University of Oslo
Department of Informatics

Communication System for a Medical Implant

Vemund Svanes Bertelsen

11th April 2006

Preface

This thesis is part of the work done to fulfill the requirements for the degree Candidatus Scientiarum at the University of Oslo, Department of Informatics. The thesis was part of a larger project aiming to develop a new glucose sensor utilizing osmotic pressure differences undertaken at Institute of Information and Communications Technology, department of Instrumentation and Microelectronics at Sintef Oslo. The project was funded by Lifecare A/S. However the resources needed for the work presented here were funded by Sintef and the University of Oslo.

The author of this thesis would like to express his sincerest appreciation of the support from my tutors Senior research scientist at Sintef ICT Oslo and Associate professor II at the University of Oslo Department of Informatics, Joar M. Østby, and scientist at Sintef ICT Oslo Morten Berg.

Also, I would like to thank for help and guidance I have received from people at Sintef and the University of Oslo.

Contents

Chapter 1

Introduction

The main focus of this thesis is two different approaches for establishing wireless communication, by means of electromagnetic fields, between an implanted glucose sensor and an external device in a glucose monitoring system for diabetics. In the first approach the implant is constructed without an internal energy source. Here the energy is transferred to the implant via an inductive link, which at the same time is used for communication. In the second approach the internal device is equipped with an internal energy source. The discussion focuses on limitations and ways of optimizing the performance of the two approaches for the application in question. The nature of the operating medium, the human body, is complex and this causes challenges not faced with short range systems in more traditional environments. Although the discussion is based on the design of a communication system for a glucose monitoring system, it will be of relevance for other applications of short range communication systems on medical implants or with similar operating environments.

1.1 The focus of this thesis

The communication system will be an important part of a medical system and should therefore ensure the degree of reliability expected from this type of equipment. As already mentioned, the main focus of this thesis will be directed against a solution for a communication system utilizing

electromagnetic fields. It is important to identify and understand parameters that are, from the communication system's point of view, vital for the performance of both the complete system and the communication system itself. One of the main challenges will be to meet demands for miniaturization without compromising the demands for performance. Two different approaches are discussed:

- **Passive Solution** - Passive refers to the fact that the system needs external stimulation to operate. It is designed this way to enable the implanted device to operate without any form of internal energy source. In this case the energy needed to perform a glucose reading is transferred via the same electromagnetic fields as used for communication. The aim is to investigate the possibilities and limitations of this approach. The discussion will focus on inductive energy transfer, more specifically on how to maximize this transfer in the given environment. Hence, the design of the transfer coils and of the internal and external communication circuits will be devoted much attention. The work done in connection with this part of the thesis will be based on a literature survey, experimental tests and computer simulations.

- **Active solution** - In this case the implant is equipped with an internal power source. The focus will be on how to minimize the power consumption of the implanted communication system. The discussion will concentrate on different techniques and schemes for transfer of information as well as choice of antenna. As the choice of transfer scheme is not independent of the electronic design of the transceivers, some attention must be drawn to this topic as well. This part of the thesis will be based on a literature survey.

To understand the operating conditions and demands for performance for the communication system in question, we need basic knowledge of other parts of the system and how they can work together to form a complete system. Interactions between the communication system and other system circuits might degrade performance.

Another important property of the internal device is it's power consumption. In an active solution, the internal power source would be decisive for both weight and size of the implanted device. Therefore, we will need to obtain an overview of different power sources and their properties. To find the limitations of a passive solution we need an estimate of the amount of power that must be transferred to do a single measurement. Hence we need an estimate of the amount of the power consumed on the internal device during a single measurement.

A medical implant is surrounded by human tissue. Hence its surface material must be bio-compatible. This implies that the material does not cause harmful reactions when in contact with biological tissue. Some materials will cause an immune reaction, others are toxic.

The communication signals will have to travel some distance through the tissue. Therefore the electromagnetic properties of human tissue must be considered to obtain an optimal design of the communication system. The EM fields connected with the communication signals and the tissue will interact and this interaction might affect the tissue. The question would be if such interactions could be harmful.

1.2 Possible System Solutions

As is discussed closer in section 2.1, to best adjust the treatment of diabetes, it is advantageous with an almost continuous monitoring of the blood glucose level. One of the most serious limitations of traditional glucose monitoring systems is the pain associated with obtaining the needed number of blood samples every day. This makes it difficult to convince diabetics to do the necessary number of tests to best treat their disease. Besides being painless, the system should be comfortable to wear. The level of user involvement necessary for the system to operate properly, is also important for user acceptance. The ideal monitoring system would be completely automated, only interfering with the users daily business when something out of the ordinary is detected. To ease data analysis, there should be extensive data acquisition and processing capabilities. This combined with a communication link and software on a PC or another

suitable device, could help extract useful information from the large amount of data.

The patient will have to undergo an invasive procedure [1] to get the implant placed inside the body. Such a procedure will always inflict a risk of infection on the patient as the skin provides an important outer barrier against foreign bodies. The procedure will require sterile environment and equipment. The physical shape and size of the implant must be designed to ease handling and make the invasive procedure as simple as possible. This will minimize both strain inflicted on the patient and workload for the medical staff involved.

To facilitate manufacturing and possible mass production it would be preferable if the system could be made independent of differences in physical appearance of the users. Individual adjustments always makes production more expensive.

A simple solution could be to place the implanted device close to the wrist of the hand. This would permit communication with an external device included in a wristwatch. The short distance between the implanted device and the watch makes it quite easy to establish a communication link. It might even turn out that the passive solution suggested above would work with this approach. The short distance between the two units may enable a reasonable transfer of energy from the external to the internal device. To switch or charge a battery in an external unit is easy compared with charging a unit inside the body. The wrist is also quite similar among humans. So, complicated individual adjustments or costly flexibility would not be necessary.

Preliminary investigations undertaken by Sintef and Lifecare indicated that the area close to the wrist was unsuitable for glucose level measurements. The measured levels in body fluid seemed to change too rapidly and was heavily dependent of muscular activity. The tests indicated that the body fluid in the abdomen is more suitable for such measurements, as the conditions were more stable here. The technical challenges concerning the communication system are greater in this location though. One of the

[1] invasive procedure - the skin has to be cut in order to carry out the procedure

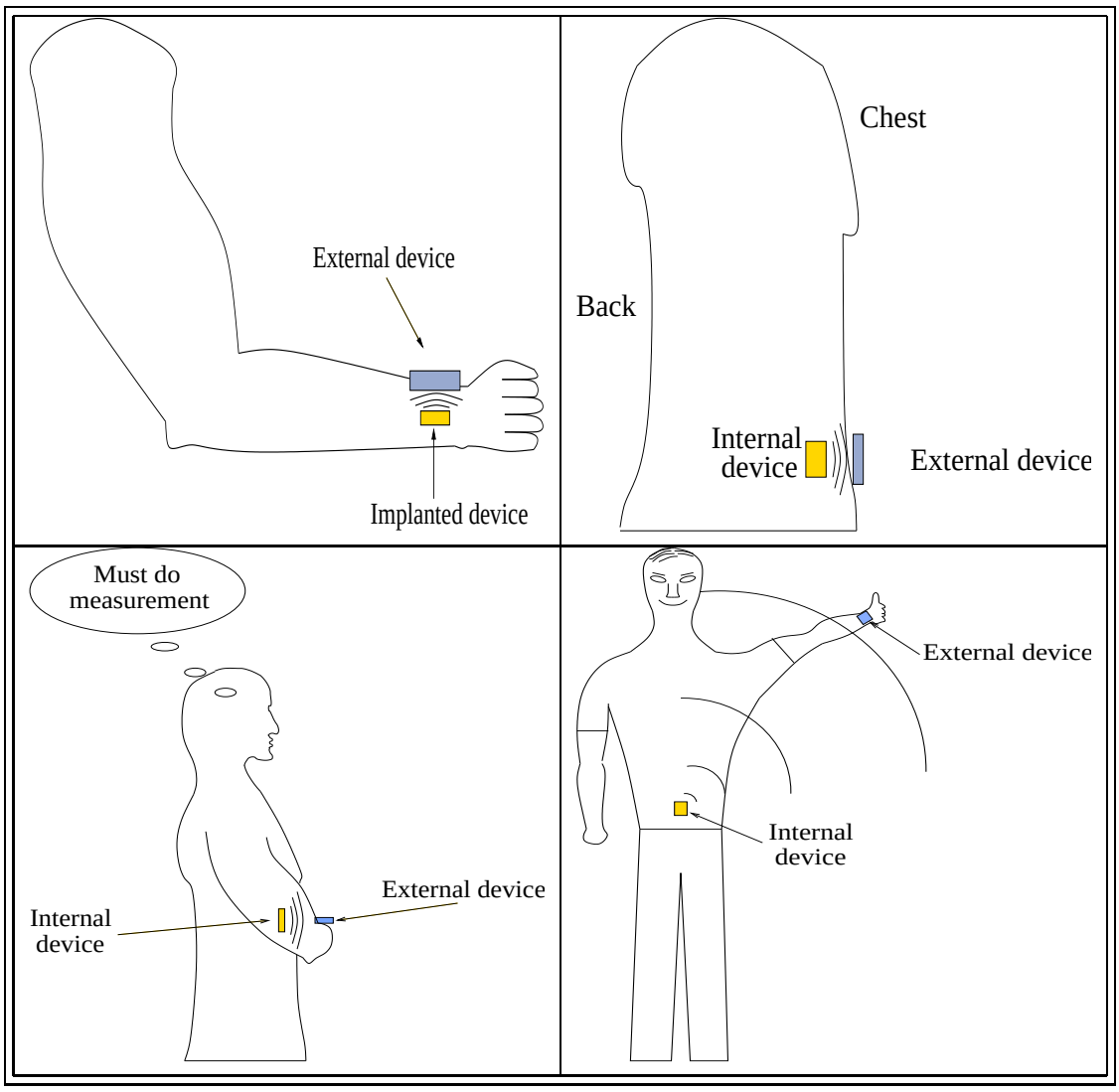


Figure 1.1: *Illustrations of different combinations of implanted device and external device placement. Upper left: Implanted device communicates with clock, short range. Upper right: Implanted device communicates with external device in a belt. Bottom left: The external device must be positioned near the implant to establish communication link. Bottom right: A system with larger operating range.*

main concerns in the design of the communication system, is the limited capacity for energy storage on the implanted device. Wireless transfer of energy would be preferable compared to equipping the implant with a battery. If wireless transfer of energy is attempted over distances larger than a few centimeters, the amount of energy received, is very small compared to the amount used to transfer it. Increased operating distance also means the energy required by the communication system will increase.

A possible solution might be to put the external device into a belt buckle. This would minimize the operating distance, but the belt buckle would be less flexible in use than a wristwatch. Further, there are large differences in the shape of people's abdomen. Type 2 diabetes is considered a lifestyle related disease. Therefore it would not be unreasonable to expect that many diabetics are overweight. A large abdomen implies a larger operating distance. The relatively large attenuation of EM fields in body tissues at radio frequencies implies even harder conditions for the communication system. The passive solution with inductive energy transfer that seemed feasible for the wristwatch solution, would probably not work in this case. The required operating distance is too large.

Another solution, still presuming the implanted device is in the abdominal region, could be to involve the user more in the execution of a glucose measurement. If the user could move a watch containing the external part of the system as close to the implanted device as possible and trigger a reading, the operating distance would be minimized and a passive solution might still be feasible. This solution demands a quick measurement. The whole process of triggering, measuring and transferring the data back to the external device must take no longer than a few seconds. An advantage is that a watch can be warn under nearly all circumstances. This would definitely be an improvement compared to many of today's systems, but does not meat the requirements of the ultimate glucose monitoring system; a completely automated system that monitors the blood sugar 24 hours a day and warns of abnormal levels or changes in the blood glucose level.

Presuming an active solution, increased complexity of the internal circuit could provide "smart" schemes to save power. If abnormal levels are

detected, the internal device could perform a very powerful transmission of a warning signal. Under normal operation the user could trig the internal device when data is needed and position the device in a more favorable position for data transmission. Although this will increase the power consumed by internal circuitry, it will result in an overall power budget gain as digital circuits can be designed to operate with very low power consumption, while the analog RF transceiver circuits requires more power.

In practice, a major shortcoming of today's standard glucose sensors is their dependence of a reactant. The reactant is little by little consumed when performing tests. This makes them less suitable for use in longterm medical implants. As mentioned earlier the goal of this project was to utilize a new idea for sensing glucose levels. The success of this project depended on whether or not such a sensor could be made with today's technology. So far, there has been lithe progress in the work with the sensor.

Chapter 2

Diabetes and Glucose Monitoring Systems Today

2.1 Short Description of Diabetes and its Treatment

Diabetes[1] is a condition where regulation of the blood sugar level is not functioning properly. Untreated the condition leads to too high blood sugar levels [27]. The illness can be caused by reduced or total loss of production of insulin or that the body cells sensitivity towards insulin is reduced. Insulin is a hormone that stimulates the cells' absorption of nutrition. Together with a couple of other hormones, the body uses insulin to regulate the absorption of nutrition. When a person eats and digests food, the secretion of these hormones increase. To put it easy, a lowered insulin level, or a reduction in the cells sensitivity toward insulin, increases the blood sugar level because less glucose is absorbed by the cells.

Diabetes has two main variants. **Type 1** was formerly known as insulin-dependent. Sufferers from this condition does not produce insulin as a result of damage to the cells in the pancreas that normally would produce

[1]The information about current treatment was mainly retrieved from the web sites of the Norwegian Diabetes Association [48], the American Diabetes Association [3] and the World Health Organization [69].

insulin. The damage is often caused by an autoimmune reaction[2]. These patients are dependent upon regular injections of insulin. Some patients carries a pump that provides a continuous supply of insulin. Without insulin the body's cells will not be able to absorb glucose. In absence of insulin, the body starts to break down fats for fuel [50]. A metabolic byproduct of fat metabolism is referred to as a ketone. The presence of elevated blood ketones in this setting, is known as diabetic ketoacidosis. In extreme, untreated cases, this can lead to coma and death.

Type 2 was formerly called non-insulin-dependent. Despite the former name, about one third of type 2 patients also need insulin injections. Type 2 diabetes is caused by a complicated interplay of genes, environment, insulin abnormalities, increased glucose production in the liver, increased fat breakdown, and possibly defective hormonal secretions in the intestine [66]. The condition can develop over a long period of time. The cells become resistant to insulin [12]. To compensate for the lower absorption of glucose in the cells, the pancreas increases production of insulin and the liver releases more glucose from its glucose storage. Eventually the pancreas becomes less able to produce insulin and the cells become even more resistant. As a result the blood glucose levels slowly start to rise. This process can go on for several years before symptoms of the disease appear. It seems there is no full understanding of how this condition arises.

Diabetes is treated with a strict diet, with a low intake of carbohydrates. In some cases the patient is given anti diabetics which increases the effect of insulin and stimulates the insulin production. To keep the right balance between the blood sugar level and the insulin level, it is recommended that patients with diabetes monitor their own glucose level. This is done with a device called a glucose monitor. One measurement means a pinprick in a fingertip to get a sample of blood. The fingertip has a large amount of pain receivers. Most patients do not think of it as a problem to carry out this procedure once or twice a day. But to be in control of the blood sugar level some diabetics should check it four to five times a day. Some patients consider the repeated pinpricks quite difficult [37].

[2]Autoimmune reaction - a malfunction in the body's immune system that causes it to attack body cells.

Long term complications of diabetes are more likely to occur if the patient is not given well adjusted treatment. Increased risk of heart disease, eye disorders, kidney disease and nerve damage are some of the complications that a diabetic may have to face. Several studies have found a connection between the average blood sugar level and the risk of long term complication. According to the American Diabetes Association, the best known study is the Diabetes Control and Complications Study conducted in the United States [3]. The results were published in 1993 in the New England Journal of Medicine. These proved among other things that among type 1 patients, improved blood glucose control prevents or delays diabetic retinopathy[3]. Therapy that kept blood sugar levels as close to normal as possible reduced damage to the eyes by 76%. Other studies have shown the best control of the glucose level is achieved by more frequent self-monitoring. In "Self-monitoring of Blood Glucose levels and Glycemic Control: the Northern California Kaiser Permanente Diabetes Registry" [30], Karter et al. concluded that "More frequent self-monitoring of blood glucose levels was associated with clinically and statistically better glycemic control regardless of diabetes type or therapy". Monitoring the blood sugar level continuously or at a sufficient rate, would give useful information for the required balancing between the intake of nutrition and the dosing of insulin.

According to the World Health Organizations (WHO) latest figures, the number of people suffering from diabetes is approximately 177 million [69]. The number is expected to increase to 300 million by the year 2025. Type 2 is the most common of the two types constituting around 90 percent of the cases. The estimated figure for diabetes related deaths is 4 million, 9 percent of the global total.

Diabetes causes human suffering and economical costs to society worldwide. Depression, anxiety, pain and other factors that result in discomfort are intangible costs of diabetes [69]. The disease often strikes people in their most productive years. The economical cost of diabetes to society is very difficult to estimate. Some estimates have been made though, and

[3]retinopathy - disorder of the retina (the part of the eye where the light sensitive cells are found)

they show that the cost of loss in production as a result of diabetes related sufferings often equals or exceeds the direct health costs. The WHO gives examples of such estimates, and for the US (in 1997) the cost of care was estimated to 44 billion USD while the loss in productivity was 54 billion USD. Similar estimates conducted for 25 Latin American countries indicated that the figure of loss in production was as much as five times the costs of care. This may have been caused by low access to good care, causing higher rates of complications and early deaths [69].

Type 2 diabetes, the most common, is regarded as a lifestyle related disease. The best way to prevent the problems related to it would be a change in lifestyle, from a passive life with fast food to a life with regular exercise and a more healthy diet. But even though there has been brought significant attention to this problem in developed countries, the problem just seems to increase. The next best solution would then be to give the affected a good treatment so that they can go on living a productive life, contributing to society. If there is found a method for monitoring blood sugar levels at higher rates than todays solutions it would be decisive for its success that the solution is cost effective. This would make it profitable for society to invest in such systems as diabetes related loss of production would decrease.

2.2 Glucose Monitoring Systems Available Today

Most currently available glucose monitoring systems designed for self-testing are based on measuring techniques that require blood samples. However, much research is now directed towards painless and automated systems. The Cygnus GlucoWatch Biographer, described at the end of this section is by many considered the first commercially available system that attempts to provide painless and automatic testing.

The equipment needed to conduct a glucose reading with a typical system of today, consists of a device called a sampler, test strips and a meter. The sampler is used to obtain a blood sample. A small needle penetrates the skin when the sampler is activated. The depth of penetration can be adjusted. Usually a finger tip is the site used for a test. Some newer meters

offer the possibility of obtaining the blood from so called alternative site testing (AST). This kind of testing is usually limited to the upper and lower arm. However, users have reported problems and there are controversies among the expertise on how reliable these measurements are, even if the recommendations and restrictions given by the manufacturers are followed [41]. A drawback of AST is the difference between tests performed on the arms and tests performed on fingertips. Research have also indicated that AST delays the detection of hypoglycemia (low blood sugar) [8,29]. Some users have reported that AST leaves needle marks [62] and therefore perhaps avoid them for esthetic reasons. Despite this, AST can offer a welcomed rest of sore fingertips under the right conditions [8].

When the blood has been obtained with the lancet, it is applied to the "test spot" of a test strip. The actual measurement can now be carried out. The test spot has a coating of chemicals that react with the blood and makes it possible to extract the blood glucose level. Two different measurement principles are used with currently available glucose meters, one optical and one electrochemical. They are discussed closer in section 2.4.

We see that glucose testing with todays system can be quite an elaborate procedure. To reduce the user "workload", Glucerna's Precision SofTact offers lancing, blood collection and glucose testing with a single press of a button, but it takes 20 seconds before the result is displayed [42].

The meters are only specified to work well within certain air temperature ranges. Common operating temperatures are from a minimum of $5 - 18^oC$ up to a maximum of $30 - 40^oC$. This complicates testing under very warm and very cold conditions.

Meters using the electrochemical principle can be manufactured in such a way that they don't require special maintenance or cleaning. The optical meters however, needs cleaning after each test. With these kind of meters the test strip is inserted into a chamber where an optical device performs the actual measurement. When a test is finished, remains after the test must be cleaned out, so they don't effect the result of the next test.

Meter manufacturers recommend that the equipment is carried in special cases for protection. The test strips are especially vulnerable. They are

kept in special bags to protect them from light, moisture and contamination, but still have limited lifetime. Especially when they are taken out of the protection bags to be applied to the meter, they are very vulnerable to contamination.

The possibility to store results from tests taken over a longer period of time is a welcomed functionality. This makes important statistical parameters, like averages, trends, maximums and minimums over a longer period of time easily accessible, which again helps keep a better long term glycemic control. Although newer meters have the capability of storing at least 100 test results, users still complain that this is not enough. If you test 6 times a day, you need a storage capability of 186 tests to be able to store results for one month and maybe you would want some in reserve. Diligent users of the meter would need to transfer the test data to a PC in order to fully utilize the value of their effort. The PC-links available use either a cable connection or a wireless IR-link.

The Cygnus Glucowatch Biographer is one of the first systems attempting to fulfill the goal of a painless and automated system. The device offers non-invasive and automatic measurements for people of age 18 and older [18]. It is warn like a watch. A sensor pad called the AutoSensor is first attached to the skin. Then the watch is placed upon it and is thereby connected to the AutoSensor. The system makes a measurement every 20th minute. Every AutoSensor lasts for twelve hours offering up to three measurements per hour, thus up to 36 per wear.

To work satisfactory the system has to go through a three hour warm-up period, before it is calibrated using the result from a traditional glucose meter as reference [18]. Even after this calibration, there is substantial uncertainty associated with the measurements. Every fourth reading will differ with more than thirty percent from a standard glucose meter. It is worth noticing that the GlucoWatch system measures not on blood, but on interstitial fluid. The fluid sample is obtained by applying a small electrical current across the skin. The glucose is then extracted from the sample and is drawn into hydrogel disks in the AutoSensor. Here it reacts with glucose oxidase. The result of this reaction is hydrogen peroxide which again reacts on a platinum plate to produce an electrical current

which is measured. The result is translated into a blood glucose value by means of a data conversion algorithm. A glucose sample obtained with the GlucoWatch will lag about 15 minutes behind a theoretical blood sample test taken at the same time.

The FDA stresses that the GlucoWatch is not meant to replace a regular glucose meter. There are several conditions under which it does not operate optimally. These include, low blood sugar levels, bumps and extensive perspiration. Because of the uncertainty associated with measurements, it is not recommend to make big adjustments in the treatment of the illness based on results obtained with the GlucoWatch. Also many users get skin irritations from wearing the AutoSensor. Despite its teething problems, the GlucoWatch is the first in a new generation of meters allowing close to continuous control and painless testing. It was approved for sale in the US in march 2001.

2.3 New solutions under development

Several ongoing projects are investigating the possibilities for developing either non-invasive, minimal invasive or implanted glucose monitoring devices. Many different approaches are attempted. They span from measuring on either blood or interstitial fluid by minimal invasive techniques and implanted techniques to non-invasive solutions where electromagnetic properties of glucose is utilized to detect it either by impedance spectroscopy or looking at how it interferes with a magnetic field. Good information on progress and technology are often hard to get hold of either because the companies want to keep others from looking into their cards or maybe they do not have a very strong hand. In this subsection, some projects that have presented technical information about what they are doing and with goals of painless and/or continuous measuring devices are presented. Most of the information was obtained from the diabetesnet homepage and the home pages of the developing companies [1, 4, 14, 54].

The Animas Corporation is developing a sensor that operates without enzyme systems. The goal is to develop a sensor that does not need regular calibration with a standard meter. In fact the project intends to read gluc-

ose values directly from blood. In a surgical procedure a C-clamp is placed around a 4 to 5 mm blood vessel. Two small probes are inserted through the walls of the vessel to allow transmission of a clean infrared light signal between them. Also placed inside the body is a larger device housing. It contains a laser generator, producing the IR signal for the probes, electronics for signal analysis and battery. At the current state of development this housing is about twice the size of a pacemaker. With a system for a rechargeable battery this size could be reduced to half. Readings would be available every 2 to 5 minutes. The company reports that tests have shown good correlation between a Hemocue glucometer and their own, based on the new technology. The company home page states that preliminary human trials were to start in late 2003. Although promising preliminary results have been found, the company stresses that much work remains.

CME Telemetrix Inc. has named their glucose monitor project GlucoNir. The detection device uses a beam of light in the Near-IR range that is focused on a person's finger for about half a minute. By applying mathematical techniques on the data extracted from the detected emerging light, the concentration of various blood analytes including glucose are determined. The company claims to have proprietary technologies to improve performance using advanced statistical analysis. Their findings await verification by independent researchers. Although this technique would be painless, the device must be manually activated. This makes continuous measuring an unavailable feature with this glucose monitor, but the user would find it acceptable to do measurements more often as they do not involve pain. CME has also initiated clinical trials, but they were delayed by the SARS outbreak in Toronto. Results were expected in late 2003.

The swiss company Pendragon Medical Ltd., has developed a glucose monitor which achieved a Medical Device Directive CE mark certification in may 2003. This led to mass production of the first 10000 units. Larger field trials were also initiated. In appearance the device is quite similar to a standard wrist watch. The detection technology is based on impedance spectroscopy. The impedance measurement is done by an open resonant circuit which lies against the skin. The different types of molecules in the skin and tissue that is excited will contribute differently to the resulting

Figure 2.1: *Left: Overview of the Animas system. Right: Picture illustrating the size of the kumetrix needle.*

impedance pattern, as every type of molecule will have a unique relaxation behavior over a relevant span in the frequency. This makes it possible to optimize the device for detecting the effects of glucose molecules on the impedance pattern. The measurements can presumably be converted into glucose concentrations by applying a specially developed mathematical transformation.

The Pendragon watch needs a two point calibration procedure to adjust to individual differences in skin and underlying tissue. This calibration process provides individual user data about absolute offset and the ratio between impedance changes and glucose changes. Temperature changes also affect the impedance patterns, but the device is designed to adjust to such changes during operation.

The system gives updated glucose values every minute and alerts the user when there is an increased risk of hypoglycemia, when glucose values are to high or low or when the rate of change in the glucose value is too high. The company's alpha series has the possibility for storing 24 hours of raw data and a USB computer interface. The company presents few test results, only some from a preliminary test which the company views as proof of concept for their approach.

A company called Kumetrix is developing a system utilizing silicon microneedels. The system consists of a handheld device and is designed to work after much the same principles as todays meters. The meter holds a

cartridge loaded with up to ten disposable sampling devices. Each disposable device is a silicon chip designed to draw blood and perform chemical measurements. The needle has a thickness approximate to that of a human hair as seen in figure 2.1. At the end of each needle is a micro-reservoir or cuvette. Presumably it is here the chemical analysis of the blood is to take place.

As Kumetrix informs on their home page, their inspiration comes from nature. A mosquito bite in itself is painless and often passes unnoticed by the victim. It is enzymes injected by the mosquito that causes irritation.

2.4 Current Glucose Sensor Technology

As mentioned in section 2.2, two different measuring principles are used with today's glucose monitor systems. One principle is based on an optical sensor while the second is based on a current sensor. In both cases a blood sample must be mixed with certain chemicals before the actual measurement can take place. The test strips mentioned above are used for this purpose. The test spot where the blood sample is applied is coated with the chemical mixture.

The optical measuring technique is the oldest and is becoming obsolete. This is partly because it requires more maintenance than meters utilizing the electrochemical sensing technique. The optical technique is based on colorimetry. Glucose oxidase and peroxidase are the main active constituents in the chemical coating of the test spot. Glucose from the blood sample and oxygen react in the presence of glucose oxidase yielding gluconic acid and hydrogen peroxide [35]. Hydrogen peroxide subsequently oxidizes dyes included in the coating. In a reaction mediated by peroxidase, a blue colored form of the dyes is produced. The intensity of this blue color is proportional to the glucose concentration in a sample and is measured by an optical device on the meter.

The electrochemical measuring technique is newer. In 1990 Kawaguri et al. [38] described such a sensing principle used in a glucose sensor. They enhanced the enzyme reaction by using potassium ferrocyanide as a mediator. First the enzyme reaction produces an amount of ferrocyanide pro-

portional to the amount of glucose, equation 2.1. Then the reduced mediator is electrochemically oxidized and the glucose concentration in the sample can be determined by the resulting electrical current, equation 2.2.

$$\text{Glucose} + 2\text{Fe(CN)}_5^{3-} \rightarrow \text{Gluconic acid} + 2\text{H}^+ + 2\text{Fe(CN)}_5^{4-} \quad (2.1)$$

$$2\text{Fe(CN)}_5^{4-} \rightarrow \text{Fe(CN)}_5^{3-} + \text{e}^- \quad (2.2)$$

At the time a $3\mu l$ blood sample was needed for the sensor to work, and it took 60 seconds before the result could be obtained. With today's meters the blood sample is reduced to $1\mu l$ and the time to wait for the result is 5 seconds.

Chapter 3

System Architecture

In this chapter possible system architectures are discussed. First a general overview of the two approaches discussed in chapter 1, the passive and active solutions, are presented. Next, a more thorough description of some important components of the system is given. The purpose of this discussion is to shed light on what limitations these components imply for the design of the communication system, especially in terms of power consumption. In this context the glucose sensor under development and possible power sources are considered most interesting. As the communication system is the focus of the thesis, it is left for discussion in chapter 4.

3.1 System Overview

The system consists of two main units; the implanted unit and the external unit. The active and passive communication system designs will require different overall system designs; consult figure 3.1. The design of a system with a passive implant is considered first. The implant will have to include a communication system circuit and a sensor. Depending on the CS design, power circuits and control logic, including A/D-conversion, are necessary for the system to work.

The external unit includes the overall system control. It will be responsible for measurement control. This implies initiating measurements

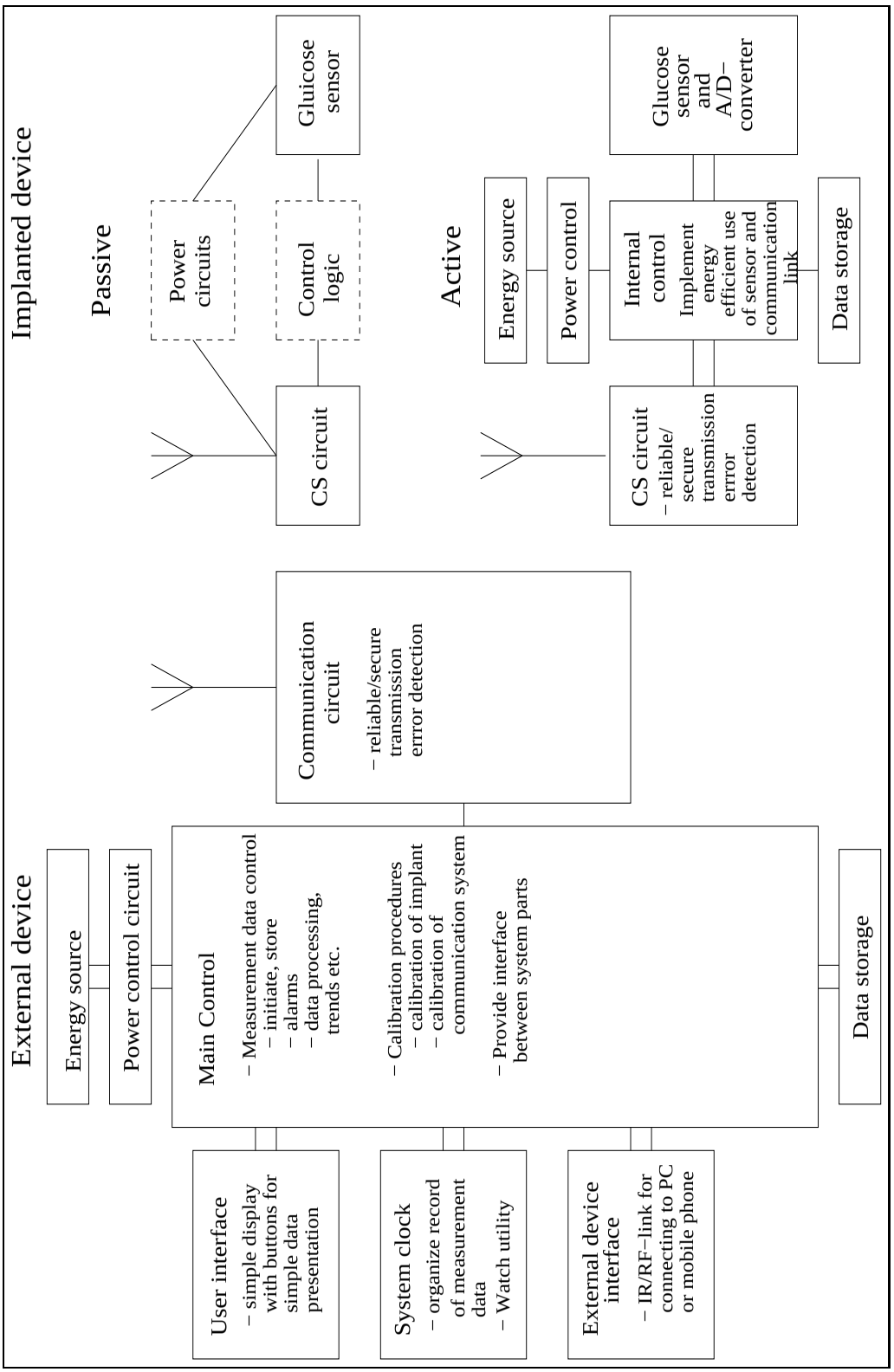


Figure 3.1: *Schematic overview of possible system designs.*

and processing and storing measurement data. In the case of a worrying glucose reading, the user must be notified by an alarm. The external unit should also include a display to give the user easy access to measured values and trends that will help the user keep glycemic control. For this utility an appropriate storage capacity and a system clock to keep record of the different measurements are needed. To be able to use measurements for further analysis and help in working out treatment plans, it should be easy to transfer measurement data to external devices like PCs or mobile phones. A standard wireless RF-chip would be a good candidate for providing this utility. Finally, the system needs a power source and the appropriate power management circuitry.

An active implant gives greater freedom in the design of the implant. The internal power source enables the implant to store measurements and do simple data processing. The implant can be programmed to decide when it is appropriate or necessary to transmit data to the external device. This could prove useful for efficient use of power for transmission. The larger operating range can cause interference with other systems using the same operating frequency. To prevent this a transmission protocol with an appropriate addressing system could prove necessary.

The external unit will in the case of an active implant not differ much form the one for the passive implant. The differences lie mostly in the communication circuitry which will be discussed in chapter 4.

3.2 The Glucose Sensor under Development

As mentioned in chapter 1 the sensor under development is based on osmotic pressure differences. The physical phenomenon of osmotic pressure is illustrated in figure 3.2. A semipermeable membrane is placed across a U-shaped tube at the bottom of the tube. The semipermeable membrane blocks for molecules larger than a given size, but lets the smaller ones pass. In the figure the small water molecules can flow through the membrane but the larger particles of the dissolved substance can not.

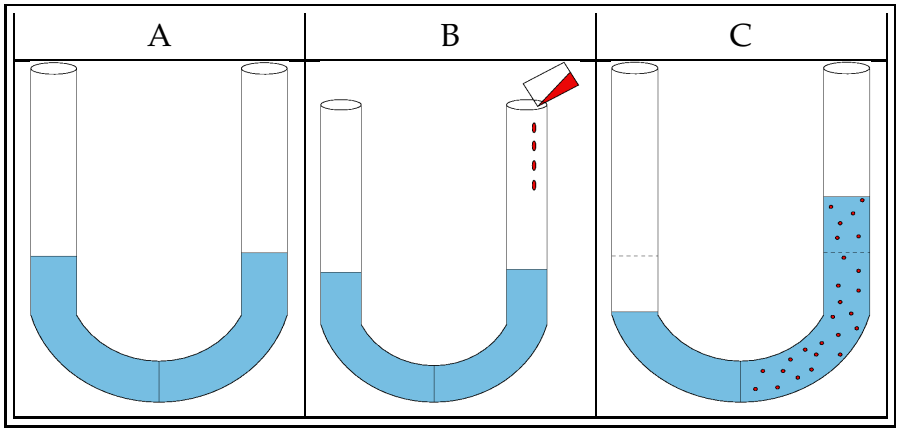


Figure 3.2: *Illustration of the principle of osmotic pressure*

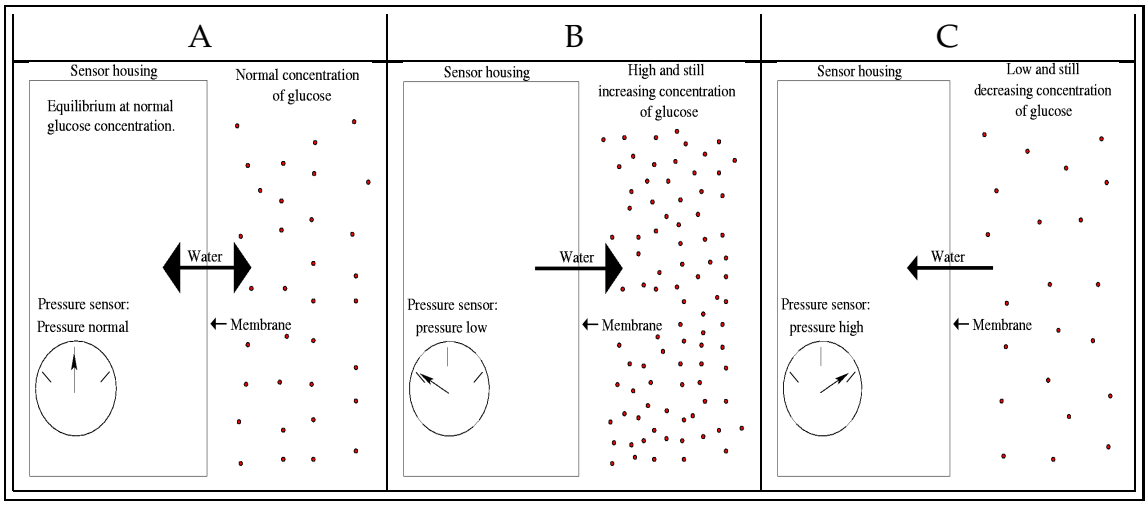


Figure 3.3: *Sketch of sensor principle*

The higher concentration of the dissolved substance on the right side of the tube causes a chemical potential to arise over the membrane. This results in net diffusion of water through the membrane, from the left side to the right, to equalize the concentration difference and hence reduce the chemical potential over the membrane. Eventually an equilibrium arises between the hydrostatic pressure due to gravity and the osmotic pressure caused by the concentration differences.

Figure 3.3 shows how the sensor would work ideally when deployed in interstitial fluid[1]. The leftmost illustration shows the situation with normal glucose concentration. The pneumometer in the figure represents a pressure sensor. It now indicates "pressure normal". In the next illustration the glucose level is higher and still increasing. Water flows out of the chamber and into the interstitial fluid implying lower pressure in the chamber. If the glucose concentration in the interstitial fluid gets very low, like in the rightmost illustration in figure 3.3, water will flow into the chamber and increase the pressure in the chamber to a level above "pressure normal" in part A of the figure. Sensing the pressure inside the chamber gives an indication of the glucose concentration in the interstitial fluid.

A problem is that electrolytes present in the interstitial fluid will affect the osmotic pressure differences over the membrane and hence affect the measured pressure. When converting the pressure measurements into glucose levels, concentration differences caused by other substances than glucose must be accounted for. This can perhaps be achieved by introducing a second sensor chamber. If one chamber is equipped with a membrane permeable only for smaller particles and the membrane on the other chamber was permeable for both smaller particles and glucose. This could enable elimination of the offset caused by the small sized particles.

However, a membrane does not have infinite selectivity with respect to particle size. When a membrane is said to have a "cut-off" at a certain particle size, it implies particles of this size will encounter a certain diffusion resistance when crossing the membrane. Smaller particles encounter less resistance, larger particles higher. The diffusion resistance of a cer-

[1]The fluid filling the space between cells

tain particle size for the sensors membrane will decide the response time the sensor will have for measuring changes in the concentration of that particle size. Hence a difficult trade off arises. To get as little disturbance in the measurement as possible from particles larger than glucose, a membrane that cuts off very close to the size of glucose is desirable. However this will imply a large diffusion resistance for glucose molecules as well and the sensor would have problems picking up rapid changes in the glucose concentration. If a larger cut-off is chosen the response time would decrease but more disturbance from larger particles would also arise. If the double chamber sensor suggested above is used, the two chambers will have membranes with different cut-offs. Hence, the two chambers will have different response times for particles of the same size. This could prove difficult to account for in the data processing.

The body will encapsulate any foreign body with tissue. This will further increase the response time of the sensor because of increased diffusion resistance. This may also affect the long term stability of the sensor.

The flexibility of the membrane can also affect the pressure measurements inside the chamber. If the membrane is flexible, different osmotic pressure differences over the membrane will not only cause pressure changes, but changes in the chamber volume as well. Hence the sensitivity of the sensor is degraded.

Gas bubbles can form in the chamber reference fluid from gases dissolved in the intracellular fluid. Bubbles occurring and dissapearing on the inside of the chamber will introduce unpredictable changes in the reference volume, which again will affect the osmotic pressure. To avoid this an overpressure is needed at the inside of the chamber. By using a reference fluid containing large protein molecules such an overpressure can be achieved. However the long term stability of this overpressure is not known.

The properties of the pressure-to-electric output signal generated by the sensor, will be significant for the system design. Miniaturized pressure sensors can be fabricated using silisium micromachining techniques. Such sensors can be made piezoresistive or capacitive. The piezoresistive sensors change resistance with varying pressure while capacitive sensors

change capacitance. Preliminary theoretical investigations done at Sintef suggested that the pressure sensor inside a chamber should withstand pressures ranging from 200kPa to about 1000kPa to be safe under extreme conditions. If the variations caused by glucose can be extracted through the use of a reference chamber, these variations would be from 2.5 to 40 kPa. 1kPa is used as an estimate for needed sensitivity of the pressure sensor.

In [15] DeHennis and Wise presents a capacitive pressure sensor designed for intracranial pressure monitoring. The sensor has a pressure range from $400mmHg$ to $1000mmHg$. Judging from the frequency shift the capacitance causes over this range in a parallel resonance LC-tank, the capacitance changes from approximately $40.5pF$ at the lowest pressure to approximately $53.6pF$ at the highest pressure; approximately 25% change in the capacitance over the pressure range in question. The size of the chip including both the sensor and the inductor is $6mm \times 6mm \times 0.5mm$. The sensitivity is $20fF/mmHg \approx 150fF/kPa$.

In [7] Berg uses a Sensonor piezoresistive pressure sensor in a demonstrator designed for intracranial pressure measurement. When the pressure is changed from 0 to 2 bar this sensor produces a bridge voltage of approximately 25mV and approximately 75mV, respectively, given a supply voltage of 5V. This corresponds to a sensitivity of $33\mu V/mmHg \approx 248\mu V/kPa$. The pressure element has a size of approximately $1cm \times 1cm$.

The sensitivities of both these sensors seems adequate for the purpose in question.

There are differences between glucose level measurements done in interstitial fluid and measurements done in blood. Not only does the level differ, but also how quick changes can be detected differ. This can cause difficulties for converting and comparing the two types of measurements. It will be necessary as blood glucose level is used as a reference standard today.

A long term medical implant for measuring glucose levels would be totally dependent of a sensor that can operate without reactant supplies. The sensors used in glucose monitoring devices today require reactants and are therefore unsuitable.

3.3 Power Supply

The power supply is an important system component. It is a critical component for the implanted device in an active solution. The capacity of the power source is a crucial factor for the performance of the system in terms of longevity and operation distance. An active implant needs a power source that is reliable, secure and small in size.

Two categories of energy sources are discussed in this section. The first is energy storing devices (for example batteries). Large capacity implies increased size and weight. As the size of the implant is important, a tradeoff involving operation distance, the rate of glucose readings and size of energy sources could prove to be critical for the active solution.

The second category of devices harvest energy from their near environment and convert it into electrical energy. This process is referred to as energy scavenging. However, the power such devices can deliver is too small to directly power other than very low power devices. If energy scavenging is found to be a feasible power solution for the system in question, small storages of energy is needed to supply the bigger bursts of power needed when measuring and transmitting data.

Today, commonly used medical implants are powered by either an internal battery or an external power source. External power sources are either direct electrical or pneumatic links [64] or wireless inductive links. Generally, wires passing through the skin should be avoided because of increased infection risks and restrictions on freedom of movement.

3.3.1 Battery

The battery represents a mature and well tested source of energy for medical implants. They have been used in pacemakers since the sixties. The most important factor for an implantable battery is reliability [64], especially for life sustaining implants like cardiac stimulators. Once the battery is mounted on the implantable device, it cannot be removed. It is hard-wired and hermetically sealed at the time of manufacture.

Because of the unusual demands concerning implanted batteries and

the highly litigious environment of the medical industry, most battery manufacturers do not allow the use of their products in an implantable medical device [64]. Only a few specialized manufacturers make batteries for medical implants. The leading provider of power sources for medical implants is Wilson Greatbatch Technologies with about 90% of the market of power sources for cardiac stimulators [9]. Some manufacturers of implantable medical devices choose to integrate specially customized batteries into their applications.

Batteries can be divided into two categories, primary (not rechargeable) and secondary (rechargeable). Primary batteries are the preferred choice for medical implants. Listed here are some of the reasons:

- Primaries provide higher energy densities than secondaries.
- Recharging batteries in an implanted device is cumbersome and requires user compliance for safe operation.
- If a primary battery can provide the energy required for the implant's lifetime, it simplifies the design; no need for battery charging circuits.
- Implants are not reused. The demand for sterility is strict. Most medical equipment with demand for this level of sterility is made disposable.

