems* and low frequency inductive systems is the frequency synthesising: the operating frequency, typically 2.45 GHz, cannot be generated directly by the quartz oscillator, but is created by the multiplication (excitation of harmonics) of a lower oscillator frequency. Because the modulation is retained during frequency multiplication, modulation is performed at the lower frequency. See Figure 11.5.

Some microwave systems employ a *directional coupler* to separate the system's own transmission signal from the weak backscatter signal of the transponder (Integrated Silicon Design, 1996).

A directional coupler (Figure 11.6) consists of two continuously coupled homogeneous wires (Meinke and Gundlack, 1992). If all four ports are matched and power P_1 is supplied to port ①, then the power is divided between ports ② and ③, with no

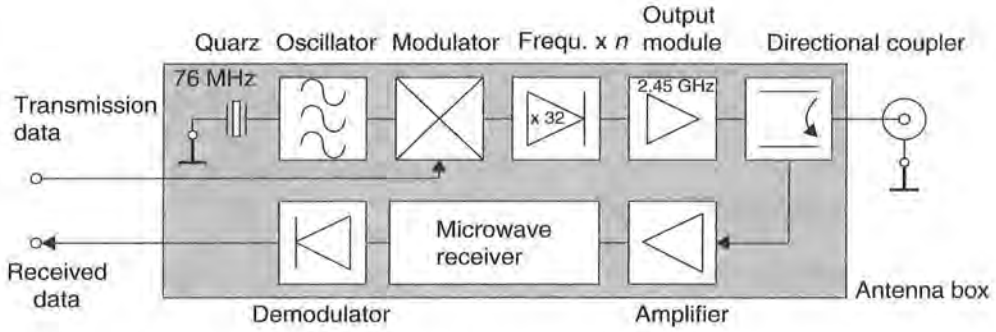


Figure 11.5 Block diagram of an HF interface for microwave systems

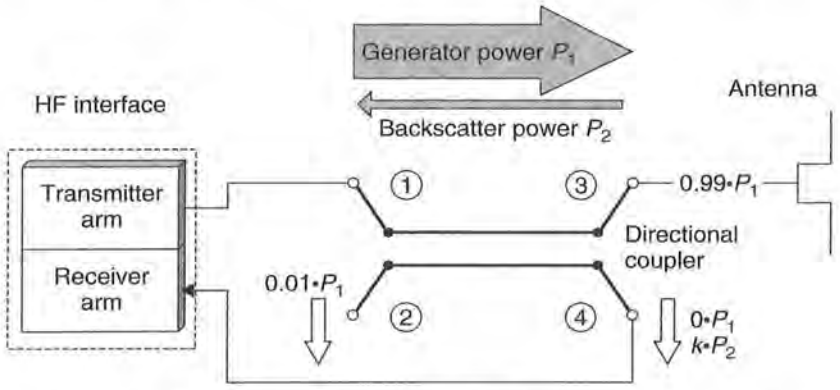


Figure 11.6 Layout and operating principle of a directional coupler for a backscatter RFID system

power occurring at the decoupled port ④. The same applies if power is supplied to port ③, in which case the power is divided between ports ① and ②.

A directional coupler is described by its *coupling loss*:

$$a_k = -20 \cdot \ln |P② / P①| \tag{11.1}$$

and *directivity*:

$$a_D = -20 \cdot \ln |P④ / P②| \tag{11.2}$$

Directivity is the logarithmic magnitude of the ratio of undesired overcoupled power P_4 to desired coupled power P_2.

A directional coupler for a backscatter RFID reader should have the maximum possible directivity to minimise the decoupled signal of the transmitter arm at port ④. The coupling loss, on the other hand, should be low to decouple the maximum possible proportion of the reflected power P_2 from the transponder to the receiver arm at port ④. When a reader employing decoupling based upon a directional coupler is commissioned, it is necessary to ensure that the transmitter antenna is well (anechoically) set up. Power reflected from the antenna due to poor adjustment is decoupled at port ④ as backwards power. If the directional coupler has a good coupling loss, even a minimal mismatching of the transmitter antenna (e.g. by environmental influences) is sufficient to increase the backwards travelling power to the magnitude of the reflected transponder power. Nevertheless, the use of a directional coupler gives a significant improvement compared to the level ratios achieved with a direct connection of transmitter output module and receiver input.

11.2.1.3 Sequential systems – SEQ

In a sequential RFID system the HF field of the reader is only ever transmitted briefly to supply the transponder with power and/or send commands to the transponder.

The transponder transmits its data to the reader while the reader is not transmitting. The transmitter and receiver in the reader are thus active sequentially, like a walkie-talkie, which also transmits and receives alternately. See Figure 11.7.

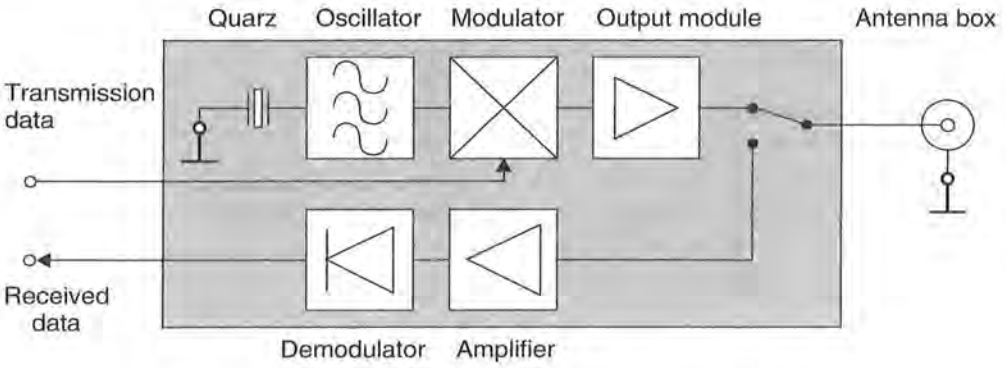


Figure 11.7 HF interface for a sequential reader system

The reader contains an instantaneous switching unit to switch between transmitter and receiver mode. This function is normally performed by PIN diodes in radio technology.

