WHITE PAPER SERIES / EDITION 1

AUTO–ID LABS

BUSINESS PROCESSES & APPLICATIONS
SOFTWARE & NETWORK
HARDWARE

AUTOIDLABS-WP-HARDWARE-013

Physics and Protocols in Radio Frequency Identification

Peter H. Cole

Professor of RFID Systems and Director
of the Auto-ID Research laboratory
The University of Adelaide

www.autoidlabs.org

1. Introduction

Radiofrequency identification systems, illustrated briefly in Figure 1, are evolving rapidly as a result of: (a) Increased awareness of the technology; (b) development of improved techniques for multiple tag reading; (c) realisation in the business community of the benefits of widespread adoption in the supply chain; (d) adoption by designers of sensible concepts in the arrangement of data between labels and databases; (e) development of efficient data-handling methodologies in the relevant supporting communication networks; (f) appreciation of the need for cost reduction, and (g) development of new manufacturing techniques that will achieve manufacture of billions of labels at acceptable costs.

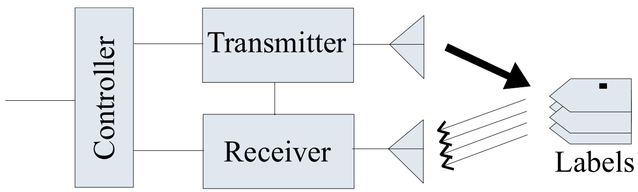


Fig. 1: An RFID system

These developments should be seen as integrated and mutually supporting. However, most RFID systems that exploit these developments are desired by the users to be *passive*. i.e. to exploit microcircuits which can both power themselves and provide power for a reply signal from the small amount of energy that they can extract from an interrogation signal.

In this effort they are subject to what must be seen as laws of physics that are not amenable to *manipulation*, but are amenable to *exploitation* when correctly understood. The importance of correct understanding is underscored by the fact that passive RFID labelling systems, in contrast to electromagnetic broadcasting or communication systems that are powered at both ends of the communication link, have *no margin for inefficiency*: if the electromagnetic propagation aspects are not understood or not optimised, the systems will just not work to acceptable performance levels.

This paper is designed to provide support to RFID system designers in what is probably the least understood aspect of system design, that of the *electromagnetic propagation to and from the label*.

2. Objectives

We pursue in this paper the following objectives.

- To outline the fruitful concepts and fundamental theory behind the development of successful RFID systems for use in the supply chain.
- To provide an analysis of the electromagnetic coupling between interrogators and labels in radio frequency identification systems that will be useful across frequency ranges from the LF to UHF, and cover both the large and small antennas operating in the near field and far field, and exploiting coupling to electric fields, magnetic fields, or electromagnetic fields.
- To encourage ways of thinking that the author believes are important in understanding and designing electromagnetic coupling systems for RFID labels.
- To assemble in one place the principal equations on which sound design may be based.

3. Outline

We will begin with a formal statement of the laws of electrodynamics, and use them to derive boundary conditions to be noted where fields interact with materials. Such relations are amplified by introduction of useful concepts of demagnetising and depolarising factors. The retarded potential solutions of Maxwell's equations and their utility in developing near and far radiated fields are developed, especially for infinitesimal electric and magnetic dipoles. The Auto-ID Center concepts behind reading of large populations of labels simultaneously present in the field of an RFID interrogator are outlined. The influence of electromagnetic compatibility constraints on the choice of operating frequency and antenna style is explored. A number of field creation structures or near and far fields are illustrated, and label antenna forms and parameters for use in various object geometries are illustrated. Coupling volume theory for electric field and magnetic field sensitive antennas is developed at length and the relation to far field antenna theory is established. Some conclusions about antenna optimisation are drawn. Appendices covering a simple introduction to electromagnetic radiation, an analysis of coupling to particular structures, generally useful formulae, and simple exercises, are attached.

4. Notation

The notation and nomenclature used in this paper for physical quantities will be as defined in ISO 1000, [1]. Sinusoidally varying quantities will be represented by peak (not r.m.s.) value phasors. Lower case variables will be used for instantaneous values of scalars, and bold calligraphic characters for instantaneous values of field vectors. Upper case variables will be used for phasors representing sinusoidally varying quantities, and bold upright Roman characters will be used for phasors representing sinusoidally varying field vectors. There are some traditional exceptions to these rules where Greek and Roman upper case symbols do not differ sufficiently.

We will reduce the results of our analysis to dimensionless ratios where possible, and try to employ concepts that can be seen as having physical meaning.

5. Electromagnetic theory

5.1 The complete laws

The complete laws of electrodynamics were first assembled correctly by Maxwell in the form enunciated in words below.

Faraday's law

The circulation of the electric field vector **E** around a closed contour is equal to minus the time rate of change of magnetic flux through a surface bounded by that contour, the positive direction of the surface being related to the positive direction of the contour by the right hand rule.

Ampere's law as modified by Maxwell

The circulation of the magnetic field vector **H** around a closed contour is equal to the sum of the conduction current and the displacement current passing through a surface bounded by that contour, with again the right hand rule relating the senses of the contour and the surface.