There are several available battery chemistries. They have different strengths and weaknesses. For a medical implant, small weight and size is important. In other words, we seek a battery with high specific energy (energy/mass $[Wh/kg]$) and high energy density (energy/volume $[Wh/dm^3]$). In figure 3.4 the energy density of some commonly available battery chemistries are shown. The lithium chemistry stands out with the highest energy density. Indeed, batteries based on lithium chemistries have the highest specific energy and energy density of any type [44] and is the preferred choice for powering medical devices.

Wilson Greatbatch Technologies Inc. (WG) provide four different cell chemistries (Information mainly from the company's internet homepage [71]):

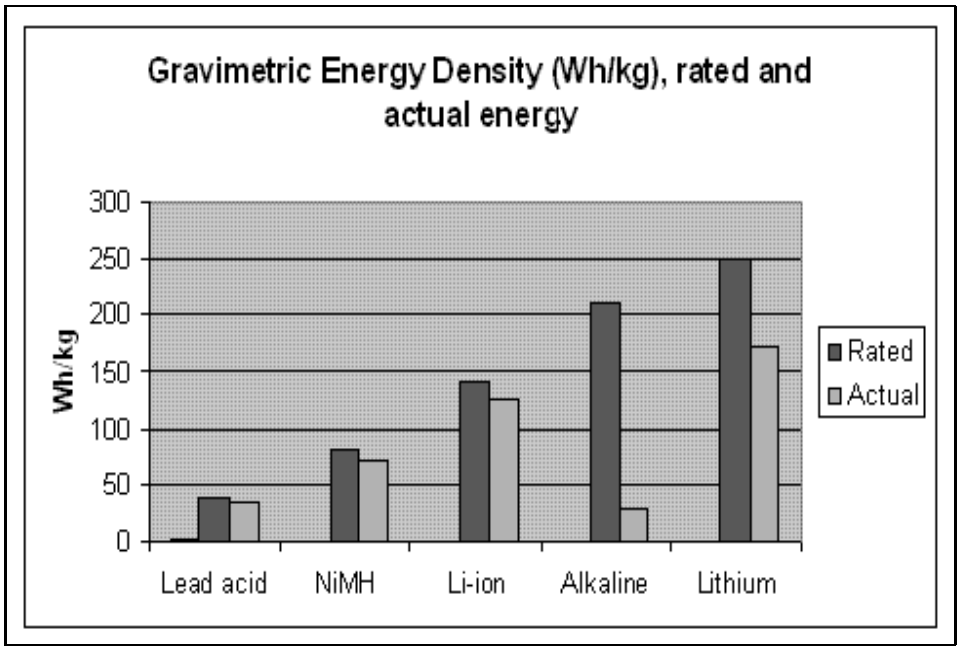


Figure 3.4: *Gravimetric energy density or specific energy for some common cell chemistries. The darker bars are the rated numbers, while the brighter ones apply for a load similar to a digital camera. Source:* `www.batteryuniversity.com` *[11].*

Property	**Model**		
	8830	9085	9086
Weight	$27.9g$	$11.0g$	$17g$
Volume	$10.3cm^3$	$2.92cm^3$	$7.9cm^3$
Capacity	$2.2Ah$	–	$2.5Ah$
OCV-BOL	$3.2 \pm 0.10V$	$2.8V$	$3.3V$
Specific energy	$0.15Wh/g$	$0.24Wh/g$	$0.38Wh/g$
Energy density	$0.41Wh/cm^3$	$0.91Wh/cm^3$	$0.82Wh/cm^3$

Table 3.1: *Specifications of some cell models produced by Wilson Greatbatch Technologies. Information retrieved from Wilson Greatbatch Technical Product Guide. OCV-BOL is short for "open circuit voltage at beginning of life".*

- **Lithium iodine** can deliver currents in the μA-range. Provides reliability and high longevity. The solid electrolyte gives the battery several advantages including no potential for leakage, increased reliability and extended shelf life of the battery itself (up to 10 years) [64]. The most serious limitation of the solid electrolyte is its low ion mobility which implies high internal impedance. This explains the limited currents the battery can deliver, and so limits its use to low-drain applications such as pacemakers. In table 3.1 specifications for the model 9085 is given (rated current $50\mu A$).

- **Lithium silver vanadium oxide (High rate)** is capable of delivering ampere level current pulses, 0.5 to 2 A. A standard configuration available is the 8830 cell model. It's specifications are listed in table 3.1. These types of cells are often customized to meet demands of a specific application.

- **Lithium silver vanadium oxide (Medium rate)** These cells are generally capable of current drains in the mA-range and they can handle moderate-rate pulse currents (50mA). Typical applications are neurostimulators and drug infusion devices.

- **Lithium/Carbon Monofluoride Cells (Medium rate)** are designed to handle currents in the μA to mA range. They have low internal impedance and high energy density. Typical applications are neurostimulators and pacemakers. Data for the model 9086 is shown in table 3.1.

The medical implant in question would require relatively large currents when the system is performing measurement and communication operations. In other words, the battery must tolerate pulsed power consumption and support larger drains of current for short periods of time. The large drains must be expected to be at least in the mA-range. Hence, of the technologies provided by WG the lithium/silver vanadium oxide cell seems to be the most suitable.

There are many topics concerning the performance of batteries. Aging causes increased self discharge. The relatively high operating temperature

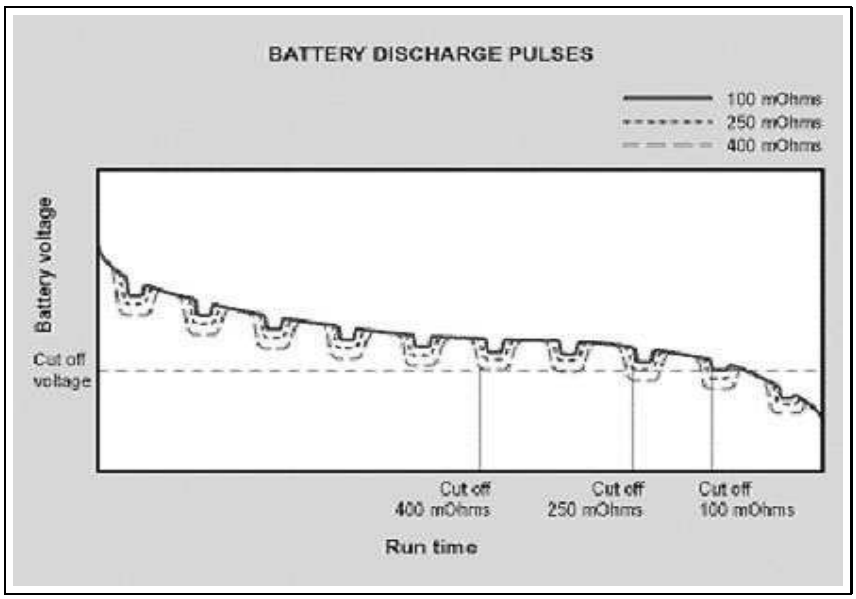


Figure 3.5: *The figure demonstrates the effect of internal resistance on the cutoff voltage of the battery. Source: www.batteryuniversity.com [11]*

of 37°C means an increased internal resistance. It is also worth while considering how a battery's internal resistance affects its capability of managing applications pulsed power consumption. Low internal resistance is one of the urgent requirements of a battery for such equipment [11]. Figure 3.5 demonstrates how different internal resistances affect the cutoff voltage of the battery. The state of charge can also affect the internal resistance of batteries. However, lithium-ion batteries have a fairly flat internal resistance as a function of state of charge [11].

Lithium-ion batteries are a good alternative for rechargeable batteries. For the application in question, they have several advantages. They provide the largest energy densities of rechargeable batteries and 300-500 discharge/charge cycles [11]. Lithium-ion batteries are low maintenance and do not require periodic discharge [11]. However, it is better to recharge before the battery is fully discharged, as a full discharge puts an unnecessary strain on the battery. Lithium-ion cannot absorb overcharge and no trickle charge is applied on full charge. They also gradually loose charge acceptance as part of aging. In general lithium-ion batteries should

Figure 3.6: *Miniature battery; the battery to the left powers the neurological stimulator to the right.*

not be stored for long periods.

The battery is still a limiting factor in many applications due to its limited capacity. The average annual gain in capacity is typically 6% [11]. Within battery technology research, a significant effort is devoted to developing the lithium chemistries.

Miniaturization is also an issue concerning low power medical implants making it possible to do the implantation by a minimally invasive procedure. One such miniature battery is developed at Quallion LLC and Argonne National Laboratory (in the US). It powers a neurological stimulator and is about 1/35 the size of a standard AA battery [5] (see figure 3.6) and a capacity of 3mAh [56].

A different approach in battery miniaturization is being pursued by an Israeli company called Paper Power [32]. They have developed a battery that can be printed onto paper. It is only 0.5 mm thick and provide 20 mAh of current at a terminal voltage of $1.5V$ for every square centi meter printed. The chemical composition of the electrolytes is a company secret, but is claimed to be non-toxic. Another advantage is that the battery can be cut into any size or shape. The battery has already been used with a

chip designed by the German company KSW Microtech to monitor the temperature of blood supplies.

Fuel Cells

A fuel cell is an electrochemical device in which a fuel and an oxidant react to directly produce electricity [20]. It consists of two electrodes and an electrolyte or ion-conducting membrane. Fuels are supplied at both electrodes. Catalysts are included in each electrode to speed up chemical reactions. To give a simplified explanation of the processes taking place in a fuel cell we can use one of the most common type of fuel cells, the Polymer Electrolyte Membrane (PEM) Fuel Cell (also called proton exchange membrane fuel cells). A simplified drawing indicates the working principle in figure 3.7. Hydrogen gas (H_2) is supplied at the anode. The catalyst, which in this case is platinum (Pt), helps separate the protons and electrons in the H_2-molecules. The protons are led through the electrolyte towards the cathode. Hence an electric potential is set up between the two electrodes. If an electrical conductor is connected between them, an electrical current will flow from the cathode to the anode. At the cathode the hydrogen reacts with oxygen from the air, producing water. Some excess heat is also produced in this process.

The relatively low energy density of hydrogen gas and the PEM fuel cell operating temperature of 80°C makes this type of fuel cell unsuitable for use in medical implants. A type of fuel cells called direct methanol fuel cells (DMFCs) uses methanol as fuel. Methanol has larger energy density than hydrogen gas and DMFCs can operate at temperatures from 20°C to 60°C. This class of fuel cells can provide energy densities up to five times that of lithium-ion batteries [11]. The chemical reactions taking place are:

$$\text{Anode: } CH_3OH + H_2O \xrightarrow{Pt/Ru} CO_2 + 6H^+6e^-$$

$$\text{Cathode: } 6H^+6e^- + \frac{3}{2}O_2 \xrightarrow{Pt} 3H_2O$$

Yen et. al [72] presents a micro methanol fuel cell (μDMFC) which can operate near room temperature. At 40°C it produces 24.8mW per square centi meter of reaction surface. When giving this output power the μDMFC

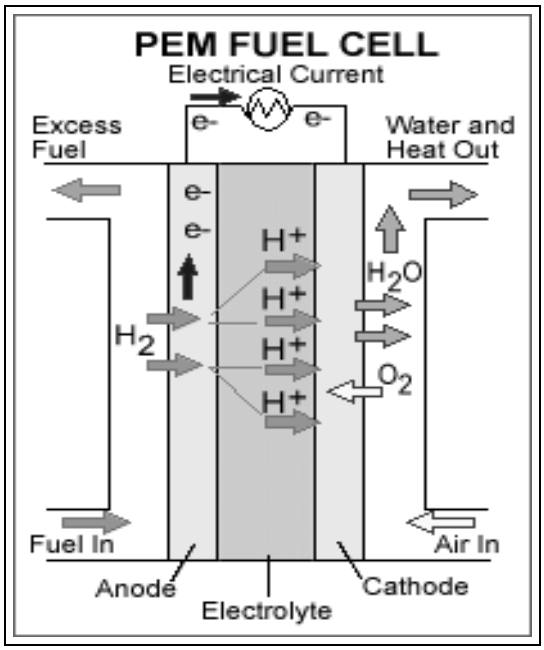


Figure 3.7: *Working principle of a PEM fuel cell. Source: U.S. Department of Energy and Renewable Energy [65].*

was fed with 0.283ml 1M aqueous methanol per minute and 88ml of air per minute. However, using this kind of fuel cell on the inside of a human body seems quite farfetched. Methanol and oxygen[2] would both have to be stored inside of the body. Methanol is toxic and oxygen in form of gas has low energy density.

The technique used in DMFC's is relatively new. Comprehensive technical information is scarce. So far this technology has only been implemented in laboratory prototypes. Recent research at Brown University (see [10]) have resulted in prototypes, utilizing the same principle as the DMFCs, that can run on other types of fuels such as glucose or formates (derivatives from methanoic acid). The Brown fuel cells do not require an ion-conducting membrane. Future technology might enable needed fuels to be extracted from the body and hence solve the problem of larger fuel storages inside the human body.

Another prototype class of fuel cells are biological fuel cells. Here the noble metal catalysts used in traditional chemical fuels cells are exchanged with micro organisms or enzymes [24]. These fuel cells can convert the chemical energy of sugars and alcohols, into electrical energy. For instance the theoretical current capacity in 1g of methanol is approximately 5000mAh. However 100% efficiency is not achievable. According to Halme et. al [24] experimental tests with bacterial fuel cells have shown that the complex reactions slows down the conversion rate to 15-25% to that of chemical fuel cells. The efficiency has in experiments been measured to 40-55 % when the substrate was glucose. The group presenting these results expects biological fuel cells based on enzymes to have a better conversion rate and an efficiency better than 50 %. Results are not yet available. Theoretical calculations presented indicate an energy density about 1/5 of a AA NiMh battery. The microbe fuel cell has even lower energy density.

Researchers at Saint Louis University have also investigated the enzyme approach. They have chosen an enzyme that converts ethanol to acetaldehyde, removing a proton in the process [2]. The proton is then added to nicotinimide adenine dinucleotide. The fuel cell's electrode strips

[2]Extracting it from blood seems quite farfetched.

the proton back off to produce electricity. The researchers claim that the enzymes can theoretically last forever given the proper environment. However their prototypes only remains active for some weeks.

Another research group has developed a biofuel cell that converts the energy produced when glucose reacts with oxygen during normal metabolism directly into electricity. The fuel cell consists of two carbon fibers, 2 cm long and 7 μm wide [6]. One electrode is coated with a polymer and glucose oxidase, which strips glucose of electrons. The polymer makes an electrical connection between the enzyme and the carbon fiber. At the other electrode, another polymer-wired enzyme adds electrons to dissolved oxygen. It produces about 1.9 mW, about the same amount of power as a wristwatch battery. Currently it looses around 6% of its power per day in model fluids. Real body fluids are more complicated than the model fluids that the researchers have used so far. Deploying them in a more realistic operating environment may introduce complications.

The fuel cell technology can be a promising technology for the future, but clearly there is much research work left before units capable of competing with batteries as the energy source in medical implants can be produced.

Energy Scavenging

As explained in the introduction to this section an energy scavenging device can harvest energy from the environment surrounding it. Solar panels are a well known and widely used example of such a device. Standard silicon panels can produce $15 mW/cm^3$ when exposed to direct sunlight [60]. However, a technique for energy scavenging inside the human body has not yet been developed. This section discusses a novel approach to energy scavenging that might have applications in medical implants.

In some research communities considerable effort is devoted to develop devices that can extract electrical energy from vibrations. The main motivation is use in self-sustaining nodes in wireless sensor networks in areas where more mature and well tested techniques for energy scavenging is not permitted by the operating environment. Vibration energy oc-

curs in the human body for instance as a result of movement and heart beats.

The energy conversion can be obtained utilizing different techniques. The movement of a magnet in the close vicinity of a coil would according to Faraday's law result in alterations in the magnetic flux through the coil. This would then induce an emf in the coil. The current resulting from this emf can be utilized to produce electrical power. An alternative approach would be to utilize a capacitor where the plates can move in relation to each other. Given an initial charge, a change in the distance between them because of absorbed vibrational energy, would induce voltage alterations and electrical energy could be drained from the device. A last approach is to utilize piezo electric crystals. When external mechanical forces are applied to such crystals, voltage differences occur across it. These voltage differences can be utilized to drain electrical energy from the device.

Of the above mentioned techniques, the one utilizing piezo electric crystals is perhaps the technique which has been devoted the most attention. Shenck and Paradiso [45] have designed and tested piezo electric shoe inserts. The solution that showed the best results produced an average of 8.4mW when the load was $250k\Omega$ and the walking pace 0.9 Hz. Roundy [60] gives a thorough discussion on the subject and has also tested different approaches in a test environment with vibrations similar to that of the casing of a microwave oven. The results show that power in the order of 200 $\mu W/cm^3$ can be achieved.

Roundy [61] has also done research on the capacitor concept mentioned above using MEMS-fabricated capacitors. Simulation results show 100 $\mu W/cm^3$ can be achieved with this approach. Meso and micro scale prototypes have been used as proof of concept.

The amount of energy that can be converted from vibrations in a given environment is dependent on the frequency spectrum of the vibrations. Information on deploying this kind of energy scavenging device inside the human body was not found in literature, nor was data on vibrations magnitude or frequency spectrum. However, it seems unlikely to achieve output powers even close to those presented above.

Energy source	Energy density	Specific energy	Comments
Battery SVO	$0.41Wh/cm^3$	$0.15Wh/g$	Commercially available for medical use
Fuel cell	-	-	Prototypes only, currently not well suited for use in implants
Vibrational Piezo Crystals Capacitors	 $200\mu W/cm^3$ $100\mu W/cm^3$	-	Prototypes only, numbers not valid for medical implant environment

Table 3.2: *Overview of the different energy sources discussed in this section that could be used in a medical implant.*

Summary

At the time of writing, batteries seems to be the only power source capable of meeting the demands of a medical implant. Other power solutions have been considered, but were not found suitable for this application. As far as could be established, the battery is the only source of power used with available wireless implants today. Other technologies are under continuous development, but the utilization of them in products available to the public seems several years ahead. Large efforts in research and development is still to be done before any of them can be utilized in medical implants.

3.3.2 Power Management Circuits

Independent of what power source is chosen, additional circuitry is necessary to supply different parts of the system with the right voltage level and provide the necessary current source capacity. The AC-voltage provided by a vibrational energy scavenging device must be rectified and converted to a suitable voltage before it can be utilized, for example to charge a battery. The voltage provided by a battery may not be suitable for all

the system circuits. DC-DC converters can be used to supply the needed voltage levels. In a circuit driving variable loads, there may be a need to keep the supply voltage level constant, independent of the load current. A voltage regulator provides this feature. The operation of each of the above parts will consume some energy. Different designs have different advantages and disadvantages. The main concern for this application will be efficiency and size without compromising other important factors. One such factor could be the noise from a switching regulator. A more thorough discussion of this part of the system was considered beyond the scope of this thesis.

Chapter 4

The Communication System

The primary objective of the communication system is wireless transfer of data from an implanted medical device to an external data acquisition and controlling system. Wireless transfer of data is achieved through RF electromagnetic fields. The propagation of such fields subject to conditions relevant to the application of interest is discussed in 4.1 giving a basis for theoretical prediction of signal transmission.

For operating ranges up to a few centimeters, the passive solution introduced in chapter 1 is of interest. This solution will be discussed closer in section 4.2. For larger operating ranges the active solution is the only suitable and it is discussed in section 4.3. To carry information the RF signals must be modulated. Relevant modulation techniques are discussed for both solutions.

4.1 Signal transmission

According to the "Antenna Engineering Handbook" [28] the field radiated from an antenna can be divided into three regions depending on the distance from the antenna. In the region close to the antenna, the field will include a reactive component. The strength of this reactive component decays rapidly and at greater distances from the antenna the radiation field dominates. The radiation field is divided into a radiating near field region and a radiating far field region. The boundaries between the differ-

ent regions are not precisely defined, but are best described as transition regions. The boundaries are among other things, dependent upon an antenna's design and electrical length.

The communication system of a passive solution has a very short operating range, typically a few centimeters. The operating frequency is often in the lower MHz region (1-30MHz) implying a free space wavelength of 10 meters or more. Hence this type of communication system operates in the reactive field region.

For a communication system based on an active solution the situation is somewhat different. Several available frequency bands of interest are located from 300 to 1000 MHz with a corresponding free space wavelength from 30cm to 1m. Supposing the system has a reasonably low power consumption, the operating range can be from several meters to several tenths of meters. Hence a study of the fields associated with an antenna in this application could involve all three regions.

To get a better understanding of the field regions discussed above, the electromagnetic fields of a single loop antenna are studied in more detail below. The loop antenna is chosen because its characteristics are similar to that of a multi-loop antenna or single layer coil. Coils with this geometry are relevant alternatives for the communication system of the passive solution. A loop antenna could also be a relevant alternative for the active solution.

Since $\nabla \cdot \mathbf{B} = 0$ under all conditions the magnetic field is derivable from the magnetic vector potential. As shown in appendix D.1 the magnetic vector potential must satisfy the following inhomogeneous differential equation:

$$\nabla^2 \mathbf{A} - \mu\epsilon \frac{\partial^2 \mathbf{A}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{A}}{\partial t} = -\mu \mathbf{J}' \tag{4.1}$$

Here $\mathbf{J}'$ is the source current density (A/m^2) of the loop. It is related to the signal current $\mathbf{I_0} = I_0 \cos \omega t$ running in the loop as $\mathbf{J}' = \frac{\mathbf{I_0}}{\Delta s}$ where *Deltas* is the cross-section of the loop. The loop is electrically small which means that the ratio a/λ is $\ll 1$, *lambda* being the wavelength of the electromagnetic fields associated with it. Hence the current is approximately constant along the loop. The wire is assumed to be of infinitesimal thickness. Thus

for this special case we can set $\mathbf{J}' = I_0 \cos \omega t$.

Figure 4.1 shows the loop antenna and an observation point **P** in the xz-plane. A spherical coordinate system is used to describe P's position relative to the loop. At the point P, a change in the electromagnetic fields due to a change in the current will be delayed by the time $R/v_p = \beta R$, where v_p is the phase velocity and β is the phase constant of the electromagnetic fields. This can be seen in the approximate solution for the magnetic vector potential presented in appendix D.1. The approximation is valid for a material medium with moderate dissipation ($\alpha a \ll 1$)given that $r \gg a$:

$$\begin{aligned} \mathbf{A} \approx & \frac{\mu a^2 I_0 e^{-\alpha r}}{4} \sin\theta \\ & \left[\left(\frac{1}{r^2} + \frac{\alpha}{r} \right) \cos(\omega t - \beta r) - \frac{\beta}{r} \sin(\omega t - \beta r) \right] \mathbf{i}_\phi \end{aligned} \tag{4.2}$$

Here α is the attenuation constant given in Np/m.

4.1.1 The electromagnetic fields from an electrically small single loop antenna in free space

An exact solution for the fields generated by a thin circular loop antenna in free space can be found as shown by Werner [67]. However, to achieve a qualitative estimation of the fields, we can use the solution in equation 4.2. In free space there are no free charges. This implies that the conductivity and hence also the attenuation constant α is zero. The result of equation 4.2 can then be written:

$$\mathbf{A} \approx \frac{\mu_0 a^2 I_0 \sin(\theta)}{4} \left[\frac{1}{r^2} \cos(\omega t - \beta r) - \frac{\beta}{r} \sin(\omega t - \beta r) \right] \mathbf{i}_\phi \tag{4.3}$$

The magnetic field strength can now be found from the definition of **A**.

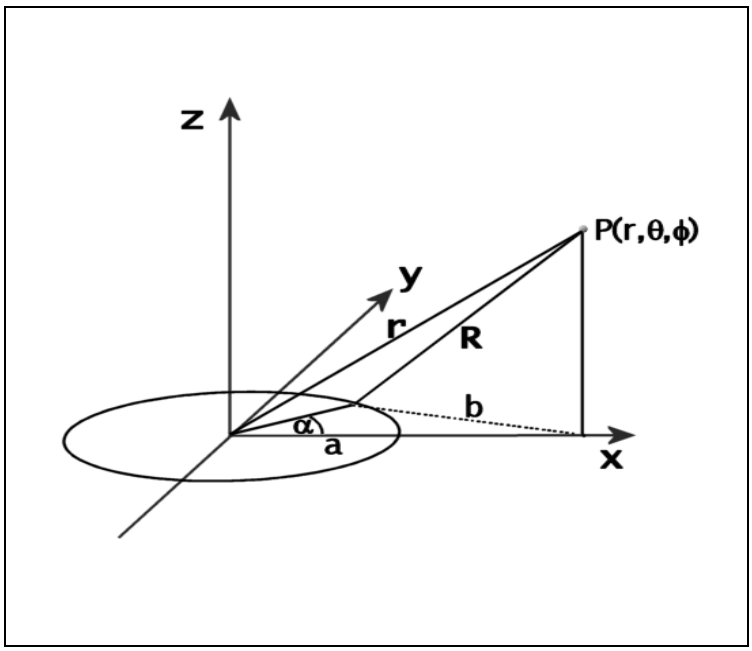


Figure 4.1: *The geometry of a single loop antenna.*

The magnetic field intensity **H** is:

$$\begin{aligned}
\mathbf{H} &= \frac{1}{\mu_0}\nabla \times \mathbf{A} \\
\mathbf{H} &= \frac{a^2 I_0}{4}\left[2\cos(\theta)\left(\frac{1}{r^3}\cos(\omega t-\beta r)-\frac{\beta}{r^2}\sin(\omega t-\beta r)\right)\mathbf{i}_r\right. \\
&\quad + \sin(\theta)\left(\frac{1}{r^3}\cos(\omega t-\beta r)-\frac{\beta}{r^2}\sin(\omega t-\beta r)\right. \\
&\quad \left.\left.-\frac{\beta^2}{r}\cos(\omega t-\beta r)\right)\mathbf{i}_\theta\right] \qquad (4.4)
\end{aligned}$$

The electric field is connected to the magnetic vector potential through Maxwell's curl equation for the electric field as given by equation D.3, which give:

$$\begin{aligned}
\mathbf{E} &= -\frac{\partial \mathbf{A}}{\partial t} \\
&= \frac{\mu a^2 I_0}{4}\left[\frac{1}{r^2}\sin(\omega t-\beta r)+\frac{1}{r}\beta\cos(\omega t-\beta r)\right]\mathbf{i}_\phi \qquad (4.5)
\end{aligned}$$

A power flow is associated with the fields. It is often expressed by Poynt-

ing's vector which describes a power density (W/m^2):

$$\begin{aligned}\mathbf{P} &= \mathbf{E}\times\mathbf{H}\\ &= \frac{\mu a^4 I_0^2\omega}{16}\left(\sin^2(\theta)\left[-\frac{1}{2r^5}\sin(2(\omega t-\beta r))+\frac{1}{r^4}\beta-\frac{1}{r^4}2\beta\cos^2(\omega t-\beta r)\right.\right.\\ &\quad\left.+\frac{1}{r^3}2\beta^2\sin(2(\omega t-\beta r))+\frac{1}{r^2}\beta^3\cos^2(\omega t-\beta r)\right]\mathbf{i}_r\\ &\quad+2\sin\theta\cos\theta\left[\frac{1}{2r^5}\sin(2(\omega t-\beta r))-\frac{1}{r^4}\beta+\frac{1}{r^4}2\beta\cos^2(\omega t-\beta r)\right.\\ &\quad\left.\left.-\frac{1}{2r^3}\beta^2\sin(2(\omega t-\beta r))\right]\mathbf{i}_\theta\right)\end{aligned} \tag{4.6}$$

In the expressions for the fields and in Poynting's vector there are terms with different dependencies of r. This is used to identify the regions introduced above; the reactive near field region, the radiating near field region and the radiating far field region.

Reactive near field

The strict IEEE definition of the reactive field region is "That portion of the near field region immediately surrounding the antenna, wherein the reactive field dominates" [40]. For electrically small antennas the boundary of this region is approximately $\lambda/2\pi$. Reactive fields are associated with an antennas reactance. They are non-radiating which means they are associated with quasi-static field components that do not propagate. In fact the reactive fields can be viewed as temporary storages of energy. For a non-resonant antenna as the small loop antenna presented above, the power can circulate between the reactive near field and the matching network or the source, or both. For a resonant antenna the power circulates within the reactive near field. The expression for Poynting's vector presented in equation 4.6 for the single loop antenna, indicate that in the time averaged power flow through a sphere enclosing the antenna, there is only one term that contributes to a net power flow away from the antenna; the $\frac{1}{r^2}\beta^3\cos^2(\omega t-\beta r)\mathbf{i}_r$ term. The time averages of all the other terms are zero. It is these terms that constitute the reactive field.

In the region $a \ll r \ll \frac{\lambda}{2\pi}$, $\beta r \ll 1$. This implies that the phase difference of two field points due to different distances to the source can be

neglected and the magnetic field can be assumed to be in phase with the source current. Based on equation 4.4, **H** has the following approximation in the reactive near field:

$$\mathbf{H} \approx \frac{a^2 I_0}{4}\left[2\cos\theta\frac{1}{r^3}\cos(\omega t)\mathbf{i}_r + \sin\theta\frac{1}{r^3}\cos(\omega t)\mathbf{i}_\theta\right] \tag{4.7}$$

This agrees with the result obtained with Biot-Savart's law for static currents, assuming that the current is $I = I_0\cos(\omega t)$ (quasi-static analysis).

In the strictest sense, every physically realizable pair of time-varying electric and magnetic fields must satisfy simultaneously both of Maxwell's curl equations; Faraday's law and the Ampere-Maxwell law [57]. This implies that there is an inherent interdependence between time-varying electric and magnetic fields. For low frequency fields however, the fields decouple approximately and in the case of the single loop antenna the magnetic fields are predominant. Low frequency approximation in signal transfer calculations, requires that both antennas are within the reactive near field region.

Signal transfer in the reactive near field region

The reactive or low frequency approximation describes signal transfer between two coils in terms of coupling of magnetic fields. Consider the situation in figure 4.2. A conduction current I_1 runs in loop 1 as result of a source emf. According to the Biot-Savart law, the magnetic field produced by loop 1 is proportional to I_1. A certain part Φ_{12} of the total magnetic flux Φ set up by loop 1 will flow through loop 2. Hence the flux flowing through loop 2 produced by loop 1, is proportional to I_1 and we may write:

$$\Phi_{12} = M_{12} I_1 \tag{4.8}$$

Here the proportionality constant M_{12} is the mutual inductance between the two coils. Faraday's law of induction states that a change in the magnetic flux Φ through a closed loop will induce an electromotive force in the loop to oppose the change in the flux. Combining Faraday's law with equation 4.8 gives:

$$V_{emf\,2} = -\frac{d\Phi}{dt} = -M_{12}\frac{dI_1}{dt} \tag{4.9}$$

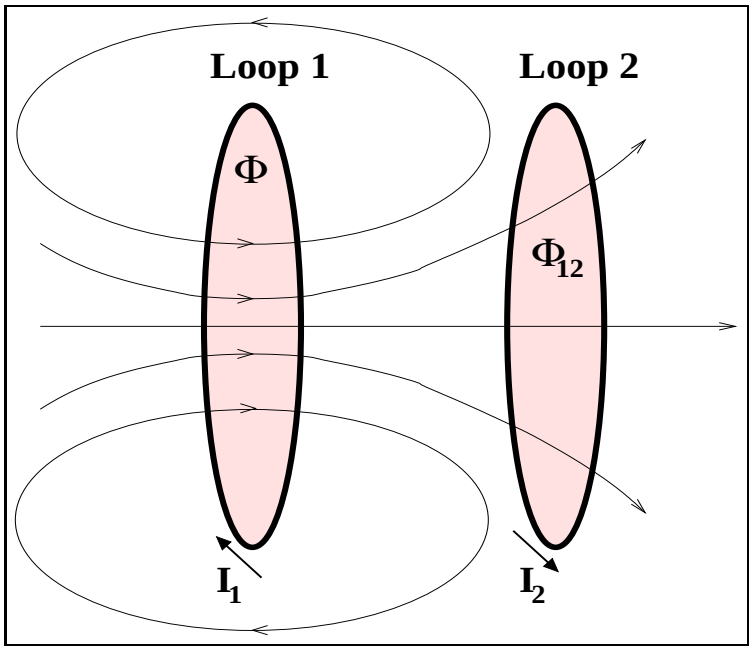


Figure 4.2: *When a current I_1 flows through loop 1, the magnetic field produced by this current will induce an electromotive force in loop 2 resulting in the current I_2*

The mutual inductance between two coils depends only on the geometry of the coils and the permeability of the intervening medium [16]. This is observed from Neumann's formula for the mutual inductance of two closed loops:

$$M_{12} = \frac{\mu}{4\pi} \oint_{C_1} \oint_{C_2} \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{R} \tag{4.10}$$

In appendix A it is shown how it can be derived from Maxwell's equations.

The concept of mutual inductance is not restricted to two separate circuits; every element of a single circuit has a mutual inductance with every other of its elements [21]. Changes in the coil current will induce electromotive forces in the coil itself. This special case of mutual inductance is called self inductance. The induced emf in a coil as a result of a current change in itself is:

$$V_{emf} = -\frac{d\Phi_{11}}{dt} = -L\frac{dI_1}{dt}, \tag{4.11}$$

where L is the self inductance.

The mutual inductance between two coils can also be expressed in terms of the self inductance of the two coils. Observe that seen from loop 1 the magnetic flux Φ_{12} defined above can be expressed as $k_{12}\Phi_{11}$, that is, the flux Φ_{12} can be expressed as being a certain fraction of the flux Φ_{11} [17]. Since $\Phi_{11} = L_1 I_1$ and $\Phi_{12} = M_{12} I_1$ we can write:

$$M_{12} = \frac{\Phi_{12}}{I_1} = \frac{k_{12}\Phi_{11}}{I_1} = k_{12}L_1 \tag{4.12}$$

In the same way, seen from loop 2, $M_{21} = k_{21}L_2$, and hence $M^2 = M_{12}M_{21} = k_{12}k_{21}L_1L_2$, so we can write [17]:

$$M = \sqrt{k_{12}k_{21}L_1L_2} = k\sqrt{L_1L_2} \tag{4.13}$$

Here k is known as the coupling constant of the loops 1 and 2 and can take on values between 0 and 1.

Radiation near field

In the radiation field the radiation pattern is an important part of a description of an antennas characteristics. The radiation pattern describes

how the antenna radiates signal power in different directions. In the radiation near field the angular distribution of radiated energy is dependent on the distance from the antenna [28]. Fresnell's approximation can give a useful description of the near field radiation pattern. The Fresnell approximation accounts for phase differences in the field due to the different delay of signals from different elements of the antenna, but amplitude differences are ignored [23]. The limit between the radiation near field region and the radiation far field region is set to $\frac{2D^2}{\lambda}$ where D is the largest dimension of the antenna. However, for electrically small antennas this field region is dominated by the reactive near field region because $\frac{2D^2}{\lambda} < \frac{\lambda}{2\pi}$. As size is an important factor for the application discussed here, the Fresnell approximation is not relevant for further discussion.

Radiation far field

In the far field the distance r to the antenna is large compared to the wavelength λ and the radiation paths from different elements of the antenna can be assumed parallel. Hence the the angular distribution of radiated energy is essentially undependent of the distance from the antenna.

As pointed out above, in the expression for Poynting's vector in equation 4.6 the only term contributing to radiation of energy from the antenna is the $\frac{1}{r^2}\beta^3 \cos^2(\omega t - \beta r)$-term and it is referred to as a radiation component.

Signal transfer in the far field can be calculated using Friis' transmission formula. It describes power transfer between two antennas with known directivty and efficiency. For a transmitting antenna fed with power P_{in} and gain G_t and a receiver antenna with gain G_r the received power P_out at the point where the power is drawn from the antenna is given by:

$$P_{out} = G_t G_r P_{in} \left(\frac{\lambda}{4\pi d}\right) \tag{4.14}$$

Other loss factors can be included depending on the conditions for signal transfer and system loss.

4.1.2 Effects of dissipative media

A comparison of the free space solution for the fields and the general solution presented in appendix D.1 reveals that the main difference is the attenuation of the fields caused by the factor $e^{(-\alpha r)}$. The sine and cosine terms including α in their amplitude factor will influence the amplitude as well as the phase of the field components.

Another phenomena that may affect communication signals transmitted via electromagnetic waves traveling through matter is frequency dispersion. The dispersion in body tissue can be noticeable for frequencies up to around 100 MHz. However, since the signals will travel a relatively short distance in the body, this should not cause major problems for a system using narrow bands. Very broad banded systems might suffer from this though.

When electromagnetic fields cross boundaries between tissues with different electromagnetic properties they must satisfy certain boundary conditions. This gives rise to reflection and scattering loss at the boundaries. The complex structure of the human body makes it difficult to give accurate assessments of this type of loss. In chapter 1 it was suggested that the wrist and the abdomen could be possible locations for the implant. Figure 4.5 shows sketches of the anatomy in these areas. The complexity is obvious.

As seen from figure 4.4 the wavelength of electromagnetic waves are also influenced by the dielectric properties electromagnetic properties of the fields. In this case it has the positive effect of decreasing the size of an antenna with given operating frequency.

4.1.3 Antenna Miniaturization

It can be difficult to find an antenna that meets the demands of an application in terms of size and frequency characteristics. The size and shape of the antenna enters the field solutions through boundary conditions. Field solutions indicate that electrically small antennas have poor efficiency. In

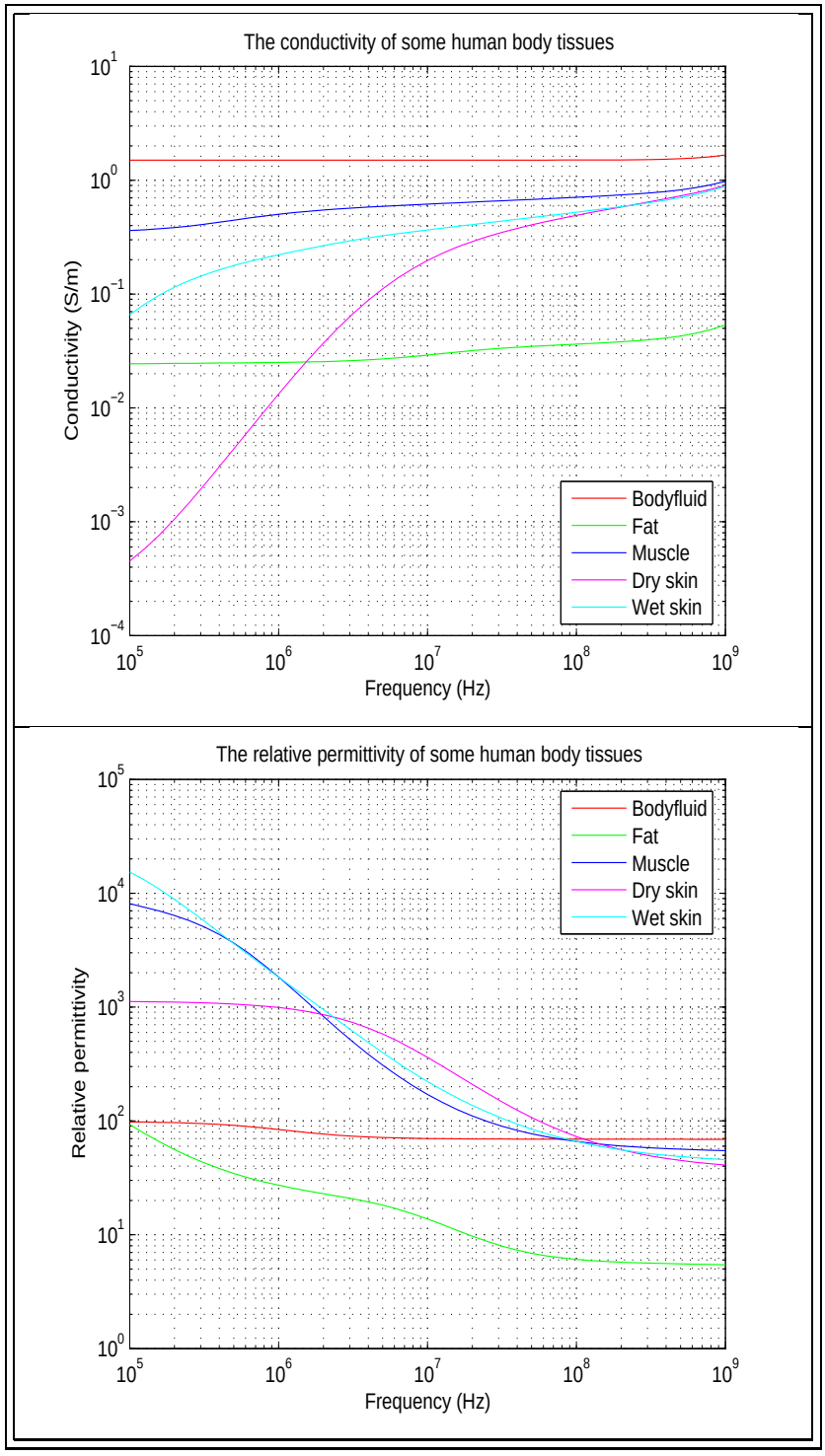


Figure 4.3: *The figure shows the frequency dependency of the conductivity σ and the relative permittivity ϵ_r of some human tissues.*

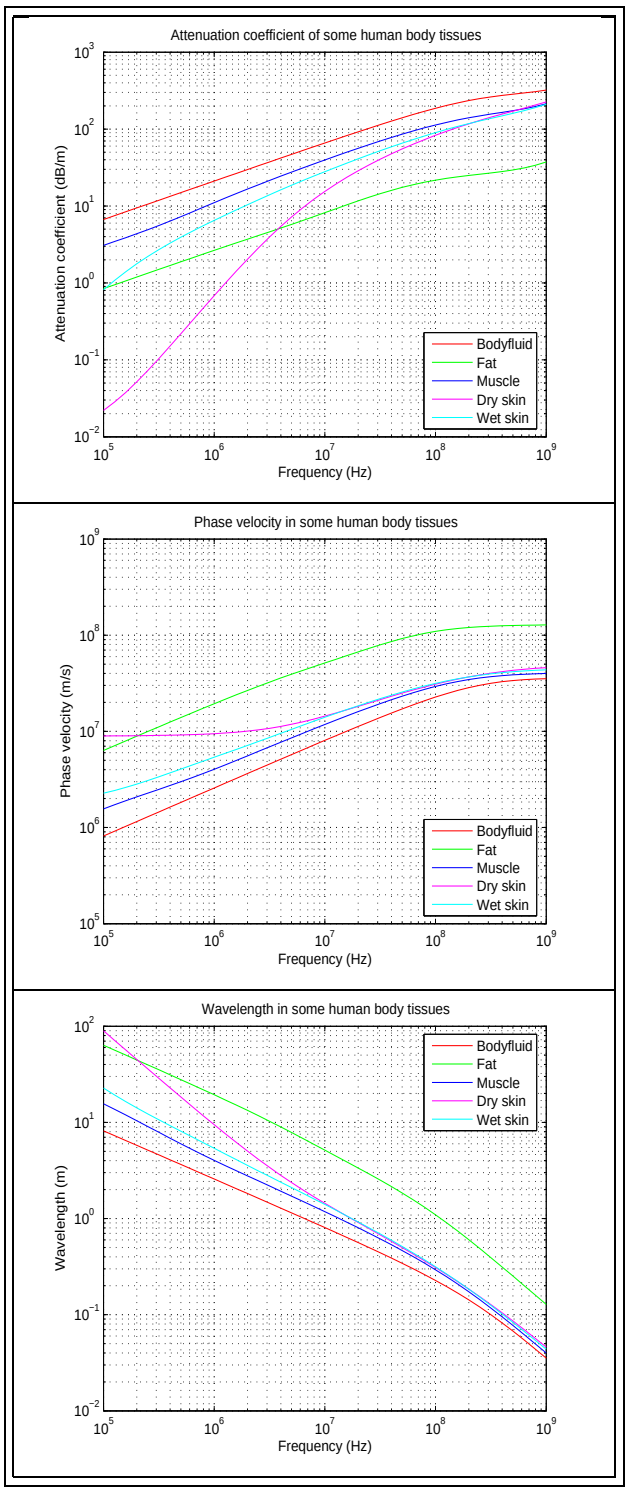


Figure 4.4: *The figure shows the frequency dependency of the attenuation, phase velocity and wavelength of some human tissues.*

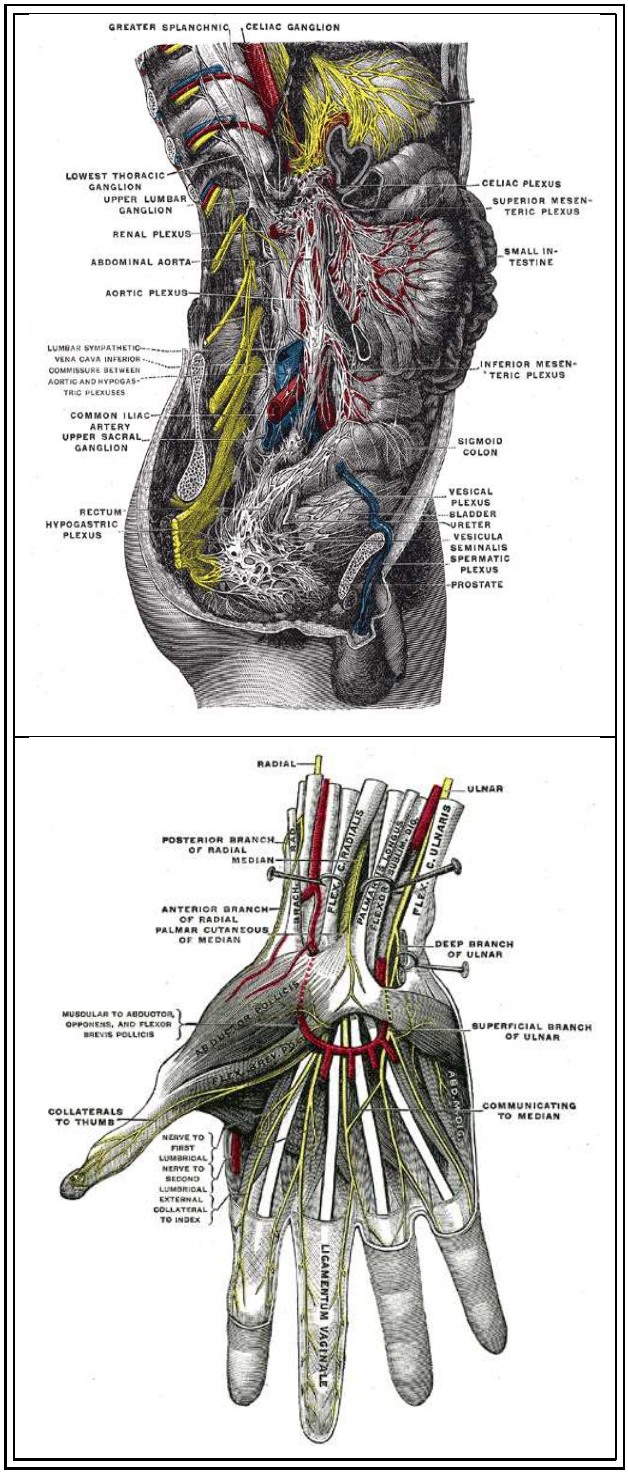


Figure 4.5: *The figure shows sketches of the male human abdomen and a human hand (Source Henry Gray (1825-1861). Anatomy of the Human Body. 1918.*

"PCS Antenna Design: The Challenge of Miniaturization", Skrivervik et al. [63] describes several techniques for miniaturizing antennas and how they affect antenna characteristics. The overview given below is to a large extent based on this article.