No special demands are made of the receiver in an SEQ system. Because the strong signal of the transmitter is not present to cause interference during reception, the SEQ receiver can be designed to maximise sensitivity. This means that the range of the system as a whole can be increased to correspond with the *energy range*, i.e. the distance between reader and transponder at which there is just enough energy for the operation of the transponder.

11.2.1.4 Microwave system for SAW transponders

A short electromagnetic pulse transmitted by the reader's antenna is received by the antenna of the *surface wave transponder* and converted into a surface wave in a piezoelectric crystal. A characteristic arrangement of partially reflective structures in the propagation path of the surface wave gives rise to numerous pulses, which are transmitted back from the transponder's antenna as a response signal (a much more comprehensive description of this procedure can be found in Section 4.3).

Due to the propagation delay times in the piezoelectric crystal the coded signal reflected by the transponder can easily be separated in the reader from all other electromagnetic reflections from the vicinity of the reader (see Section 4.3.3). The block diagram of a reader for surface wave transponders is shown in Figure 11.8.

A stable frequency and phase oscillator with a surface wave resonator is used as the high-frequency source. Using a rapid HF switch, short HF pulses of around 80 ns duration are generated from the oscillator signal, which are amplified to around 36 dBm (4 W peak) by the connected power output stage, and transmitted by the reader's antenna.

If a SAW transponder is located in the vicinity of the reader it reflects a sequence of individual pulses after a propagation delay time of a few microseconds. The pulses

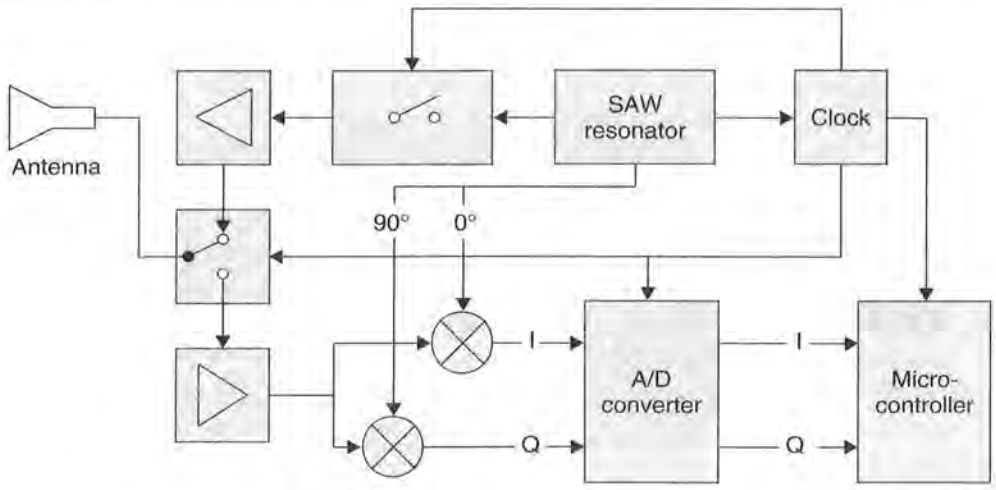


Figure 11.8 Block diagram of a reader for a surface wave transponder

received by the reader's antenna pass through a low-noise amplifier and are then demodulated in a quadrature demodulator. This yields two orthogonal components (I and Q), which facilitate the determination of the phase angle between the individual pulses and between the pulses and the oscillator (Bulst *et al.*, 1998). The information obtained can be used to determine the distance or speed between SAW transponder and reader and for the measurement of physical quantities (see Section 10.4.3).

To be more precise, the reader circuit in Figure 11.8 corresponds with a *pulse radar*, like those used in flight navigation (although in this application the transmission power is much greater). In addition to the pulse radar shown here, other radar types (for example FM-CW radar) are also in development as readers for SAW transponders.

11.2.2 Control unit

The reader's control unit (Figure 11.9) performs the following functions:

- communication with the application software and the execution of commands from the application software;
- control of the communication with a transponder (master–slave principle);
- signal coding and decoding (Figure 11.10).

In more complex systems the following additional functions are available:

- execution of an anticollision algorithm;
- encryption and decryption of the data to be transferred between transponder and reader;
- performance of authentication between transponder and reader.

The control unit is usually based upon a microprocessor to perform these complex functions. Cryptological procedures, such as stream ciphering between transponder and reader, and also signal coding, are often performed in an additional ASIC module to relieve the processor of calculation intensive processes. For performance reasons the ASIC is accessed via the microprocessor bus (register orientated).

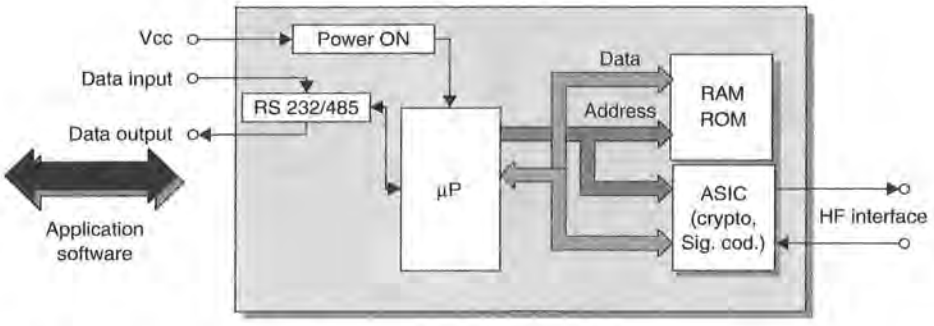


Figure 11.9 Block diagram of the control unit of a reader. There is a serial interface for communication with the higher application software