Gauss' law for the electric flux

The total electric flux (defined in terms of the **D** vector) emerging from a closed surface is equal to the total conduction charge contained within the volume bounded by that surface.

Gauss' Law for the magnetic flux

The total magnetic flux (defined in terms of the **B** vector) emerging from any closed surface is zero.

With the aid of Gauss' and Stokes' laws of mathematics and the definitions

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \quad \text{and} \quad \mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M})$$

these laws may be expressed, when the fields are spatially continuous, in the differential form

$$\begin{aligned} \nabla \times \mathrm{E} &= -\frac{\partial \mathrm{B}}{\partial t} \\ \nabla \times \mathrm{H} &= \mathrm{J} + \frac{\partial \mathrm{D}}{\partial t} \\ \nabla \cdot \mathrm{D} &= \rho \\ \nabla \cdot \mathrm{B} &= 0 \end{aligned}$$

5.2 Source and vortex interpretation

We remain firmly committed to the source and vortex interpretation of those equations, [2]. In that interpretation, the above equations state that the electric field vector **E** can have vortices caused by changing magnetic flux, the magnetic field **H** can have vortices caused by conduction or displacement currents; the electric flux density **D** can have sources caused by conduction charge density; and the magnetic flux density vector **B** can have no sources.

In linear media, some of the statements about **D** and **B** can be extended to **E** and **H** as well, but when non-uniform fields and boundaries are considered, it can be shown that **E**, **D**, and **H** can have both sources and vortices, but **B** is alone in that it can have no sources.

Figures 2 and 3 below provide archetypical illustrations of the source nature of the electrostatic field and the vortex nature of a magnetodynamic field, as well as illustrations of two of the most important boundary conditions which apply when any electric field **E** or a magnetodynamic field **H** approaches a conducting surface.

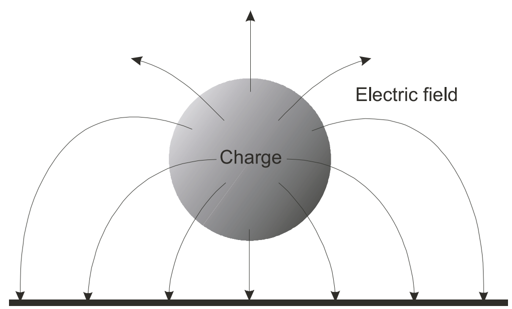


Fig. 2: Electric field near a conducting surface

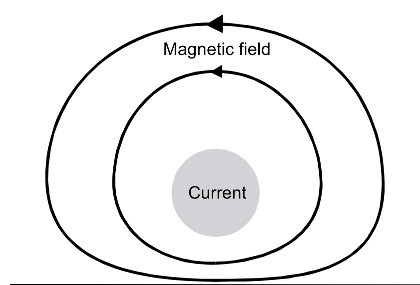


Fig. 3: Oscillating magnetic field near a conducting surface

5.3 Boundary conditions

We have already given the simplified illustrations for the most important cases of sinusoidal electric and magnetic fields occurring adjacent to conductors.

A full statement of the electromagnetic boundary conditions is provided in the two paragraphs below. All of those results may be derived from the statement in words of the basic laws given earlier.

What we can deduce from those laws, without taking to account the properties on any materials involved, is that the tangential component of **E** is continuous across any boundary; the normal component of **B** is continuous across such a boundary; the normal component of **D** maybe discontinuous across a boundary, with a discontinuity being equal to any conduction charge density p_v^c per unit area on the surface; and the tangential component of **H** may be discontinuous across a boundary, with the discontinuity being equal to in magnitude and at right angles in direction to a surface current density flowing on the surface.

When we take into account the restrictions imposed by the properties of the materials which may exist on one or other side of the boundary, we may further conclude that: the electric field is continuous across the boundary for all materials and time variations; that there are no electric fields or fluxes, or time-varying magnetic field or flux densities inside a good conductor; that a surface current density can exist only on the surface of a perfect conductor; and that time-varying charge density cannot exist on the surface of a perfect insulator, although in a feline world, a static surface charge density can.

5.4 Demagnetising and depolarising factors

The theory of demagnetising and depolarising factors, [3] is useful in gaining an appreciation of how the interior fields of an object differ from the fields exterior to that object, and for calculating useful properties of magnetic cores. When an ellipsoidal shape as illustrated in Figure 4 below of dielectric or soft magnetic material is introduced into a region in which there was previously a spatially uniform electric field **E** or magnetic field **H** caused by some distribution of sources (charges or currents), and these sources do not vary as a result, the material becomes polarized with a polarisation vector **P** or magnetised with a magnetisation vector **M**. That polarisation or magnetisation causes on the surface of the shape, induced surface charge densities or magnetic pole densities which make an additional contribution $\mathbf{E}^d$ or $\mathbf{H}^d$ to the fields interior to the shape. These fields are in a direction opposite to the original applied fields $\mathbf{E}^a$ or $\mathbf{H}^a$ and tend to depolarise or demagnetise the material, and hence reduce (but not reverse in direction) the polarisation or magnetisation in the interior of the shape. The internal fields E and H are also reduced.