A relatively simple approach to reduce the resonance frequency of an antenna is to load it with a lumped element. A well known example of this, is to connect a capacitive element to a loop antenna. This can give a desired reduction of the resonance frequency. However, undesirable effects like reduction in the antennas radiation efficiency as a consequence of losses in the lumped element and increase in the circuits Q-factor reducing the bandwidth of the antenna will occur.

To encapsulate the antenna in a material with certain electromagnetic properties is another technique to achieve antennas with desired size and frequency characteristics. If the material has a higher permittivity or permeability than in free space, the wavelength is shorter in the material than in free space. Dielectrics are most frequently used as the package material. The reduction in antenna size then depends on the permittivity and the dimensions of the dielectric. Drawbacks encountered using this technique is reduction in bandwidth and dielectric loss. The dielectric material will also cause a concentration of the electrical field. This causes problems with the transmission of the signal from the dielectric and into the surrounding medium. The technique is most often used with patch antennas.

The introduction of a ground plane can also be used to reduce the size of an antenna. The best known example of this is the monopole. When it operates over a ground plane a virtual copy of the monopole arises in the ground plane. The two monopoles appears as a dipole. The dimensions of the ground plane are crucial for this approach. This technique can be used on a patch antenna as well. If one end of the antenna is connected to the ground plane the length can be cut down to half the original.

Changing and optimizing the geometry of an antenna is yet another way to miniaturize antennas. L and F antennas are examples of this technique. By making slots and holes in a patch antenna the resonance frequency can also be reduced. This can be explained by the fact that the current is forced to follow paths around the obstacles inserted. Larger

current densities means higher ohmic losses, which again reduces the antenna efficiency. Changes of this kind, ground planes, short circuiting and changing the geometry of the antenna often makes it frequency sensitive. Compared with standard antennas, the bandwidth is thus reduced.

Lately, within the field of antenna design, there has been some interest for fractal design. This means fractal structures are used in the construction of antennas. The technique is primarily used to construct multi or broad banded antennas. An overview of the research done in this field is given in "An Overview of Fractal Antenna Engineering Research" (Werner and Ganguly, 2003 [68]).

The last method discussed here is to let the surroundings of the antenna contribute to the antennas radiation characteristics. Often the casing of the antenna can be utilized as a strong contributer to the radiation. In some cases the main purpose of the antenna is to act as a resonator to set the appropriate working frequency.

The interest of miniaturizing in the field of antenna design has mainly focused on Personal Communication Systems (PCS). These types of communication systems often operate at higher frequencies than what is desirable for medical implants. Results achieved for PCS is therefore not necessarily directly transferable to the application in question.

4.2 Communication system for Passive Solution

In this section we discuss a system design where the implant can operate without an internal power source. The absence of the power source can make the implant smaller, lighter and give it extended lifetime as batteries, currently the most suitable energy source for medical implants, have limited capacity, energy density and lifetime. The most serious limitation of these systems appears to be the limited operating range. This type of wireless communication is often referred to as passive telemetry and has two main branches, passive load modulation and resonant peak monitoring [15]. The first method relies on active on-chip circuitry to modulate the load. The other method monitors the resonance frequency of an LC-tank. Both rely on power transfer through inductive coupling.

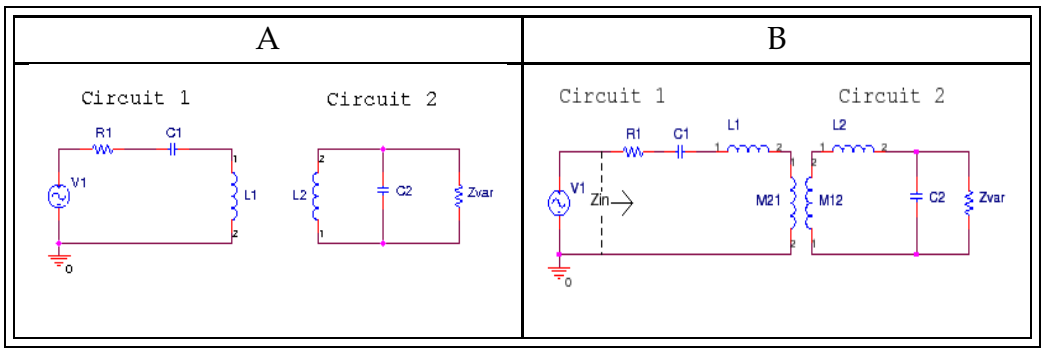


Figure 4.6: *The figure shows two schematic drawings illustrating a basic impedance modulation circuit. Circuit B shows an equivalent circuit used for impedance calculations.*

4.2.1 Impedance modulation - Basic Principles

Impedance modulation is a technique that can be used in a two node communication system with one passive node and one active node to achieve a (full duplex) data link as well as provide the passive node with operating power. For the system in question, the medical implant is designed to be a passive node and the external device the active node.

In section 4.1.1, we saw how a connection between the external and the implanted communication circuit could be established through an inductive coupling. The basic principle of impedance modulation is to convert the desired data, in this case sensor data, in the passive node into impedance variations that can be sensed in the active node through the inductive coupling.

A principle schematic of an impedance modulation circuit is shown in figure 4.6 A. Here the external circuit is denoted circuit 1 and the implanted circuit is denoted circuit 2. Circuit 1 is a series resonance circuit consisting of the source resistance R_1 in series with a capacitance C_1 and the coil L_1. In circuit 2 L_2 and C_2 form a parallel resonance circuit. Z_{var} represents the impedance that is modulated according to the data that is transmitted.

To describe the behavior of the circuit it is appropriate to introduce the equivalence circuit in figure 4.6 B. The inductive coupling of the two

circuits is represented by the mutual inductance of L_1 and L_2; denoted M_{21} in circuit 1 and M_{12} in circuit 2. Then the input impedance seen from the signal source becomes:

$$Z_{in} = R_1 + \frac{1}{j\omega C_1} + j\omega L_1 + Z_M \tag{4.15}$$

Here the last term is the impedance seen in circuit 1 as a result of the inductive coupling with circuit 2. It can be shown ([51]) that Z_M is:

$$Z_M = \frac{M^2\omega^2}{Z_2} \tag{4.16}$$

Z_2 in equation 4.16 is the impedance in circuit 2 seen from M_{12} and is:

$$Z_2 = j\omega L_2 + \frac{1}{j\omega C_2 + \frac{1}{Z_{var}}} = j\omega L_2 + \frac{Z_{var}}{j\omega C_2 Z_{var} + 1} \tag{4.17}$$

Equations 4.15, 4.16 and 4.17 shows that a change in Z_{var} will change the impedance of circuit 1.

To transfer data from the passive circuit to the active circuit it would be advantageous if the effect of Z_{var} on the impedance of circuit 1 is maximized. By examining equation 4.15, we understand that the effect of Z_M and hence Z_{var} is maximized if:

$$\frac{1}{j\omega C_1} + j\omega L_1 = 0$$

This occurs at resonance when $\omega = \omega_r = \frac{1}{\sqrt{L_1 C_1}}$ and the impedance of the coil and capacitance cancel.

We now presume that circuit 1 works at its resonance frequency. For simplicity it is assumed that $L_1 = L_2$, $C_1 = C_2$ and that Z_{var} is a pure resistance R_2. This implies that $\omega_r = \frac{1}{\sqrt{L_1 C_1}} = \frac{1}{\sqrt{L_2 C_2}}$. Z_2 will now take the form:

$$Z_2 = \frac{1}{j\omega_r C_2 - \omega_r^2 C_2^2 R_2} \tag{4.18}$$

At resonance Z_{in} becomes:

$$Z_{in} = R_1 + M^2\omega_r^2 \left(j\omega_r C_2 - \omega_r^2 C_2^2 R_2 \right) \tag{4.19}$$

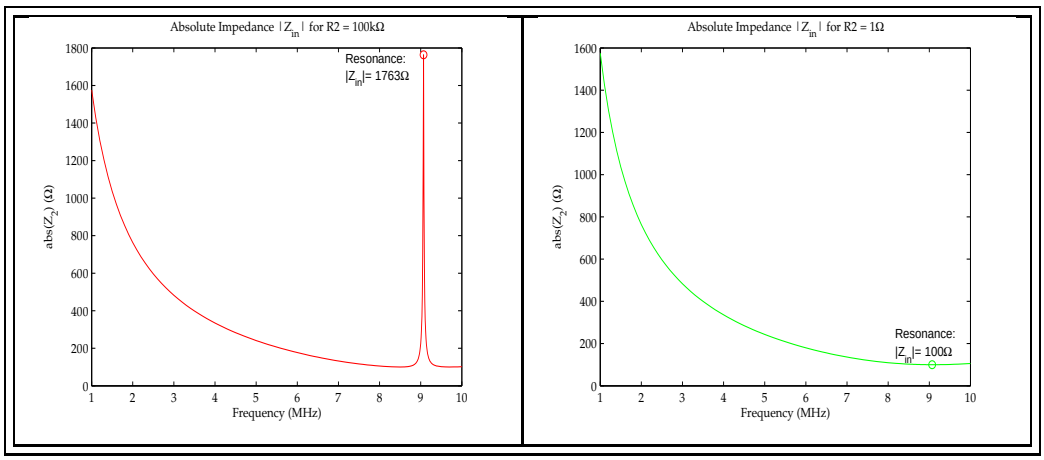


Figure 4.7: *The figure shows the absolute values of the impedance Z_{in}. The calculations were done with $L_1 = L_2 \approx 3.1\mu H$ $C_1 = C_2 = 100pF$ and $R1 = 100\Omega$. The coupling constant was high; $k \approx 0.1$.*

The series LC network of circuit 1 has an impedance minimum and hence the current has a maximum at resonance. This implies that maximum magnetic field is generated by the coil. Correspondingly, at resonance the parallel connection in circuit 2 causes a large voltage drop over the capacitor making R_2 a dominating factor in the impedance Z_2. Figure 4.7 shows a plot of Z_{in} for $R_2 = 100k\Omega$ and when $R_2 = 1\Omega$. At resonance the impedance Z_{in} is strongly affected by the large valued R_2 and results in the spike in Z_{in}. A small R_2 results in a small absolute value. If sensor data can be converted to impedance changes in the passive node these impedance changes can be measured in the active node and the data retrieved.

An interesting observation regarding power transfer is that the strongest coupling k, does not necessarily imply maximum voltage over R_2. From [17] we have an expression for $|V_2/V_1|$:

$$\left|\frac{V_2}{V_1}\right| = \frac{kR_2\sqrt{\frac{L_1}{L_2}}}{\sqrt{\left(R_1 + k^2R_2\frac{L_1}{L_2}\right)^2 + (\omega_r k^2 L_1)^2}} \tag{4.20}$$

If this expression is differentiated with respect to k, we will find that V_2/V_1 has a maximum point known as the critical coupling [17]. This behavior is

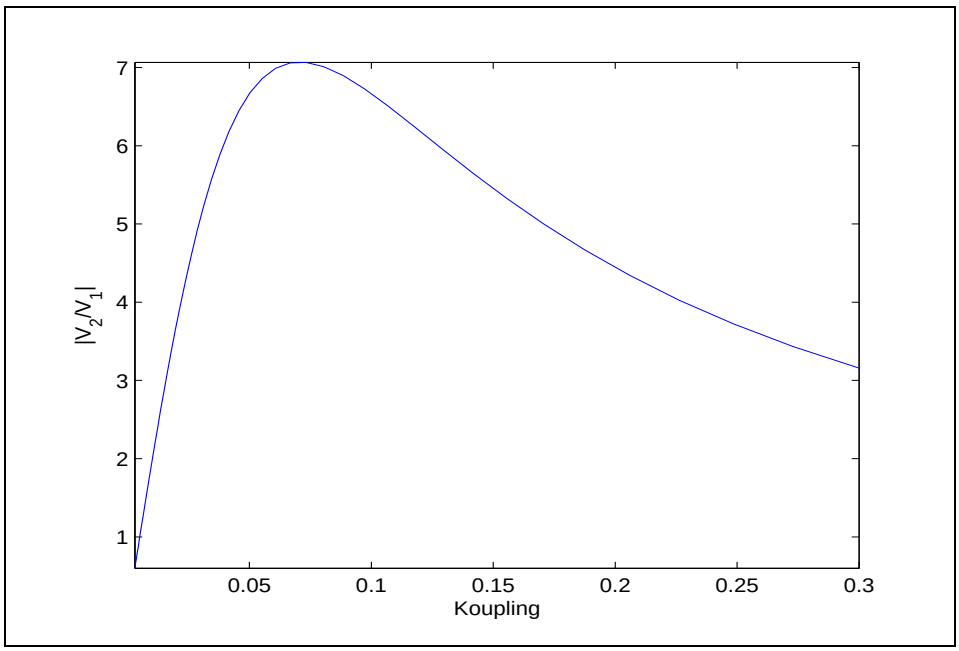


Figure 4.8: *The ratio of* $\left|\frac{V_2}{V_1}\right|$ *of two coils plotted against the coupling constant k of two coils.*

shown in figure 4.8 where V_2/V_1 is plotted against the coupling k. The behavior can be explained by examining equation 4.19. When the coupling between the coils are strong the term Z_M will correspond to a large series impedance. Hence it limits the current in circuit 1. A small current implies a weaker magnetic field and hence a smaller induced voltage in circuit 2. A weaker coupling (smaller M in equation 4.19) reduces the series impedance Z_M. Hence the current increases and the induced voltage in circuit 2 increases. For coupling values smaller than the maximum point in figure 4.8 the coupling gets so weak that the induced voltage drops.

4.2.2 Resistive impedance modulation

The simplest modulation technique used is amplitude shift keying (ASK). This is achieved by switching the resistance R_2 in figure 4.6 between two values. As demonstrated by figure 4.7 this will switch the impedance Z_1 of the active circuit between two values. The voltage over the L_1C_1-series

connection will appear as an ASK-modulated signal.

To read the sensor status an excitation signal is set up in the active circuit. The passive circuit extracts and stores power from the excitation signal through the inductive coupling of the two circuits. When a sufficient amount of energy has been stored the logic circuits on the passive implant initiates a reading of the sensor value and does an A/D conversion of the measured value. The digital value representing the measured value is now fed as a serial digital signal to the switch controlling the resistance R_2.

Berg [7] and Due-Hansen [17] has suggested a solution for this kind of system based on Zoubir Hamici et al [25] for an intra cranial pressure sensor. A sketch giving an overview of the implanted part of the system is shown in figure 4.9. Starting to the left in the figure, we have the resonance circuit; the coil and the capacitance. In parallel with this is a series combination of a diode, resistance and a switch. This is where the resistance modulation takes place. The switch is controlled by the digital signal 'DATA' which is data to be transmitted. The diode ensures that the switch operates properly. In the negative half of a period in the sinusoidal excitation signal, the substrate on the switch would be exposed to a negative potential and conduct current independent of 'DATA'. Next a standard rectifier bridge and capacitance supplies a DC-voltage to the digital control circuitry and the analog sensor circuitry. Z_{dig} is the impedance of the voltage regulator for the digital electronics and the electronics itself. The impedance Z_{analog} is the impedance of the sensor circuit and its regulator. The signal 'ANALOG' turns the analog circuit on and off. To maximize the effect of R_{mod} the analog circuitry is turned off during data transmission. This is because we can choose a smaller R_{mod} but still ensure the needed voltage level for the digital electronics.

If two-way communication is desired, the external circuit can produce an amplitude modulated excitation signal. This can quite easily be decoded on the implanted circuit. A problem with this approach is that the power supply circuitry on the implanted circuit has to cope with the changes in the supply voltage during decoding of such a message from the external circuit.

The potential problem of lack of power during data transmission from

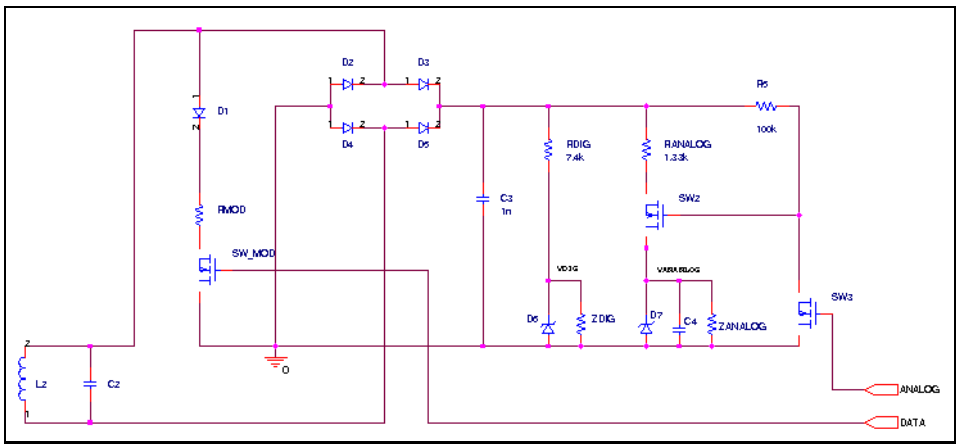


Figure 4.9: Schematic of part of the implanted device in the intracranial pressure monitoring system presented by Berg [7] and Due [17]. This figure corresponds to figure 3.5 in [7].

the external to the implanted circuit, can be addressed by utilizing PSK-modulation instead of ASK-modulation in the signal generated at the external circuit. This approach is presented by Marschener et al in [39]. When using PSK-modulation the amplitude of the excitation signal will be constant. However, since the resonance circuits have high Q-factors (this implies sharp resonance frequencies), the amplitude of the signal will still be affected exactly when a phase transition takes place. This is because the phase transition will change the instant frequency, and the frequency of the signal will for a short time fall outside the sharp resonance top of the circuits. By utilizing a PSK-technique with smoother phase transitions, such as differential PSK (DPSK), the reduction in signal amplitude and duration of this can be minimized. The technique demands for a more complex demodulation circuitry than ASK-modulation.

Reactive impedance modulation

If the sensor element is capacitive the impedance variations of the internal circuit will be reactive in stead of resistive. Hence, to observe the state of the capacitive sensor a different technique is required. A simplified circuit diagram of the circuit on the implant and the external read

circuit is shown in figure 4.10. The expression for the impedance of such a circuit using the equivalent circuit shown in figure 4.11 is:

$$Z_{in} = \frac{1}{j\omega C_{C1} + \frac{1}{R_{C1} + j\omega L_1 + \frac{M^2\omega^2}{Z_2}}} \tag{4.21}$$

where:

$$Z_2 = j\omega L_2 + R_{C2} + \frac{1}{j\omega(C_{C2} + C_s)} \tag{4.22}$$

The sensor circuit's resonance frequency is shifted when the capacitance of the sensor is changed and can therefore be used to observe the state of the sensor. The resonance frequency of the sensor circuit is the frequency where the impedance contributions from L_2 and $(C_2 + C_s)$ cancel, leaving R_2 as the only contributer to Z_2. A possible way to detect resonance frequency changes that has been verified by use in prototypes, see [15] and [26], is by logging the phase of the impedance . The resonance of the sensor circuit will create phase dips in the systems total phase provided that the coupling between the circuits is sufficient.

$L_1 =$	$1.16\mu H$
$C_{C1} =$	$1pF$
$R_1 =$	0.3Ω
$L_2 =$	$4.33\mu H$
$C_{C2} =$	$1.5pF$
$R_2 =$	1.52Ω

Figure 4.12 shows plots of the phase of the impedance of the circuit in figure 4.10 for four different values of sensor capacitance. In the calculation the external coil was assumed to be a solenoid of 3 turns with diameter 5cm and wire diameter 0.5mm while the implanted coil was assumed to be a flat spiral coil of 23 turns, wire diameter $100\mu m$, $70\mu m$ of turn space and an outer radius of 5mm. The components as shown in figure 4.10 were assumed to have the values as shown in the table to the left.

The self inductance of the coils was calculated with program [43] found on the internet that showed good correspondence with measurements. The mutual inductance was calculated with the program described later in chapter 6. The plot to the left is for $M = 1.56 \cdot 10^{-8}H$ corresponding to a distance of approximately 3cm between the coaxially aligned coils.

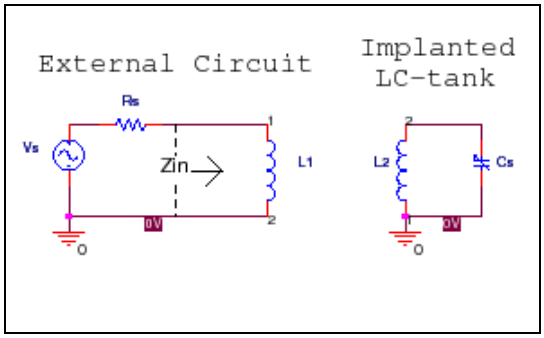


Figure 4.10: The figure shows a schematic of an impedance modulation circuit with a capacitive sensor element

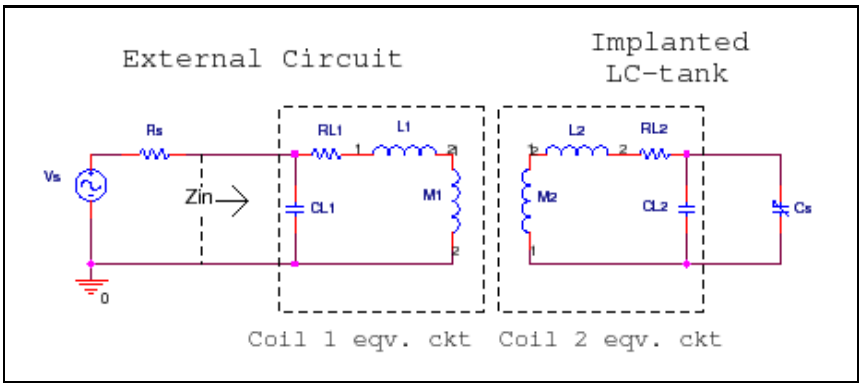


Figure 4.11: *This schematic shows equivalent circuit used in impedance calculations for the schematic in figure 4.10.*

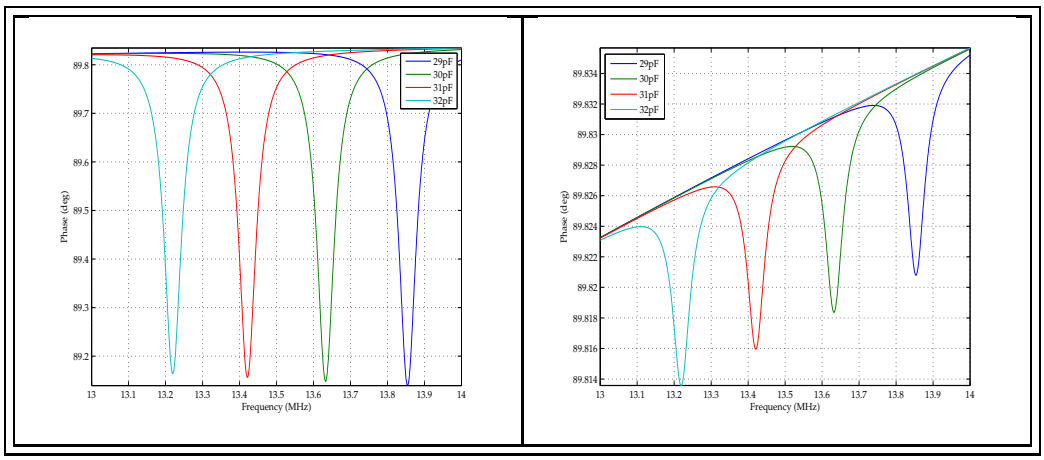


Figure 4.12: *The figure shows calculated phase variation of the system shown in figure 4.10 for strong coupling (left) and weak coupling (right) of L_1 and L_2.*

The plot to the right has $M = 2.15 \cdot 10^{-9} H$ corresponding to a distance of about $7cm$. The phase dips in this figure results from resonances in the sensor circuit. They can be used in the detection process. Phase differences can be measured with high accuracy[1]. We see that the change in the plot to the right of the figure indicates variations of approximately 0.01 degrees, which is expected to be the limit of detection.

Observe that when the sensor circuit is at resonance the total impedance of the system can be approximated by:

$$Z_{in} \approx R_1 + j\omega_0 L_1 + \frac{M^2\omega_0^2}{R_2} \tag{4.23}$$

Here the capacitance C_{C1} of coil L_1 can be neglected as Z_{in} will be dominated by the reactance ωL_1 if the external circuit works well below the self resonance of the coil [2]. Due to the dominance of $j\omega L_1$, the phase of Z_{in} will

[1] For example the HP4195A used in the measurements described in chapter 5, has a resolution of 0.01 degrees for the phase in impedance measurements.

[2] L_1 should work well below resonance or else much of the current will go through the capacitance and less magnetic field is generated by L_1. At the same time, if $j\omega L_1$ is relatively small, the resonance of the sensor circuit will cause a more significant change in the total impedance Z_1.

be close to 90°. At resonance the term $\frac{\omega^2 M^2}{R_2}$ causes a phase dip. The magnitude of this term will decide the depth of the phase dips and is therefore crucial for the performance of the system. However, how to obtain a maximum magnitude can be hard to realize because of the interdependence of the parameters M, ω and R_2.

An expression for the magnitude of the phase dip can be found by subtracting the phase of the system when $M \neq 0$ from the phase for $M = 0$ [26]:

$$\begin{aligned} |\Delta\phi| &= |\angle(R_1 + j\omega_0 L_1) - \angle(Z_{in}(\omega_0))| \\ &= \left| \arctan\left(\frac{\omega_0 L_1}{R_1}\right) - \arctan\left(\frac{\omega_0 L_1}{R_1 + \frac{\omega_0^2 M^2}{R_2}}\right) \right| \\ &\approx \left| \frac{\pi}{2} - \arctan\left(\frac{1}{\frac{\omega_0 M^2}{R_2 L_1}}\right) \right| \\ &= \arctan\left(\frac{\omega_0 M^2}{R_2 L_1}\right) \end{aligned} \tag{4.24}$$

The value of M is found from Neumann's formula (equation 4.10). M can be maximized by increasing the number of turns in the coils N_1 and N_2, and by increasing the diameters d_1 and d_2. However, this will also imply an increase of the self inductances L_1 and L_2. From the result above an increase in L_1 implies a reduction of $|\Delta\phi|$. An increase in L_2 implies a reduction of the resonance frequency ω_0 of the sensor circuit and hence also a reduction of $|\Delta\phi|$.

An increase in either d or N will also imply an increase in R_2, which also would decrease $|\Delta\phi|$. At the resonance frequencies in question here, the resistance increases with frequency due to the skin effect. How significant the effect is, is dependent on the ratio between the wire diameter and the skin depth, see section 4.2.2 for a more thorough discussion of this issue.

The geometrical constraints given by a systems size restrictions decides the design tradeoffs of the coils. To indicate how the coil design affects this kind of system, impedance calculations were done for different sets of L_1 and L_2. The results are shown in figure 4.13. An interesting observation

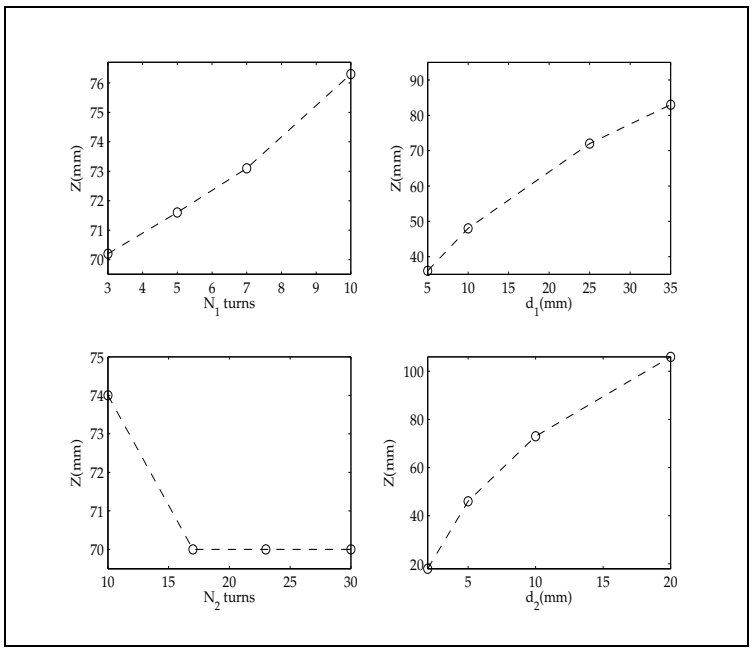


Figure 4.13: *The figure indicates how different design parameters affect the maximum operating distance of a system utilizing reactive impedance modulation.*

is that an increase in N_2 causes a decrease in maximum operating range. This is probably caused by the reduction of ω_0, resulting from the increase in L_2. The increase in M is not large enough to compensate the reduction in ω_0. For L_1 the situation is quite opposite. Increasing N_1 can increase the operating range significantly. Also of interest is the large dependence of the coil diameters.

The dependence upon ω_0 also suggests reducing both coil capacitance C_{C1} and sensor capacitance C_s would be advantageous for increasing operating range. This will however have implications for the sensor sensitivity and its dynamic range.

In [51] Ong et al. reported achieving an increase in operating range with a system design using a combination of two loop antennas in the read circuit. One antenna transmitted an RF signal which was read on the receiver side with a lock-in amplifier. The antennas were placed on opposite sides of an LC-sensor circuit. The LC-sensor circuit affected the voltage

Reference	max. operating range	Sensor coil dimensions	Read coil dimensions	Mod. Tech.
[7] and [17]	$0.5cm$	$(2 \cdot 2)cm^2$	$(2 \cdot 2)cm^2$	R
[26]	$1cm$	$(1.5 \cdot 0.5 \cdot 8)mm^3$	$d = 15mm$	P
[51]	$9cm$	$(4 \cdot 4)cm^2$	d=4cm	Z
[15]	$3.1cm$	$(6 \cdot 6 \cdot 0.5)mm^3$	$d = 48mm$	P

Table 4.1: *The table gives an overview over reported systems and their maximum operating range. The column marked reference refers to the bibliography at the end of this text. For the "Mod. Tech." column R equals resistive impedance modulation, P is reactive impedance modulation with phase logging as described above and finally Z is reactive impedance modulation logging of absolute impedance.*

measured on the receiving side and enabled detection of the LC-sensor resonance frequency. This technique was compared to a one antenna approach using a variant of reactive impedance modulation measuring the absolute value of the impedance. The LC-sensor was equipped with a $4cm \times 4cm$ quadratic coil. Using a "background subtraction" routine to eliminate the intrinsic antenna impedance of the loop antenna, the maximum reported read distance was 13.5 cm. This was increased to 32 cm for the two antenna approach. In the tests of the two antenna approach, each antenna consisted of two wire loops of radius 4cm, i.e. a total of four wire loops.

Comparison of performance of reported prototypes

Ending the discussion of communication systems using impedance modulation techniques, table 4.1 presents an overview of operating ranges for some reported prototypes. The prototype presented by Berg [7] and Due [17] demands a relatively large energy transfer to power the electronics required for reading the resistive sensor. This limits the operating range seriously. The technique involving LC-tank resonance frequency monitoring, is not dependent of such energy transfer, but depends heavily on the sensitivity of the phase detection device used in the monitoring circuit.

The skin effect and stray capacitance

As pointed out above in the description of the LC-tank resonance frequency monitoring technique, the resistance of the coils affect the system performance. A large coil resistance reduces the operating range. This is also the case for the resistive modulation circuit, which can be seen if the high frequency equivalence circuit with coil resistance R_{L2} is included in equation 4.18:

$$Z_2 = R_{L2} + \frac{1}{j\omega_r C_2 - \omega_r^2 C_2^2 R_2} \tag{4.25}$$

This implies that Z_in can be written:

$$Z_{in} = R_1 + \frac{M^2\omega_r^2}{R_{L2} + \frac{1}{j\omega_r C_2 - \omega_r^2 C_2^2 R_2}} \tag{4.26}$$

We see that a large R_{L2} will reduce the magnitude of Z_M, and hence reduce the operating range.

As was also mentioned above, at the operating frequencies in question the coil resistance is frequency dependent due to the skin effect. The skin effect confines the current to the outer sections of a conductor at high frequencies. This increases the resistance the current experiences when it flows through the conductor.

The magnitude of this resistance increase for a straight conductor is decided by the ratio of the skin depth to the wire diameter. The skin depth is given by:

$$\delta = \frac{1}{\sqrt{\pi f \mu \sigma}} \tag{4.27}$$

The skin depth is the depth at which the electric field causing the current flow is attenuated a factor e^{-1} from the conductor surface. A high frequency approximation often used, assumes that the current density is uniformly distributed in a cross-section from the conductor surface at $r = a$ into $r = a - \delta$. That is, the expression:

$$A_{eff} = \pi a^2 - \pi(a-\delta)^2 = 2\pi a\delta - \pi\delta^2 \approx 2\pi a\delta \tag{4.28}$$

is taken to be the cross-sectional area with uniform current density. Inserting this into the formula for DC-resistance in a wire, we get:

$$R(\delta) = \frac{l}{A_{eff}}\rho = \frac{l}{2\pi a\delta}\rho = R_{DC}\frac{a}{2\delta} \tag{4.29}$$

This formula works well provided that $a \gg \delta$.

The current density is actually not uniformly distributed in this cross section, but is dependent upon the distance r for given values of ω, μ and σ. In Appendix B on page 164 there is presented an expression, equation B.21 for the resistance taking account for this r-dependency. Using MAPLE a 25-term series expansion of R in ω has been used to generate data for the R_{Bessel} curves in figure 4.14. In fact the skin depth resistance formula presented above is an asymptotic approximation of this expression. This is now used to indicate the regional validity of the R_{DC} $R(\delta)$ for the largest and smallest wire diameter at relevant frequencies used in the calculations in section 4.2.2. Figure 4.14 shows plots of R_{DC}, $R(\delta)$ and R_{Bessel}. This figure indicates that the $R(\delta)$ gives the best approximation for all combinations wire diameters and frequencies used. (Observe that R_{Bessel} has a finite radius of convergence. The use of Hankel functions at high frequencies would address this problem.)

It is also to be noted that the skin effect in a coil acts differently than in a straight wire. The skin effect is modified by the mutual effects of the turns one on to the other, so that the current in the cross section is crowded toward the axis of the winding [21]. This both the resistance and capacitance as well as the inductance of the coil.

The high frequency resistance of a coil may be several times the resistance of the same wire stretched out straight and measured at the same frequency [21][3]. The self inductance of a coil is reduced because of the reduction in effective radius the coil experiences due to the current crowding toward the coil axis.

The inductance of a coil is also altered by its self-capacitance. The self capacitance is a distributed capacitance between the windings of the coil

[3]This indicates that the operating ranges presented in section 4.2.2 are a bit optimistic.

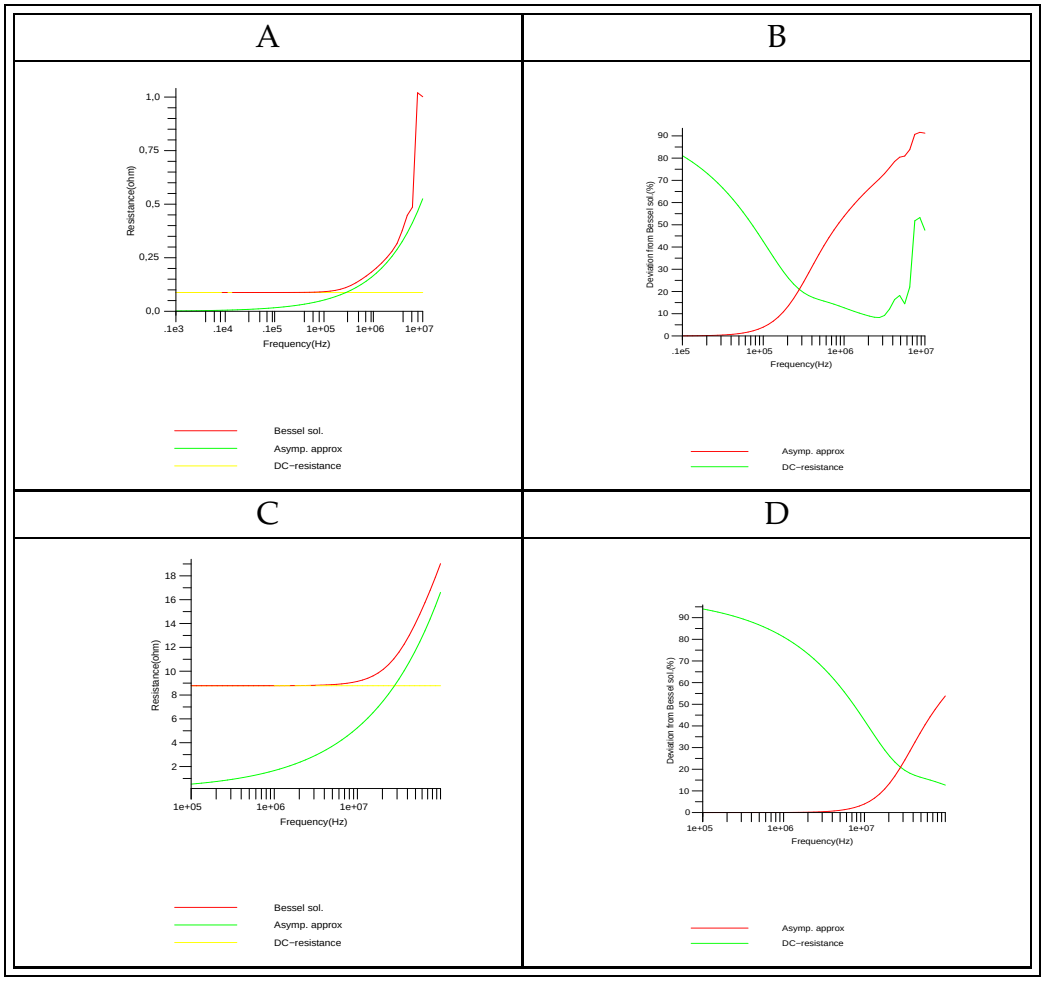


Figure 4.14: *The two plots to the left show R_{DC}, $R(\delta)$ and R_{Bessel} for a straight cylindrical copper wire of diameter* $0.5mm$ *(A) and* $50\mu m$ *(C). The two plots to the right show the deviation of R_{DC} and $R(\delta)$ from R_{Bessel} in % of R_{Bessel}.*

and acts much as its equivalent lumped capacitance in parallel with the inductance. It is especially important when considering multilayer coils. The inductance is increased by the self-capacitance and this effect may overcome the opposite result of the skin effect [21].

Stray capacitances will also affect the characteristics of the resonance circuit. This can be of interest considering the operating environment. Body tissues have relatively large electrical permittivity at the frequencies of interest. This means stray capacitances are larger than in air. In figure 4.15 A the results of a simulation with $L_1 = L_2 = 3.12\mu H$, $C_1 = C_2 = 100pF$, $R_s = R_1 = 50\Omega$ and $R_2 = 10k\Omega$ is shown. In 4.15 B the results of a simulation with a parasitic capacitance of $6pF$ in both the implanted and external circuits are shown. The altering of the frequency of maximum voltage in the implanted circuit when the two circuits are close to each other is easily observed.

4.3 Communication System for Active Implanted Device

To achieve greater operating ranges than for passive implants, a source of energy is needed on the implant to power signal amplification and processing. This enables an active system to handle very weak signals. The design of such communication systems involves more "degrees of freedom" and is therfore often more complex than design of passive systems.

4.3.1 Basic Building Blocks

The basic building blocks in an active communication system for transferring measurement data is shown in figure 4.16. The signal that carries the digital measurement information the system is to transfer, is called *the baseband signal*. Digital modulation schemes are preferred to analog schemes because the are more robust against noise. A digital modulation scheme implies that the baseband signal is digital. Important aspects to consider is the quantity of information that is to be transmitted and how

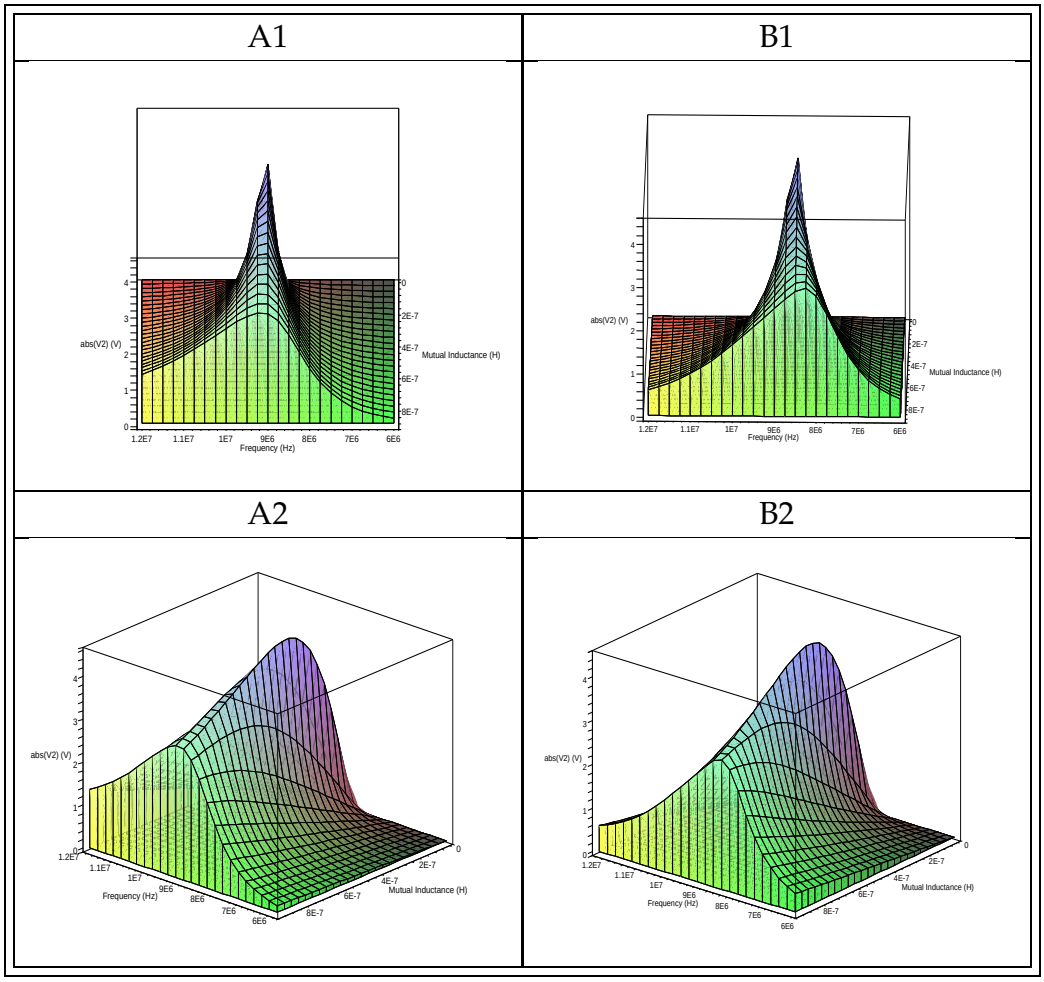


Figure 4.15: *3D-plots of* $\left|\frac{V_2}{V_1}\right|$ *as a function of both frequency and coupling. The figures show how parasitic capacitances affect the resonance of the circuit. An ideal system is shown on the two figures to the left (A1 and A2) and a system with parasitic capacitances is shown to the right (B1 and B2).*

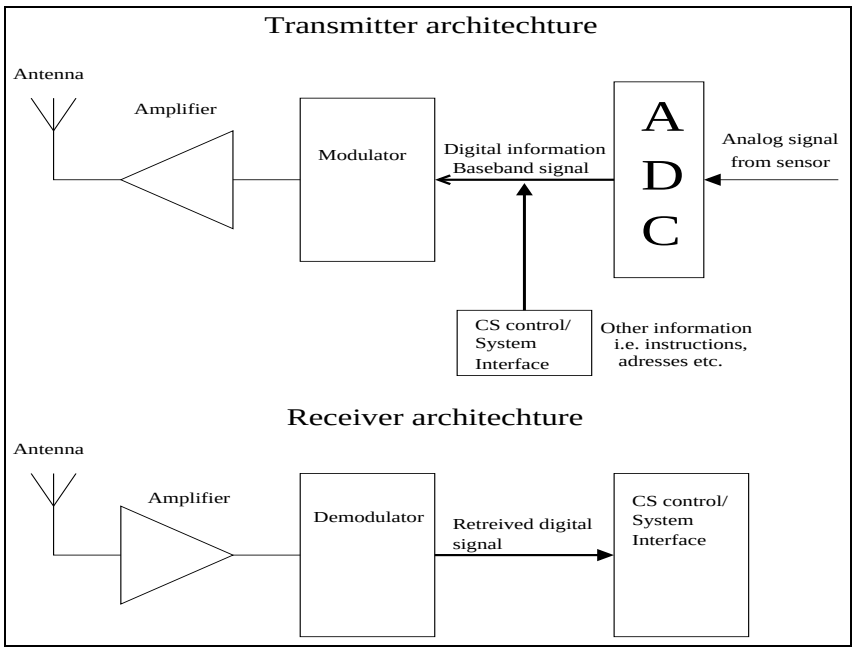


Figure 4.16: *A block diagram of the basic building blocks in the transmitter and receiver*

fast it is necessary for the user to access it. This will decide what bandwidth is needed to support the demanded data rate and what transmission schemes can be utilized.