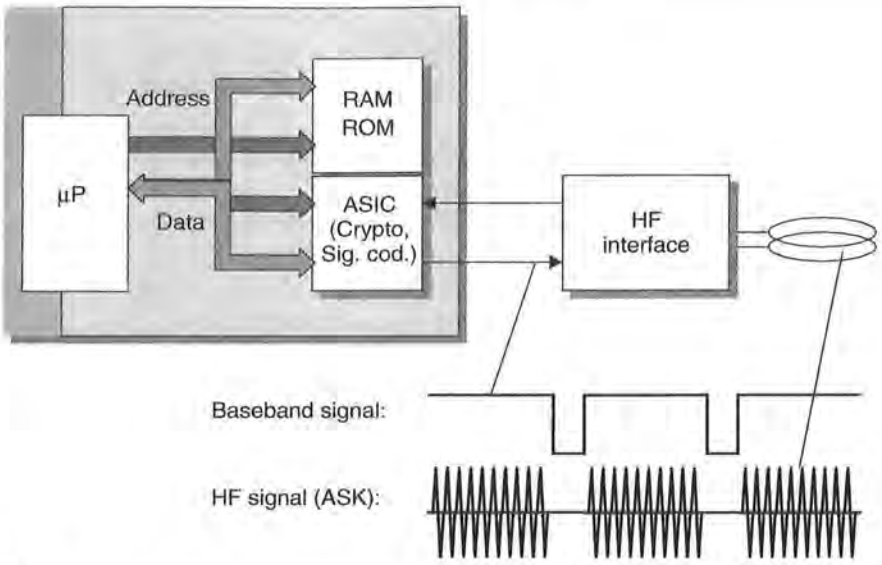


Figure 11.10 Signal coding and decoding is also performed by the control unit in the reader

Data exchange between *application software* and the reader's control unit is performed by an RS232 or RS485 interface. As is normal in the PC world, NRZ coding (8-bit asynchronous) is used. The baud rate is normally a multiple of 1200 Bd (4800 Bd, 9600 Bd, etc.). Various, often self-defined, protocols are used for the communication protocol. Please refer to the handbook provided by your system supplier.

The interface between the HF interface and the control unit represents the state of the HF interface as a binary number. In an ASK modulated system a logic '1' at the modulation input of the HF interface represents the state 'HF signal on'; a logic '0' represents the state 'HF signal off' (further information in Section 10.1.1).

11.3 Low Cost Configuration – Reader IC U2270B

It is typical of applications that use contactless identification systems that they require only a few readers, but a very large number of transponders. For example, in a public transport system, several tens of thousands of contactless smart cards are used, but only a few hundred readers are installed in vehicles. In applications such as animal identification or container identification, there is also a significant difference between the number of transponders used and the corresponding number of readers. There are also a great many different systems, because there are still no applicable standards for inductive or microwave RFID systems. As a result, readers are only ever manufactured in small batches of a few thousand.

Electronic *immobilisation* systems, on the other hand, require a vast number of readers. Because since 1995 almost all new cars have been fitted with electronic immobilisation systems as standard, the number of readers required has reached a completely new order of magnitude. Because the market for powered vehicles is also very price sensitive, cost reduction and miniaturisation by the integration of a small number of

functional modules has become worth pursuing. Because of this, it is now possible to integrate the whole analogue section of a reader onto a silicon chip, meaning that only a few external components are required. We will briefly described the *U2270B* as an example of such a reader IC.

The reader IC U2270B by TEMIC serves as a fully integrated HF interface between a transponder and a microcontroller (Figure 11.11).

The IC contains the following modules: on-chip oscillator, driver, received signal conditioning and an integral power supply (Figure 11.12).

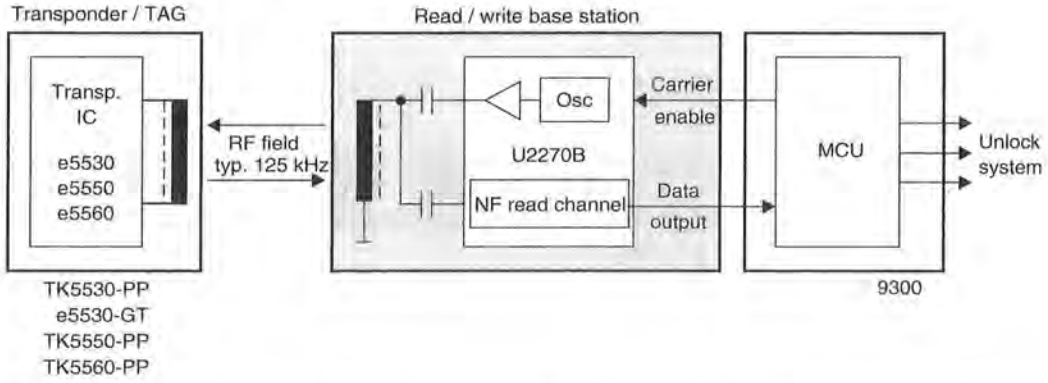


Figure 11.11 The low-cost reader IC U2270B represents a highly integrated HF interface. The control unit is realised in an external microprocessor (MCU) (reproduced by permission of TEMIC Semiconductor GmbH, Heilbronn)

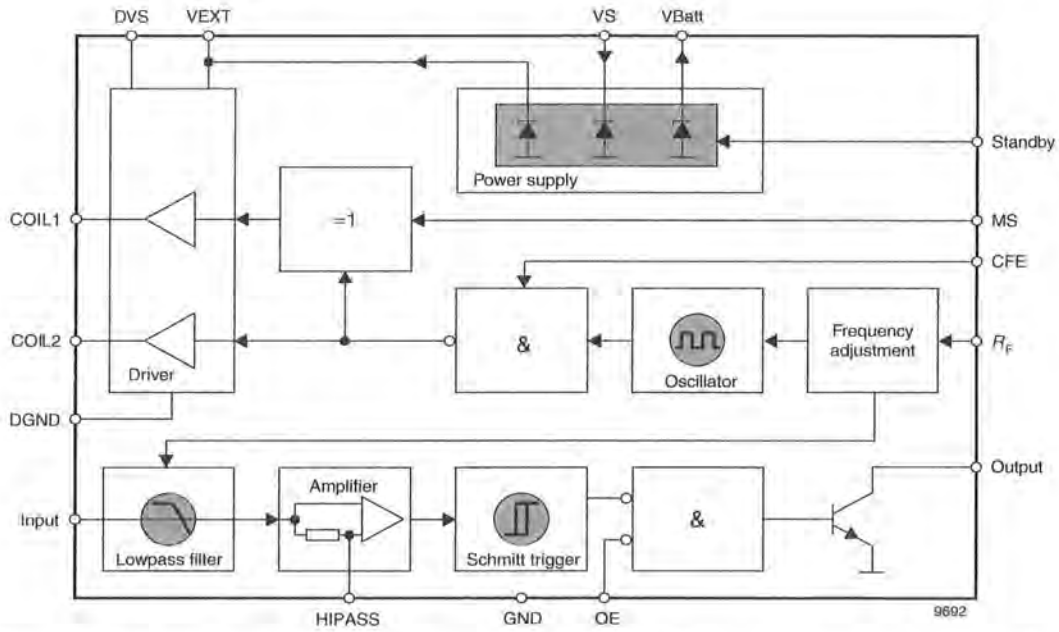


Figure 11.12 Block diagram of the reader IC U2270B. The transmitter arm consists of an oscillator and driver to supply the antenna coil. The receiver arm consists of filter, amplifier and a Schmitt trigger (reproduced by permission of TEMIC Semiconductor GmbH, Heilbronn)

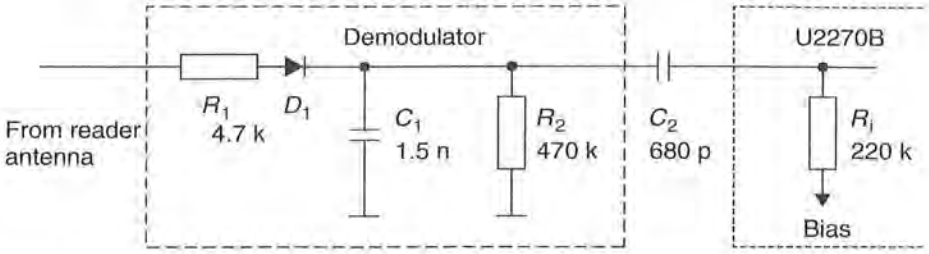


Figure 11.13 Rectification of the amplitude modulated voltage at the antenna coil of the reader (reproduced by permission of TEMIC Semiconductor GmbH, Heilbronn)

The on-chip oscillator generates the operating frequency in the range 100–150 kHz. The precise frequency is adjusted by an external resistor at pin R_F. The downstream driver generates the power required to control the antenna coil as push–pull output. If necessary, a baseband modulation signal can be fed into pin CFE as a TTL signal and this switches the HF signal on/off, generating an ASK modulation.

The *load modulation* procedure in the transponder generates a weak amplitude modulation of the reader's antenna voltage. The modulation in the transponder occurs in the baseband, i.e. without the use of a subcarrier. The transponder modulation signal can be reclaimed simply by demodulating the antenna voltage at the reader using a diode. The signal, which has been rectified by an external diode and smoothed using an RC low-pass filter, is fed into the 'Input' pin of the U2270B (Figure 11.13). Using a downstream Butterworth low-pass filter, an amplifier module and a Schmitt trigger, the demodulated signal is converted into a TTL signal, which can be evaluated by the downstream microprocessor. The time constants of the Butterworth filter are designed so that a Manchester or bi-phase code can be processed up to a data rate of $f_{osc}/25$ (approximately 4800 bit/s) (TEMIC, 1977).

A complete application circuit for the U2270B can be found in the following chapter.

11.4 Connection of Antennas for Inductive Systems

Reader antennas in inductively coupled RFID systems generate magnetic flux Φ, which is used for the power supply of the transponder and for sending messages between the reader and the transponder. This gives rise to three fundamental design requirements for a reader antenna:

- maximum current i_1 in the *antenna coil*, for maximum magnetic flux Φ;
- power matching so that the maximum available energy can be used for the generation of the magnetic flux;
- sufficient bandwidth for the undistorted transmission of a carrier signal modulated with data.

Depending upon the frequency range, different procedures can be used to connect the antenna coil to the transmitter output of the reader: direct connection of the antenna coil to the power output module using power matching or the supply of the antenna coil via coaxial cable.

11.4.1 Connection using current matching

In typical low cost readers in the frequency range below 135 kHz, the HF interface and antenna coil are mounted close together (a few centimetres apart), often on a single printed circuit board. Because the geometric dimensions of the antenna supply line and antenna are smaller than the wavelength of the generated HF current (2200 m) by powers of ten, the signals may be treated as stationary for simplification. This means that the wave characteristics of a high frequency current may be disregarded. The connection of an antenna coil is thus comparable to the connection of a loudspeaker to an NF output module from the point of view of circuitry.

The reader IC U2270B, which was described in the preceding section, can serve as an example of such a low cost reader (Figures 11.14–11.16).

Figure 11.14 shows an example of an antenna circuit. The antenna is fed by the push–pull bridge output of the reader IC. In order to maximise the current through the antenna coil, a *serial resonant circuit* is created by the serial connection of the antenna coil L_S to a capacitor C_S and a resistor R_S. Coil and capacitor are dimensioned such that the resonant frequency f_0 is as follows at the operating frequency of the reader:

$$f_0 = \frac{1}{2\pi\sqrt{L_s \cdot C_s}} \tag{11.3}$$

The coil current is then determined exclusively by the series resistor R_S.