In order to achieve synchronization and reliable transmission, the data bit stream must be modified. This is done by mapping it into a code which enables synchronization and then adding extra bits of error detecting/correcting code. Coding can also provide data compression at the cost of more pre- and post-processing.

The baseband signal must be multiplied with an RF carrier to be transmitted in the desired channel. This operation is called modulation. Modulation can be done in a number of different ways, for example amplitude modulation, frequency modulation and phase modulation.

Before the signal is fed to the antenna it is amplified by a power amplifier (PA). To waist as little of the signal energy as possible in this stage, matching networks are needed at both ends of the amplifier.

The receiver will do something close to the inverse operation of the

transmitter. Due to the low signal level, noise problems are more severe in the receiver than in the transmitter. This makes receiver architectures more complex and more difficult to design than transmitters. To recover the base band signal and the information sent by the transmitter, extensive signal processing must be carried out on the received signal.

The purpose of the discussion in this section will therefore be to see how selected parts of the communication system should be build and put together to provide the needed operating range and data rate with as low power consumption as possible.

4.3.2 Performance demands

In the simplest case, the performance of transceivers can be quantified in terms of maximum operating distance [58]. However, the maximum operating distance alone, is in most cases not sufficient to characterize or analyze the performance of a CS. Different systems are designed for operation in different environments and to meet different combinations of demands to data rate, bandwidth, power consumption, operating distance and so forth. Fulfilling one demand can make it difficult to meet others. Tradeoffs between these parameters must be made according to each application's set of demands. To comply with different sets of demands, a number of different system architectures, component designs and schemes for transferring information have been developed.

In the case studied here, an operating range of a few meters is needed. The required data rate is small. Todays glucose meters give a measured value of glucose with an accuracy of 3 decimal digits, which is sufficient. Thus 10 bits would be enough to represent the value from one measurement. Reliable transmission is needed because the actions taken or not taken by the user as a consequence of wrongly read values, could cause severe medical complications. The error correcting/detecting code mentioned in the previous section will involve some redundant bits. In order to avoid interference between units operating close to each other and reading other users' glucose level, an address system could be included in the transfer protocol. Sampling the glucose value more than once a minute

does not seem reasonable as changes in the glucose value takes place over relatively long periods of time. Hence, the amount of data to be transferred per data sample is still very small and from this point of view the system requires a very low data rate.

An important performance parameter in communication systems is the bandwidth efficiency. It is a measure for how large data rate a system supports within the given bandwidth. It is measured in *bits* $s^{-1}Hz^{-1}$ and can be found for a system by dividing the data rate r_b by the bandwidth W [13]:

$$\eta = \frac{r_b}{W} \tag{4.30}$$

The bands suggested for use with this application in section 4.4 have all bandwidths of at least 20kHz. The low data rate, should make it quite easy to find a modulation scheme that can meet the demand for bandwidth efficiency.

Also decisive for a communication system's performance is the bit error rate (BER). It is defined as [13]:

> *The ratio of erroneous bits received to the total number of bits received.*

It is closely related to the ratio E_b/N_0, that is the signal energy per bit divided by the noise density. Different modulation/coding schemes and system architectures require different E_b/N_0 to give a certain BER. This is decisive for the system's power consumption.

There is an inherent trade-off between the bandwidth efficiency and the bit-energy-to-noise-density requirement of a modulation/coding scheme [13]. Shannon gave the following bound on capacity for a given E_b/N_0:

$$\frac{E_b}{N_0} \geq \frac{2^{\eta} - 1}{\eta} \tag{4.31}$$

This bound is plotted in figure 4.17. We see from the figure that we can trade less bandwidth efficiency for lower E_b/N_0. E_b/N_0 depends among other parameters on the transmitted power. It is therefore one of the main parameters in the power consumption budget in the communication system. Hence a low bandwidth efficiency allows for greater power efficiency.

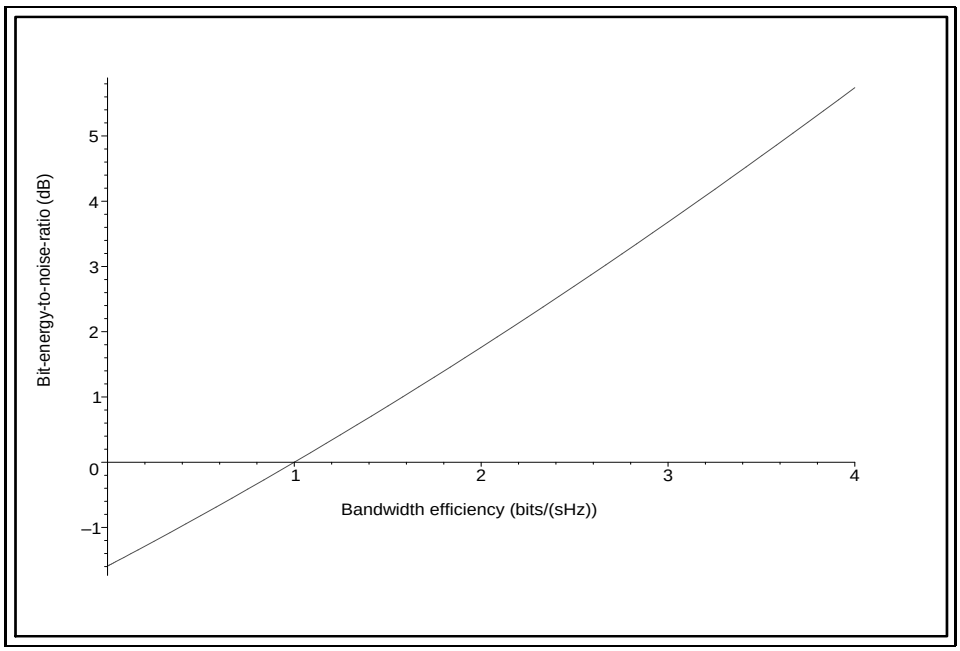


Figure 4.17: *Plot of the Shannon bound on capacity.*

4.3.3 Modulation Techniques

Modulation[4] is the technique used to make the transmitted signal carry the information in a baseband signal. Modulation can be defined as the process, or result of the process, of varying a characteristic of a carrier, in accordance with an information-bearing signal [19]. *The carrier signal* is a signal suitable for transmission in the desired RF channel. The carriers amplitude, frequency or phase is varied as a function of the information in the baseband signal. The result is *the modulated signal*, and this is transmitted through the channel. When detected by the receiver, the variation in one or more of the characteristics mentioned above can be detected and the information retrieved.

[4]This section is based on chapter 2 and 3 of the book "Modulation and Coding for Wireless Communications" written by Alister Burr [13] and chapter 2 and 3 of the book "RF Microelectronics" written by Behzad Razavi [58].

Analog and Digital Modulation

Amplitude modulation, frequency modulation and phase modulation are three different techniques often used for analog transfer of information over a communication link. In amplitude modulation it is the amplitude of the carrier wave that is information carrying, see figure 4.18. For example, if the information is sound waves transformed into voltage variations in a microphone, the amplitude of the carrier varies as a function of these voltage variations. The receiver can detect these amplitude variations, transform them into voltage variations again and reproduce the sound. Amplitude modulation can be described mathematically by [58]:

$$s_{AM}(t) = A_c[1 + mi_{BB}(t)]\cos(\omega_c t) \tag{4.32}$$

Here s_{AM} is the AM-modulated signal, A_c the amplitude of the carrier, m the modulation index, i_{BB} the baseband signal and t is time. The modulation index indicates to which degree the carrier signal is modulated. If $m = 1$, then the amplitude of the modulated signal varies between 0 and some maximum amplitude. The equation for m in amplitude modulation is:

$$m = \frac{A_{max} - A_{min}}{A_{max} + A_{min}} \tag{4.33}$$

Here A_{max} is the maximum amplitude of the signal and A_{min} is the minimum amplitude. In figure 4.18 the amplitude varies between 2 and 0 and hence $m = 1$.

In frequency modulation the frequency is varied between two limits centered around a carrier frequency. Here the variation in frequency is information carrying. Frequency modulation is described mathematically by [58]:

$$s_{FM}(t) = A_c\cos(\omega_c t + m\int_{-\infty}^{\tau} i_{BB}(t)d\tau) \tag{4.34}$$

An FM-modulated signal is shown in figure 4.19. In phase modulation it is the total phase of the signal that is varied as a function of the information. It's mathematically described as [58]:

$$s_{PM}(t) = A_c\cos(\omega_c t + mi_{BB}(t)) \tag{4.35}$$

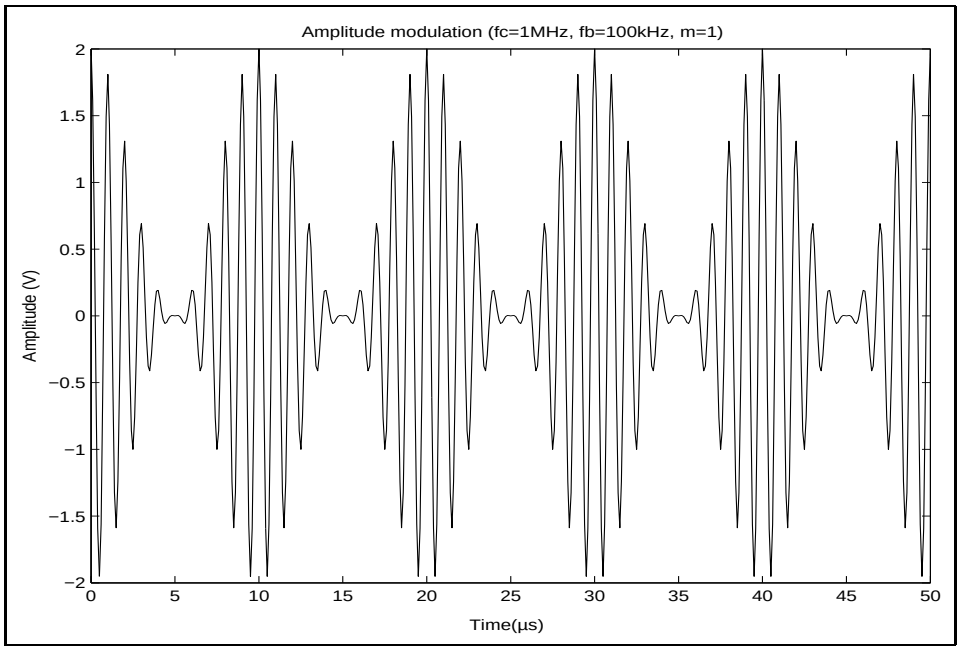


Figure 4.18: *An AM-modulated 1MHz carrier in the time domain.*

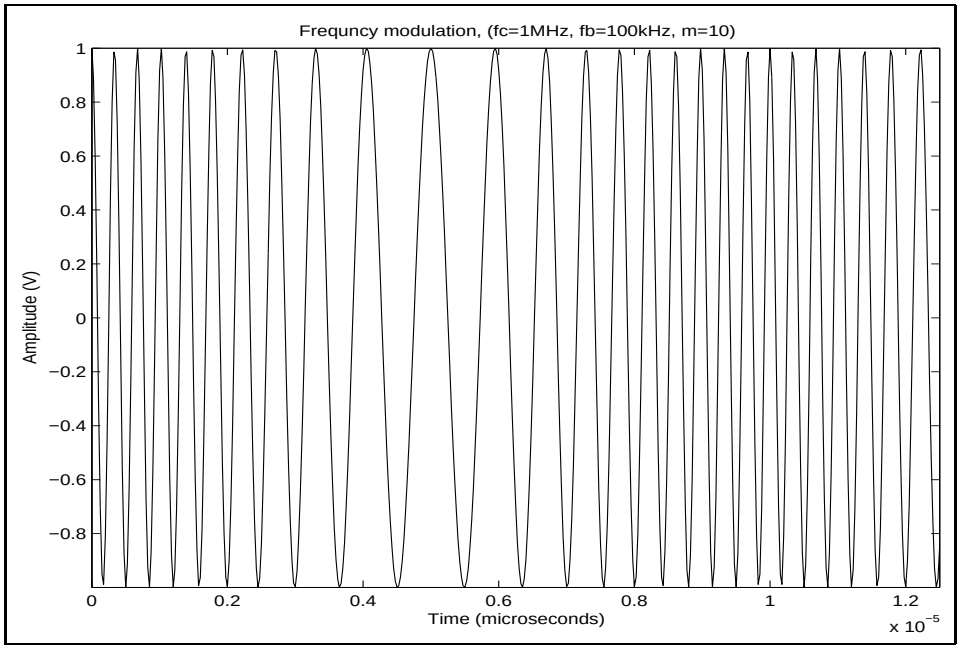


Figure 4.19: *An FM-modulated 1MHz carrier in the time domain.*

When the baseband signal is digital, the modulation technique is called digital. The technique makes it possible to transmit a bit stream of data. This means any data in a binary format can be transmitted using the same transmitter. A mobile phone for instance can use the same transmitter to transfer sound, text messages (SMS) and pictures (MMS).

Another advantage over analog modulation is the robustness of digital modulation schemes. An analog modulation scheme is to provide the receiver with a close to exact reconstruction of the analog signal at the transmitter. This only allows for small errors before the signal is too corrupted to recreate the original signal in a satisfactory way (how small is dependent on the application). If the baseband is digital, it is enough to distinguish between two values. In addition binary coded data can be accompanied by error correcting or error protecting code to ensure a reliable data transfer.

The digital counterparts of the analog modulation schemes mentioned above are amplitude shift keying (ASK), frequency shift keying (FSK) and phase shift keying (PSK). Except for in combination with PSK, in schemes known as quadrature amplitude modulation (QAM), amplitude modulation schemes have rather limited use. Pure ASK is seldom used. Although it allows for a very simple system architecture, the power efficiency is very poor [13]. Amplitude modulation has little robustness against noise. A special problem arises if zero signal is to be interpreted as transfer of the bit zero. It gets difficult to separate situations where the signal does not reach the receiver or no signal is actually being sent, and when the transmitter is actually transmitting the bit zero. Hence, pure ASK is not discussed in the following presentation of different digital modulation schemes.

Binary Phase Shift Keying

In binary PSK (BPSK), the phase of the carrier is alternated between $\phi = \pi$, symboling the bit '0', and $\phi = 0$, symboling the bit '1'. A principle schematic of a BPSK-modulator is shown in figure 4.20. The carrier is multiplied with -1 when the baseband signal is '0' and with 1 when the baseband signal is '1'.

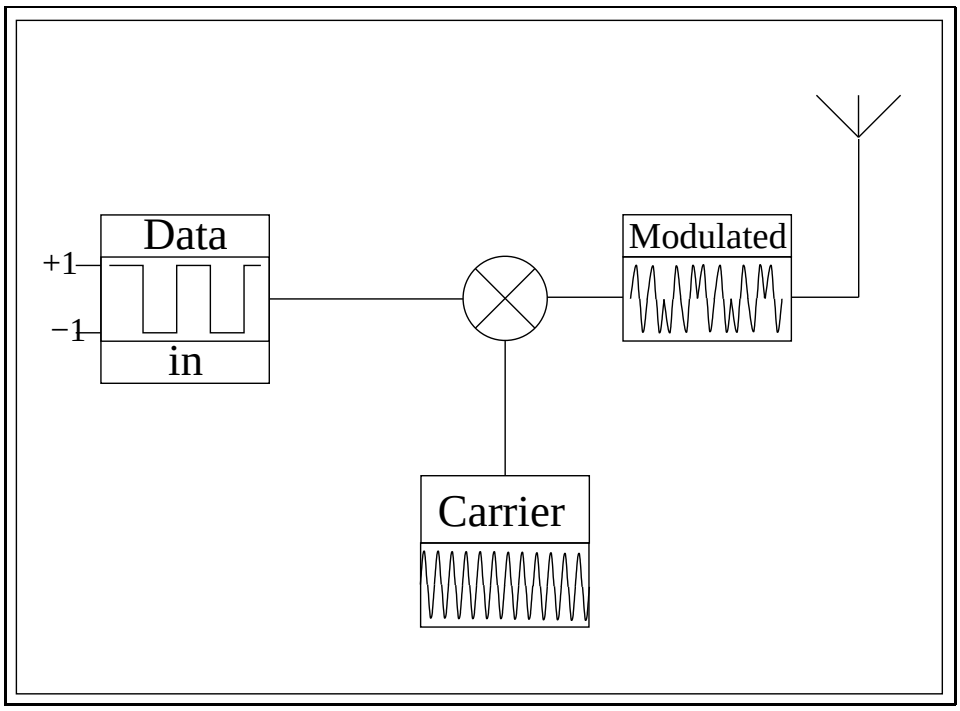


Figure 4.20: *A principle schematic of a BPSK-modulator.*

Consider a baseband signal starting at $t = -T$ and ending at $t = T$, composed of two bits centered around zero. The spectrum of a BPSK modulated version of this signal is found by taking the Fourier transform:

$$F_{BPSK} = \int_{-T}^{0} -\cos(\omega_c t)e^{-j\omega_c t}dt + \int_{0}^{T} \cos(\omega_c t)e^{-j\omega_c t}dt \quad (4.36)$$

$$= \frac{j(1-\cos^2(\omega_c T))}{\omega_c} \quad (4.37)$$

The purely imaginary result is due to the asymmetry of the baseband signal around zero. If we instead considered one pulse placed symmetrically around zero, we would get a $\sin(x)/x$ form on the spectrum. If we know the spectrum of one such pulse, we can use the time-shifting property of the Fourier transform, find the spectrum of all other similar pulses although they are shifted in time. This spectrum is often used for theoretical analysis. However, if we consider signals with long and complex bit sequences, the spectrum's shape is altered.

When Nyqiust filtering[5] is applied, BPSK has a bandwidth of [13]:

$$W = r_s(1+\beta) = r_b(1+\beta) \quad (4.38)$$

β is known as the rolloff factor and is related to the shape of the frequency response and the time domain properties of the Nyquist filter. Note also that in BPSK the modulation rate[6] r_s equals the information rate[7] r_b. By equation 4.30 this gives a bandwidth efficiency of:

$$\eta = \frac{1}{(1+\beta)} \quad (4.39)$$

In [13] Burr presents a theoretical estimate of the BER for BPSK and other modulation techniques applied in an additive white Gaussian noise (AWGN) channel[8] based on their constellation diagrams. Such a diagram

[5]Nyquist filtering - filtering done to limit the bandwidth of the baseband signal without causing intersymbol interference

[6]Modulation rate - symbols per time unit, also called symbol rate or baud rate

[7]Information rate - bits per time unit, also called bit rate

[8]The noise amplitude follows a Gaussian distribution. Additive means that it is added to the signal amplitude.

for BPSK is shown in figure 4.21 A. A bit would be wrongly interpreted if noise displaces the value across the quadrature axis. The probability of error is:

$$P(\text{error}) = P('0')P('0' \rightarrow' 1') + P('1')P('1' \rightarrow' 0') \tag{4.40}$$

Here $P('0')$ represents the probability of the bit '0' being sent and $P('0' \rightarrow' 1')$ the probability of the bit being misinterpreted as a '1'. Correspondingly $P('1')$ is the probability of '1' being the bit sent while $P('1' \rightarrow' 0')$ is the probability that a sent '1' is misinterpreted as '0'. Let d denote the distance or difference in signal amplitude between the two points representing the signals '1' and '0' in the constellation diagram.

To find the probability $P('0')P('0' \rightarrow' 1')$ the probability of the noise amplitude displacing the signal with a distance greater or equal to $d/2$ in the direction of increasing in-phase amplitude is needed. Hence we need to find the probability of the noise amplitude being $d/2$ or more. For an AWGN channel the probability in question, is determined by (see appendix C.3 equation C.13 for the probability density function of AWGN):

$$P(n \geq d/2) = \int_{d/2}^{\infty} p(n)dn = \frac{1}{2}\text{erfc}\left(\frac{d/2}{\sqrt{2}\sigma}\right) \tag{4.41}$$

where *erfc* denotes the complementary error function[9]. Because of the symmetry in the Gaussian distribution around the average value, $P('1' \rightarrow' 0')$ will be equal to $P('0' \rightarrow' 1')$. If we assume that there is an equal number of '0"s and '1"s in the transmitted data sequence, then $P('0') = P('1') = 0.5$. By applying these results to equation 4.40 we end up with:

$$P(\text{error}) = \frac{1}{2}\cdot\frac{1}{2}\text{erfc}\left(\frac{d/2}{\sigma}\right) + \frac{1}{2}\cdot\frac{1}{2}\text{erfc}\left(\frac{d/2}{\sigma}\right) = \frac{1}{2}\text{erfc}\left(\frac{d}{2\sigma}\right) \tag{4.42}$$

We can relate the BER to the SNR of the output of a matched filter (see appendix C.4) using equation C.30 and the fact that in the case of BPSK $d_p = d/2$ and $d_q = 0$:

$$\frac{d_p^2 + d_q^2}{\sigma^2} = \frac{d^2}{4\sigma^2} = \frac{2E_s}{N_0} = \frac{2E_b}{N_0} \tag{4.43}$$

[9] $erfc(z) = 1 - erf(z) = \frac{2}{\sqrt{\pi}}\int_z^{\infty} e^{-t^2}dt$

Inserting this into equation 4.42 we get:

$$P_{\text{error}} = \frac{1}{2}\text{erfc}\left(\frac{1}{\sqrt{2}}\sqrt{\frac{2E_b}{N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{4.44}$$

The last form in this equation is introduced to have a short form to represent this error function.

Quadrature Phase Shift Keying

QPSK represents bits with four different phases instead of BPSK's two as seen in the constellation diagram given in figure 4.21 B. This implies doubling the number of bits sent per symbol and that we can send twice as many bits within the same bandwidth:

$$W = r_s(1+\beta) = \frac{r_b}{2}(1+\beta) \tag{4.45}$$

This gives an improvement in the bandwidth efficiency of a factor two compared to BPSK:

$$\eta = \frac{r_b}{W} = \frac{r_b}{\frac{r_b}{2}(1+\beta)} = \frac{2}{1+\beta} \tag{4.46}$$

The spectrum of a QPSK-modulated signal will have the same shape as the BPSK-spectrum given the same symbol rate.

To find the relationship between the BER and E_b/N_0 in QPSK we can use a similar approach to that shown for BPSK. The result turns out to be the same as for BPSK [13]:

$$P(\text{biterror}) = \frac{P_{error}}{2} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{4.47}$$

Other Linear Modulation Schemes

Both BPSK and QPSK belong to a class of modulation schemes known as linear. Within this class of schemes the baseband information is communicated through phase and amplitude alterations of the carrier. Such

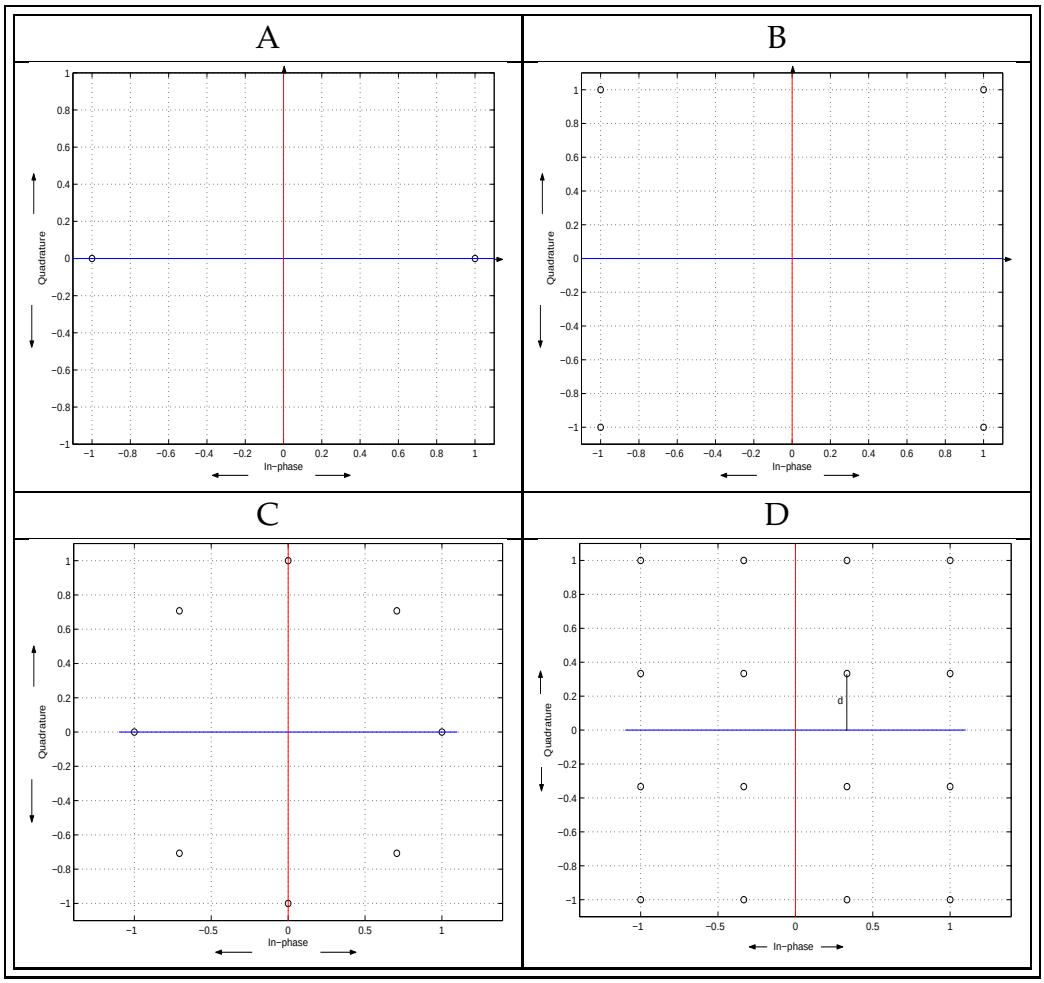


Figure 4.21: *Constellation diagram for several digital modulation schemes; A: BPSK, B: QPSK, C: 8-PSK, D: 16-QAM. The constellation points are marked with black circles.*

alterations can be expressed mathematically by linear combinations of a signaling pulse $g(t)$. In a linear modulation scheme we can therefore write the baseband signal as [13]:

$$b(t) = \sum_{i=-\infty}^{\infty} d_i g(t - iT) \tag{4.48}$$

Here d_i represents the i'th bit of the baseband bit stream and T the sampling period.

It is possible to increase the number of constellation points in a scheme. This will increase the bandwidth efficiency. Two widely used techniques in bandwidth limited applications are multi phase shift keying (MPSK) and quadrature amplitude modulation (QAM). An 8-PSK scheme has eight constellation points each representing three bits as shown in figure 4.21 C. A 16-QAM scheme has 16 constellation points, each point represents 4 bits as shown in figure 4.21 D. The drawback of increasing the number of constellation points, is that more energy is needed per bit to achieve the same BER thus reducing power efficiency. Hence, BPSK and QPSK are the most power efficient linear modulation schemes.

In table 4.2 a list of characteristics of the modulation schemes presented in this discussion is found. The same approach that was lined out above for BPSK and QPSK is used to arrive at the formulas given for MPSK and QAM.

Frequency Shift Keying

When exposed to nonlinear effects[10] , the linear schemes discussed above encounter effects corrupting the signal. Nonlinear schemes do not have the same limitations in these regards. Nonlinear or exponential modulation involves modulation of the rate of change of phase. This can not be expressed by equation 4.48.

Frequency shift keying (FSK) is a nonlinear modulation scheme. In FSK-schemes the digital baseband signal determines the frequency of the

[10]Nonlinear effects - can arise for instance from electronics or signal interaction with operating environment or signal interference

signal. In the binary case, the signal will alter between two frequencies. We can express this mathematically by [13]:

$$\begin{aligned} s(t) &= \begin{cases} Re\{e^{j\omega_d t}e^{j\omega_c t}\} \\ Re\{e^{j-\omega_d t}e^{j\omega_c t}\} \end{cases} \\ &= \begin{cases} \cos((\omega_c+\omega_d)t), \ d_i = +1('1') \\ \cos((\omega_c-\omega_d)t), \ d_i = -1('0') \end{cases} \end{aligned} \tag{4.49}$$

Here the baseband signal is represented by the ω_d-terms. The frequency shifts between $\omega_c - \omega_d$ and $\omega_c + \omega_d$ which correspond to the bits '0' and '1' respectively. An important parameter in FSK modulation schemes is the modulation index. It relates the difference in frequency between the two frequencies with the symbol rate T_s [13]:

$$m = \frac{\omega_d T_s}{\pi} = 2f_d T_s = \frac{f_d}{r_s} \tag{4.50}$$

This parameter influences both demodulation and spectral properties.

Phase continuity also plays an important role for the spectral properties. Abrupt phase changes, that is discontinuities in the time domain, leads to large side lobe levels in the frequency domain. This can be avoided by ensuring smooth transitions between symbols, in other words keeping the phase continuous when switching between the two operation frequencies. FSK-schemes with such phase properties are called continuous phase FSK (CP-FSK).

It can be shown[11] that optimum detection of FSK is achieved when:

$$m = \frac{k}{2}, \ k \in \mathbb{N} \tag{4.51}$$

If this condition is fulfilled, the signals for bit '0' and bit '1' are said to be orthogonal. The special case of binary continuous phase frequency shift keying with $m = 0.5$ is called minimum shift keying (MSK).

By increasing the number of possible frequencies to shift between, we can increase the number of bits sent per symbol. We can write the baseband signal in a more general form than the way it was represented in equation 4.49:

$$b(t) = e^{jd_i\omega_d t} \tag{4.52}$$

[11] See appendix C

For binary FSK d_i would vary between $+1$ and -1. For M-ary FSK, d_i can take on the values of the uneven numbers [-(M-1), -(M-3) ...-1, 1 ..., (M-3), (M-1)], ensuring orthogonal signals. An MFSK scheme occupies a bandwidth of

$$BW = Mmr_b. \tag{4.53}$$

With a demodulator structure with M branches, it can be shown that the relationship between the BER and bit-energy-to-noise is [13, 55]:

$$\begin{aligned} P_b &= \frac{2^{k-1}}{2^k-1}\int_{-\infty}^{\infty}\left(1-Q(-y)^{M-1}\right)\cdot \\ &\quad \frac{1}{\sqrt{2\pi}}\exp\left(-0.5\left(y-\sqrt{2kE_b/N_0}\right)^2\right)dy \\ &\approx \frac{M}{2}Q\left(\sqrt{2k\frac{E_b}{N_0}}\right) \end{aligned} \tag{4.54}$$

As is seen in figure 4.22, the M-FSK schemes using M orthogonal signals are the schemes that can obtain the best power efficiency, but at the same time they have very poor bandwidth efficiency.

To achieve better bandwidth efficiency and lower sidelobe levels we can change the baseband signal of MSK. By applying a Gaussian filter on the baseband signal, a scheme called Gaussian minimum shift keying (GMSK) is obtained. A scheme called tamed FM (TFM) is achieved when using a baseband signal filter optimized to achieve the best possible bandwidth efficiency. The increase in bandwidth efficiency again has the drawback of reducing power efficiency. TFM is less used than GMSK because it is more complicated to generate and receive. In addition, the spectrum of GMSK can be shaped as desired by altering the Gaussian filter characteristics.

Concluding remarks to the Theoretical Discussion of Modulation Techniques

An overview of the theoretical results for bandwidth efficiency in the above discussion is summarized in table 4.2. Plots of the BER versus E_b/N_0 is shown in figure 4.22 for some of the schemes discussed. If we compare

Modulation scheme	BW efficiency η
BPSK	$\frac{1}{1+\beta}$
QPSK	$\frac{2}{1+\beta}$
M-PSK	$\frac{log_2(M)}{1+\beta}$
M-QAM	$\frac{log_2(M)}{1+\beta}$
MSK	1
M-FSK	$1/Mm$

Table 4.2: *Overview of the bandwidth efficiency of some modulation schemes discussed in this section.*

the power efficiency and bandwidth efficiency of the schemes in the plot, it confirms the inherent trade-off of these two important pararmeters described by the Shannon bound in equation 4.31.

A theoretical discussion of modulation techniques alone is not sufficient to give full insight into how these techniques will perform when implemented in a real system. A communication channel suffering from non-linear effects degrades the performance of linear modulation schemes. Such degradations are [13]:

- Generation of self-interference due to in-band intermodulation products. It causes spreading of the constellation points seen in figure 4.21 even without the presence of noise.
- Variations in received amplitude will result in non-linear variations of output amplitude and phase shifts; so called AM-AM and AM-PM conversions. It will cause displacment of the constellation points.
- Spectral regrowth; sidelobes in the spectrum filtered out with Nyquist filters reoccur.

Sources of non-linear effects in a communication system can be:

- Non-linear power amplifiers. (Are offend preferred in power limited systems due to their superior power efficiency).

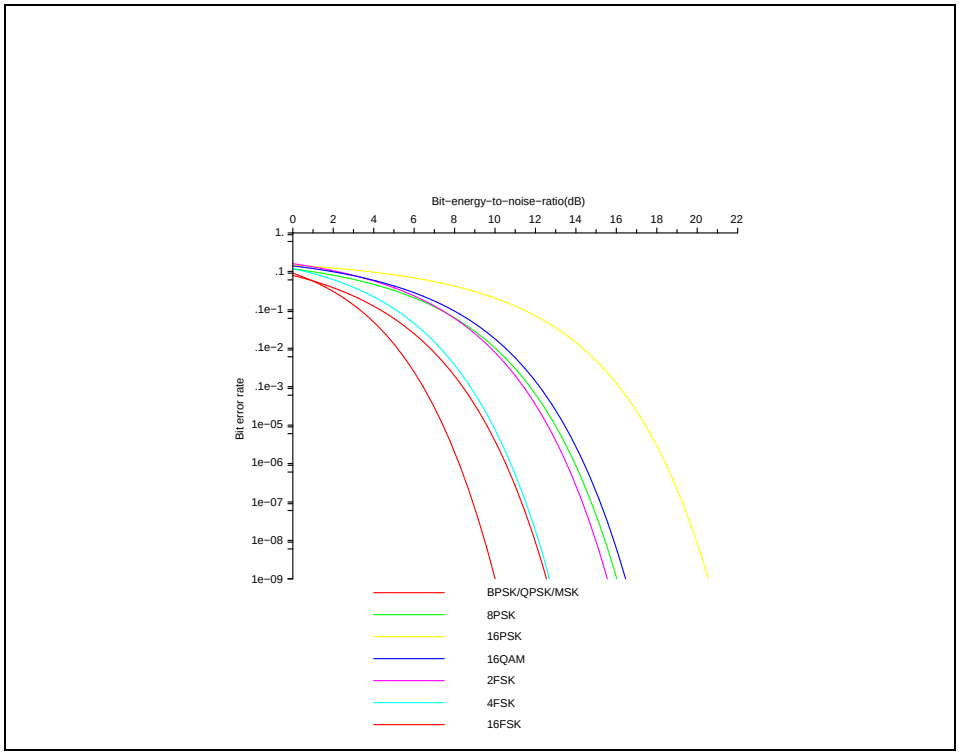


Figure 4.22: *BER versus E_B/N_0 for some of the modulation schemes discussed here. This enables us to compare the power efficiency of these schemes.*

- Multipath signal propagation.

This gives nonlinear modulation schemes an advantage over linear schemes regarding power efficiency and are thus used in power limited systems. The linear schemes therefore have restricted application [13] and are most often used in systems requiring high bandwidth efficiency and where power is not of major concern, however exceptions exist.

4.3.4 An example based on a commercially available transceiver

The architecture of a radio transceiver can be quite complicated and a discussion regarding all aspects of an active communication system design could easily become too extensive. Therefore, to indicate what performance can be expected from an active system it was decided to try to set up a link budget based on Nordic Semiconductor's single chip transceiver nRF905 using the 433MHz ISM-band. Referring to the situations illustrated in figure 1.1 on page 5, the example discussed here would correspond to the bottom right situation.

The receiver antenna is taken to be a $35mm \times 25mm$-loop antenna with an antenna efficiency of $\eta_{dB} = -14.5dB$ described in [46] and suggested as a suitable antenna for the nRF905 in [47]. For simplicity the transmitter antenna is assumed to have the same efficiency.

The transmission loss in the body was estimated with using the calculation procedure of section 4.5 in "Antennas in Matter" [31]. The physical model is sketched in figure 4.23. The barrier is to simulate the outer layers of an animal body. In this example we want to calculate the attenuation of a wave propagating from the muscle layer, medium 3 in figure 4.23 toward medium 0. Medium 2 has the electromagnetic properties of fat, medium 1 skin and medium 0 is air. We assume that the fat layer has a thickness of $a = 1.5cm$, the skin $b = 0.1cm$ and the air medium extends to ∞. By solving the coefficient equations found from the boundary conditions for both the electric and magnetic field components of the wave at each of the three medium interfaces, the amplitude of the normalized E-field was found for

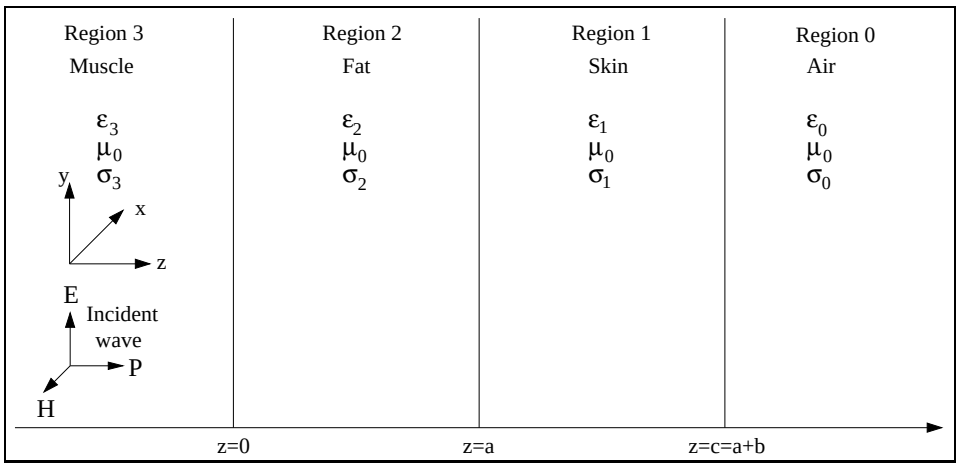


Figure 4.23: *A schematic drawing of a plane wave incident on a three layer barrier.*

all three sections at $433MHz$. We find the power loss, using the time averaged Poynting vector for a plane wave traveling in the z-direction [57] (using symbols corresponding to figure 4.23) and get:

$$< P_z >= -\frac{1}{2}\Re\left\{E_y H_x^*\right\} \tag{4.55}$$

The power lost is then found by

$$L_{body} = -10\log\frac{< P_z >_{out}}{< P_z >_{in}} \approx 30dB \tag{4.56}$$

The basic free space transmission loss describes the geometrical losses due to the spreading out of the transmitted power on a spherical surface of radius equal to the transfer distance. In free space a small loop antenna can be shown to have a maximum directivity of approximately 1.5 in the plane of the loop and it has a null in the directions of the loop axis. We assume that the antennas are aligned in such a way that the directivity is 1. Judging from the situation in the bottom right illustration in 1.1 on page 5 an operating range of approximately of $d = 2m$ would be sufficient. From equation (4.14) we have that the basic free space transmission loss is:

$$L_{FS} = -20\log\left(\frac{\lambda}{4\pi d}\right) \approx 31dB \tag{4.57}$$

LINK BUDGET	
System Parameter	Value
Transmit power	$-10dBm$
- Transmitter antenna loss	$14.5dB$
- Loss body	$30dB$
- Basic free space transmission loss	$31dB$
- Receiver antenna loss	$14.5dB$
Received signal power	$-100dBm$

The nRF905 has a sensitivity of $-100dBm$ given a 0.1% BER. From the link budget we see that 2m in fact is the limit of the operating range. However, it should be emphasized that the budget is only a rough estimate.

Now, based on information in the product specification of the nRF905 chip [47], we want to estimate the energy consumption and the needed battery capacity. According to [47], the nRF905 has a current consumption of 9mA for the transmitting mode producing an output power of $-10dBm$; given that the supply voltage is 3.0V. This implies a power consumption of $27mW$. In the receiver mode the current consumption is $12.5mA$ giving a power consumption of $37.5mW$ and in power down mode $2.5\mu A$ which gives $7.5\mu W$ power consumption.

Assuming we need to transmit in the order of 100 bits once a minute a rough estimate of the energy consumption is shown in the table below.

Energy consumption 1 min		
Power down mode	$7.5\mu W \cdot 60s =$	$450\mu J$
Transmit mode	$27\mu W \cdot 10ms =$	$270\mu J$
Receiver mode	$37.5\mu W \cdot 10ms =$	$375\mu J$
Correction for BER	$(0.1) \cdot (270 + 375)\mu J =$	$64.5\mu J$
Total	$\approx$	$1.2mJ$
Energy consumption one year	$1.2mJ \cdot 60 \cdot 24 \cdot 365 =$	$630J$
	$=$	$0.175Wh$

Based on the number for the energy consumption for one year we can

estimate the needed battery capacity to power the transceiver for one year. Given a battery that can supply 3.0V for the current levels given above, a rough estimate of the needed battery capacity would be:

$$\frac{0.175Wh}{3.0V} = 58mAh \tag{4.58}$$

Reviewing the battery capacities and energy densities of table 3.1 such capacities would demand a battery size in the order of $1cm^3$.

4.4 Choosing the operating Frequency

The radio frequency bands of the electromagnetic spectrum is in great demand. International cooperation to harmonize standards takes place within the International Telecommunication Union (ITU) and also regionally, for instance in Europe by the European Conference of Postal and Telecommunications Administration (CEPT). National authorities sets up frequency plans to try to ensure an effective utilization of the spectrum. In these plans the frequency spectrum is divided into bands, each band designated for a specific use. Type of communication, duty cycles, channel separation, out-of-band power levels and radiated power can be restricted within each band. The bandwidth of a band is defined such that the signal power falling outside the band is below some specified threshold [13].

The choice of operating frequency for a communication system is subject to the restrictions mentioned above and other available design criteria. For the application discussed here, the relationship between antenna size and operating frequency is of special importance. The antenna is most efficient when it operates at it's resonance frequency. Small sized antennas have high resonance frequencies. However, both geometric losses and the losses due to signal attenuation in body tissues increase with higher operating frequency. Special techniques for miniaturization of antennas have been developed, but these imply degrading other often important characteristics. This will be investigated closer in the following sections.

The power consumption of the electronics of the system, also depends on frequency; the higher the frequency, the higher the power consumption.

In Norway the legislative authorities have set aside a number of bands for free use. Frequency bands that could be of interest for the communication system discussed here, are listed in table 4.3.

International agreements on frequency allocation can be found in the ITU Radio Regulations, Article S5, Frequency allocations. In these agreements some parts of the spectrum have been set aside for Industrial, Scientific and Medical use (ISM-bands). These bands have recently also been used for license free wireless applications like Bluetooth. The bands $6.765 - 6.795$MHz, $13.553 - 13.567$MHz and $26.957 - 27.283$MHz are ISM-bands. Despite that some effort has been put down in international harmonization, it is not unlikely that the system must adjust for different national frequency plans around the world. However, special exceptions are often made for medical equipment as long as it does not cause unacceptable interference with the service designated for the band in question.

4.5 Biological Effects of Electromagnetic Radiation

The knowledge of electromagnetic magnetic properties of non-biological material is quite extensive, but the knowledge of such properties in biological material is more limited. Biological material is a complex matter. It is both inhomogeneous and anisotropic. This makes it difficult to give good description of how biological tissue and electromagnetic fields interact.