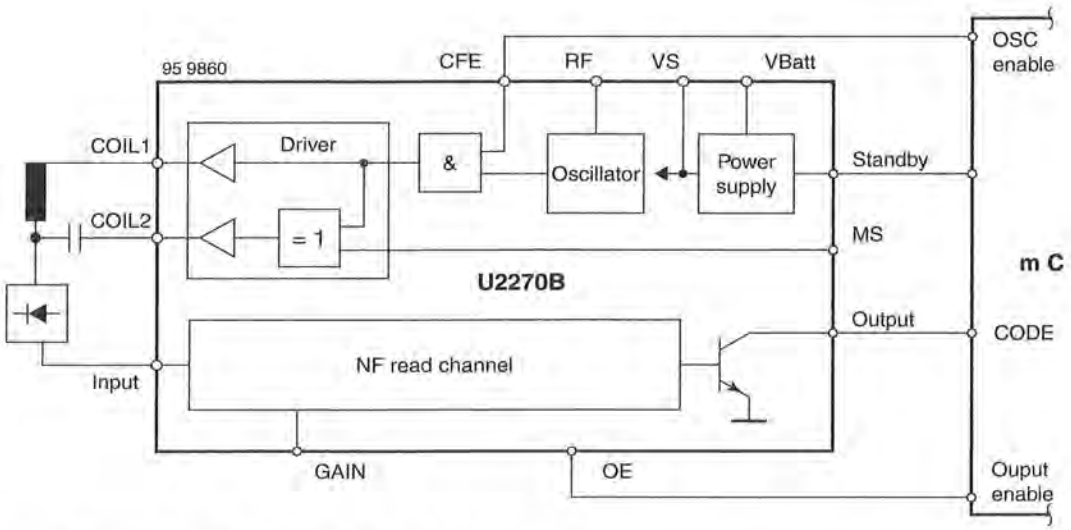


Figure 11.14 Block diagram for the reader IC U2270B with connected antenna coil at the push–pull output (reproduced by permission of TEMIC Semiconductor GmbH, Heilbronn)

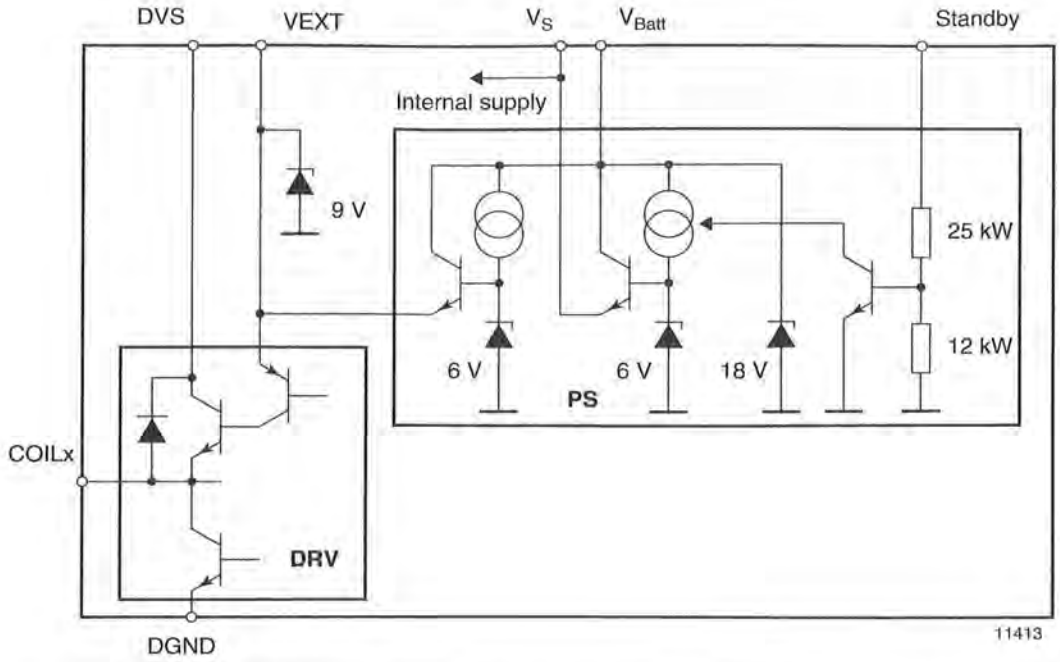


Figure 11.15 Driver circuit in the reader IC UU2270B (reproduced by permission of TEMIC Semiconductor GmbH, Heilbronn)