4.5.1 The Structure of the Biological Medium

To understand the effects of electromagnetic radiation on biological matter, basic knowledge of the structure of the biological tissues, the cell and the interstitial fluid, is necessary. The cell membrane is isolating. Normally there is a voltage difference between the inside and outside of a cell of about 70mV for an excitable cell and 20 mV for other cell types (The inside is negative compared to the outside). The cell consists of a number of smaller structures which flows freely in the cytoplasm. The cytoplasm consists of the cell organelles and the cytosol which again consists of water, electrolytes, nutrients and a network of proteins. This network is called the

Frequency band	Maximum radiated effect/ Maximum field strength	Allowed for
9 - 59.750kHz	72 dBmA/m at a distance of 10m Reduced by 3dB every octave from 30kHz	Inductive applications
60.250kHz - 70kHz	72dBmA/m at a distance of 10m Reduced by 3dB every octave from 30kHz	Inductive applications
119 - 135kHz	72dBmA/m at a distance of 10m Reduced by 3dB every octave from 30kHz	Inductive applications
59.750 - 60, 250kHz	42dBmA/m at a distance of 10m	Inductive applications
70 - 119kHz	42dBmA/m at a distance of 10m	Inductive applications
6.765 - 6.795MHz	42dBmA/m at a distance of 10m	Inductive applications
13.553 - 13.567MHz	42dBmA/m at a distance of 10m	Inductive applications
26.957 - 27.283MHz	42dBmA/m at a distance of 10m	Inductive applications
7.4 - 8.8MHz	9dBmA/m at a distance of 10m	Inductive applications
402 - 405MHz	25 mW e.r.p (Channel separation up to 300 kHz)	Medical implants
9 - 315kHz	-	Medical implants in accordance with ERC/REC70-03 appendix 12

Table 4.3: *Frequency bands that have potential interest for medical implants in the Norwegian national frequency plan. (e.r.p stands for effectively radiated power)*

cell skeleton. K^+, Na^+ and Ca^{2+} are important ions in the cell and in the interstitial fluid. This means that ions and polarizeable components are important elements in biological tissue. It should be reasonable to assume that this is of great importance for the electromagnetic properties of biological tissue. A description of interaction between biological tissue and electromagnetic radiation on a microscopic level would be the most fundamental. Such interaction averaged in space and time can be depicted by the electromagnetic quantities ϵ, μ and σ. The inhomogeneity and anisotropy of biological material makes it hard to retrieve well defined values of these parameters.

Currents induced by low frequency electric fields will be confined to the interstitial fluid due to its high conductivity. The cells isolating membrane and highly conductive inner makes it a series combination of the membranes capacitance and the cytoplasm's resistance. The dielectric properties of cells, and thereby their ability to act as capacitances depends on frequency. Cells typically have a large dielectric constant at low frequencies which decreases in more or less distinctive steps, as frequency increases [73].

4.5.2 Interaction between Biological tissue and EM-fields

Biological tissue is generally considered to be non-magnetic, that is the relative permeability is 1 or $\mu = \mu_0$. This means that the interactions are mainly due to the electric field. On the microscopic level there are three kinds of interactions that are essential to describe macroscopic effects [22]:

- Polarization of bound charges
- Orientation of permanent electric dipoles
- Drift of conduction charges.

Polarization of bound charges takes place when a they are exposed to an external electric field. The field forces the charges out of their original position. The charges are displaced. The electrical field induces forces with opposite direction on positive and negative charges. Hence, positive

and negative charges will create an induced dipole. From a macroscopic view, the movement of charges is small. Assuming the wavelength of the radiation is not much smaller than the dimensions of the medium, the movement of many charges will result in polarization of the medium.

When permanent dipoles are exposed to an electric field, they tend to align with the applied field. If the field is constantly switching its polarity, the dipole will move back and forth to maintain the position of minimum energy. The moving of the particles sets up new fields which can interact with the surroundings.

Loosely bound particles will move when exposed to an electric field. The movement is called drift. The particles moves rapidly until they collied with another particle and continues in another direction. The net movement is in the direction of the electric field, and a current arises.

4.5.3 Absorption and transmission

The exchange of energy between electromagnetic waves and tissue can take place in two different ways, dissipation and storage. Dissipation of radio frequency radiation comes mainly form the induction of currents in tissue. As mentioned above, the resulting moving charges collide with other particles. The electrical energy is then converted into thermal energy.

Storage of energy can arise as a result of polarization of bound charges. Polarization demands movement of charges, causing a displacement current. The movement is dissipative.

On a boundary between to different media, electromagnetic fields must, according to Maxwell's equations, fulfill certain boundary conditions. Dependent on the differences in the electromagnetic properties of the two media, the angle of incidence and the dimensions of the boundary compared to the wavelength, the wave will be reflected, transmitted, scattered or some combination of these.

Absorption of energy from EM radiation in the human body is frequency dependent. The body is relatively transparent to low frequency radiation. Here low frequency means frequencies below 1MHz. A frequency of 1MHz gives an attenuation constant of about 4Np/m. The cor-

responding skin depth is 25cm. This means a wave travels 25cm before it is attenuated by a factor of e^{-1}.

Absorption is considerably higher close to resonance frequencies. Habash [22] claims that isolated from ground, a grown up has a resonance frequency between 70 and 80MHz. Grounded, the resonance frequency is between 35-40MHz. The resonance frequency of children is some what higher than in adults.

Resonance can also arise in body parts. As an example, the resonance frequency of an adults head is about 400MHz, for a baby's head about 700MHz. A new resonance top is found at 1 GHz.

4.5.4 Biological Effects of the Interactions

The biological effects of the interactions between biological tissue and EM-field, can be considered form different points of view. The effects can occur on molecular level, cell level, organ or system level or in the entire body. Habash [22] follows Repacholi [59] and divides the effects in three different categories:

- Effects caused by high intensity radiation↔ Thermal effects
- Effects caused by medium intensity radiation↔ Athermal effects
- Effects caused by low intensity radiation↔ Non-thermal effects

Thermal effects can be defined as a energy transfer from the radiation to the body. The transfer is so big that the body's thermo-regulation system can't handle it. Areas with low blood circulation are more vulnerable to such effects than others. Examples of such areas are the lens of the eye and the testicles. Another element of risk is metal objects situated inside or close to the body such as wrist watches and spectacle frames. This is because the fields tend to concentrate around such objects. This makes the shape of an implant important. It is important that it does not have sharp edges.

The two remaining categories includes effects that occur even though the body's thermo-regulation system manages to keep the body temperature unchanged. There exists confirmed research results that such effects

do occur (see [22]. Confirmed here means experiments with reproduceable results. Today one is not able to explain the underlying mechanisms for these results.

4.5.5 Modulation

To transmit information, EM-signals has to be modulated. The modulation introduces new frequencies in the spectrum in addition to the carrier wave. One may have to use ELF-field theory to explain the effects that arises as a result of this modulation. Some investigations done indicate that human and animals easier adapts to a continues stable signal than one that is modulated or pulsed [22].

Chapter 5

Experimental Setups: Design and Integration

The primary objective of the experimental investigation was to obtain signal transmission data relevant to a short range communication system applicable to medical implants. Two different procedures for investigating the conditions for signal transmission were conducted. The first was relevant to the passive communication system discussed in chapter 4. The second could have relevance for both the passive and the active communication system solution. The effect of body fluids on signal transmission was also investigated. Here an aqueous sodium chloride solution was used to simulate the electromagnetic properties of body fluids.

The design and integration of the components needed for the experimental setup became an important and time-consuming preliminary work for these investigations.

5.1 The experimental setup

Two slightly different experimental setups were used for the transmission measurements. One setup was designed to measure the conditions in the transmitter for detection of impedance variations in the receiver circuit and hence was directly related to the design of the passive solution discussed in section 4.2. The second setup was designed to measure the at-

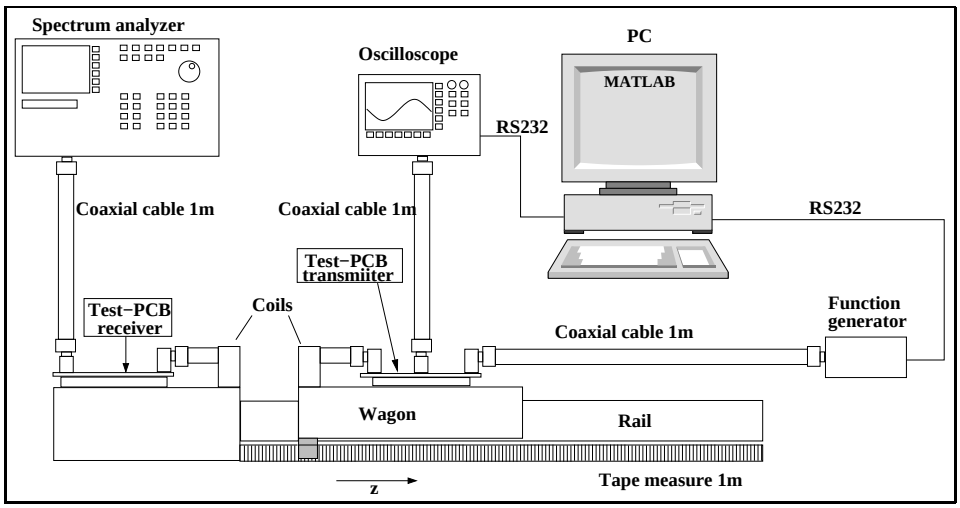


Figure 5.1: *Experimental setup used for determining how signal transmission limits the maximum operating range of a passive system.*

tenuation of the transmitted signal. This could have relevance to both the active and passive solution.

5.1.1 Signal transmission setup for impedance modulation

A schematic drawing of the experimental setup used with these measurements are shown in figure 5.1. The setup comprises the following main parts:

- Signal source; Agilent 33250A 80MHz
- Transmitter measuring system, a Tektronix TDS 3012 oscilloscope and a specially designed printed circuit board (PCB) supplied by a HP6236B tripple output power supply.
- PC for measurement control and data acquisition.
- PCB for the receiving part of the system.
- A HP4195A 10Hz-500MHz network/spectrum analyzer configured as a spectrum analyzer.
- Coordinate board (build from Lego)

The transmitter resonance circuit was fed with a signal from the function generator. The voltage over a resistor in series with the function generator was measured through an instrumentation amplifier implemented on the test PCB. The output signal of this amplifier was recorded on an oscilloscope. The scope was instructed to take RMS-voltage samples of the amplifier output and the result was transferred to a PC via an RS232-link. The PC was running a MATLAB-script that also controlled the settings of the function generator. The script was programmed in such a way that it first set a given frequency on the function generator, then adjusted the settings of the oscilloscope appropriately and finally queried the scope for the RMS-voltage. By repeating this procedure for several frequencies, the frequency characteristics of the system could be achieved.

At the receiver side the signal from the resonance circuit was analyzed using the HP4195A as a spectrum analyzer. The relative positions of the antennas were read on the coordinate board.

This experimental setup was modified to make it possible to do similar measurements when one of the coils was submerged in water as is shown in figure 5.2. The water tank is made of polypropylene, can contain 8l and it measures $(335 \times 225 \times 155)mm^3$. The extension cable connecting the coil feed cable and the SMA-contact on the receiver test PCB was an ALPHA 9158 58C/U 50Ω cable, 24 cm long with mounted SMA-connectors included. It was measured to have an inductance of $57.10nH$ and a capacitance of $24.41pF$.

For freshwater measurements tap water tapped at Sintef was used. The sodium chloride solution, was made by adding a sodium chloride amount of 1 % of the mass of the water in the tank. That is the $8l$ of water was added 80g of sodium chloride. As a control test, a small plastic cylinder with brass electrodes was filled with the solution and the conductivity was measured. The measurement was done with the HP4195A for a frequency sweep in the MHz region. By using the equivalent circuit function the resistance could be extracted. The corresponding conductivity of the solution was found to be 16.2 mS/cm which was in good agreement with values listed in "The Handbook of Chemistry and Physics" [34]. The listed values for 1.0% salinity seawater is 15mS/cm at $20^\circ C$ and $17mS/cm$ at $25^\circ C$.

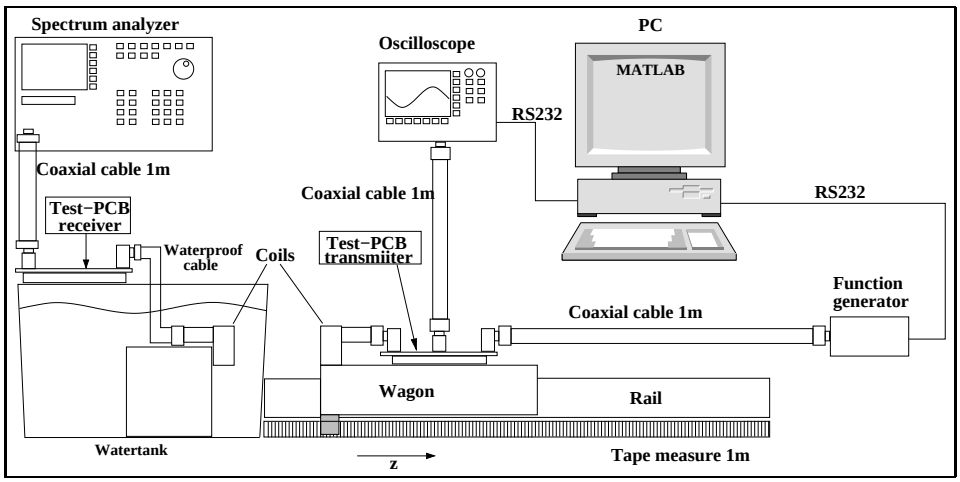


Figure 5.2: *Experimental setup used for determining how signal transmission is affected by water and aqueous sodium chloride solution.*

The room temperature is expected to have been $23 - 24°C$ at the time of measurement.

The function generator, the oscilloscope and the spectrum analyzer are all commercially available instruments with well defined specifications. The antenna coils and coordinate board system were hand made from parts of local or private origin without given specifications. The test-PCB was designed specifically for this experimental setup. Although the components used on the PCB were standard components with well defined specifications, the performance was not known before it was experimentally verified. The next paragraph gives a more detailed presentation of the coordinate board and PCB-design. As the assessment of the antenna characteristics in form of impedance measurements was time-consuming and important part of the preliminary investigations it is devoted a subsection of its own; section 5.2.

Coordinate board

The coordinate board was constructed from Lego. The positions of the antennas were controlled with a linear guide and a rotating device. The linear guide consisted of two rails and a wagon that could be manually

pushed and pulled along the rails into the desired position (The direction defined by the linear guide will later be referred to as the Z-direction). To read the position in the Z-direction a tape measure was mounted right besides one of the rails of the linear guide with double sided tape. A small Lego bar pointing out from the wagon front indicated the position on the tape measure. One could tell no difference between the tape measure and a more accurate ruler up to 50 cm. The position read from the tape measure can reasonably be expected to be accurate within some tenths of a mm.

A device that could be used to rotate either the transmitter antenna or the receiver antenna was also constructed in Lego. The device was designed so that the angle could be read on a protractor. The accuracy of this reading is expected to have been roughly $\pm 1^\circ$.

Preliminary experimental setups which were taken on tables with metal bars along the sides showed that the received signal strength would only decay with distance up to some point and then be approximately constant with further increasing distance. This can have been caused by eddycurrents induced in the metal and it was decided to construct a coordinate board in a material with high resistivity. Lego is made from ABS-plastic which has very high resistivity and a dielectric constant of approximately 3. Hence it was expected that it would not seriously affect the measurements.

PCB design

In preliminary tests it was found that standard axial leaded components mounted on a veroboard due to their relatively big stray impedances, especially resistors and some capacitors, did not have a satisfying performance at frequencies from about 5MHz and above. Motivated by Linear Technology Corporation's application note 47 [70] a breadboard design with an instrumentation amplifier measuring the voltage over a 51 Ω resistor was constructed see figure. These preliminary tests provided an opportunity to test different grounding configurations. Proper circuit grounding design was found to be crucial to achieve a system configuration that produced predictable measurement results. In agreement with [52] a single

point connection to AC ground was found to be the only suitable approach. These preliminary tests also showed the importance of a ground plane to give low inductance connections to the connection point of the AC ground, the components on the breadboard, the DC-power supply and the function generator.

On the basis of the lessons learned in the preliminary tests a dedicated PCB was designed to measure what effect a resistor in the sensor circuit had on the total impedance in the combined circuit. This was done by letting an instrumentation amplifier measure the voltage over a 51Ω resistor in the reader circuit.

To minimize loading of the resonance circuit, the first stage in the amplifier acts as a differential buffer. It consists of two OPA656 operational amplifiers configured as voltage followers. (The OPA656 has a FET-input stage which provides a very high differential input resistance of 1GΩ and low input capacitance of 0.2 pF. Together with a unity-gain bandwidth of 500 MHz this results in good high frequency performance, and sought specifically in this case, high input impedance also at high frequencies.

5.1.2 Transmission measurements with network analyzer

An important factor when considering whether or not it is possible to use a passive system design for the application in question is the amount of energy that can be transferred from the external to the implanted device. Applied with the 41952A Transmission Reflection set the HP4195A has a transmission measurement capability when it is configured as a network analyzer. In transmission measurements the amplitude ratio between a selected test input and a selected reference input is measured.

In order to achieve the maximum measurement accuracy for transmission measurements the instrument was calibrated by performing a **Normalize** calibration, which eliminates frequency response errors and sets a reference level, and an **Isolation** calibration which is used to eliminate crosstalk error.

The setup is shown in figure 5.4. This experimental setup was also modified to make it possible to measure the signal transmission when one

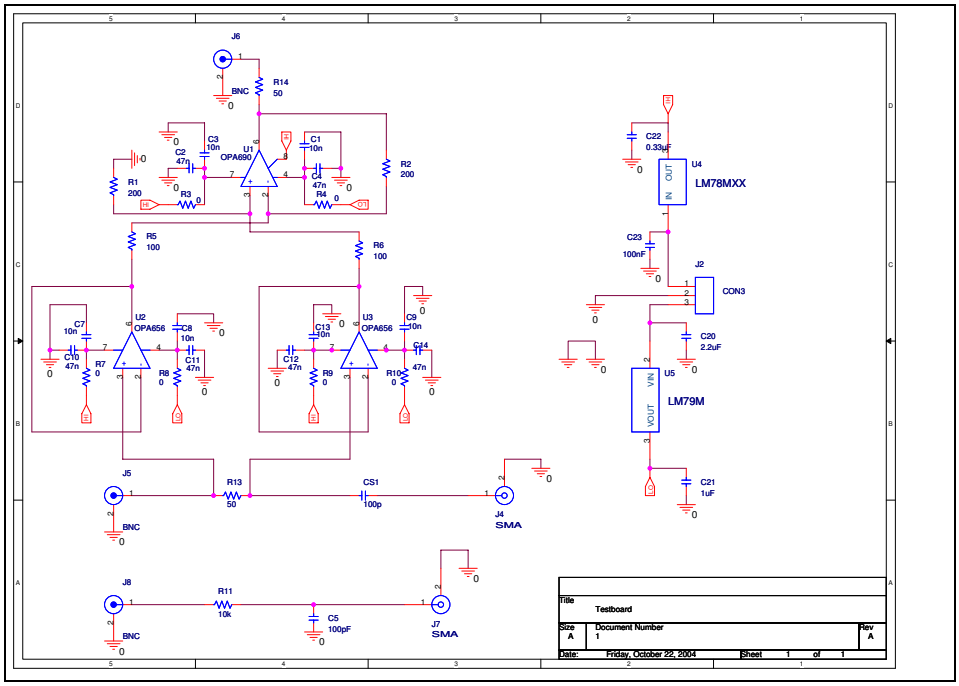


Figure 5.3: *Schematic of the PCB designed to measure frequency characteristics of the reader circuit in the experimental setup.*

of the coils was submerged in water as is shown in figure 5.5. The coordinate board, test PCBs and the water tank was the same as for the signal transmission setup for impedance modulation.

5.2 Coil impedance measurements

To measure coil impedance a HEWLETT PACKARD 4195A 10Hz-500MHz Network/Spectrum Analyzer equipped with the 41951A impedance kit, was used[1]. The HP4195A, when configured as an Impedance Analyzer, employees a measuring method known as the RF-IV method. This method is used for characterizing IF and HF devices, that is, it can be applied with useful results from 100 kHz to 3 GHz (see Agilent Technologies Impedance Measurement Handbook for details [49]). To increase the accuracy of the impedance measurements the instrument is calibrated with open circuit, short circuit and load standards. The intrinsic impedance of these standard devices are stored in registers in the HP4195A and can be adjusted when using different calibration standard sets. The purpose of this calibration is to eliminate stray impedance in the measuring system (see the Agilent Impedance Measurement Handbook for a more thorough description). When connecting test-fixtures and extension cables to the calibration point, additional stray impedances are introduced to the measurement setup. The HP4195A has 0S- and 0Ω- compensation and port extension compensation functions designed to compensate for such errors. However these functions proved to be inadequate for the measurement setup needed; the residual circuit associated with the measurement setup was to complicated.

[1]The HP4195A used in these measurements, purchased in 1989, has not followed the recommended maintenance, and the accuracy of its operation has not been controlled by an appropriate institution since. The use of the instrument appears to have been very limited. However, measurements on passive devices from well known manufacturers have shown good agreement with their given nominal values. Also measurements on the stray impedance and admittance of Agilents 16194A High Temperature test fixture were found to be very close to the given values and comfortably inside the given error limits.

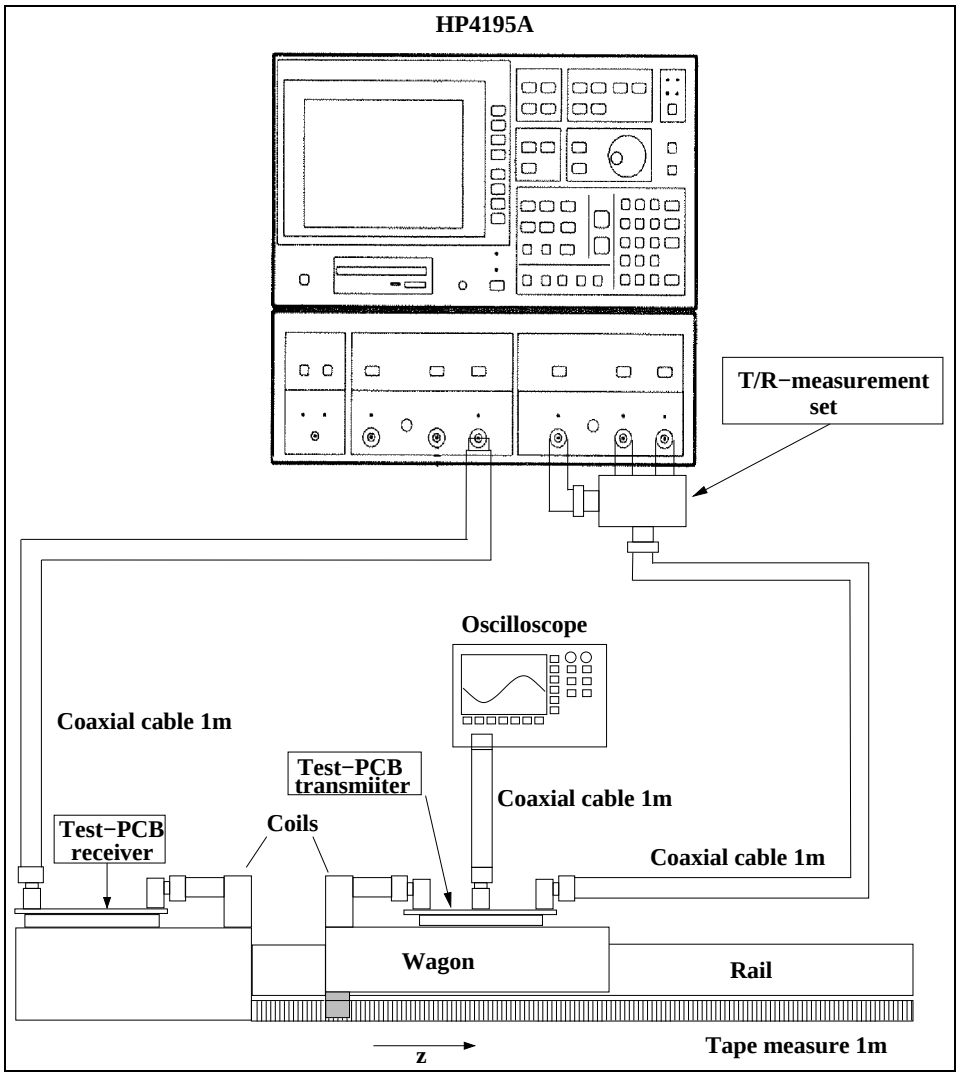


Figure 5.4: *Experimental setup used for determining how power transfer from the external to the implanted device limits the operating range of a passive system. This type of measurements can also have relevance for determining the maximum operating range of an active system design.*

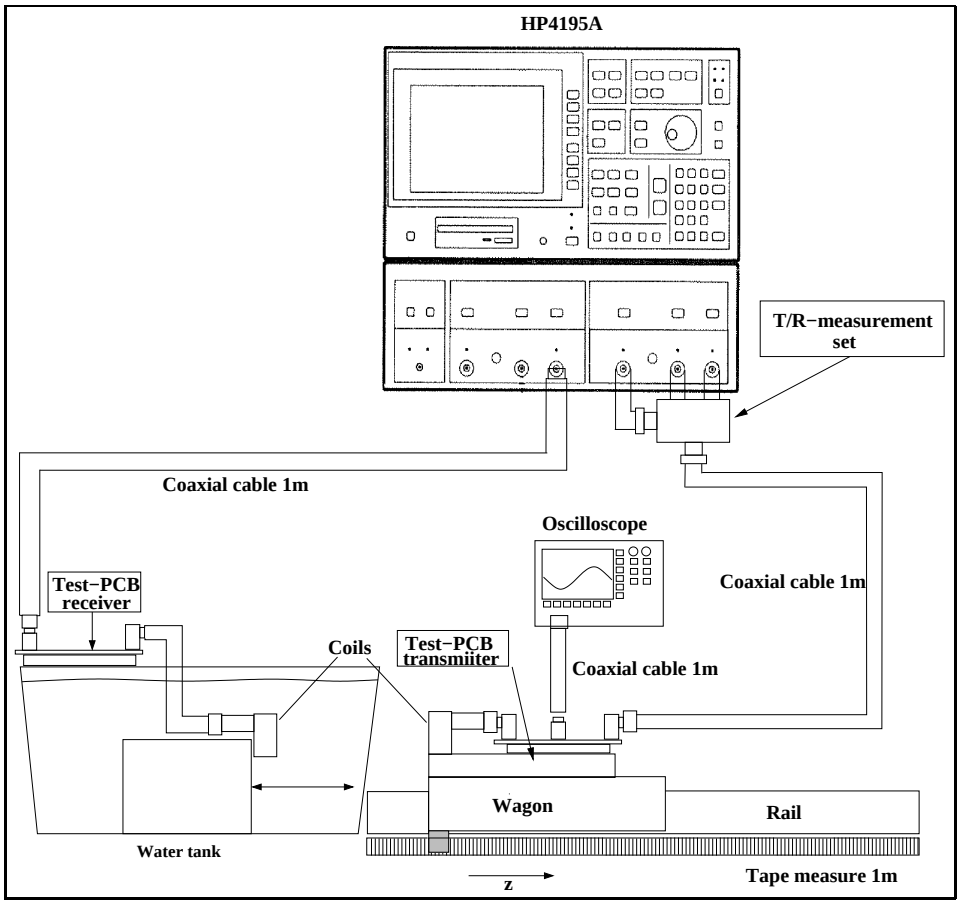


Figure 5.5: *Modified experimental setup used for transmission measurements in water.*

To be able to measure the impedance of coils submerged in water, the setup included a 50 Ω waterproof RG58 cable from Alpha of approximately 70 cm as an extension of the APC7 measurement port of the impedance kit. To ensure that the cable did not change properties when part of it was submerged in water, the cable was tested for a 12 hour period partly submerged in the NaCl-solution used in the coil measurements. No significant change could be measured.

The coils were equipped with feeding cables to connect to the extension cable of the impedance kit and eventually encapsulated in silicon rubber to be tested in the NaCl-solution. A two component silicon elastomer called Sylgard 170 from Dow Corning was used for encapsulation of the coils. The two components was first thoroughly mixed together and evacuated. Various small glasses and plastic containers were used as molds. A small amount of Sylgard 170 was poured into the mold, evacuated and then heated for 20-25 minutes at 70°C leaving it still adhesive. The coil was now placed in the mold and more Sylgard 170 was applied until the coil was completely submerged. Again it was evacuated and heated. Because the molds were not designed for the coils, the position of the coils inside the mold could not be controlled.

To get the encapsulation of the coil watertight at the cable entry point it was necessary to use a primer (Loctite 770) on the cable jacket. However, when the cable was subjected to bending forces, the encapsulation had a tendency of loosening around the cable entry and water entered the encapsulation. The effect of such leakages was immediately detected in the measurements. To avoid such leakages silicon grease was applied at the cable entry point.

The cable used as coil feeding cable was either the same RG58 cable as used for the port extension cable or a high temperature RG316 cable from Multicomp. The port extension cable was equipped with a coaxial male SMA connector. A female-female connector was used to connect the coil feeding cables to the port extension cable. To keep the connection watertight when submerged in water it was smeared in with silicon grease and a shrink hose was shrunk around it. The same approach was later used in the transmission measurement water setup. A picture of the setup

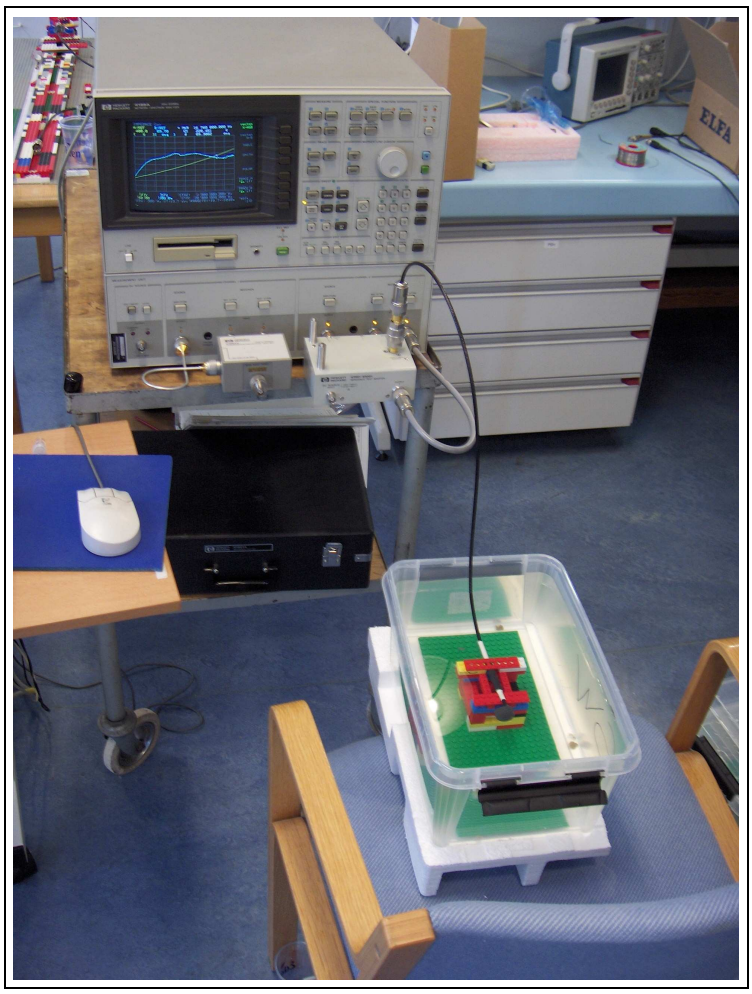


Figure 5.6: *The picture shows the experimental setup for the impedance measurements of encapsulated coils in water.*

is shown in figure 5.6.

To get the calibration plane as close to the coil as possible, in other words eliminate the feeding cable impedance from the measurements, there were made "calibration cables" that were to simulate the coil feeding cable when performing calibration. This calibration procedure was not ideal as the calibration cables could not be cut to the exact same length as the coil feeding cables. Small differences in length would cause different residuals in the cables which again would lead to errors in the results. This

is especially true for the coil capacitance measurements as the coil capacitance is very small, in the order of $100fF$ to $1pF$. The cables used have a characteristic capacitance of approximately $96pF/m = 960fF/cm$ and $105pF/m = 1.05pF/cm$. Hence a misfit of only 1mm can introduce residuals in the same order as the capacitance that is sought measured.

The calibration was carried out in the following manner:

- The open calibration was carried out without a standard device, ie. the end of the cable was left open. It is difficult to measure or calculate the capacitance of such a geometry so the cal. std value for the open measurement was left at the $82fF$ of the HP-APC7 open standard.
- The short calibration was carried out with a Huber&Suhner SMA shorting device of type 64SMA50-0-1/111.
- The load calibration was carried out with a Huber&Suhner SMA 50Ω RF termination Type 65_SMA-50-0-1/111.

The test setup and calibration procedure was confirmed to give accurate results for the coil inductance, within the HP4195A's inherent measurement accuracy for impedance measurements. In this confirmation procedure 132-SMG10 and 132-SMG20 coils from Coilcraft with nominell values of respectively $111nH \pm 2\%$ and $538nH \pm 2\%$ were used. The coils were first measured on the Agilent 16196A High Temperature test fixture. The measurements on the test fixture were in excellent agreement with the nominal values given by Coilcraft, only 0.09% deviation from nominal value for 132-SMG10 and 0.28% for 132-SMG20. The results from the measurements are presented as 132-SMG10 TF and 132-SMG20 TF in table 5.1.

Then calibration was done at the end of the cable that was designed to match the electrical properties of the coil feed cable. The result of these measurements are presented as 132-SMG10 CP and 132-SMG20 CP in table 5.1. Now the deviation from the nominal value was 1.00% and 0.08%. However, the capacitance and resistance changed drastically compared to the test fixture measurements. The capacitance deviation is thought to

	$L(nH)$	$C(fF)$	$R(m\Omega)$
132-SMG10 TF			
Series #1	110.963	233.919	355.438
Series #2	110.720	213.815	309.495
Series #3	111.611	237.998	348.275
Tot. mean	111.098	228.578	337.736
σ mean	0.376	10.571	20.182
diff(%) nom val	0.088	-	-
132-SMG10 CP			
Series #1	112.130	317.763	124.280
Series #2	112.119	322.287	124.550
Series #3	112.075	316.186	114.021
Tot. mean	112.108	318.746	120.951
σ mean	0.024	2.586	4.901
diff(%) nom val	0.998	-	-
132-20SMG TF			
Series #1	539.306	207.317	722.551
Series #2	539.653	198.978	736.912
Total average:	539.479	203.147	729.732
σ mean	0.174	4.170	7.180
diff(%) nom val	0.275	-	-
132-SMG20 CP			
Series #1	538.310	266.238	667.683
Series #2	538.041	274.112	646.959
Series #3	538.886	273.528	640.836
Total average:	538.412	271.293	651.826
σ mean	0.352	3.582	11.488
diff(%) nom val	0.077	-	-

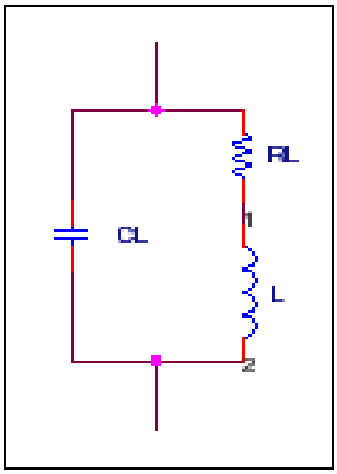
Table 5.1: *Results from the test of the calibration procedure used in the measurements. TF is short for test fixture, CP is short for calibration procedure meaning the procedure using SMA std. devices described above.*

have been caused by the above described inability to exactly reproduce the coil feeding cable. The reason for the reduction of resistance is probably less contact resistance in the solder connection between coil and feed cable than between coil and connection points of the test fixture.

The coils were potted in silicon rubber and measured again in air. The port extension cable to feed the coil submerged in water was then attached and to confirm the setup the coils were again measured in air. From figure 5.8 to 5.13 we see that there is good agreement between the measurements done for the encapsulated coils in air when connected to the measuring port of the impedance kit and when connected to the end of the port extension cable. These are the measurements marked "Encapsulated ref" and "Port Ext Ref".

The consequence of these insufficiencies of the measurement setup is that the coil capacitance values measured can not be used to find the value of the coil capacitance. We can however see the relative change of the coil capacitance when exposed to different operating media.

5.2.1 The Equivalent circuit function



To characterize different types of components the HP4195A has a utility called equivalent circuit calculation. This function calculates the values of ideal R, L and Cs in a chosen equivalent circuit based on the impedance data obtained in a frequency sweep. The exact calculation procedure was said to be confidential by Agilent. For a coil the standard equivalent circuit is a capacitor in parallel connection with a series connection of an inductance and a resistance as shown to the left.

A main goal of these measurements was to extract coil capacitance. Even when measurements were carried out on the Agilent 16194A test fixture the equivalent circuit function produced negative results for the coil capacitance when choosing certain frequency sweeps. Choosing different

Freq. sweep	$L(nH)$	$C(fF)$	$R(m\Omega)$
132-SMG10 TF			
20-120MHz 1	111.499	243.516	246.106
20-120MHz 2	111.532	235.448	246.081
35-135MHz 1	111.349	228.156	282.132
35-135MHz 2	111.299	235.662	278.601
75-175MHz 1	111.532	242.739	370.622
75-175MHz 2	111.485	248.322	369.478
132-SMG20 TF			
20-120MHz 1	535.432	206.997	1049.19
20-120MHz 2	535.113	209.460	1049.38
5-90MHz 1	538.696	205.320	482.903
5-90MHz 2	538.585	207.230	489.956
5-90MHz 3	539.509	210.319	495.785

Table 5.2: *Results from measurements on coils 132-SMG10 and 132-SMG20 with the 16194A test fixture for different frequency sweeps.*

frequency sweeps also produced different results, see examples in table 5.2. Observe also the increase of resistance with increasing frequency. This suggests that the skin effect is influencing the coils and that the equivalent circuit function is not sufficient. This may also have affected the capacitance values obtained with the equivalent circuit function.

A closer inspection of the phase data revealed sinusoid like fluctuations, see figure 5.7. The calculation method used by the 4195A only uses two frequency points; the $1/\sqrt{2}\cdot$ (Max frequency) and the $\sqrt{2}\cdot$ (Min frequency) which is claimed to be the two points were maximum accuracy is obtained [49]. However, one could be tempted to belive that the selection of these two points may not be ideal considering the phase fluctuations. In the worst case, if one point is selected at a local phase minimum and the other point at a local phase maximum the resulting capacitance value could become negative.

To try to compensate for these errors further calculations on the equi-

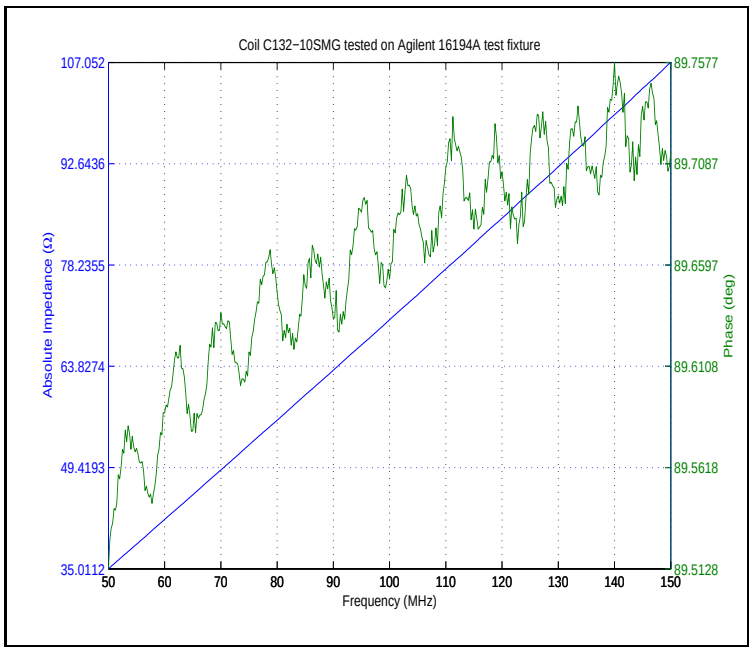


Figure 5.7: *The figure shows an example of an impedance measurement done at the 16194A test fixture. The absolute value of the impedance shown on the left axis and blue line and its phase angle right axis, green line.*

valent circuit used for HF-coils were done. The idea was to transfer complex impedance data from the measurements on the HP4195A to a computer and a dedicated data reduction procedure was written to try improve the prediction of L, C and R in the equivalent circuit. The idea underlying the procedure was to put available physical knowledge into the correlation analysis of the data.

The total impedance of the equivalent circuit schematic can be written as:

$$Z_{tot} = \frac{\frac{-j}{\omega C}\left(R + j\omega L\right)}{\frac{-j}{\omega C} + \left(R + j\omega L\right)} \tag{5.1}$$

By further calculation the real and imaginary parts can be extracted. They are given by:

$$Re(Z_{tot}) = Z_{re}(\omega) = \frac{R}{\left(1 - \omega^2 CL\right)^2 + \omega^2 R^2 C^2} \tag{5.2}$$

and

$$Im(Z_{tot}) = Z_{im}(\omega) = \frac{\omega L - \omega^3 CL^2 - \omega R^2 C}{\left(1 - \omega^2 CL\right)^2 + \omega^2 R^2 C^2} \tag{5.3}$$

Subject to the conditions

$$1 \gg \omega^2 CL \gg \omega^2 R^2 C^2 \tag{5.4}$$

a series expansion of equation 5.3 gives

$$Z_{im}(\omega) = \omega L + \omega^3 CL^2 + \ldots \tag{5.5}$$

Therefore an adequate correlation function in this case is:

$$Z_{im}(\omega) = a_1 \omega + a_3 \omega^3 \tag{5.6}$$

The coefficients a_1 and a_3 were determined using least squares fitting of equation 5.6 to the data $[\omega_i, Z_{im}(\omega_i)]$ obtained from measurements on the HP4195A in a MATLAB-script. Comparing equations 5.5 and 5.6 we obtain:

$$L = a_1, \qquad C = \frac{a_3}{L^2} \tag{5.7}$$

This gave reasonable values for L and C provided the data was obtained from a relatively wide frequency sweep allowing smoothed values

of the phase data ϕ to be calculated. The choice of smoothening function was in some cases found to be important. From 5.2 and 5.3, we have:

$$\tan(\phi) = \frac{\omega L - \omega^3 C L^2 - \omega C R(\omega)^2}{R(\omega)} \tag{5.8}$$

The frequency dependence of R is caused by the skin effect. For plane boundaries the skin depth is given by:

$$\delta(\omega) = \sqrt{\frac{2}{\sigma \mu \omega}} \tag{5.9}$$

Inserting 5.9 in 5.8 we formally get

$$\tan(\phi) = b_1 \omega^{1/2} + b_1 \omega^{3/2} + b_1 \omega^{5/2} \tag{5.10}$$

which can be used to correlate the ϕ-data. Using this approach increased the correlation of coil impedance measurements and impedance calculations on the basis of the equivalent circuit values.

A way to improve the measurements of the coil capacitance could be to first measure the impedance of all coil cables individually. The results from these measurements could then be used to find the impedance caused by the coil cable residuals when connected to the coil and the impedance caused by the coil itself could be extracted from the total impedance measured.

5.2.2 Impedance measurement results

Below the results obtained using the calibration procedure with "calibration cables" described in the beginning of this section is presented. An overview of the coils used in these measurements is given in table 5.3. From figures 5.8 and 5.11 we observe that there is very little influence of the encapsulation and water on the coil inductance. A slight increase is seen from the "Air ref" measurement to the encapsulated in figure 5.11. The "Port Ext Ref", "Freshwater" and "Saltwater" measurements show no clear tendency and the variations are so small that the conclusion must be that no change was observed. That the inductance variations are small

Coil	$D(mm)$	$d_w(mm)$	N	$b(mm)$	$l_c(mm)$	Cable
132-SMG10	~ 5	~ 0.6	10	7.98	57.00	RG58
132-SMG20	~ 5	~ 0.3	20	7.98	57.20	RG58
10-1	19.8	0.33	10	3.85	64.05	RG58
10-2	19.8	0.33	10	3.85	64.05	RG58
10-3	20.0	0.5	10	5.85	63.75	RG58
10-4	20.0	0.5	10	5.85	63.75	RG58
10-5	14.7	0.33	10	3.85	55.40	RG58
10-6	14.7	0.33	10	3.85	55.40	RG58
10-7	14.7	0.33	10	3.85	55.40	RG58
10-9	10.3	0.33	10	2.15	50.70	RG316
10-10	10.2	0.18	10	2.15	49.70	RG316
10-12	16.3	0.33	10	3.85	56.45	RG58
10-13	32.3	0.33	10	3.85	56.45	RG58

Table 5.3: *All the coils have only one layer. The coil length b and cable length l_c was measured with a slide caliper with resolution 0.05mm. The diameter D is the distance between the center of the wire on one side to the center of the wire on the other side. It was obtained by measuring the external diameter of the coil bobbin (all circular except for 10-12 which is quadratic) and then adding one wire diameter d_w. Coils 10-1, 10-2, 10-3, 10-5 and 10-13 were only used for air measurements.*

is expected as neither the silicon rubber nor the water and NaCl-solution has electromagnetic properties that indicates otherwise. Hence inductance changes is not expected to affect the performance of a passive system with the passive resonance circuit implanted.

In figures 5.9 and 5.12 we see that the capictance can increase up to 450% from air to saltwater for the coils tested here. The silicon rubber is expected to have a relative electrical permittivity $\epsilon_r \approx 3.6$. Water is expected to have approximately 80. Not surprisingly the material closest to the coil has relatively larger effect on the coil capacitance. This can be of interest when choosing encapsulation material.

The increase in capacitance will reduce the resonance frequency of an implanted resonance circuit. According to the discussion in section 4.2.2 this would reduce operating range of the passive LC-tank system. It also reduces the relative effect of capacitance changes in the sensor. For the resistive impedance modulation approach it causes detuning of the resonance circuits. This detuning could also be serious for the coil in the external resonance circuit. If the coil is placed near the surface of the body the capacitance will be changed depending on the distance from the surface.