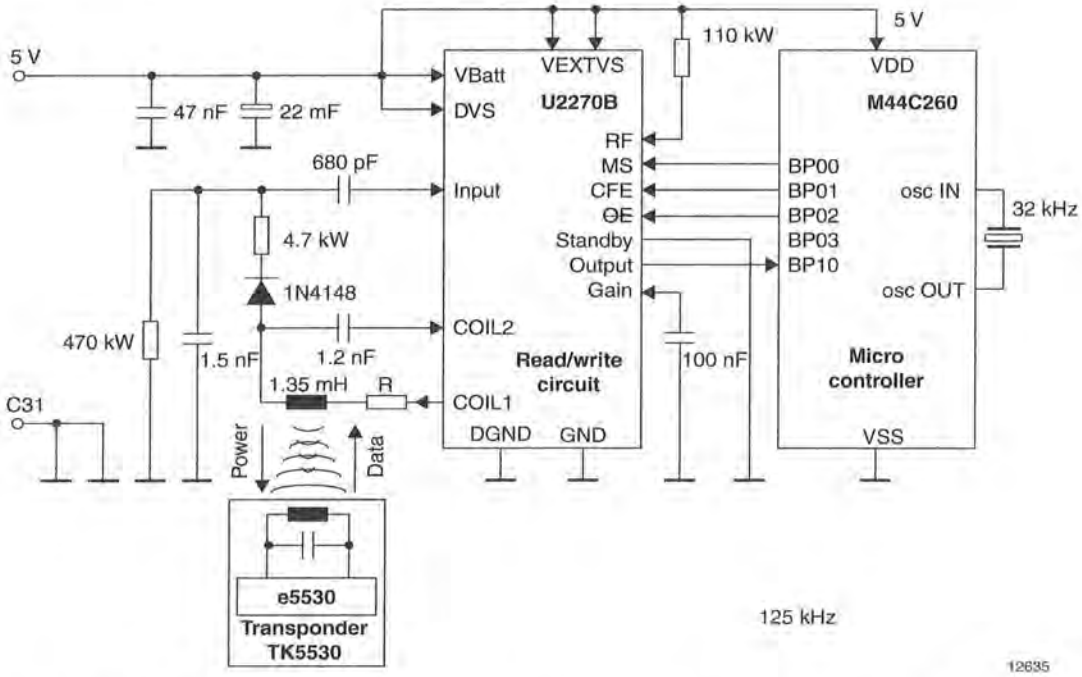


Figure 11.16 Complete example application for the low cost reader IC U2270B (reproduced by permission of TEMIC Semiconductor GmbH, Heilbronn)

11.4.2 Supply via coaxial cable

At frequencies above 1 MHz, or in the frequency range 135 kHz if longer cables are used, the HF voltage can no longer be considered stationary, but must be treated as an *electromagnetic wave* in the cable. Connecting the antenna coil using a long, unshielded two core wire in the HF range would therefore lead to undesired effects, such as power reflections, impedance transformation and parasitic power emissions, due to the wave nature of a HF voltage. Because these effects are difficult to control when they are not exploited intentionally, shielded cable — so-called *coaxial cable* — is normally used in radio technology. Sockets, plugs and coaxial cable are uniformly designed for a cable impedance of 50 Ω and, being a mass produced product, are correspondingly cheap. RFID systems generally use 50 Ω components.

The block diagram of an inductively coupled RFID system using 50 Ω technology shows the most important HF components (Figure 11.17).

The antenna coil L_1 represents an impedance Z_L in the operating frequency range of the RFID system. To achieve power matching with the 50 Ω system, this impedance must be transformed to 50 Ω (matched) by a passive *matching circuit*. Power transmission from the reader output module to the matching circuit is achieved (almost) without losses or undesired radiation by means of a coaxial cable.

A suitable matching circuit can be realised using just a few components. The circuit illustrated in Figure 11.18, which can be constructed using just two capacitors, is very simple to design (Suckrow, 1997). This circuit is used in practice in various 13.56 MHz RFID systems.

Figure 11.19 shows a reader with an integral antenna for a 13.56 MHz system. Coaxial cable has not been used here, because a very short supply line can be realised

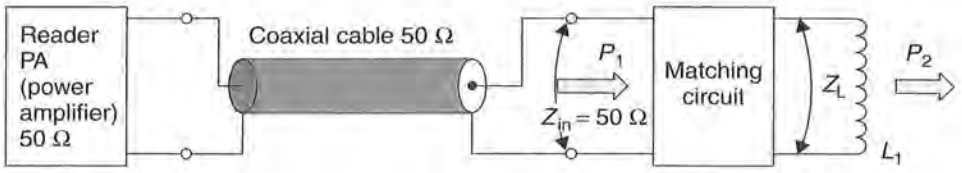


Figure 11.17 Connection of an antenna coil using 50 Ω technology

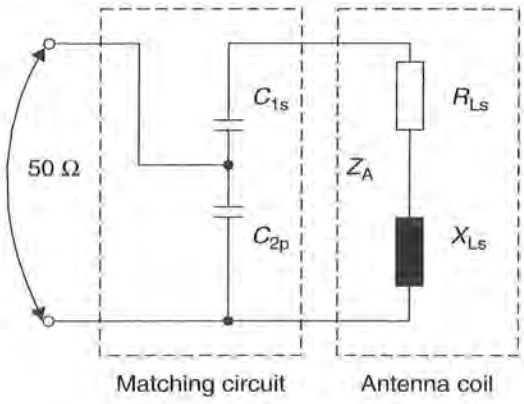


Figure 11.18 Simple matching circuit for an antenna coil

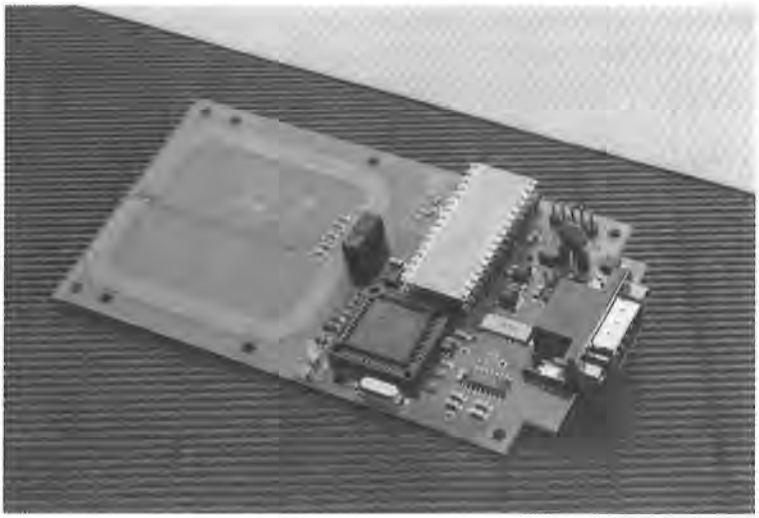


Figure 11.19 Reader with integral antenna and matching circuit (MIFARE®-reader, reproduced by permission of Philips Electronics N.V.)

by a suitable layout (stripline). The matching circuit is clearly visible on the inside of the antenna coil (SMD component).

Before we can dimension the circuit, we first need to determine the impedance Z_A of the antenna coil for the operating frequency by measurement. It is clear that the impedance of a real antenna coil is generated by the serial connection of the coil inductance L_S with the ohmic wire resistance RL_S of the wire. The serial connection from XL_S and RL_S can also be represented in the impedance level.