The coil resistance measurements in figures 5.10 and 5.13 also show an increase from the "Air ref" to the "Encapsulated ref" measurements and further increase for the "Freshwater" and "Saltwater" measurements. This is probably due to increased losses in the reactive fields when exposed to lossy media. These losses will appear as resistance increases in the equivalent circuit of the coil. We saw in section 4.2.2 that the operating range of the passive systems were dependent on this resistance. Here this resistance increase is more correctly described as a transmission signal loss.

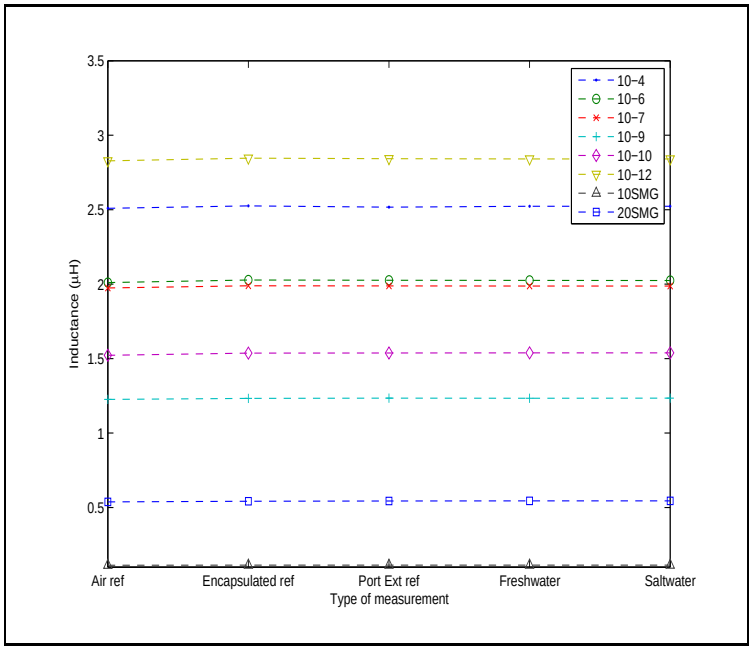


Figure 5.8: *Coil inductance measurements*

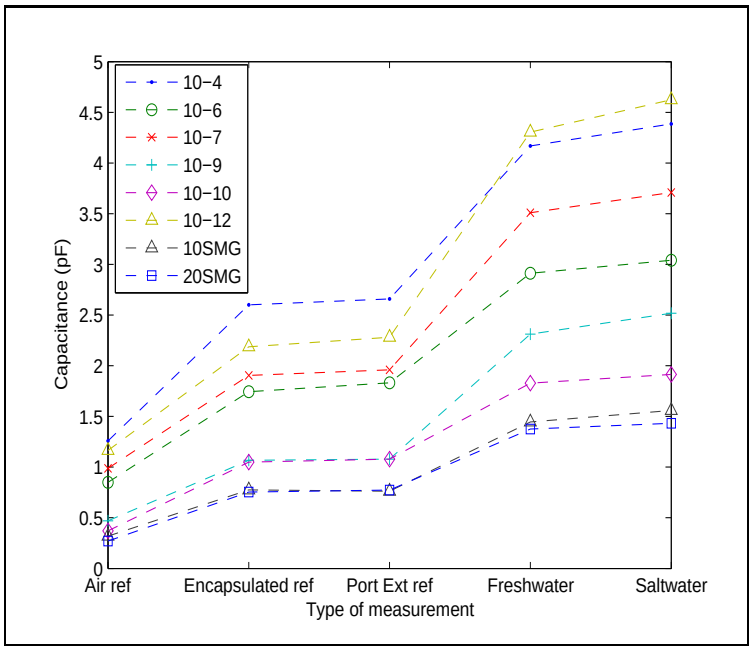


Figure 5.9: *Coil capacitance measurements*

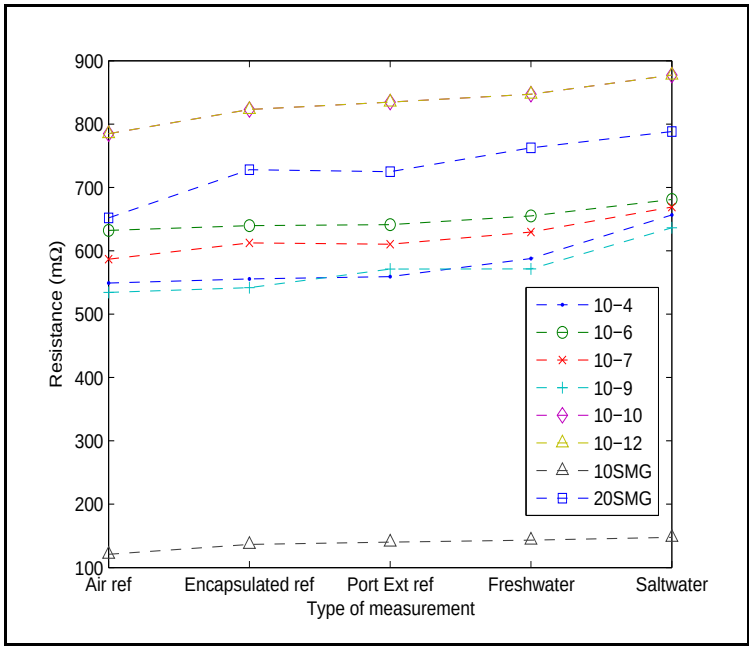


Figure 5.10: *Coil resistance measurements*

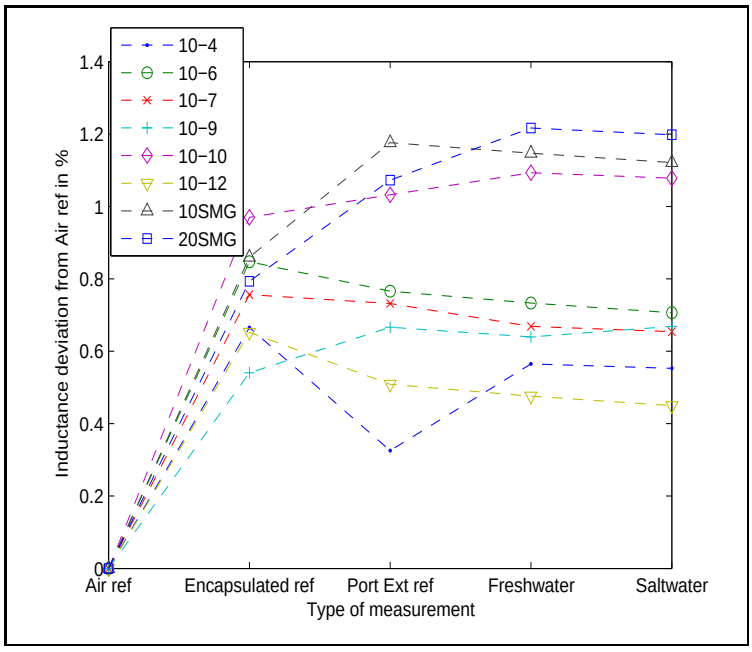


Figure 5.11: *Coil Inductance measurements*

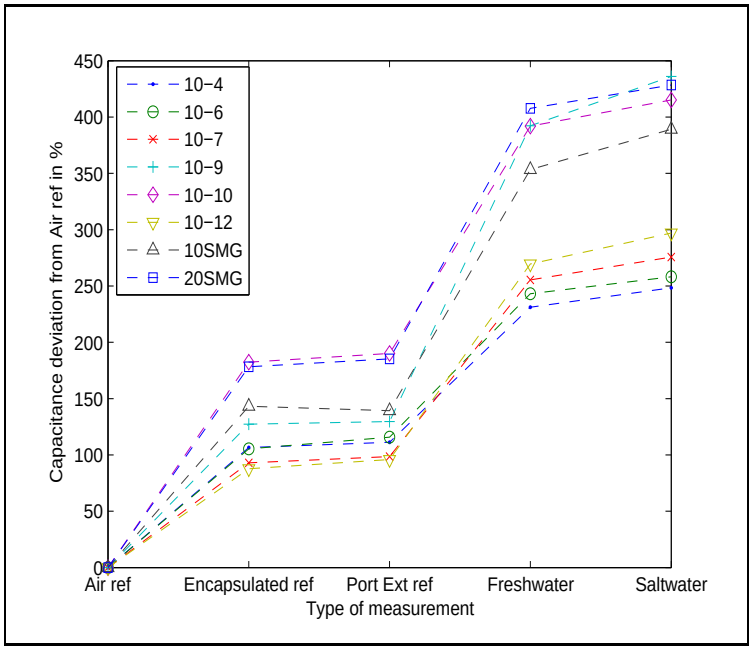


Figure 5.12: *Coil capacitance measurements*

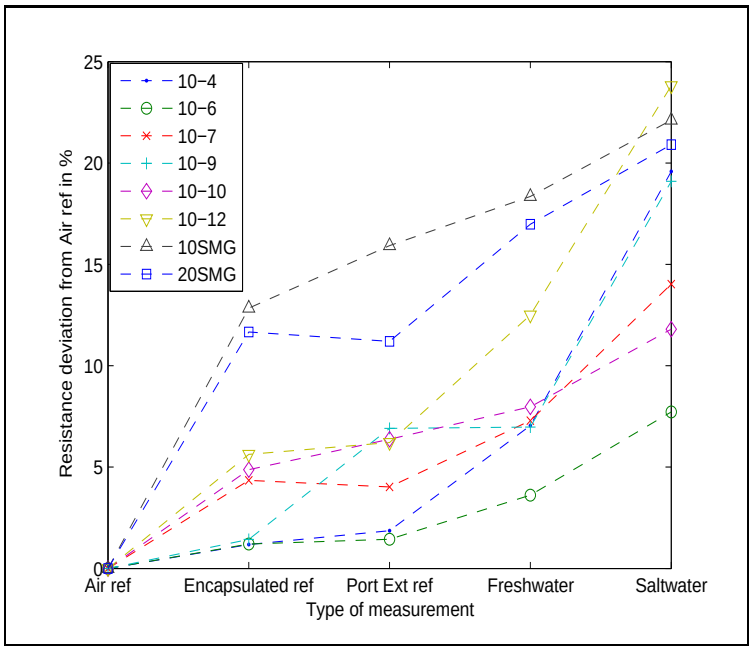


Figure 5.13: *Coil resistance measurements*

Chapter 6

Transmission Predictions

This chapter presents the calculation method used to predict near field signal transmission between two coils. The method is based on the theory presented in section 4.1.1 regarding mutual inductance and Neumanns formula. To simulate the electrical behavior of the experimental setup, PSPICE-simulation was used.

6.1 Mutual inductance calculation

Based on Neumann's formula, equation 4.10, a calculation program for circular coils was written in C. A desired functionality was that the program would work for any possible positioning of the coils. To present the calculations we first introduce a well arranged coordinate system.

We start by defining an XYZ-system with origo in the center of coil 1, see figure 6.1. In addition we introduce an xyz-system and a x'y'z'-system, both with origo in the center of coil 2. The xyz-system is always parallel with the XYZ-system. The x'y'z'-system is rotated with coil 2. We define u to be the angle between the y- and y'-axis and v to be the angle between the x- and x'-axis.

Neumann's formula is as presented in equation 4.10:

$$M = \frac{\mu_0}{4\pi} \oint_{C_1} \oint_{C_2} \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{|\mathbf{R}|} \tag{6.1}$$

Since the coils are circular we need to introduce the means to describe

this geometry in the integral. This can be done by describing $d\mathbf{l}_1$ and $d\mathbf{l}_2$ as arclengths with the aid of cylindrical systems decomposed on the Cartesian systems we defined above. Figure 6.3 shows the geometry and we see that we can express $d\mathbf{l}_1$ and $d\mathbf{l}_2$ as:

$$\begin{aligned} d\mathbf{l}_1 &= \mathbf{i}_\theta r_1 d\theta \\ &= \left(-\sin(\theta)\mathbf{i}_X + \cos(\theta)\mathbf{i}_Y\right) r_1 d\theta \qquad (6.2) \\ d\mathbf{l}_2 &= \mathbf{i}_\phi r_2 d\phi \\ &= \left(-\sin(\phi)\mathbf{i}_{x'} + \cos(\phi)\mathbf{i}_{y'}\right) r_2 d\phi \qquad (6.3) \end{aligned}$$

To relate the positions of every element of coil 2 to coil 1 we need to transform the coordinates of the x'y'z'-system to the XYZ-system. We do this in two steps. To transform from the x'y'z'-system to the xyz-system the following coordinate transformations apply:

$$\mathbf{i}_{x'} = \mathbf{i}_x \cos(\angle(x,x')) + \mathbf{i}_y \cos(\angle(y,x')) + \mathbf{i}_z \cos(\angle(z,x')) \qquad (6.4)$$

$$\mathbf{i}_{y'} = \mathbf{i}_x \cos(\angle(x,y')) + \mathbf{i}_y \cos(\angle(y,y')) + \mathbf{i}_z \cos(\angle(z,y')) \qquad (6.5)$$

$$\mathbf{i}_{z'} = \mathbf{i}_x \cos(\angle(x,z')) + \mathbf{i}_y \cos(\angle(y,z')) + \mathbf{i}_z \cos(\angle(z,z')) \qquad (6.6)$$

Transforming $d\mathbf{l}_2$ into xyz-coordinates results in:

$$\begin{aligned} d\mathbf{l}_2 \;=\; & (\mathbf{i}_x(-\sin(\phi)\cos(v)) \\ & + \mathbf{i}_y(-\sin(\phi)\sin(u)\sin(v) + \cos(\phi)\cos(u)) \\ & + \mathbf{i}_z(\sin(\phi)\cos(u)\sin(v) + \cos(\phi)\sin(u)))\, r_2 d\phi \qquad (6.7) \end{aligned}$$

Because the xyz- and XYZ-systems are completely parallel there is no need for further transformations to calculate the dot product between $d\mathbf{l}_1$ and $d\mathbf{l}_2$. It becomes:

$$\begin{aligned} d\mathbf{l}_1 \cdot d\mathbf{l}_2 \;=\; & (\sin\theta\sin\phi\cos v - \cos\theta\sin\phi\sin u\sin v \\ & + \cos\theta\cos\phi\cos u)\, r_1 r_2 d\phi d\theta \qquad (6.8) \end{aligned}$$

The distance R_{ik} between one element $d\mathbf{l}_1$ in loop i in coil 1 and $d\mathbf{l}_2$ in loop 2 can be expressed as:

$$R_{ik} = \left[(X_{2k} - X_{1i})^2 + (Y_{2k} - Y_{1i})^2 + (Z_{2k} - Z_{1i})^2\right] \qquad (6.9)$$

The result is:

$$\begin{aligned} R_{ik} \;=\; & \Big[(X_{2k} - r_2\cos\phi\cos v)^2 \\ & + (Y_{2k} - r_2\cos\phi\sin u\sin v + r_2\sin\phi\cos u \\ & \quad - z'_k\sin u\cos v - Y_{1i})^2 \\ & + (Z_{2k} - r_2\cos\phi\cos u\sin v + r_2\sin\phi\sin u \\ & \quad + z'_k\cos u\cos v - Z_{1i})^2\Big]^{1/2} \end{aligned} \tag{6.10}$$

The double integral over θ and ϕ is calculated by using numerical integration after Simpson's formula. When the coils have more than one winding, the mutual inductance between all windings in coil 2 and all windings of coil 1 must be added together, i.e. the integration procedure has to be carried out $N_1 \cdot N_2$ times.

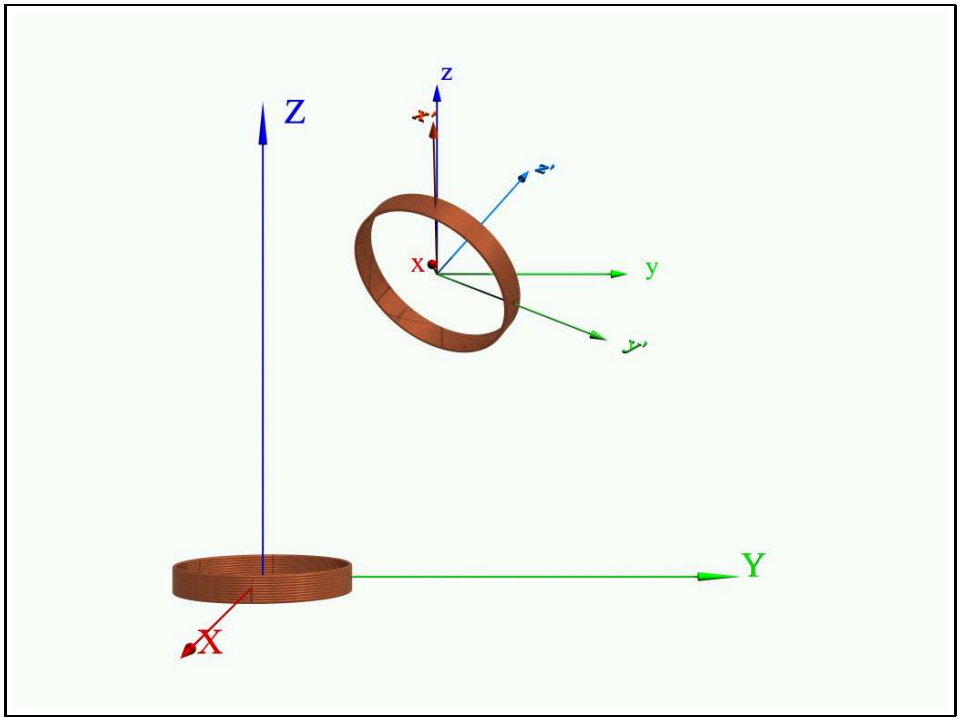


Figure 6.1: *The figure shows the nomenclature related to the geometry of the coils used in the calculation program. Coil 1 has center in origo of the XYZ-system. The center of Coil 2 is translated to a point (X,Y,Z) where the xyz-system has it's origo. Coil 2 is tilted by first rotating it an angle u with the x-axis as the rotation axis and then rotating it an angle v with the y'-axis as the rotation axis.*

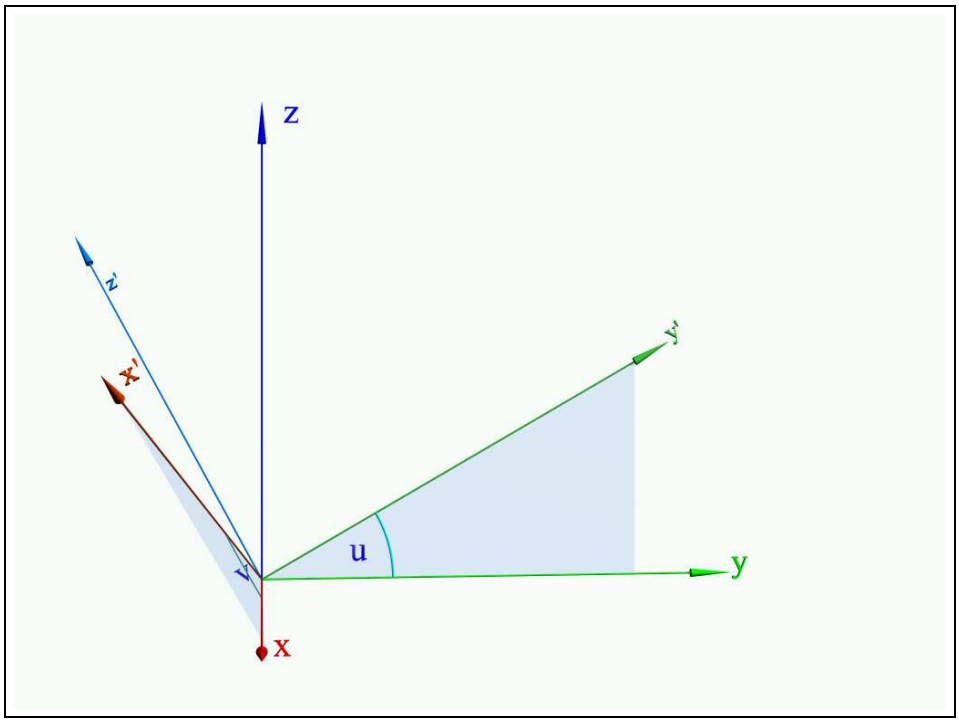


Figure 6.2: *The geometry used to define coil 2's position in the XYZ-system. The coordinate transformations between the x'y'z'-system and the xyz-system coil 2 are found by the projection of coordinates from one system to the other.*

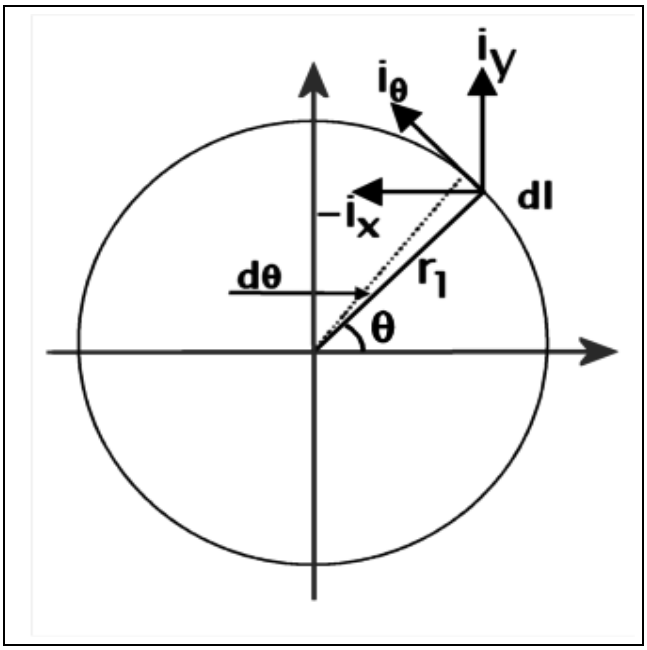


Figure 6.3: *The figure shows a winding in coil 1 and how the transformation from* $d\mathbf{l}_1$ *to* $d\theta$ *is done. The unit vector* $\mathbf{i}_\theta$ *must be decomposed onto the fixed unit vectors* $\mathbf{i}_x$ *and* $\mathbf{i}_y$ *in order to carry out the integral.*

Chapter 7

Signal Transmission Predictions Compared with Observations

The results from the experimental investigation using the signal transmission setups discussed in chapter 5 for a short range communication system is presented below. They are compared to results from PSPICE-simulations done in OrCAD 10.0. These were based on the impedance measurements of chapter 5 and calculations performed with the mutual inductance calculation program described in chapter 6.

The primary goal of these measurements were to provide relevant information on the performance limits of the passive solution discussed in chapter 4. A secondary goal was to investigate how well the combination of the calculation program and PSPICE-simulations could predict the system behavior, both in air and with one coil submerged in water or aqueous sodium chloride solution.

7.1 Resistance voltage modulation through inductive coupling

Figure 7.1 on page 133 shows the schematic for the circuit used in the PSPICE-simulations of the experimental setup. The transmission line tagged "Coil feed 3" was the extension cable used for measurements in water.

This was removed from the simulation when simulating air measurements. The experimental setups used was shown in figures 5.1 on page 102 and 5.2 on page 104. The coils were coaxially aligned during the measurements of sections 7.1 and 7.2. Figure 7.2 on page 134 shows plots of frequency sweeps obtained from both simulations (green solid line) and from measurements (red dots) done in air with coils 10-1 and 10-2.

Figure 7.3 on page 135 shows results obtained with both coils located in air. The graphs show a voltage difference plotted against the distance between the centers of the two coils in question. This voltage difference will be referred to as V_{diff}. The frequency sweeps shown in figure 7.2 on page 134 provide the basis for finding V_{diff}. The voltage level $V_{'1'}$ is the voltage measured at the test point of the transmitter PCB (refer to figure 5.1 on page 102) at the resonance frequency ω_0 when the resistance R2=10kΩ has its greatest impact on the impedance of circuit 1. The voltage level $V_{'0'}$ is the voltage measured at the test point at ω_0 when R2=1Ω. In this case the resistance R_2 is so small that it has no significant effect on the transmitter circuit and the minimum at the resonance disappears completely. The same voltage would be read when no coupling exists between the two circuits. This was how the voltage $V_{'0'}$ was obtained for the measured results shown.

This experimental setup is meant to simulate a simple implementation of an ASK scheme for an implanted communication system based on resistive impedance modulation. V_{diff} is the difference between the voltage amplitude level $V_{'1'}$ read in the external circuit when a logical '1' is transferred from the implanted circuit to the external circuit and the amplitude voltage level $V_{'0'}$ read when a logical '0' is transferred. The depth of $V_{'1'}$ will decrease when the coupling of the coils decrease and hence decide the limit for the operating distance.

We see that there is relatively good correspondence between the measured and simulated voltage levels. The simulation was carried out using values of coil inductance, coil capacitance and coil resistance obtained using the results from the impedance measurements presented in chapter

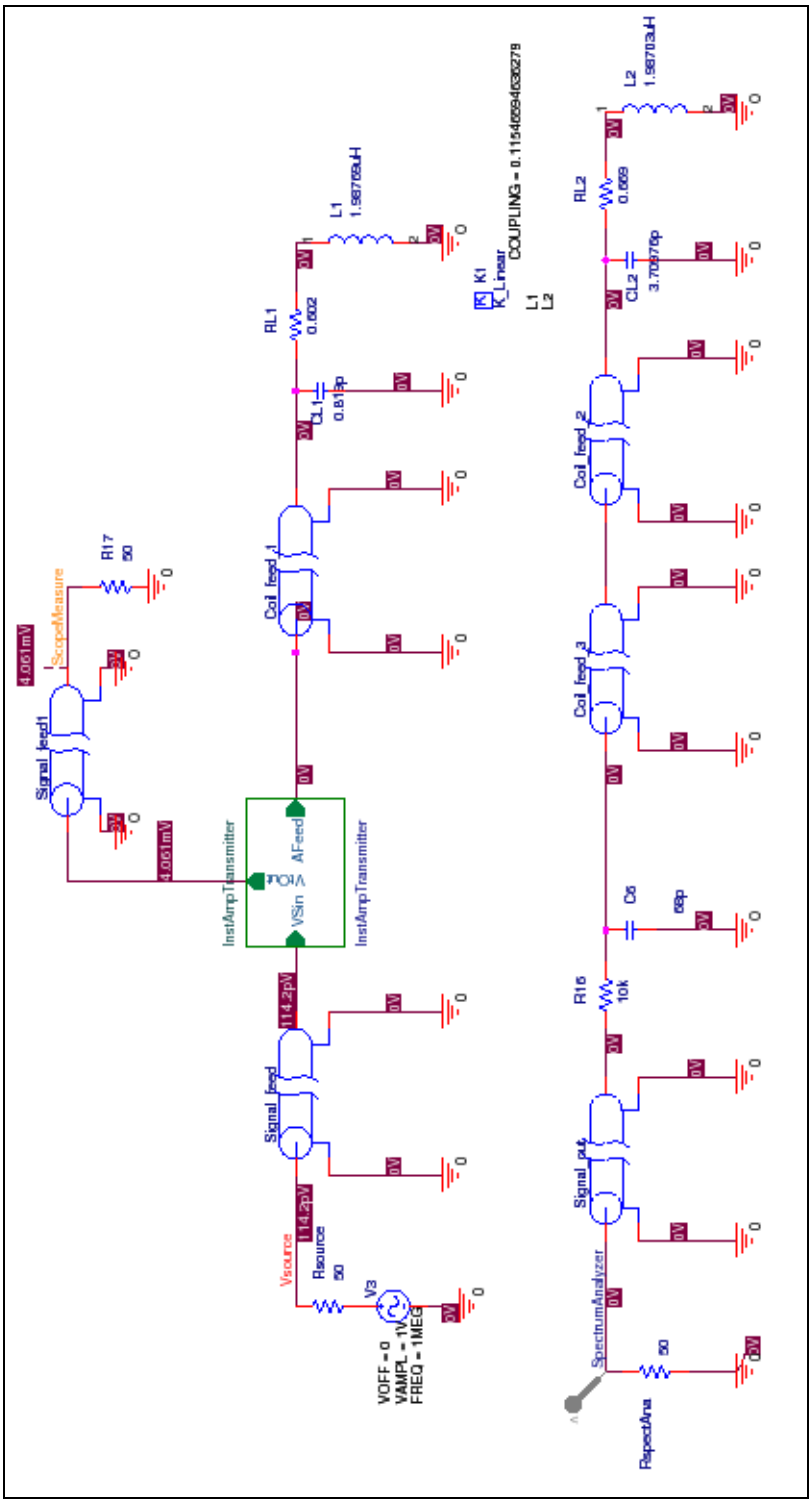


Figure 7.1: *Schematic of the experimental setup from the PSPICE-simulation.*

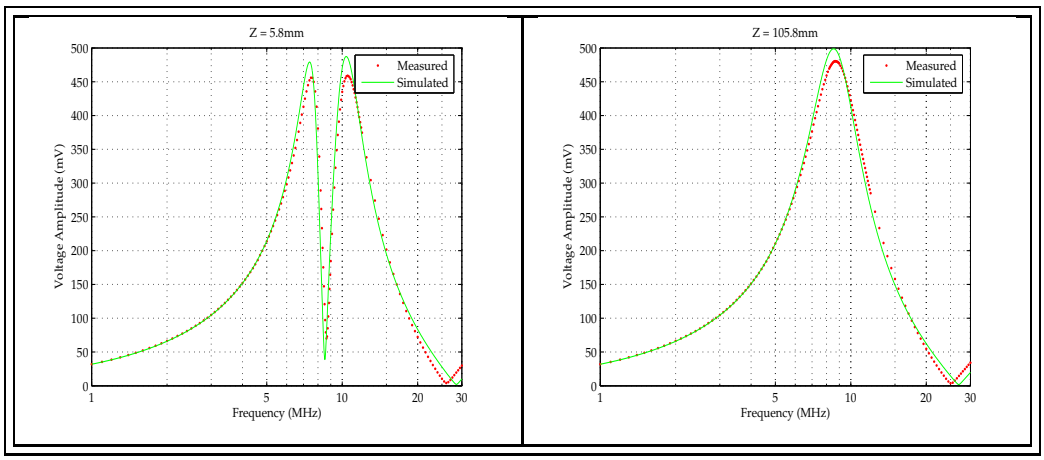


Figure 7.2: *Frequency sweeps providing the basis for finding* V_{diff}

5.1. The simulated voltage levels lie slightly above the measured. As seen in the frequency sweep of figure 7.4 on page 136 for coils 10-1 and 10-2 in air, it is possible to further improve the simulation by including parasitic series resistance and shunt capacitance. The extra resistance is probably due to resistance in cables and PCB leads, while the extra capacitance is caused by the PCB leads. The tuned simulation shows very good correspondence with the measured results.

Figure 7.5 on page 139 compares results obtained for measurements carried out with the coil corresponding to L2 in figure 7.1 submerged in water and aqueous sodium chloride solution with simulations. The results from impedance measurements carried out in water and sodium chloride solution was used in the simulations. As seen in section 5.2.2 on page 119 the coil capacitance was altered by the water and sodium chloride solution. The extension cable necessary for these measurements as described in section 5.1 introduced additional stray capacitance. This caused detuning of the resonance circuits and demanded changing the tuning capacitances of the resonance circuits. The ohmic losses due to the conductivity of the sodium chloride solution was simulated by the increased resistance measured in the impedance measurements. This approach shows good agreement with measurements in figure 7.5 on page 139 B and D. Also C has satisfying agreement.

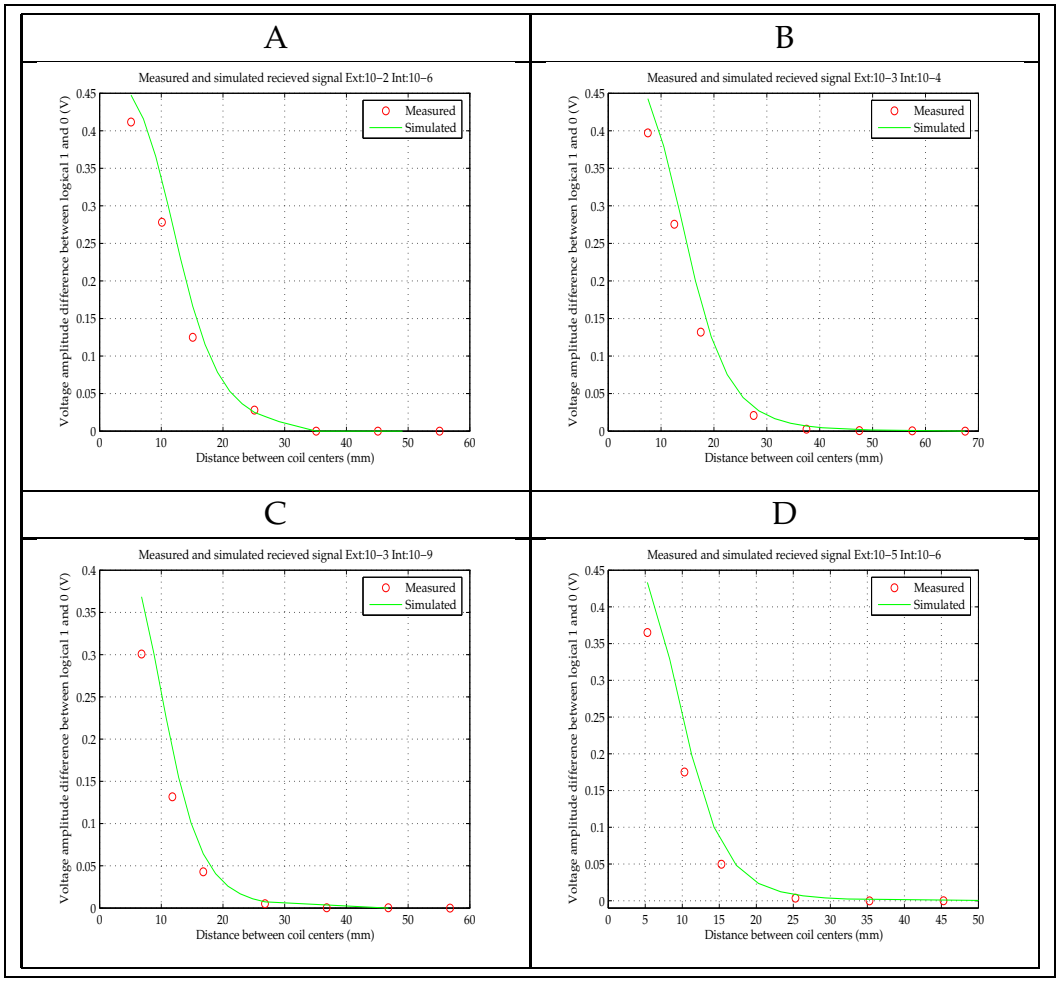


Figure 7.3: *Signal transmission measured with experimental setup for impedance modulation in air*

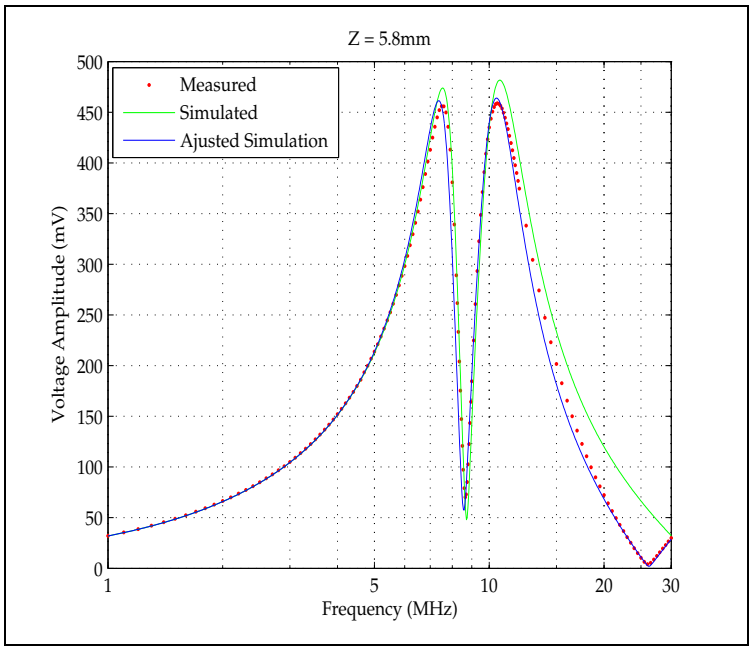


Figure 7.4: *The effect of parasitics in the experimental setup for the coils 10-1 as (L1) and 10-2 (L2) in air, is shown. The "Adjusted Simulation" curve was achieved by manually tuning the values of parasitic shunt capacitance and series resistance.*

Coil	$Z_w(mm)$ pos. 1	$Z_w(mm)$ pos. 2	$Z_w(mm)$ pos. 3	$Z_w(mm)$ pos. 3B
10-4FW	9.2	17.3	33.5	57.8
10-4SW	8.3	16.4	32.6	56.9
10-6SW	9.8	17.9	34.1	58.4
10-7SW	10.7	18.8	35.0	59.3
10-9SW	10.4	18.5	34.7	59.0
10-10SW	11.2	19.4	35.5	59.8

Table 7.1: *Overview of distance from coil center to wall of water tank for measurements with coaxial alignment.*

However, in figure 7.5 on page 139 plots E and F the correspondence between measurements and simulation is rather poor. This can at least partly be explained by the plot shown in figure 7.6 on page 140. We see that the resonance circuits are detuned and that the simulation does not predict the circuit behavior well. The measured values indicate a more severe mismatch than predicted by the simulation.

Considering figure 7.5 plots A, B, C and D where the system was properly tuned, the effect that the water and saltwater has on the system performance is very limited. Note that the measurements are tagged "Measured pos. 1" and so on. These tags indicate different positions of the coil in the water tank as shown in table 7.1. This implies that the fields travel greater distances through the water tank when a coil is placed in position 2 and 3. The ohmic losses of the sodium chloride solution appear to have limited effect on the operating distance at these frequencies. Although the saltwater solution is a dissipative medium with an attenuation constant $\alpha = 0.69dB/cm$ at 10 MHz, the fields travel only very short distances in water.

In figure 7.7 on page 140 tuned simulations for air, freshwater and sodium chloride solution are compared. We see that the changes in coil properties measured in section 5.2 does not effect the system performance noticeably.

The measurements indicate that a passive system based on resistive impedance modulation could be designed to work in the application in question if the detuning of the resonance circuits are compensated for. The operating distance is however limited to a few centimeters.

7.2 Transmitted signal

In this section the measurements conducted with the experimental setup presented in section 5.1.2 on page 106 and the corresponding simulations are presented and discussed. To achieve simulation results comparable to the amplitude ratio level presented by the HP4195A in dB, the following equation was used:

$$A = 20 \log \left(2 \frac{V_{out}}{V_{in}} \right) \tag{7.1}$$

Here V_{out} corresponds to the voltage probe marker tagged "SpectrumAnalyzer" in figure 7.1 and V_{in} is "Vsource". The factor 2 comes from a difference in voltage reference levels.

Figure 7.8 on page 142 presents results from measurements conducted with the coils 10-2 and 10-6 in air. The simulated results are in good agreement with the measurements. The maximum observed at around 13mm is due to the critical coupling discussed in section 4.2.1 on page 56. In addition to the simulation based on the coil parameters found in the air impedance measurements of the coils, a simulation with adjusted parasitics is shown.

Figure 7.9 on page 143 shows results obtained in the freshwater and saltwater experiments. The simulations are still based on the impedance measurements of section 5.2.2. The agreement between simulation and measurements is good in the case of the freshwater measurement series carried out with coil 10-3 and 10-4 shown in plot A. The difference between measured and simulated values are 1-2dB. Also in plot B, for the same coil configuration in saltwater, the difference is approximately 2dB, and likewise for plot C for the coils 10-5 and 10-6.

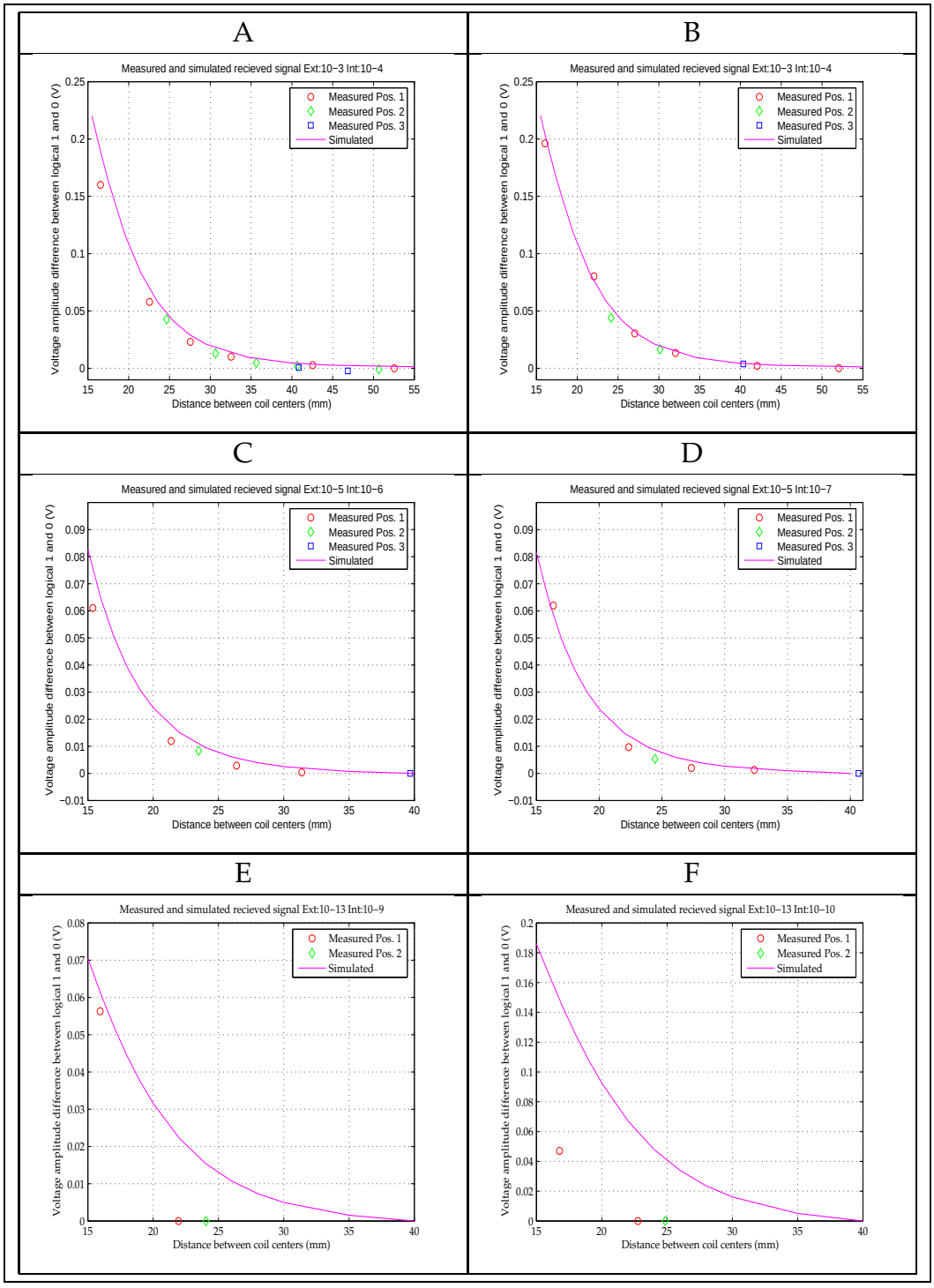


Figure 7.5: *Signal transmission measured with passive communication experimental setup in freshwater (only A) and saltwater.*

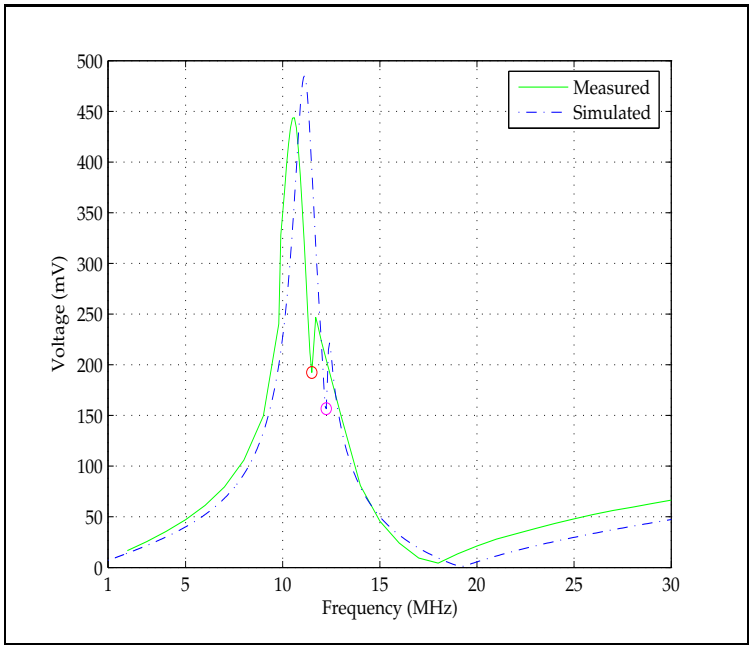


Figure 7.6: *The mismatch between the two resonance circuits is clearly seen for in the measured data. The mismatch was not well predicted by the simulations.*

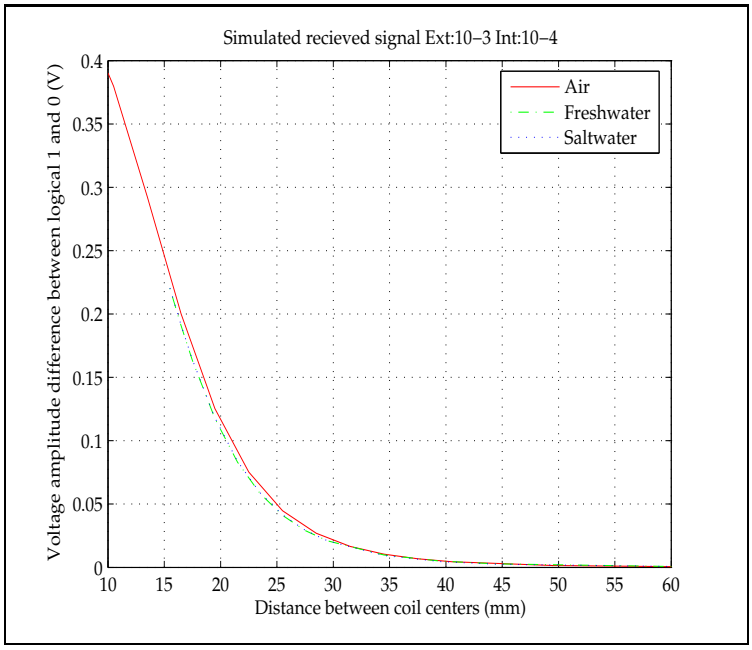


Figure 7.7: *Simulated signal transmission for passive communication experimental setup in air, freshwater and saltwater*

However in figure 7.9 plot D the differences increase. Even the simulation with adjusted parasitics does not give as good results as the plots A, B and C. The reason for this is not known. It is possible that the coil in the water tank has been moved out of its initial position on the water tank coil mount. This seems perhaps to be the most reasonable explanation as the additional deviation from the simulation results is the same for all points.

The figure 7.10 on page 144 is based on data from an attempt to measure and simulate signal transmission over slightly greater distances than earlier. In an attempt to increase the signal level at the input port of the HP4195A and thus increase the distance the signal could be measured the $10k\Omega$ resistance was substituted for a 51Ω resistor. However, it was not considered what implications this would have for the parallel resonance circuit on the PCB. A broadening of the resonance top was observed as well as an increase in the signal level. At large distances residual reactance caused other resonance peaks to become larger than the main resonance of the resonance circuits, which for the greatest distance was no longer observable. Hence the measurements taken at the three largest distances should be disregarded. The validity of the measurements presented in this figure is uncertain.

Summarizing the observations presented in this section, we can say that the power that can be transferred from an active excitation circuit through an inductive coupling to a passive circuit, is fairly limited. Even at very short distances, the signal level was observed to be attenuated 30dB and more. Hence a reduction of the power put in top the system is quickly reduced with a factor 1000 at the output. The measurements conducted in water did not show significant decrease of power transferred. This is in agreement with short operating ranges observed in the work of Berg [7] and Due [17].

7.3 Inclination angle dependence

An interesting parameter to investigate is how the system performance is affected by coil misalignment. With the same experimental setup as

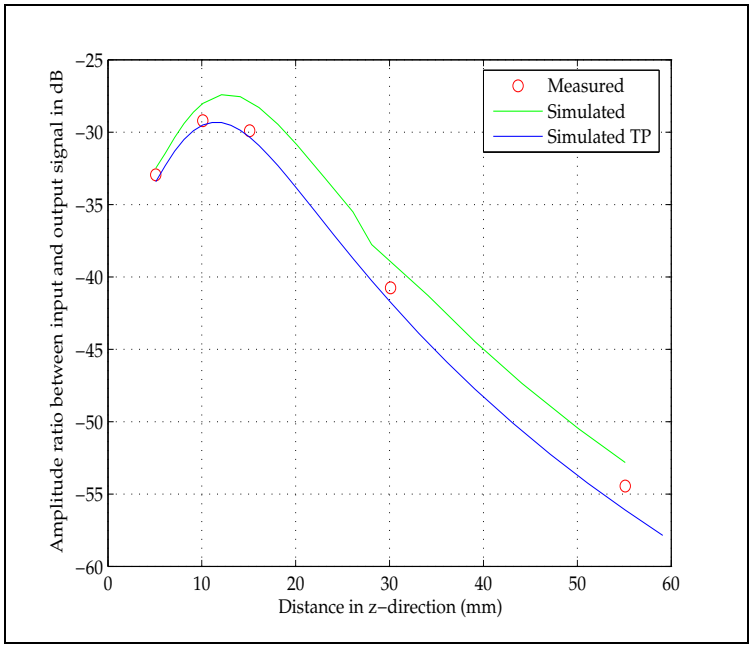


Figure 7.8: *The figure compares measured and simulated received signal level at the test point of circuit 2. In this case the both coils, 10-2 and 10-6 were situated in air.*

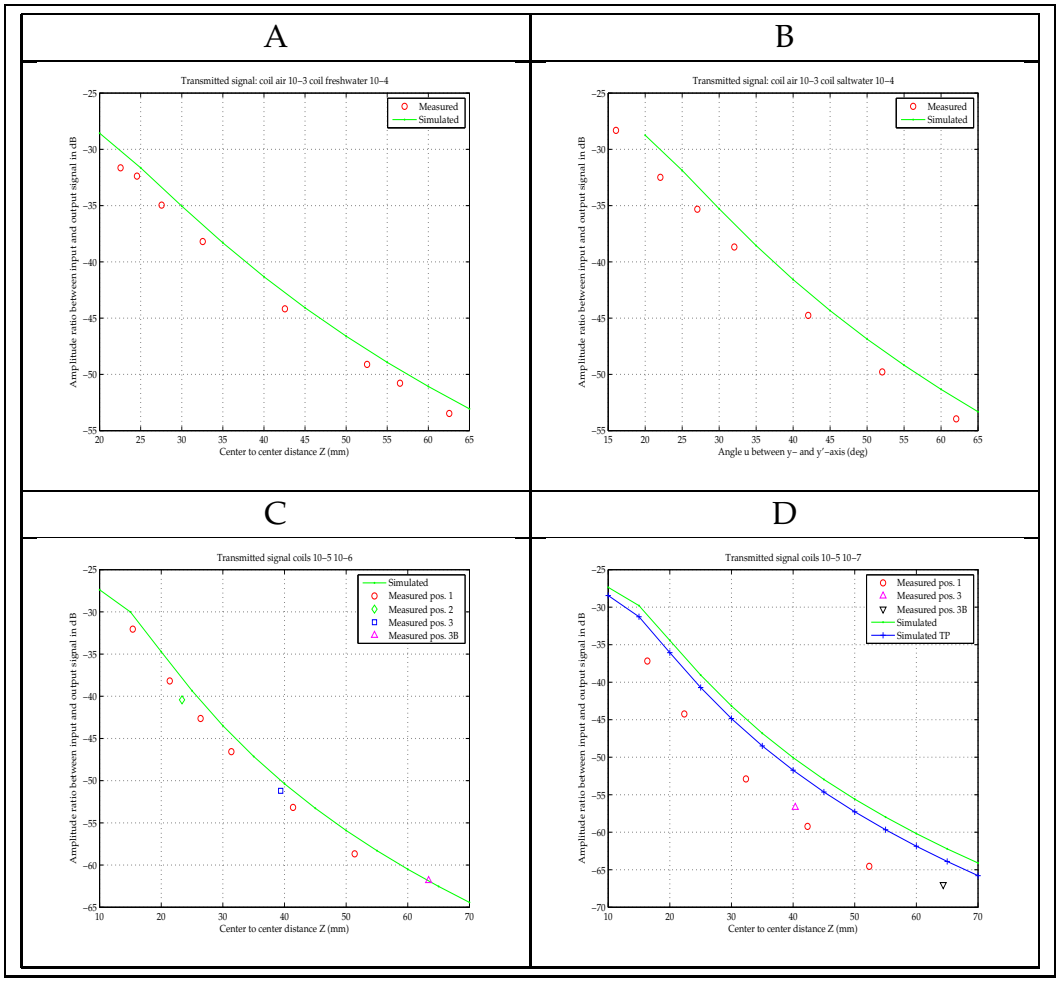


Figure 7.9: *Comparison of transmitted signal level (dB) measured with experimental setup with TR experimental setup. A1 shows results from a freshwater experiment. The remaining plots are saltwater experiments.*

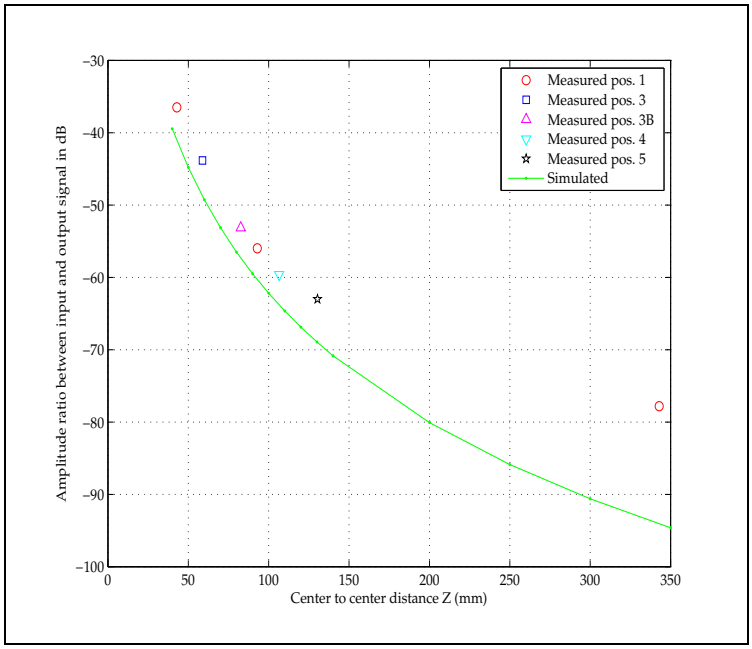


Figure 7.10: *The figure compares measured and simulated received signal level at the test point of circuit 2. In this case the coil 10-3 was situated in air while 10-4 was situated in the saltwater tank. This time the series resistance is reduced to 50Ω. Postions 4 and 5 are positions 23.8mm and 47.7 mm further from the water tank wall than position 3B*

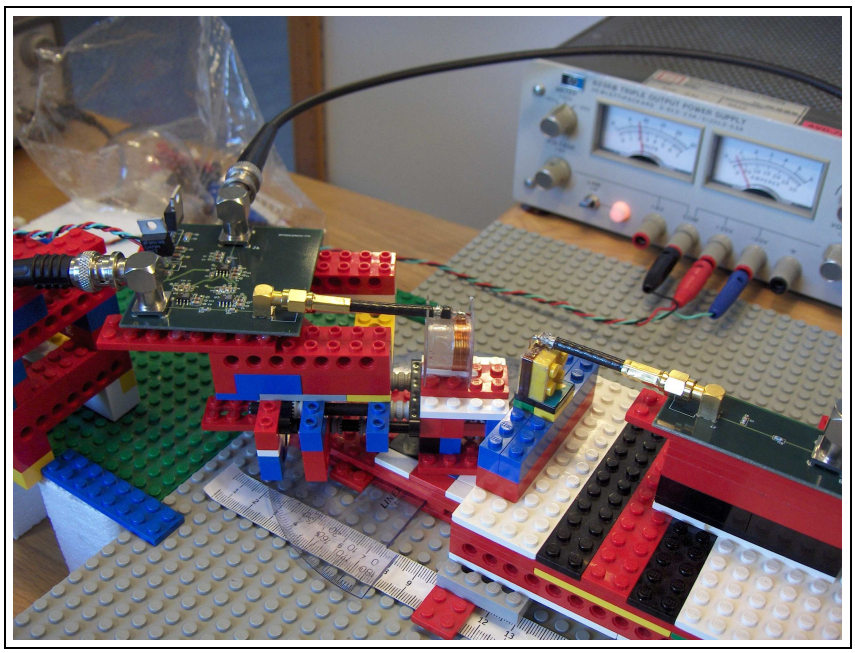


Figure 7.11: *A picture of the coil rotating device build in Lego.*

in the preceding section including the coil rotating device described in section 5.1 on page 101 and shown in the figure 7.11, such an investigation was undertaken. Using the same geometry as in the presentation of the calculation program in Chapter 6 on page 125, the coil rotator controls misalignments in the angle u.

Figure 7.12 on page 147 presents results of measurements and simulations done with this setup. The top and center plot of figure 7.12 shows results for air measurements. The increasing deviation for angles approaching 90° is expected as the null of the coil with calculation method used, theoretically should be zero. In the mutual inductance calculation program the coil is approximated to consist of adjacent rings in stead of the helix form its physical implementation has. Thus at $u = 90°$ the dot product $d\mathbf{l}_1 \cdot d\mathbf{l}_2$ in equation (6.1 on page 125) becomes zero in the theoretical case, but never really becomes zero in the real world. Actually this is a well known limitation of the loop antenna when used in simple direction finding applications where the null of the loop antenna is used to determine the direction from which a signal is transmitted. The points marked (*)

are plots achieved by adjusting for a displacement of the rotation axis of the coil rotator and the coil center in the mutual inductance calculations. The measurements were done at initial coil center to coil center distances of 30.8mm (top plot) and 28.45mm (center plot).

The measurements conducted in saltwater is presented in the bottom plot in figure 7.12. They were done with a coil center to center distance of 34.95mm at $u = 0^\circ$. The measured values do not correspond as well with simulations as in the case above. Difficulties in determining the initial center to center distance is thought to be a source of error.

From these measurements we can conclude that the power transfer between the coils is not significantly affected for misalignment angles less than 50 degrees. Also the simulations did not predict signal transfer as well as when the coils were coaxially aligned.

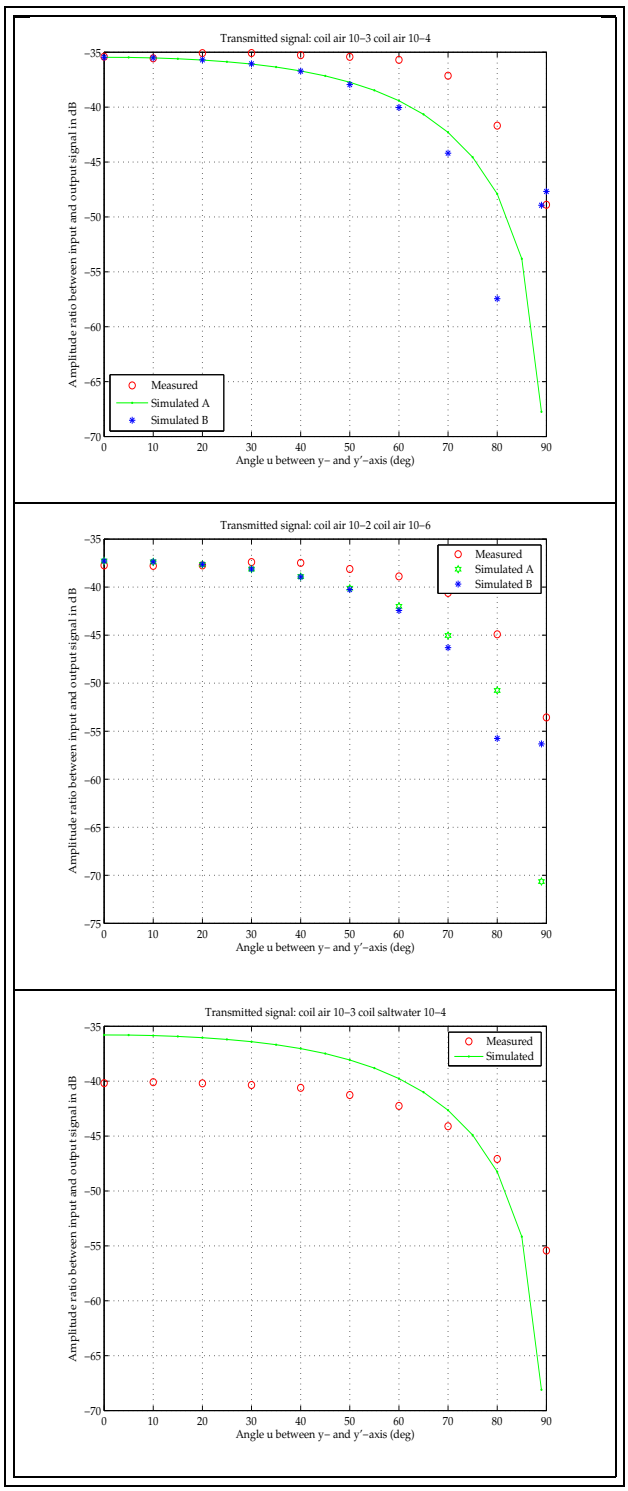


Figure 7.12: *The figure shows the results from measurements of the signal transmission dependency of the inclination angle of the axes of the coils.*

Chapter 8

Summary and Conclusions

The subject of this thesis is wireless communication by means of electromagnetic fields, between an implanted glucose sensor and an external device in a glucose monitoring system for diabetics. Two different design approaches have been investigated:

- The passive approach; the implanted device is designed without an internal energy source.
- The active approach; the implanted device is designed with an internal energy source.

An experimental setup primarily relevant for an investigation of the passive approach was designed. It consisted essentially of an RF transmission system with signal monitoring instruments connected. An implant is situated in a medium with electromagnetic properties differing from those in air. A main objective of the measurements were to investigate effects of the medium on the communication system. Antenna coils with a design relevant for that of a passive system design was therefore constructed. These coils were later encapsulated and the change in their properties when submerged in aqueous sodium chloride solution was measured. In fact the encapsulation material itself was measured to change the properties.

The measured coil properties were used in PSPICE-simulations to predict signal transmission relevant for the transmission experiment. Predicted and observed signal levels were compared and the agreement was

found generally good.

The experimental results indicate that a passive system is a possible solution for wireless communication between an implanted glucose sensor and an external monitoring system.

Theoretical estimates of signal attenuation in human tissue indicate that an active system also is a possible solution using state of the art radio transceiver chips and batteries.

The operating range is limited to a few centimeters for a passive system and the order of one meter for the active system. The implanted communication system will in both cases demand a volume in the order of a few cubic centimeters.

Further investigations should include transmission measurements on active systems, with focus on antenna design and operating frequencies. More accurate measurements on passive system performance require a stable coordinate system and improved coil construction and encapsulation equipment. This could also make it possible to study the effect of the encapsulation closer, that is its thickness and how well suited different materials are for this purpose.

Bibliography

[1] New advances in glucose monitoring, 2004. Article at World Wide Web Site: http://www.diabetesnet.com/.

[2] Alcohol powers fuel cell, September 2003. News Article at World Wide Web Site: http://www.trnmag.com/.

[3] American diabetes association. World Wide Web Site: http://www.diabetes.org/.

[4] Animas corporation website, 2003. World Wide Web Site: http://www.animascorp.com/.

[5] Battery powers tiny implants that aid neurological disorders, September 2003. World Wide Web Site: http://www.anl.gov/OPA/news03/news030919.htm.

[6] Phillip Ball. Chemists build body fluid battery, November 2002. World Wide Web Site: http://www.nature.com/nsu/021111/021111-1.html.

[7] Morten Berg. Sensor og utlesningskrets for måling av intrakranielt trykk. Master's thesis, Fysisk institutt, Universitet i Oslo, 2002.

[8] Dawn M. Bina, Robyn L. Anderson, Mary L. Johnson, Richard M. Bergenstal, and David M. Kendall. Clinical impact of prandial state, exercise, and site preparation on the equivalence and alternative-site blood glucose testing. *Diabetes Care*, 2003.

[9] Wilson greatbatch, January 2006. World Wide Web Site: http://www.biophan.com/greatbatch.php.

[10] New advance in fuel cell technology may help power medical implants, March 2002. World Wide Web Site: http://www.brown.edu/Aministration/News_Bureau/2002-03/02-098.html.

[11] Isidor Buchmann. Battery university, January 2006. World Wide Web Site: http://www.batteryuniversity.com.

[12] Bupa diabets 2 fact sheet, November 2003. World Wide Web Site: http://hcd2.bupa.co.uk/fact_sheets/html/diabetes2.html.

[13] Alister Burr. *Modulation and Coding for Wireless Communications*. Prentice Hall, 2001.

[14] Cmetelemetrix website, 2003. World Wide Web Site: http://www.cmetele.com/.

[15] Andrew DeHennis and Kensall D. Wise. A double-sided single-chip wireless pressure sensor. *Digest IEEE Conference on MicroElectroMechanical Systems*,, pages 252–255, 2002.

[16] Henry E. Duckworth. *Electricity and Magnetism*. Holt, Rinehart and Winston, Inc, 1960.

[17] Jon Due-Hansen. Systemløsning for måling av intrakranielt trykk. Master's thesis, Fysisk institutt, Universitet i Oslo, 2002.

[18] Summary of safety and effectiveness data, March 2001. FDA approval of the Cygnus GlucoWatch Biographer, available from World Wide Web Site: http://www.fda.gov/cdrh/pdf/p990026.html.

[19] Alliance for Telecommunications Industry Solutions. American national standard for telecommunications - telecom glossary 2000, January 2001. World Wide Web Site: http://www.atis.org/tg2k/.

[20] Fuel cell, January 2006. World Wide Web Site: http://www.ecn.nl/bct/fuelcellinfo/principle.en.html.

[21] Frederick W. Grover. *Inductance Calculations - Working Formulaes and Tables*. D. Van Nostrand Company, Inc, 1947.

[22] Riadh W. Y. Habash. *Electromagnetic Fields and Radiation. Human Bioeffects and Safety*. Marcel Dekker, Inc, 2002.

[23] M. P. M. Hall, L. W. Barcly, and M. T. Hewitt. *Propagation of Radiowaves*. IEE, 1996.

[24] Aarne Halme, Xia-Chang Zhang, and Anja Ranta. Study of biological fuel cells, January 2006. World Wide Web Site: http://www.automation.hut.fi/research/bio/sfc00pos.htm.

[25] Z. Hamici, R. Itti, and J. Champier. A high-efficiency power and data transmission system for biomedical implanted electronic devices. *Meas. Sci. Technol*,, pages 192–201, 1996.

[26] Timothy J. Harpster, Brian Stark, and Khalil Najafi. A passive wireless integrated humidity sensor. *Sensors and Acutators*, 2002.

[27] Egil Haug, Olav Sand, and Øystein V. Sjaastad. *Menneskets fysiologi*. Gyldendal, 2000.

[28] Richard C. Johnson. *Antenna Engineering Handbook*, chapter 1 Introduction to Antennas. McGraw-Hill, third edition, 1993.

[29] Karsten Jungheim and Theodor Koschinsky. Risky delay of hypoglycemia detection by glucose monitoring at the arm. *Diabetes Care*, 2001.

[30] Andrew J. Karter, Lynn M. Ackerson, Jeanne Darbinian, Ralph B. D'Agostino Jr, Assiamira Ferrara, Jennifer Liu, and Joe V. Selby. Self-monitoring of blood glucose levels and glycemic control: the northern california kaiser permanente diabetes registry. *The American Journal of Medicine*, 2001.

[31] Ronold W. P. King and Glenn S. Smith. *Antennas in Matter*. MIT Press, 1981.

[32] Will Knight. Printable battery rolls of the presses, July 2001. World Wide Web Site: http://www.newscientist.com/news/news.jsp?id=ns99991069.

[33] N. N. Lebedev. *Special Functions and their Applications*. Prentice Hall, 1965.

[34] David R. Lide, editor. *The Handbook of Chemistry and Physics*. CRC press, 82 edition, 2001-2002.

[35] Test strips for blood glucose testing, 2003. Patient information from LifeScan concerning OneTouch Test Strips for Blood Glucose Testing.

[36] Reinhold Ludwig and Pavel Bretchko. *RF Circuit Design - Theory and Applications*. Prentice Hall, 2000.

[37] Painless test for blood sugar under development, 2000. World Wide Web Site: http://webmd.lycos.com/content/article/21/1728_55291.

[38] Toshiko Yoshioka Mariko Kawaguri and Shiro Nankai. Disposable glucose sensor employing potassium ferricyanide as a mediator. *Denki Kagaku Kyokai*, 58(12):1119–1124, 1990.

[39] Claudia Marschner, Sven Rehforss, Dagmar Peters, Hilmar Bolte, and Rainer Laur. A novel circuit concept for psk-demodulation in passive telemetric systems. *Microelectronics Journal*, 2002.

[40] James McLean, Robert Sutton, and Rob Hoffmann. Interpreting antenna performance parameters for emc applications - part 2. DJM electronics article, 2003. World Wide Web Site: http://www.djmelectronics.com/articles.html.

[41] Rick Mendosa. Lag time in alternativeland. are alternative site meters dangerous?, 2001. World Wide Web Site: http://www.diabetesmonitor.com/alt_lags.htm.

[42] Rick Mendosa. Blood glucose meters, 2004. World Wide Web Site: http://www.mendosa.com/meters.htm.

[43] Steve Moshier. coil02, 2002. Coil inuctance calculation program avialable at http://www.moshier.net/coildoc.html.

[44] Stephen J. Mraz. Powering up implants - lithium batteries for mediacal implants. *Machine Design*, 1995.

[45] Joseph A. Paradiso Nathan S. Shenck. Energy scavenging with shoemounted piezoelectrics. *IEEE Micro*, pages 30–42, Jun 2001.

[46] Nordic Semiconductor. *Small loop antennas Application Note nAN400-03*, 1.2 edition, feb 2000.

[47] Nordic Semiconductor. *Single chip 433/868/915 MHz Transceiver nRF905*, 1.3 edition, dec 2005.

[48] Norges diabetesforbund. World Wide Web Site: http://www.diabetes.no/.

[49] Kazunari Okada and Toshimasa Sekino. *Agilent Technologies Impedance Measurement Handbook*. Agilent Technologies, 2003. Available at www.agilent.com.

[50] On-line medical dictonary. World Wide Web Site: http://cancerweb.ncl.ac.uk/cgi-bin/omd.

[51] K.G. Ong, C.A. Grimes, C.L. Robbins, and R.S. Singh. Design and application of a wireless, passive, resonant-circuit environmental monitoring sensor. *Sensors and Actuators*, 2001.

[52] Henry W. Ott. *Noise Reduction Techniques in Electronic Systems*. John Wiley & Sons, Inc., 1988.

[53] Wolfgang K. H. Panofsky and Melba Phillips. *Classical Electricity and Magnetism*. Addison-Wesley, 2 edition, 1962.

[54] Pendragon website, 2003. World Wide Web Site: http://www.pendragonmedical.com/.

[55] J.G. Proakis. *Digital Communications*. McGraw-Hill, 3 edition, 1995.

[56] Quallion - products and applications, 2004. World Wide Web Site: http://www.quallion.com/prod_i.html.

[57] Nannapaneni Narayana Rao. *Elements of Engineering Electromagnetics*. Prentice Hall, 5 edition, 2000.

[58] Behzad Razavi. *RF Microelectronics*. Prentice Hall, 1998.

[59] M. H. Repacholi. Low-level exposure to radiofrequeny fields: Health effects and research needs. *Bioelectromagnetics 19, pp 20-32*, 1998.

[60] Shadrach Joseph Roundy. *Energy Scavenging for Wireless Sensor Nodes with a Focus on Vibration to Electricity Conversion*. PhD thesis, The Unniversity of California, Berkeley, 2003.

[61] Shadrach Joseph Roundy. Energy scavenging for wireless sensor nodes with a focus on vibration to electricity conversion, April 2003. Power Point Presentation found at: http://engnet.anu.edu.au/DEpeople/Shad.Roundy/.

[62] Gene Siscoe. Ultrasmart® glucose monitor, three strikes and out, January 2006. World Wide Web Site: http://www.mendosa.com/otus.htm.

[63] A.K. Skrivervik, J.-F. Zurcher, O. Staub, and J. R. Mosig. Pcs antenna design: The challenge of miniaturization. *IEEE Antennas and Propagation Magazine*, 2001.

[64] Orhan Soykan. Power sources of implantable medical devices. In *Medical Devices Manufacturing & Technology 2002 - June 2002*, pages 76–79. World Markets Research Centre Ltd, 2002.

[65] U.s. department of energy and renewable energy - types of fuel cells, January 2006. Article found at World Wide Web Site: http://www.eere.energy.gov/.

[66] What causes type 2 diabetes, January 2006. Article at World Wide Web Site: http://www.well-connected.com/report.cgi/000060_2.htm.

[67] Douglas H. Werner. An exact integration procedure for vector potentials of thin circular loop antennas. *IEEE Transactions on Antennas and Propagation*, 44(2):157 to 165, feb 1996.

[68] Douglas H. Werner and Suman Ganguly. An overview of fractal antenna engineering research. *IEEE Antennas and Propagation Magazine, Vol. 45, No. 1*, 2003.

[69] World health organization. World Wide Web Site: http://www.who.int/health_topics/diabetes_mellitus/en/.

[70] Jim Williams. High speed amplifier techniques - a designer's companion for wideband circuitry. Linear Technology, Application Note 47, 1991.

[71] Wilson greatbatch technologies inc. World Wide Web Site: http://www.greatbatch.com.

[72] T.J. Yen, N. Fang, Zhang X, and G. Q. Lu amd C. Y. Wang. A micro methanol fuel cell operating at near room temperature. *Applied Physics Letters*, 2003.

[73] Ørjan G. Martinsen, Sverre Grimnes, and Herman P. Schwan. Interface phenomena and dielectric properties of biological tissue. *Encyclopedia of Surface and Colloid Science*, 2002.

Appendix A

Neumann's Formula

Given two loops of conductive wire as shown in figure 4.2 on page 47. If a current $\mathbf{I_1}$ is present in loop 1, it sets up a magnetic field $\mathbf{B_1}$ around the loop. Loop 2 will then enclose a certain magnetic flux. A change in $\mathbf{B_1}$ will induce an electromotive voltage in loop 2. Faraday's law gives this voltage as:

$$V_{emf\,2} = \oint_C \mathbf{E} \cdot d\mathbf{l}_2 = -\frac{\partial}{\partial t} \int_S \mathbf{B}_1 \cdot dS_2 \tag{A.1}$$

The magnetic field can be represented by the magnetic vector potential $\mathbf{A}$:

$$\mathbf{B}_1 = \nabla \times \mathbf{A}_1 \tag{A.2}$$

By combining A.1 and A.2, we get

$$V_{emf\,2} = -\frac{\partial}{\partial t} \int_S (\nabla \times \mathbf{A}_1) dS \tag{A.3}$$

By applying Stoke's Theorem this can be rewritten:

$$-\frac{d}{dt} \int_S (\nabla \times \mathbf{A}_1) dS = -\frac{d}{dt} \oint_C \mathbf{A}_1 \cdot d\mathbf{l}_2 \tag{A.4}$$

In vacuum $\mathbf{A}_1$ is given as:

$$\mathbf{A}_1 = \frac{\mu_0 I_1 \cdot d\mathbf{l}_1}{4\pi R}, \tag{A.5}$$

where μ_0 is the permeability in vacuum, R is the distance from the differential length vector $d\mathbf{l}_1$ to the point where $\mathbf{A}_1$ is evaluated. If we rewrite

$\mathbf{A}_1$ in equation A.4 with the expression in equation A.5, we get:

$$V_{emf\,2} = -\frac{d}{dt}\oint_{C_1}\oint_{C_2}\frac{\mu_0 I_1}{4\pi R}\cdot d\mathbf{l}_1 d\mathbf{l}_2 \tag{A.6}$$

Combining this with the definition of mutual inductance results in:

$$M = \frac{V_{emf\,2}}{\frac{dI_1}{dt}} = \frac{\mu_0}{4\pi}\oint_{C_1}\oint_{C_2}\frac{d\mathbf{l}_1\cdot d\mathbf{l}_2}{R}, \tag{A.7}$$

which is Neumann's formula for mutual inductance between two current conducting loops.

Appendix B

Skin Effect

The combination of frequency and wire thickness encountered in the simulations and measurements lies in an area where neither DC-estimates nor the high frequency skin depth approximation may give a satisfactory estimate of the wire resistance. Hence, the skin effect is studied more closely.

The case of a straight cylindrical wire in free space is studied here. We seek to find an expression for the impedance of the wire in figure B.1. The wire is of length $l \ll \lambda$ and has a diameter of $2a$. A voltage source applies an excitation voltage of $V = V_0 \cos(\omega t)$ to the conductor. A relevant solution is given in [36] appendix B, and similar calculations are carried out here.

If end-effects are neglected the induced electric field strength can be approximated by:

$$\mathbf{E}(r,t) = E_0(r) \exp(j\omega t)\mathbf{i}_z \tag{B.1}$$

We note at this point that the part of the E-field carrying physical significance is:

$$E_p(r,t) = \Re\{E_0(r) \exp(j\omega t)\} \tag{B.2}$$

From Maxwell's curl equations we get the following differential equation for the E-field inside the conductor:

$$\frac{\partial^2 \mathbf{E}}{\partial r^2} + \frac{1}{r}\frac{\partial \mathbf{E}}{\partial r} - \sigma\mu\frac{\partial \mathbf{E}}{\partial t} = 0 \tag{B.3}$$

Here the displacement current term in Maxwell's curl equation for the **H**-field has been neglected as $\sigma \gg \omega\epsilon$ in a good conductor. Using the equa-

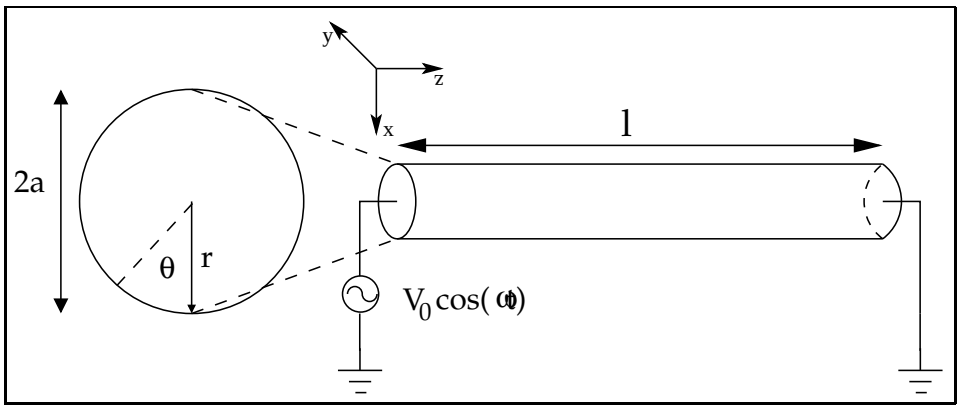


Figure B.1: *The figure shows the geometry of the cylindrical conductor referred to below.*

tions B.1, B.3 and the separation of variables technique we get:

$$\frac{\partial^2 E_0(r)}{\partial r^2} + \frac{1}{r}\frac{\partial E_0(r)}{\partial r} - j\omega\sigma\mu E_0(r) = 0 \tag{B.4}$$

We rewrite this equation by introducing the variable $\eta = r \cdot \sqrt{-j\omega\sigma\mu}$:

$$\frac{d^2 E_0(\eta)}{d\eta^2} + \frac{1}{\eta}\frac{dE_0(\eta)}{d\eta} + E_0(\eta) = 0 \tag{B.5}$$

The solution of this linear second order differential equation is given by [33] as:

$$E_0(\eta) = A_0 J_0(\eta) + B_0 Y_0(\eta) \tag{B.6}$$

where $J_0(\eta)$ and $Y_0(\eta)$ are Bessel functions of the first and second kind, respectively.

Now, the solution for the electric field strength must fulfill certain conditions to be a valid physical solution. The function Y_0 approaches ∞ as the argument approaches zero. Hence we demand $B_0 = 0$. A_0 is a complex variable; $A_0 = A_{0r} + jA_{0i}$. The function $J_0(\eta)$ is also complex and can be rewritten $J_0(\eta) = J_{0r}(\zeta(r)) + jJ_{0i}(\zeta(r))$ where $\zeta(r) = r\sqrt{\omega\sigma\mu}$. A_{0r}, A_{0i}, J_{0r}, J_{0i} and ζ all have real values.

The voltage over the conductor was given as $V = V_0 \cos(\omega t)$. This implies that the E-field at the surface of the conductor is:

$$E_p(a,t) = \frac{V_0}{l}\cos(\omega t) \tag{B.7}$$

This boundary condition can be used to find A_{0r} and A_{0i} by setting $r = a$ in the Bessel solution for the E-field:

$$\begin{aligned}
\Re\{E_p(a,t)e^{j\omega t}\} &= \Re\{(A_{0r} + jA_{0i})(J_{0r}(\zeta(a)) + J_{0i}(\zeta(a)))e^{j\omega t})\} \\
&= \Re\{[(A_{0r}J_{0r}(\zeta(a)) - A_{0i}J_{0i}(\zeta(a))) \\
&\quad +j(A_{0r}J_{0i}(\zeta(a)) + A_{0i}J_{0r}(\zeta(a)))] \\
&\quad [\cos(\omega t) + j\sin(\omega t)]\} \\
&= (A_{0r}J_{0r}(\zeta(a)) - A_{0i}J_{0i}(\zeta(a)))\cos(\omega t) \\
&\quad -(A_{0r}J_{0i}(\zeta(a)) + A_{0i}J_{0r}(\zeta(a)))\sin(\omega t) \qquad \text{(B.8)}
\end{aligned}$$

Applying the boundary condition gives:

$$\begin{aligned}
A_{0r}J0r(\zeta(a)) - A_{0i}J_{0i}(\zeta(a)) &= \frac{V_0}{l} \\
A_{0r}J0i(\zeta(a)) - A_{0i}J_{0r}(\zeta(a)) &= 0
\end{aligned}$$

Finally the results for A_{0r} and A_{0i}

$$A_{0r} = \frac{V_0\, J_{0r}(\zeta_a)}{l\,[J_{0r}(\zeta_a)^2 + J_{0i}(\zeta_a)^2]} \qquad \text{(B.9)}$$

$$A_{0i} = -\frac{V_0\, J_{0i}(\zeta_a)}{l\,[J_{0r}(\zeta_a)^2 + J_{0i}(\zeta_a)^2]} \qquad \text{(B.10)}$$

$$\text{(B.11)}$$

where $\zeta_a = \sqrt{\omega\sigma\mu}$. This results in the following electric field:

$$\begin{aligned}
E_p(r,t) &= \frac{V_0}{l\,[J_{0r}(\zeta_a)^2 + J_{0i}(\zeta_a)^2]}\,([J_{0r}(\zeta_a)J_{0r}(\zeta) + J_{0i}(\zeta_a)J_{0i}(\zeta)]\cos(\omega t) \\
&= +[J_{0i}(\zeta_a)J_{0r}(\zeta) - J_{0r}(\zeta_a)J_{0i}(\zeta)]\sin(\omega t)) \qquad \text{(B.12)}
\end{aligned}$$

To find the impedance of the wire we need to find the current running in the wire. The current density is given by:

$$J_p(r,t) = \sigma E_p(r,t) \qquad \text{(B.13)}$$

This indicates the radial distribution of the current density. The total current is:

$$I(t) = \int_0^{2\pi}\int_0^a J_p(r,t)r\,dr\,d\theta \qquad \text{(B.14)}$$

which formally give

$$I(t) = I_1 \cos \omega t + I_2 \sin \omega t = I_0 \cos(\omega t + \phi) \tag{B.15}$$

where

$$I_0 = \sqrt{(I_1^2 + I_2^2)}, \;\; \cos \phi = \frac{I_1}{I_0}, \;\; \sin \phi = -\frac{I_2}{I_0} \tag{B.16}$$

We can finally calculate the impedance $Z = \|Z\| e^{j\Phi}$ from $V = ZI$. Introducing complex notation for the current we get:

$$I(t) = \Re\{I_0 \exp(j(\omega t + \phi)\} \tag{B.17}$$

which give

$$V_0 \cos \omega t = \|Z\| \, I_0 \, \Re\{\exp(j(\omega t + \phi + \Phi)\} \tag{B.18}$$

and hereby

$$V_0 = \|Z\| \, I_0 \tag{B.19}$$

$$\Phi = -\phi \tag{B.20}$$

The ohmic resistance in the conductor is

$$R = \|Z\| \, \cos \phi \tag{B.21}$$

and the reactance, which is purely inductive, is

$$X = -\|Z\| \, \sin \phi \tag{B.22}$$

High order series expansions of R using maple has been used to generate the R_{Bessel} curves in figure 4.14 on page 70.

Appendix C

Modulation Techniques - Theoretical Background

In this appendix some theoretical background necessary to get a better understanding of the discussion in section 4.3.3. The presentation of the theory is based on the books "Modulation and Coding" by Alistair Burr [13] and "RF Microelectronics" by Behzad Razavi [58].

C.1 Analytical Representation of Digitally Modulated Signals

In the quest for power efficient modulation schemes, we need a suitable representation of modulation schemes and good tools for analyzing them. A general analytical form for modulated signals can be achieved with what is called the complex baseband representation. It is based on a structure known as an I/Q-modulator which can be used to generate a signal of any given phase and amplitude. A rough sketch of the principles in this structure is given in figure C.1. Here an in-phase signal $b_p(t)$ and a quadrature signal $b_q(t)$ modulates two carriers, one of them phase shifted by $\pi/2$ compared to the other:

$$\begin{aligned} s(t) &= b_p(t)\cos(\omega_c t) + b_q(t)\cos\left(\omega_c t + \frac{\pi}{2}\right) \\ &= b_p(t)\cos(\omega_c t) - b_q(t)\sin(\omega_c t) \end{aligned} \tag{C.1}$$

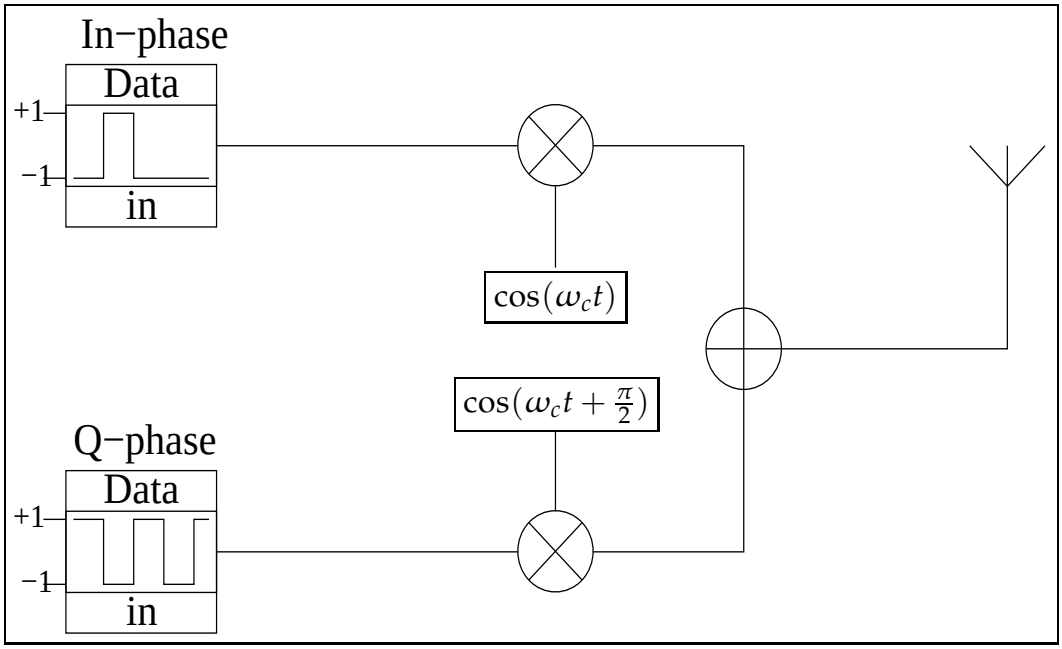


Figure C.1: Principles of an IQ-modulator

This motivates for a complex representation of the baseband signal:

$$b(t) = b_p(t) + jb_q(t) \tag{C.2}$$

Now we may write the modulated signal as:

$$\begin{aligned} s(t) &= Re\{b(t)e^{j\omega_c t}\} = Re\{b(t)(\cos(\omega_c t) + j\sin(\omega_c t))\} \\ &= b_p(t)\cos(\omega_c t) - b_q(t)\sin(\omega_c t) \end{aligned} \tag{C.3}$$

We can show that in general a modulated signal can be represented by:

$$\begin{aligned} s(t) &= A(t)\cos(\omega_c t + \phi(t)) \\ &= A(t)\cos\phi(t)\cos(\omega_c t) - A(t)\sin\phi(t)\sin(\omega_c t) \\ &= Re\{b(t)e^{j\omega_c t}\} \end{aligned} \tag{C.4}$$

A useful visual representation of the complex baseband representation is the constellation diagram. It is a plot of the signal constellation points in the complex plane or in an I-Q plane.

Another important tool in analyzing modulation schemes is the signal spectrum. It provides amplitude information for each frequency present in a signal. It is obtained by taking the Fourier transform of the modulated signal:

$$\begin{aligned} S(\omega) &= \int_{-\infty}^{\infty} \left(Re\{b(t)e^{j\omega_c t}\} \right) e^{-j\omega t} dt \\ &= \int_{-\infty}^{\infty} \frac{1}{2} \left(b(t)e^{j\omega_c t} + b^*(t)e^{j\omega_c t} \right) e^{-j\omega t} dt \\ &= \frac{1}{2} \left(\int_{-\infty}^{\infty} b(t)e^{-j(\omega-\omega_c)t} dt + \int_{-\infty}^{\infty} b(t)e^{-j(-\omega-\omega_c)t} dt \right) \\ &= \frac{1}{2} \left(B(\omega - \omega_c) + B^*(-\omega - \omega_c) \right) \end{aligned} \tag{C.5}$$

The terms $B(\omega - \omega_c)$ and $B^*(-\omega - \omega_c)$ are frequency shifted versions of the baseband signal. The first term is centered around ω_c, while the second term is centered, but flipped, around $-\omega_c$. So all essential information of the modulated signal, except the carrier frequency lies in the spectrum of the baseband signal.

The power density spectrum of a signal shows how much power a signal carries in a unit bandwidth around frequency f [58] . The formal definition for a continuous time signal is:

$$S_x(f) = \lim_{T \to \infty} \frac{\overline{|\ X_T(f)\ |^2}}{T}, \tag{C.6}$$

where

$$X_T(f) = \int_0^T x(t)e^{(-j2\pi ft)} dt \tag{C.7}$$

which is the Fourier transform when the signal is evaluated in the time interval from 0 to T. By squaring this, a measure for the power dissipated by a unit resistor is obtained. This value is then averaged over time, that is, the Fourier transform is calculated for a large number of consecutive time intervals (as indicated in equation C.6 by the averaging over infinite time).