The function of the matching circuit is the transformation of the complex coil impedance Z_A to a value of 50 Ω real. A reactance (capacitance, inductance) in series with the coil impedance Z_A shifts the total impedance Z in the direction of the jX axis, while a parallel reactance shifts the total impedance away from the origin in a circular path (Figure 11.20).

The values of C_{2p} and C_{2s} are dimensioned such that the resulting coil impedance Z_A is transformed to the values desired to achieve 50 Ω.

The matching circuit from Figure 11.18 can be mathematically represented by equation 11.4:

$$Z_0 = 50\,\Omega = \frac{1}{-j\omega C_{2p} + \left(\dfrac{1}{\dfrac{1}{-j\omega C_{1s}} + R_{Ls} + j\omega L_s} \right)} \tag{11.4}$$

From the relationship between resistance and conductance in the complex impedance plane (Z-level), we find the following relationship for C_{2p}:

$$C_{2p} = \sqrt{\frac{Z_0 \cdot R_{Ls} - R_{Ls}^2}{\omega Z_0 R_{LS}}} \tag{11.5}$$

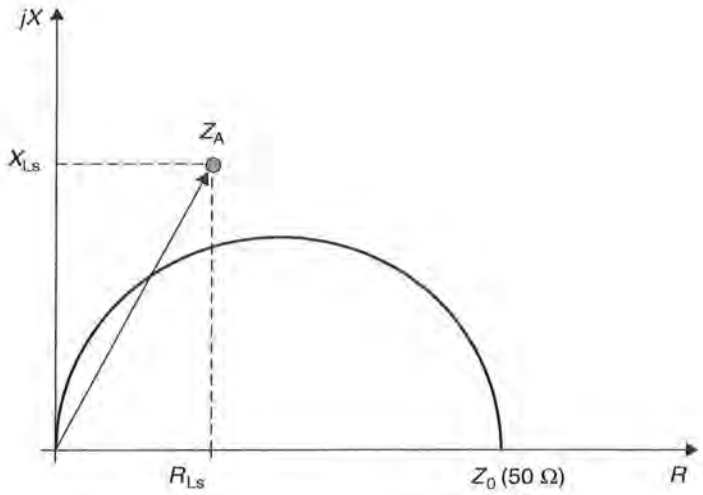


Figure 11.20 Representation of Z_A in the impedance level (Z plane)

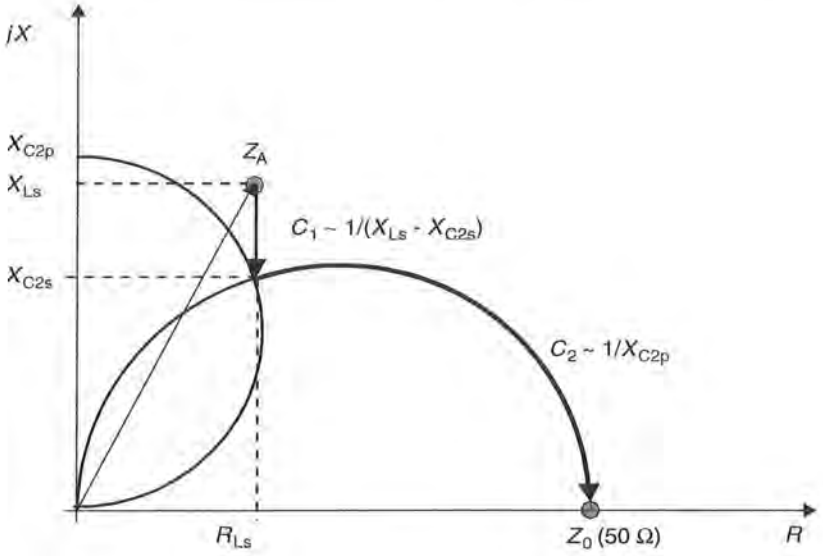


Figure 11.21 Transformation path with C_{ls} and C_{2p}

As is clear from the impedance plane in Figure 11.21, C_{2p} is determined exclusively by the serial resistance R_{ls} of the antenna coil. For a serial resistance R_{LS} of precisely 50 Ω, C_{2p} can be dispensed with altogether; however greater values for R_{ls} are not permissible, otherwise a different matching circuit should be selected (Fricke *et al.*, 1979).

We further find for C_{ls}:

$$C_{1s} = \frac{1}{\omega^2 \cdot \left(L_s - \dfrac{\sqrt{Z_0 R_{LS} - R_{LS}^2}}{\omega} \right)} \tag{11.6}$$

The antenna current i_{LS} is of interest in this context, because this allows us to calculate the magnetic field strength H that is generated by the antenna coil (see Chapter 4).

To clarify the relationships, let us now modify the matching circuit from Figure 11.18 slightly (Figure 11.22).

The input impedance of the circuit at operating frequency is precisely 50 Ω. For this case, and only for this case(!), the voltage at the input of the matching circuit is very simple to calculate. Given a known transmitter output power P and known input impedance Z_0, the following is true: $P = U^2/Z_0$. The voltage calculated from this equation is the voltage at C_{2p} and the serial connection of C_{ls}, R_{ls} and X_{LS}, and is thus known. The antenna current i_2 can be calculated using the following equation:

$$i_2 = \frac{\sqrt{P \cdot Z_0}}{R_{Ls} + j\omega L_s - j\dfrac{1}{\omega C_{1s}}} \tag{11.7}$$

11.4.3 The influence of the Q factor

A reader antenna for an inductively coupled RFID system is characterised by its resonant frequency and by its *Q factor*. A high Q factor leads to high current in the antenna coil and thus improves the power transmission to the transponder. In contrast, the transmission bandwidth of the antenna is inversely proportional to the Q factor. A low bandwidth, caused by an excessively high Q factor, can therefore significantly reduce the modulation sideband received from the transponder.