C.2 Band limiting of the baseband signal

To meet the restrictions mentioned in 4.4 from regulatory authorities, filtering of the baseband signal can be necessary to limit the bandwidth of the main lobe and suppress the side lobes. When such band limiting filters are applied, the signal is dispersed in time and can cause interference of one signaling pulse with the next. This is called intersymbol interference (i.s.i.) and can cause an increase in the BER of a system.

It is possible, by careful design of the baseband bandpass filter, to eliminate i.s.i. If the received signal is sampled at the end of each period T of the digital baseband signal, it is possible to force the i.s.i. to be zero at these instances in time. Each signaling pulse can be adjusted so that it passes through zero at all sampling instances except when that pulse itself is being sampled. That is the sum of all other signaling pulses is zero when one pulse is sampled. This condition for eliminating i.s.i is known as the Nyquist criterion in the time domain. If each signaling pulse is denoted $g(t)$, we may write:

$$g(iT) = \begin{cases} 1 & \text{if} \quad i = 0 \\ 0 & \text{if} \quad i \neq 0 \end{cases} \tag{C.8}$$

Here T is the sampling period and $g(t)$ is sampled at $i = 0$. In the frequency domain this takes the form, [13, 55]:

$$\sum_{k=-\infty}^{\infty} G(f + \frac{k}{T}) = T, \qquad |f| \leq \frac{1}{2T} \tag{C.9}$$

This means that the double-sided bandwidth of such a signal will never be less than 1/T.

To achieve this kind of filtering, a raised cosine filter is most often used; described mathematically by:

$$|G(f)| = \begin{cases} 1 & \text{if} \quad |f| \leq \frac{1}{2T}(1-\beta) \\ \frac{1}{2}\left(1 - \sin\left(\frac{\pi T}{\beta}(|f| - \frac{1}{2T})\right)\right) & \text{if} \quad \frac{1}{2T}(1-\beta) < |f| < \frac{1}{2T}(1+\beta) \\ 0 & \text{if} \quad |f| \geq \frac{1}{2T}(1+\beta) \end{cases} \tag{C.10}$$

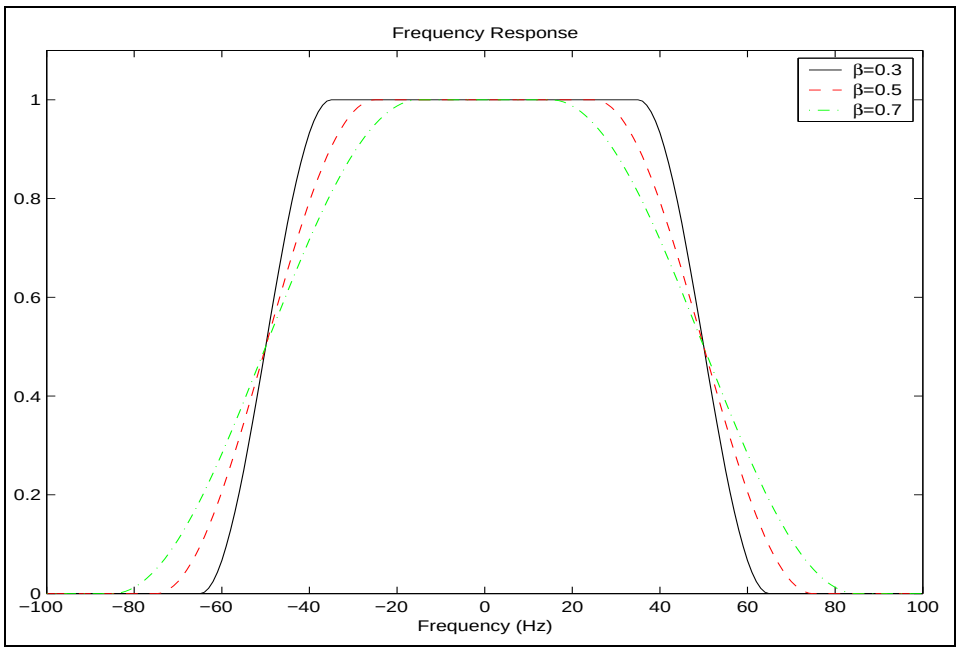


Figure C.2: *The frequency response of a raised cosine filter shown for different values of β. The symbol period is in this case 10ms.*

or in the time domain by:

$$g(t) = \frac{2\cos\left(\frac{\pi t\beta}{T}\right)\sin\left(\frac{\pi t}{T}\right)}{t\left(1-\frac{4t^2\beta^2}{T^2}\right)}. \tag{C.11}$$

When such a filter is used, the bandwidth is given by:

$$BW = \frac{1+\beta}{T} = r_s(1+\beta) \tag{C.12}$$

Here T is still the period of the digital baseband signal, and r_s the symbol rate (=modulation rate). β is called the roll-off factor and may assume values between 0 and 1, corresponding to an excess bandwidth of $0-100\%$. Choosing the value for β is a trade-off between narrow bandwidth and more desirable time-domain properties. How β affects the filter frequency response is illustrated in figure C.2.

C.3 Representing Noise with the Complex Baseband Representation

To be able to give a more realistic description of signals encountered in real world situations, we need to include noise into the signal representation given in section C.1. Considering the frequency of operation, operating distance and the fact that the conditions for transmission in the channel in question will be close to constant, there is little need for including multipath and other fading effects. Therefore, an adequate way to describe the channel of operation should be to apply a white Gaussian noise model. It has the following probability density function for the noise amplitude assuming the time average of the noise is zero:

$$p(n) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{n^2}{2\sigma^2}} \tag{C.13}$$

The model assumes the noise is frequency independent, that is the spectrum is flat with power spectral density N_0. The noise average power will be determined by the standard deviation:

$$P_{noise} = \overline{n^2} = \sigma^2 \tag{C.14}$$

This can be confirmed by

$$\overline{n^2} = \int_{-\infty}^{\infty} n^2 p(n) dn = \sigma^2 \tag{C.15}$$

The noise will be exposed to the same bandlimiting filters as the baseband signal. The noise within the bandwidth of the baseband in the transmitter will therefore be transferred up to the band occupied around the carrier. The noise can be represented as a random modulation of amplitude and phase of the carrier:

$$n(t) = A_n(t)\cos(\omega_c t + \Phi_n(t)) \tag{C.16}$$

We can also use the complex baseband representation:

$$n(t) = n_p(t)\cos(\omega_c t) - n_q(t)\sin(\omega_c t) \tag{C.17}$$

It can be shown [13,55] that if $n(t)$ has a Gaussian probability density function, then n_p and n_q must be independent and is also Gaussian distributed:

$$\overline{n_p^2(t)} = \overline{n_q^2(t)} = \sigma^2 = P_{noise} \tag{C.18}$$

Equivalently the phase angle Φ_n is uniformly distributed, and the amplitude A_n follows a Rayleigh distribution.

n_p and n_q are baseband noise signals, which after filtering has the double-sided bandwidth W. Now if $n_p(t)$ and $n_q(t)$ are assumed to be real signals, $N_p(f)$ and $N_q(f)$ are even functions of f. Intact the total power carried in one of these noise components is

$$\int_{-f_{max}}^{-f_{min}} N_x(f)df + \int_{f_{min}}^{f_{max}} N_x(f)df = \int_{f_{min}}^{f_{max}} 2N_x(f)df \tag{C.19}$$

The double-sided PSD is N_0. Hence the single-sided PSD of each is $2N_0$.

C.4 Matched Filters

To maximize the performance of the transmit and receive operation, a matched filter should be applied. The matched filter is designed to maximize the signal-to-noise-ratio (SNR) at its output. Now the SNR at the output of the filter at the sampling instant T is given by

$$SNR_{out} = \frac{s^2(T)}{\overline{n^2}} \tag{C.20}$$

where n^2 is the noise power and s(T) is the resulting signal after letting a signal g(t) pass through the matched filter with impulse response h(t). If $G(f)$ and $H(f)$ are the frequency domain representation of g(t) and h(t), we can represent s(T) as the inverse Fourier transform:

$$s(T) = \int_{-\infty}^{\infty} G(f)H(f)e^{(j2\pi fT)}df \tag{C.21}$$

The noise power, assuming white noise (one-sided spectrum), is:

$$\overline{n^2} = \frac{1}{2}\int_{-\infty}^{\infty} N_0|H(f)|df \tag{C.22}$$

It can be shown by use of the Cauchy-Schwartz inequality that the best SNR is achieved for

$$H(f) = G^*(f)e^{-j2\pi fT}. \tag{C.23}$$

This is what is called a matched filter. In the time domain the condition transforms to $h(t) = g(T-t)$, the time reversed version of g(t). By combining C.23 with equations C.20, C.21 and C.22 the output signal-to-noise-ratio becomes:

$$SNR_{out} = \frac{s^2(T)}{\overline{n^2}} \leq \frac{E\int_{-\infty}^{\infty}|H(f)|^2df}{\frac{1}{2}N_0\int_{-\infty}^{\infty}|H(f)|^2df} = \frac{2E}{N_0} \tag{C.24}$$

C.5 Combining Nyquist and Matched Filtering

To achieve both Nyquist and matched filtering in a radio system we need to consider both the transmitter, channel and receiver when designing such filters. Because it is at the input of the sampling device the i.s.i must be zero, the system must be designed so that the cascade of the filter in the receiver and the filter in the transmitter obey the Nyquist criterion. Hence to obey both the Nyquist criterion and to achieve matched filters, the filters must fulfill the criterions:

$$T(f)R(f) = RC(f)$$
$$T(f) = R^*(f)$$

$RC(f)$ denoting a raised cosine filter. Solving the set of equations leads to:

$$|R(f)| = |T(f)| = \sqrt{|RC(f)|} \tag{C.25}$$

To be able to calculate how the noise will effect the radio system's performance, we take a closer look on the noise picked up in the receiver. The output noise spectral density, assuming that the noise at the input is white, is:

$$\eta(f) = N_0|R(f)|^2 = N_0RC(f) \tag{C.26}$$

This means the total noise picked up is:

$$N = \int_0^{\infty}\eta(f)df = N_0\int_0^{\infty}RC(f)df = N_0W_{eff} = \frac{N_0}{T} \tag{C.27}$$

Not surprisingly we see that the wider the bandwidth, the more noise is picked up. In other words the noise present inside the filter bandwidth is the one distorting the signal at this point.

When the signal is received, the signal-to-noise-ratio is S/N_0, that is the received signal power to the noise power. Next the signal is demodulated and divided into one in-phase part b_p and one quadrature-phase part b_q by an IQ demodulator. On each phase part the p.s.d of the noise is $2N_0$ as seen in section C.3. Both phase components are run through a matched filter providing data estimates d_p and d_q for a decision device. At the output of each matched filter the signal energy to noise energy will be (applying the result of equation C.24):

$$\text{In} - \text{phase}: \quad SNR_{out} = \frac{2E_{in}}{N_{in}} = \frac{2E_p}{2N_0} = \frac{E_p}{N_0} = \frac{\overline{b_p^2}T_s}{N_0} = \frac{d_p^2}{\sigma^2} \tag{C.28}$$

$$\text{Q} - \text{phase}: \quad SNR_{out} = \frac{E_q}{N_0} = \frac{\overline{b_q^2}T_s}{N_0} = \frac{d_q^2}{\sigma^2} \tag{C.29}$$

where E_{in} is the signal energy at the input to the filter, N_{in} is the noise energy at the input to the filter, E_p and E_q the in-phase and quadrature-phase signal energy, respectively, and T_s the symbol period of the signal. $\overline{b_p^2}$ will be the power in the in-phase signal, so when multiplied by the period T, we get the energy per symbol in the in-phase signal part. Finally d_p^2 will be the power and σ the noise power in the signal fed to the decision device.

If we combine the to phases we see that the whole structure acts as a matched filter:

$$\frac{d_p^2 + d_q^2}{\sigma^2} = \frac{A^2}{\sigma^2} = \frac{(\overline{b_p^2} + \overline{b_q^2})T}{N_0} = \frac{2E}{N_0} \tag{C.30}$$

where $2E = E_p + E_q$.

C.6 Detection of FSK

The optimum detection of FSK is achieved by applying a matched filter for each of the signal frequencies, see figure. When ω_1 is the frequency of

the received signal, the output on the ω_1-branch would be:

$$\begin{aligned} s_{\omega_1}(\omega_{in} = \omega_1) &= \int_0^{T_s} A_{in} A_{osc} \cos^2(\omega_1 t) dt \\ &= A_{in} A_{osc} \left(\frac{1}{2} T_s + \frac{1}{4\omega_1} \sin(2\omega_1 T_s) \right) \\ &\approx A_{in} A_{osc} \frac{1}{2} T_s \end{aligned} \tag{C.31}$$

Here A_{in} and A_{osc} are the amplitude of the incoming signal and the amplitude on the signal from the local oscillator, respectively. For the ω_2-branch it would be:

$$\begin{aligned} s_{\omega_2}(\omega_{in} = \omega_1) &= \int_0^{T_s} A_{in} A_{osc} \cos(\omega_1 t) \cos(\omega_2 t) dt \\ &= \int_0^{T_s} A_{in} A_{osc} \left(\cos(\omega_1 t + \omega_2 t) + \cos(\omega_1 t - \omega_2 t) \right) dt \\ &= \frac{1}{2} A_{in} A_{osc} \left(\frac{\sin((\omega_1 + \omega_2)t)}{(\omega_1 + \omega_2)} + \frac{\sin((\omega_1 - \omega_2)t)}{(\omega_1 - \omega_2)} \right) \\ &\approx A_{in} A_{osc} \frac{\sin((\omega_1 - \omega_2)T_s)}{2(\omega_1 - \omega_2)} \end{aligned} \tag{C.32}$$

The approximation in the last step is based on the assumption that $(\omega_1 + \omega_2) \gg (\omega_1 - \omega_2)$. It would be preferable if the output on the '0' branch was close to zero when the bit '1' was received. We see that this would be the case if $(\omega_1 - \omega_2)T = k\pi, \ k \in \mathbb{N}$. If we apply this and that $(\omega_1 - \omega_2) = 2\omega_d$ to equation 4.50, we get:

$$m = \frac{\omega_d T_s}{\pi} = \frac{(\omega_1 - \omega_2)T_s}{2\pi} = \frac{k\pi}{2\pi} = \frac{k}{2} \tag{C.33}$$

The same would be valid in the case when ω_2 is the frequency of the incoming signal, but then of course the ω_1-branch would be zero.

Appendix D

Calculation of Electromagnetic Fields from a Loop Antenna

D.1 The fields of a single loop antenna in a moderately dissipative medium

It is slightly more complicated to derive the electromagnetic fields from a single loop antenna in a material medium than for the same antenna in free space. The relation of these fields and their sources is most readily found in terms of potential functions [53]. The loop is assumed to be located in an infinite space of a medium with electromagnetic properties given by the permittivity ϵ, the permeability μ and the conductivity σ. The geometry of the problem is the same as that shown in figure 4.1. Since $\nabla \cdot \mathbf{B} = 0$ under all conditions and a vector field is completely specified by its curl and divergence (Helmholtz's Theorem) the magnetic field in a material medium is still derivable from the magnetic vector potential $\mathbf{A}$:

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{D.1}$$

The electric field $\mathbf{E}$ can be derived from the scalar and vector potentials by:

$$\mathbf{E} = -\nabla\Phi - \frac{\partial \mathbf{A}}{\partial t} \tag{D.2}$$

as is seen from Maxwell's curl equation for the electric field:

$$\begin{aligned} \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} &= -\frac{\partial}{\partial t} \nabla \times \mathbf{A} \\ \nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) &= 0 \\ \mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} &= -\nabla \phi \end{aligned} \tag{D.3}$$

The $\nabla\phi$-term is included because the curl of a gradient field is zero.

If equations D.1 and D.2 are introduced into Maxwell's curl equation for the magnetic field we obtain:

$$\begin{aligned} \nabla \times \left(\frac{1}{\mu} \nabla \times \mathbf{A} \right) &= \mathbf{J} + \frac{\partial \mathbf{A}}{\partial t} \\ -\nabla^2 \mathbf{A} + \nabla(\nabla \cdot \mathbf{A}) &= \mu(\mathbf{J}_s + \sigma \mathbf{E}) + \mu\epsilon \frac{\partial \mathbf{E}}{\partial t} \\ -\nabla^2 \mathbf{A} + \nabla(\nabla \cdot \mathbf{A}) &= \mu \left(\mathbf{J}_s + \sigma \left(-\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi \right) \right) \\ &= +\mu\epsilon \left(-\frac{\partial^2 A}{\partial t} + \nabla \frac{\partial \phi}{\partial t} \right) \end{aligned} \tag{D.4}$$

Here $\mathbf{J}$ denotes current density in this case consisting of the source current density in the loop $\mathbf{J}_s$ and the conduction current density in the surrounding medium $\sigma\mathbf{E}$. If the divergence of $\mathbf{A}$ is defined according to the Lorenz condition [53]:

$$\nabla \cdot \mathbf{A} + \mu\epsilon \frac{\partial \phi}{\partial t} + \mu\sigma\phi = 0, \tag{D.5}$$

equation D.4 can be written:

$$\nabla^2 \mathbf{A} - \mu\epsilon \frac{\partial^2 \mathbf{A}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{A}}{\partial t} = -\mu \mathbf{J}_s \tag{D.6}$$

A solution for $\mathbf{A}$ satisfying D.6 can be found using the Green's function technique. The total $\mathbf{A}$-field is then found by integrating the contributions from infinitesimal current elements along the loop.

The Green's function is the field generated by a point source of unit strength which in our case is:

$$\mathbf{J}_p = \cos(\omega t) \mathbf{i}_\nu (\nu = \phi) \tag{D.7}$$

Since the time differentiation in equation D.6 implies phase shift, complex notation will be used in the following calculations:

$$\mathbf{J'}_p = e^{j\omega t}\mathbf{i}_\nu(\nu = \phi) \tag{D.8}$$

However, only the real part has physical significance.

The Green's function can be specified in a local spherical coordinate system (R, Θ, Φ) with it's origin in a source point. In this coordinate system the source current density can be specified by:

$$\mathbf{J'}_p = e^{j\omega t}\mathbf{i}_\nu(\nu = \phi)\delta(\mathbf{R}) \tag{D.9}$$

where $\delta(\mathbf{R}) = \delta(x)\delta(y)\delta(z)$ is the delta function defined by $\int_{-\infty}^{\infty}\delta(x)dx = 1$. The Green's function of this problem can be written:

$$\mathbf{G}(R, t; \nu) = G(R)e^{j\omega t}\mathbf{i}_\nu \tag{D.10}$$

Inserting D.9 and D.10 into equation D.6 gives:

$$\frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2\frac{\partial G}{\partial R}\right) + \omega^2\mu\epsilon G - j\omega\mu\sigma G = -\delta(\mathbf{R}) \tag{D.11}$$

The solution of this differential equation for $R > 0$ is:

$$G(R) = C_1\frac{e^{j\bar{\gamma}R}}{R} + C_2\frac{e^{-j\bar{\gamma}R}}{R} \tag{D.12}$$

where $\bar{\gamma} = \sqrt{j\omega\mu(\sigma + j\omega\mu)}$ and is called the propagation constant. Only outgoing waves are of interest. Therefore we can set $C_1 = 0$ and the result for $G(R)$ reduces to:

$$G(R) = C_2\frac{e^{-j\bar{\gamma}R}}{R} \tag{D.13}$$

The solution has a singularity in $R = 0$. We seek to determine C_2 so that it makes the solution in D.13 satisfy D.11 also in the limit $R \to 0$. This can be done by substituting G in equation D.11 with the solution in D.13 and integrate over a small sphere with radius $R_0 \to 0$ surrounding the source element:

$$\iiint_V \nabla^2\left(C_2\frac{e^{-j\bar{\gamma}R}}{R}\right)dV + (\omega^2\mu\epsilon - j\omega\mu\sigma)\iiint_V\left(C_2\frac{e^{-j\bar{\gamma}R}}{R}\right)dV = -\iiint_V \delta(R)dV \tag{D.14}$$

Evaluation of the integrals gives:

$$C_2 = \frac{1}{4\pi} \tag{D.15}$$

which makes the solution in D.13 satisfy D.11 also in the limit $R \rightarrow 0$. Therefore the particular solution for $G(R)$ is:

$$G(R) = \frac{1}{4\pi}\frac{e^{-j\bar{\gamma}R}}{R} \tag{D.16}$$

The next step is to link these results to a fixed coordinate system (x, y, z) with origin in the center of the loop. We will also use a spherical coordinate system (r, θ, ϕ) with origin in the center of the loop. It is sufficient to consider an observation point in the xz-plane because the field will be the same for all other points with the same r and θ due to the loop's symmetry and that the current is constant along the loop. The distance from a source point to an observation point P in the xz-plane is denoted R. The radius of the loop is denoted a. We can then express R as $\mathbf{R} = \mathbf{r} - \mathbf{a}$ and it is found to be:

$$\begin{aligned} R &= |\mathbf{r} - \mathbf{a}| \\ &= \sqrt{(\mathbf{r} - \mathbf{a}) \cdot (\mathbf{r} - \mathbf{a})} \\ &= \sqrt{r^2 - 2\mathbf{r} \cdot \mathbf{a} + a^2} \end{aligned}$$

In the (x, y, z)-system $\mathbf{r}$ and $\mathbf{a}$ will be

$$\mathbf{r} = r(\mathbf{i}_x \sin\theta + \mathbf{i}_z \cos\theta)$$
$$\mathbf{a} = a(\mathbf{i}_x \cos\nu + \mathbf{i}_y \sin\nu)$$

The result for R then becomes:

$$R = \sqrt{r^2 - 2ra\sin\theta\cos\nu + a^2} \tag{D.17}$$

and inserted into D.16 gives

$$G(r) = G(r, \theta; \nu) = \frac{1}{4\pi}\frac{e^{-j\gamma\sqrt{r^2 - 2ra\sin\theta\cos\nu + a^2}}}{\sqrt{r^2 - 2ra\sin\theta\cos\nu + a^2}} \tag{D.18}$$

The solution of equation D.6 can now be written:

$$\mathbf{A} = \int_0^{2\pi}\int_{\Delta s} G(r,\theta;\nu)\frac{\mu I_0 e^{j\omega t}ds}{\Delta s}a\mathbf{i}_\nu d\nu \tag{D.19}$$

The source current density can be assumed to be constant across the cross section of the conductor and the diameter of the loop wire can be assumed to be infinitesimally small. Hence the source current density can be approximated by:

$$\mathbf{J}_s = I_0 e^{(j\omega t)}\mathbf{i}_\nu \tag{D.20}$$

and we can rewrite D.19:

$$\mathbf{A} = \mu a I_0 e^{j\omega t}\int_0^{2\pi} G(r,\theta;\nu)\mathbf{i}_\nu d\nu$$

Relating $\mathbf{i}_\nu$ to the fixed unit vectors $\mathbf{i}_x$ and $\mathbf{i}_y$ and substituting G with the solution in equation D.16 we get:

$$\mathbf{A} = \frac{\mu a I_0 e^{j\omega t}}{4\pi}\int_0^{2\pi}\frac{e^{-j\bar{\gamma}R}}{R}(-\mathbf{i}_x\sin\nu + \mathbf{i}_y\cos\nu)d\nu$$

The complex propagation constant $\bar{\gamma}$ can be expressed in terms of its real part $\Re\{\bar{\gamma}\} = \alpha$ called the attenuation constant, and its imaginary part $\Im\{\bar{\gamma}\} = \beta$ the phase constant. Substituting for $\bar{\gamma}$ we get:

$$\mathbf{A} = \frac{\mu a I_0}{4\pi}\int_0^{2\pi}\frac{e^{-\alpha R}e^{j(\omega t-\beta R)}}{R}(-\mathbf{i}_x\sin\nu + \mathbf{i}_y\cos\nu)d\nu$$

Only the real part has physical significance. Hence we can leave out the imaginary part of $e^{j(\omega t-\beta R)}$, leaving the physical significant part:

$$\mathbf{A}_p = \frac{\mu a I_0}{4\pi}\int_0^{2\pi}\frac{e^{-\alpha R}\cos(\omega t-\beta R)}{R}(-\mathbf{i}_x\sin\nu + \mathbf{i}_y\cos\nu)d\nu \tag{D.21}$$

An exact solution is not necessary for a qualitative discussion of the fields associated with the loop antenna. Therefore linear approximations are introduced in the expression for R in equation D.17. The series expansion $\sqrt{1+x} = 1 + \frac{1}{2}x + \ldots$ is used:

$$R \approx r - a\sin\theta\cos\nu \tag{D.22}$$

To further simplify the integration the denominator $\frac{1}{R}$ is approximated using the series expansion $\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \ldots$:

$$\frac{1}{R} \approx \frac{1}{r}(1 + \frac{a}{r}\sin\theta\cos\nu) \tag{D.23}$$

Introducing these approximations into equation D.21 gives:

$$\begin{aligned}\mathbf{A}_p &= \frac{\mu a I_0}{4\pi}\int_0^{2\pi}\left(e^{-\alpha(r-a\sin\theta\cos\nu)}\cos(\omega t - \beta(r - a\sin\theta\cos\nu))\right) \\ &\quad \left(\frac{1}{r}(1+\frac{a}{r}\sin\theta\cos\nu)\right)(-\mathbf{i}_x\sin\nu + \mathbf{i}_y\cos\nu)d\nu\end{aligned}$$

Further approximations are necessary to ease the integration. First we rewrite the exponential attenuation term with the aid of a first order series expansion of $e^x = 1 + x + \ldots$:

$$\begin{aligned}e^{-\alpha(r-a\sin\theta\cos\nu)} &= e^{-\alpha r}e^{\alpha(a\sin\theta\cos\nu)} \\ &\approx e^{-\alpha r}(1 + \alpha a\sin\theta\cos\nu)\end{aligned} \tag{D.24}$$

Second, we want to rewrite the cosine term describing the propagation:

$$\begin{aligned}\cos(\omega t - \beta(r - a\sin\theta\cos\nu)) &= \cos(\omega t)\cos(\beta r - \beta a\sin\theta\cos\nu) \\ &\quad + \sin(\omega t)\sin(\beta r - \beta a\sin\theta\cos\nu)\end{aligned}$$

The terms $\cos(\beta r - \beta a\sin\theta\cos\nu)$ and $\sin(\beta r - \beta a\sin\theta\cos\nu)$ can be rewritten with the aid of a first order series expansions of $\cos(x)$ and $\sin(x)$, that is $\cos(x) = 1$ and $\sin(x) = x$:

$$\cos(\beta r - \beta a\sin\theta\cos\nu) \approx \cos(\beta r) + \sin(\beta r)\beta a\sin\theta\cos\nu \tag{D.25}$$

$$\sin(\beta r - \beta a\sin\theta\cos\nu) \approx \sin(\beta r) + \cos(\beta r)\beta a\sin\theta\cos\nu \tag{D.26}$$

This gives a good approximation provided that $\beta a \ll 1$. $\beta = \frac{2\pi}{\lambda}$ and because we have already assumed that $a \ll \lambda$ (electrically small loop antenna) the condition $\beta a \ll 1$ is fulfilled. With this approximation we can rewrite D.1:

$$\begin{aligned}&\cos(\omega t - \beta(r - a\sin\theta\cos\nu)) \approx \\ &\quad \cos(\omega t - \beta r) - \beta a\sin\theta\cos\nu\sin(\omega t - \beta r)\end{aligned} \tag{D.27}$$

Now, inserting the approximations in D.24 and D.1 into the expression for $\mathbf{A}_p$ gives:

$$\begin{aligned}\mathbf{A}_p &= \frac{\mu a I_0 e^{-\alpha r}}{4\pi}\Bigg[\\ &\int_0^{2\pi}(1+\alpha a\sin\theta\cos\nu)(\cos(\omega t-\beta r)-\beta a\sin\theta\cos\nu\sin(\omega t-\beta r)) \\ &\left(\frac{1}{r}(1+\frac{a}{r}\sin\theta\cos\nu)\right)(-\mathbf{i}_x\sin\nu+\mathbf{i}_y\cos\nu)d\nu\Bigg]\end{aligned}$$

In the following calculations we leave out the terms that result from products of the linear approximations. To include such terms would be inconsistent as higher order terms from the series expansions used includes terms that would be of the same order. The term $-\mathbf{i}_x\sin\nu$ can also be excluded as products of $\sin\nu\cos\nu$ will be zero when integrated from 0 to 2π as the product of an antisymmetric and a symmetric function is antisymmetric. The other terms will include either $\cos\nu$, $\cos^2\nu$ or $\cos^3\nu$. Only the terms of order $\cos^2\nu$ will contribute to the final result as $\int_0^{2\pi}\cos\nu d\nu = 0$ and $\int_0^{2\pi}\cos^3\nu d\nu = 0$. The final expression to be integrated is:

$$\begin{aligned}\mathbf{A}_p &\approx \frac{\mu a I_0 e^{-\alpha r}}{4\pi r}[\\ &\int_0^{2\pi}(\frac{1}{r}+\alpha)a\sin\theta\cos^2\nu\cos(\omega t-\beta r) \\ &-\beta a\sin\theta\cos^2\nu\sin(\omega t-\beta r))d\nu\Big]\,\mathbf{i}_y\end{aligned}$$

$\int_0^{2\pi}\cos^2\nu d\nu = \pi$ and because of the symmetry of the loop the solution for xz-plane is valid also when swept around the angle phi. Hence we can change the unit vector $\mathbf{i}_y$ with the unit vector $\mathbf{i}_\phi$ and the solution is valid for all ϕ. The final expression for A_p is:

$$\begin{aligned}\mathbf{A} &\approx \frac{\mu a^2 I_0 e^{-\alpha r}}{4}\sin\theta \\ &\left[\left(\frac{1}{r^2}+\frac{\alpha}{r}\right)\cos(\omega t-\beta r)-\frac{\beta}{r}\sin(\omega t-\beta r)\right]\mathbf{i}_\phi \qquad \text{(D.28)}\end{aligned}$$

The H-field is now found from:

$$\begin{aligned}
\mathbf{H} &= \frac{1}{\mu}\nabla \times \mathbf{A}_p \\
&= \frac{a^2 I_0}{4} e^{-\alpha r} \\
&\quad \left[2\cos\theta \left[\frac{1}{r^3}\cos(\omega t - \beta r) + \frac{1}{r^2}\alpha\cos(\omega t - \beta r) \right.\right. \\
&\quad \left. -\frac{1}{r^2}\beta\sin(\omega t - \beta r) \right] \mathbf{i}_r \\
&\quad + \sin\theta \left[\frac{1}{r^3}\cos(\omega t - \beta r) + \frac{1}{r^2}\alpha\cos(\omega t - \beta r) \right. \\
&\quad -\frac{1}{r^2}\beta\sin(\omega t - \beta r) + \frac{1}{r}\alpha^2\cos(\omega t - \beta r) \\
&\quad \left.\left. -\frac{1}{r}2\alpha\beta\sin(\omega t - \beta r) - \frac{1}{r}\beta^2\cos(\omega t - \beta r) \right] \mathbf{i}_\theta \right]
\end{aligned}$$

The E-field is found from D.2:

$$\begin{aligned}
\mathbf{E} &= -\frac{\partial \mathbf{A}}{\partial t} \\
&= \frac{\mu a^2 I_0}{4}\omega\sin\theta e^{-\alpha r} \\
&\quad \left[\frac{1}{r^2}\sin(\omega t - \beta r) + \frac{1}{r}\alpha\sin(\omega t - \beta r) + \frac{1}{r}\beta\cos(\omega t - \beta r) \right] \mathbf{i}_\phi
\end{aligned}$$

(The $-\nabla\phi$-term is zero if we assume there are no sources of charge present. The energy flow out from the antenna can be described with aid of the Poynting vector ($\mathbf{P} = \mathbf{E} \times \mathbf{H}$). In this case it will contain a large number of terms. Therefore we choose to look at what terms will be dominant in the reactive field an what terms will dominate in the radiation far field. The near field approximation is:

$$\begin{aligned}
\mathbf{P}_{rct} &= \frac{\mu a^4 I_0^2}{16} e^{-2\alpha r} \\
&\quad \left(\frac{1}{r^5}\sin^2\theta\omega\sin(\omega t - \beta r)\cos(\omega t - \beta r)\mathbf{i}_r \right. \\
&\quad \left. + \frac{1}{r^5}\sin 2\theta\omega\sin(\omega t - \beta r)\cos(\omega t - \beta r)\mathbf{i}_\theta \right)
\end{aligned}$$

The far field approximation is:

$$\begin{aligned}
\mathbf{P}_{far} &= \frac{\mu a^4 I_0^2 \omega}{16} e^{-2\alpha r} \\
&\quad \Bigg(\sin^2\theta \left[\frac{1}{r^2}\alpha^3 \sin(\omega t - \beta r)\cos(\omega t - \beta r) - 2\frac{1}{r^2}\alpha\beta |\sin^2(\omega t - \beta r) \right. \\
&\quad -\frac{1}{r^2}\alpha\beta^2 \sin(\omega t - \beta r)\cos(\omega t - \beta r) + \frac{1}{r^2}\alpha^2\beta \cos^2(\omega t - \beta r) \\
&\quad \left. -2\frac{1}{r^2}\alpha\beta^2 \sin(\omega t - \beta r)\cos(\omega t - \beta r) - \frac{1}{r^2}\beta^3 \cos^2(\omega t - \beta r) \right] \mathbf{i}_r \\
&- \quad \sin 2\theta \left[\frac{1}{r^2}\alpha\beta \sin^2(\omega t - \beta r) \right. \\
&\quad \left. -\frac{1}{r^2}\alpha\beta \sin(\omega t - \beta r)\cos(\omega t - \beta r) \right] \mathbf{i}_\theta \Bigg)
\end{aligned}$$

This completes the field calculations.

Appendix E

List of Symbols

Quantity	Symbol
Antenna Gain	G
Attenuation constant	α
Angular frequency	ω
Bandwidth	W
Bandwidth efficiency	η
Bit Energy	E_b
Capacitance	C
Cartesian coordinate system	(x, y, z)
Coupling constant	k
Conductivity	σ
Current	I
Current density	$\mathbf{J_c}$
Differential length element	$\mathrm{d}\mathbf{l}$
Directivity	D
Frequency	f
Greens function	$\mathbf{G}$
Impedance	Z
Inductance	L
Information rate = bit rate	r_b

Quantity	Symbol
Magnetic field intensity	$\mathbf{H}$
Magnetic flux intensity	$\mathbf{B}$
Magnetic vector potential	$\mathbf{A}$
Modulation rate = symbol rate	r_s
Mutual inductance	M
Magnetic flux	Φ
Noise Density	N_0
Probability	P
Permeability	μ
Permittivity	ϵ
Phase constant	β
Phase of impedance	ϕ
Poynting vector	$\mathbf{P}$
Propagation constant	γ
Reactance	X
Resistance	R
Skin depth	δ
Spherical Coordinate system	(r, θ, ϕ)
Time	t
Voltage	V
Unit normal vector	$\mathbf{i}_n$
Wavelength	λ

Comments and errata to V. S. Bertelsen: *Communication System for a Medical Implant*

J. A. Aas, 16.03.2006

Simple misprints

p. 6, line 6 from bottom:	"meat" should be "meet"
p. 7, last line:	"lithe" should be "little"
p. 22, Fig. 3.1:	"Gluicose" should be "Glucose"
p. 26, line 3 from bottom:	"silisium" should be "silicon"
p. 42, line 5 from bottom:	"*Deltas*" should be "Δs"
p. 42, line 3 from bottom:	"*lambda*" should be "λ"
p. 45, line 7 from bottom:	"expression …indicate" should be "expression ...indicates"
p. 49, lines 3-11:	"Fresnell" should be "Fresnel"
p. 49, line 17:	"undependent" should be "independent"
p. 49, line 1 above (4.14):	"$P_o ut$" should be "P_{out}"
p. 58, Fig. 4.7:	Axis label "abs(Z_2)(Ω)" should be "abs(Z_{in})(Ω)"
p. 59, Fig. 4.8:	"Koupling" should be "Coupling"
p. 59, line 2 from bottom:	"Z_1" should be "Z_{in}"
p. 61, line 3 from bottom:	"in stead" should be "instead"
p. 68, line 1 below (4.25):	"$Z_i n$" should be "Z_{in}"
p. 69, line 9 from bottom:	"This both the" should be "This influences both the"
p. 71, line 3 from bottom:	"the" should be "they"
p. 74, line 4:	"base band" should be "baseband"
p. 99, line 2 from bottom:	"animals … adapts" should be "animals … adapt"
p. 108, line 1 from bottom:	"to complicated" should be "too complicated"
p. 180, line 3 from bottom:	"rewrite D.1" should be "rewrite the cosine term in $\mathbf{A}_p$"
p. 181, line 1:	"in D.24 and D.1 into" should be "in D.24 and D.27 into"

Chapter and table numbering

p. 23, line 7 from bottom:	"in chapter 1" should be "in the preface"
p. 31, line 13:	"table 3.3.1" should be "table 3.1"
p. 31, line 23:	"table 3.3.1" should be "table 3.1"
p. 67, line 7 from bottom:	"table 4.2.2" should be "table 4.1"
p. 94, line 3:	"table 4.4" should be "table 4.3"
p. 113, line 7 from bottom:	"table 5.2" should be "table 5.1"
p. 113, line 3 from bottom:	"table 5.2" should be "table 5.1"
p. 116, line 2:	"table 5.2.1" should be "table 5.2"
p. 119, line 7 from bottom:	"table 5.2.2" should be "table 5.3"

Bad wording

p. 38, lines 4-5: That the movement of a magnet relative to a coil changes the magnetic flux through the coil is obvious and not a result of Faraday's law. That it induces an emf is according to Faraday's law.

p. 45, lines 3-4 from bottom: It is claimed that time averages of all terms except one are zero. This would be more obvious if the two sums $r^{-4}\beta - r^{-4}2\beta\cos^2(\omega t-\beta r)$ were replaced by $-r^{-4}\beta\cos(2(\omega t-\beta r))$. Each term in the sums has a time average different from zero.

p. 49, line 10 from bottom: The radiation term, here written $r^{-2}\,\beta^3\cos^2(\omega t-\beta r)$, should contain the important factor $\sin^2\theta$.

pp. 51-52, Figs 4.3-4.4: The sources of these two figures are not referenced. Fig. 4.3 is not even mentioned in the text.

p. 98, lines 13-14 from bottom: Differences between "Athermal effects" and "Non-thermal effects" should be clarified.

p. 125 Chapter 6: The chapter heading "Transmission Predictions" is not descriptive for its content. It contains only mathematical details for a numerical evaluation of Neumann's formula and no results.

pp. 151-157: A great number of references are taken from various web pages. These pages change with time so it is hard or impossible to check the information at a later time. Moreover, the order in which the references appear in the text seems arbitrary. Books, journals, and conferences numbered as they appear in the text are preferred references.

Mathematical errors and "misprints"

p. 43, line 1: The equation $\mathbf{J}' = I_0\cos\omega t$ is meaningless. Left hand side is a vector that should be infinite large on the loop while the right hand side is a scalar. $\mathbf{J}' = \mathbf{i}_\phi\,\delta(z)\,\delta(r\text{-}a)$ would be more appropriate.

p. 43, line 5: The equation $R/v_p = \beta R$ is meaningless. Left hand side is a time delay [s] while the right hand side is a phase [rad]. β is ω/v_p and not $1/v_p$.

p. 44, eq. (4.5): A factor $\omega\sin\theta$ is missing in the last expression for **E**. (This factor is present at page 182).

p. 45, eq. (4.6): The third term in the $\mathbf{i}_r$ direction should not contain the factor 2 in front of β^2.

p. 46, eq. (4.9): $-d\Phi/dt$ should be $-d\Phi_{12}/dt$

p. 48: The mutual inductances M_{12} and M_{21} are treated separately. The only stated relation between them is that $M^2 = M_{12}M_{21}$, saying that M is the geometric mean value. It aught to be stated that $M_{12} = M_{21}$ as can be readily seen from (4.10). Thus it is no need to distinguish between the two mutual inductances. The quantity R in eq. (4.10) should be defined.

p. 49, eq. (4.14): $(\lambda/4\pi d)$ should be $(\lambda/4\pi d)^2$. The dB value calculated in eq. (4.57) is correct, however, because 10 log x^2 = 20 log x.

p. 57, eqs. (4.18)-(4.19): There is a sign error in these two equations. They should be
$Z_2 = -1/(j\omega_r C_2 - \omega_r^2 C_2^2 R_2)$
$Z_{in} = R_1 - M^2\omega_r^2(j\omega_r C_2 - \omega_r^2 C_2^2 R_2)$
Will this also affect eqs. (4.25)-(4.26)?

p. 58, Fig. 4.7: Due to the above sign error the resonance peak of the left figure should be 1100 Ω instead of 1763 Ω.

p. 63, Fig. 4.11 There seems to be inconsistent notation in eqs. (4.21)-(4.22) versus Fig. 4.11, see C_{C1} and CL1, R_{C1} and RL1, R_{C2} and RL2, C_{C2} and CL2. C_2 and R_2 on lines 4-5 below eq. (4.22) should also be changed accordingly.

p. 86, eq. (4.49): $e^{j-\omega_d t}$ should be $e^{-j\omega_d t}$
p. 159, eq. (A.1): dS_2 should be $d\mathbf{S}_2$ (a vector to complete the dot product)
p. 159, eq. (A.3): dS should be replaced by $\cdot d\mathbf{S}$ (dot product again)
p. 159, eq. (A.4): dS should be replaced by $\cdot d\mathbf{S}$ (dot product again)
p. 159, eq. (A.5): $\mathbf{A}_1$ should be replaced by $d\mathbf{A}_1$ (differential contribution from $d\mathbf{l}_1$)
p. 160, eq. (A.6): $d\mathbf{l}_1 d\mathbf{l}_2$ should be $d\mathbf{I}_1 \cdot d\mathbf{I}_2$ (as in (A.7))
p. 163, eq. (B.8): The beginning of this equation should be
$E_p(a,t) = \Re\{E_0(r=a)e^{j\omega t}\} = \Re\{\cdots\} = \cdots$
(To take the real part of $(V_0/l)\cos(\omega t)$ times $\exp(j\omega t)$ is nonsense.)

p. 163: The second equation below (B.8) should read
$A_{0r}J_{0i}(\zeta(a)) + A_{0i}J_{0r}(\zeta(a)) = 0$ (sign error)
p. 163, eq. (B.11): This relation (eq. (B.11)?) should read $\zeta_a = a\sqrt{\omega\sigma\mu}$ (a is missing)
p. 163, eq. (B.12): The second equal sign (=) should be removed

p. 167, eq. (C.5): The derivation of (C.5) should be

$$S(\omega) = \int_{-\infty}^{\infty} \left(\mathrm{Re}\left\{ b(t) e^{j\omega_c t} \right\} \right) e^{-j\omega t} dt$$

$$= \int_{-\infty}^{\infty} \frac{1}{2} \left(b(t) e^{j\omega_c t} + b^*(t) e^{-j\omega_c t} \right) e^{-j\omega t} dt$$

$$= \frac{1}{2} \left(\int_{-\infty}^{\infty} b(t) e^{-j(\omega-\omega_c)t} dt + \left(\int_{-\infty}^{\infty} b(t) e^{-j(-\omega-\omega_c)t} dt \right)^* \right)$$

$$= \frac{1}{2} \left(B(\omega - \omega_c) + B^*(-\omega - \omega_c) \right)$$

p. 176, eq. (D.4): $\frac{\partial \mathbf{A}}{\partial t}$ in the first line should be replaced by $\frac{\partial \mathbf{D}}{\partial t}$.

p. 176, eq. (D.4): The last equal sign should be removed.

p. 176, eq. (D.4): The term $\mu\varepsilon \left(-\frac{\partial^2 A}{\partial t} + \nabla \frac{\partial \phi}{\partial t} \right)$ should be replaced by

$$\mu\varepsilon \left(-\frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla \frac{\partial \phi}{\partial t} \right)$$

p. 176, eq. (D7): There is no definition of the symbols $\mathbf{i}_v$, v and ϕ. They are not shown in Figure 4.1 either. It is also unfortunate that the same symbol ϕ is used both as an angle and as a scalar potential at the same page.

p. 177, line 1 below (D.12): The equation for the propagation constant should read $\bar{\gamma} = \sqrt{j\omega\mu(\sigma + j\omega\varepsilon)}$

p.178, line 8 below (D16): "express R as **R** = **r**–**a**" should be "express R as |**R**| = |**r**–**a**|"

p.178, eqs. (D.17)-(D.19): The angle v used in these equations seems to be the same angle as α in Fig. 4.1, but this should be stated clearly.

p. 180, eq. (D.26): The right hand side of this equation should read $\sin(\beta r) - \cos(\beta r)\beta a \sin\theta \cos v$ (sign error)

p. 182, last eq.: The component of $\mathbf{P}_{\mathrm{rct}}$ in the $\mathbf{i}_r$ direction should have a negative sign because it is $-E_\phi H_\theta$.

p. 183: The component of $\mathbf{P}_{\mathrm{far}}$ in the $\mathbf{i}_r$ direction should have a negative sign because it is $-E_\phi H_\theta$. (The sign is correct in eq. (4.6).) Moreover, the second term should contain $\alpha^2\beta$ instead of $\alpha\beta$.

p. 183: The component of $\mathbf{P}_{\mathrm{far}}$ in the $\mathbf{i}_\theta$ direction should be zero because H_r does not contain a component with $1/r$ dependence. (Eq. (4.6) shows that the theta component of **P** only contains a term with $1/r^3$ dependence.)