The Q factor of an inductive reader antenna can be calculated from the ratio of the inductive coil resistance to the ohmic loss resistance and/or series resistance of the coil:

$$Q = \frac{2\pi \cdot f_0 \cdot L_{coil}}{R_{total}} \tag{11.8}$$

The bandwidth of the antenna can be simply calculated from the Q factor:

$$B = \frac{f_0}{Q} \tag{11.9}$$

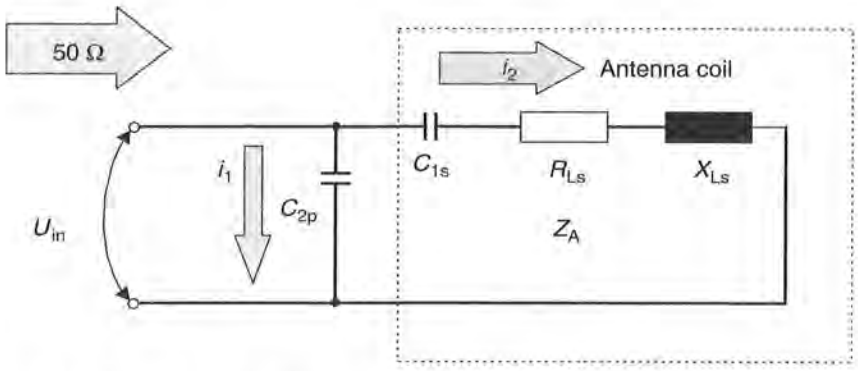


Figure 11.22 The matching circuit represented as a current divider

The required bandwidth is derived from the bandwidth of the modulation sidebands of the reader and the load modulation products (if no other procedure is used). As a rule of thumb, the following can be taken as the bandwidth of an ASK modulated system.

$$B \cdot T = 1 \tag{11.10}$$

where T is the turn-on-time of the carrier signal, where modulation is used.

For many systems, the optimal Q factor is 10–30. However, it is impossible to generalise here because, as already mentioned, the Q factor depends upon the required bandwidth and thus upon the modulation procedure used (e.g. coding, modulation, subcarrier frequency).

11.5 Reader Designs

Different types and designs of readers are available for different applications. Readers can be generally classified into OEM readers, readers for industrial or portable use and numerous special designs.

11.5.1 OEM readers

OEM readers are available for integration into customers' own data capture systems, BDE terminals, access control systems, till systems, robots, etc. OEM readers are supplied in a shielded tin housing or as an unhoused board. Electrical connections are in the form of soldered, plug and socket or screw-on terminals. See Figure 11.23.

Figure 11.23 Example of an OEM reader for use in terminals or robots (photo: Long-Range/High-Speed Reader LHRI, reproduced by permission of SCEMTEC Transponder Technology GmbH, Reichshof-Wehnrath)

Table 11.2 Typical technical data

Supply voltage:	Typically 12 V
Antenna:	External
Antenna connection:	BNC box, terminal screw or soldered connection
Communication interface:	RS232, RS485
Communication protocol:	X-ON/X-OFF, 3964, ASCII
Environmental temperature:	0–50 °C

Table 11.3 Typical technical data

Supply voltage:	Typically 24 V
Antenna:	External
Antenna terminal:	BNC socket or terminal screw
Communication interface:	RS485, RS422
Communication protocol:	3964, InterBus-S, Profibus, etc.
Ambient temperature:	−25–+80 °C
Protection types, tests:	IP 54, IP 67, VDE

Table 11.4 Typical technical data

Supply voltage:	Typically 6 V or 9 V from batteries or accumulators
Antenna:	Internal, or as "sensor"
Antenna terminal:	—
Communication interface:	Optional RS232
Ambient temperature:	0–50 °C
Protection types, tests:	IP 54
Input/output elements	LCD display, keypad

Figure 11.24 Reader for portable use in payment transactions or for service purposes. (Photo of LEGIC® reader reproduced by permission of Kaba Security Locking Systems AG, CH-Wetzikon)

11.5.2 Readers for industrial use

Industrial readers are available for use in assembly and manufacturing plant. These usually have a standardised field bus interface for simple integration into existing systems. In addition, these readers fulfil various protection types and explosion protected readers (EX) are also available.

11.5.3 Portable readers

Portable readers are used for the identification of animals, as a control device in public transport, as a terminal for payments, as an aid in servicing and testing and in the commissioning of systems. Portable readers have an LCD display and a keypad for operation or entering data. An optional RS232-interface is usually provided for data exchange between the portable readers and a PC.

In addition to the extremely simple devices for system evaluation in the laboratory, particularly robust and splash-proof devices (IP 54) are available for use in harsh industrial environments.

RFID HANDBOOK

Fundamentals and Applications in Contactless Smart Cards and Identification

Second Edition

KLAUS FINKENZELLER

Giesecke & Devrient GmbH, Munich, Germany

Translated by Rachel Waddington, *Swadlincote, UK*

Developments in RFID (Radio-Frequency Identification) are yielding larger memory capacities, wider reading ranges and quicker processing, making it one of the fastest growing sectors of the radio technology industry.

RFID has become indispensable in a wide range of automated data capture and identification applications, from ticketing and access control to industrial automation. The second edition of Finkenzeller's comprehensive handbook brings together the disparate information on this versatile technology. Features include:

- Essential new information on the industry standards and regulations, including ISO 14443 (contactless ticketing), ISO 15693 (smartlabel) and ISO 14223 (animal identification).
- Complete coverage of the physical principles behind RFID technologies such as inductive coupling, surface acoustic waves and the emerging UHF and microwave backscatter systems.
- A detailed description of common algorithms for anticollision.
- An exhaustive appendix providing listings of RFID associations, journals and standards.
- A sample test card layout in accordance with ISO 14443.
- Numerous sample applications including e-ticketing in public transport systems and animal identification.

End users of RFID products, electrical engineering students and newcomers to the field will value this introduction to the functionality of RFID technology and the physical principles involved. Experienced ADC professionals will benefit from the breadth of applications examples combined within this single resource.

WILEY
wiley.com