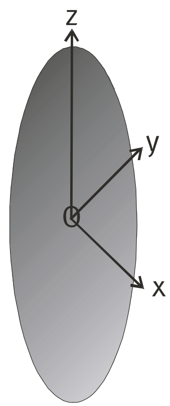


Fig. 4: An ellipsoid of dielectric or magnetic material

The depolarising or demagnetising fields are given by:

$$\begin{bmatrix} \mathrm{E}_x^d \\ \mathrm{E}_y^d \\ \mathrm{E}_z^d \end{bmatrix} = -\frac{1}{\varepsilon_0}\begin{bmatrix} N_x & 0 & 0 \\ 0 & N_y & 0 \\ 0 & 0 & N_z \end{bmatrix}\begin{bmatrix} \mathrm{P}_x \\ \mathrm{P}_y \\ \mathrm{P}_x \end{bmatrix}$$

$$\begin{bmatrix} \mathrm{H}_x^d \\ \mathrm{H}_y^d \\ \mathrm{H}_z^d \end{bmatrix} = -\begin{bmatrix} N_x & 0 & 0 \\ 0 & N_y & 0 \\ 0 & 0 & N_z \end{bmatrix}\begin{bmatrix} \mathrm{M}_x \\ \mathrm{M}_y \\ \mathrm{M}_x \end{bmatrix}$$

where N_x, N_y or N_z are known as depolarising or demagnetising factors. These dimensionless constants depend upon the shape of the ellipsoid, and are subject to the condition

$$N_x + N_y + N_z = 1$$

This constraint, and relations between N_x, N_y or N_z for certain symmetrical shapes, allow conclusions to be drawn for the cases of the sphere, a long thin rod, or a flat thin desk.

5.5 Retarded potentials

The retarded potentials listed below may be regarded as integral solutions for Maxwell's equations which are available when charge and current distribution are known.

For the calculation of electric and magnetic fields at a point r_2 caused by a distribution of charge and current at points *r1* over a volume *v* we may make use of the retarded potentials

$$\Phi(r_2) = \frac{1}{4\pi\varepsilon_0}\int_v \frac{\rho(r_1)e^{-j\beta r_{12}}}{r_{12}}dv$$

and

$$\mathbf{A}(r_2) = \frac{\mu_0}{4\pi}\int_v \frac{\mathbf{J}(r_1)e^{-j\beta r_{12}}}{r_{12}}dv$$

The fields in the sinusoidal steady-state can be derived from these potentials by the equations

$$\mathbf{E} = -\mathbf{grad}\Phi - j\omega\mathbf{A}$$
$$\mathbf{B} = \mathbf{curl}\mathbf{A}$$

These formulae may be used in the calculation of the electromagnetic fields generated by oscillating infinitesimal electric or magnetic dipoles as reported below.

5.6 Reciprocity

The integral form of the Lorenz reciprocity relation for two solutions $\mathbf{E}_1$, $\mathbf{H}_1$ and $\mathbf{E}_2$, $\mathbf{H}_2$ of Maxwell's equations in the same region and at the same angular frequency ω is under appropriate conditions

$$\int_S (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1)\cdot \mathbf{ds} = \int_v (\mathbf{J}_1 \cdot \mathbf{E}_2 - \mathbf{J}_2 \cdot \mathbf{E}_1)\cdot dv$$

where the integral is over a closed surface S bounding a volume *v*. In the derivation, which proceeds for Maxwell's equations, it is allowed that the material parameters be characterized by complex (to allow for losses) and possibly tensor (to allow for anisotropy) dielectric permittivities and magnetic permeabilities, but gyromagnetic behaviour of magnetic materials is not permitted.

In certain cases the right-hand side becomes zero. These cases include those where the surface is a conducting surface, where the surface encloses all sources, where the surface encloses no currents, and where currents flow only by the mechanism of drift, (but not by diffusion, as they can in a semiconductor).

The theorem has in electromagnetic theory, and on the simplification of electromagnetic theory known as lumped circuit theory, some profound consequences. These include: the symmetry of impedance, admittance and suitably defined scattering matrices; the equivalence of transmitting and receiving antennas; the gain of a lossless transmitting antenna being related to the effective area of the same antenna used as a receiver; and the propagation loss from an interrogator to label being equal to propagation loss label to receiver when antennas have single ports. Use of these results will be made in data sections dealing with coupling between interrogators and labels.

6. Infinitesimal dipole fields

6.1 Electric dipole

In spherical polar coordinates at a point $P(\rho,\theta,\phi)$ the non-zero field components of an oscillating small electric dipole of length L carrying a current I and of moment P where $j\omega P = IL$ are

$$E_r = \frac{\beta^2 j\omega P\eta}{4\pi}\left(\frac{2}{(\beta r)^2} - \frac{2j}{(\beta r)^3}\right)e^{-j\beta r}\cos\theta$$

$$E_\theta = \frac{\beta^2 j\omega P\eta}{4\pi}\left(\frac{j}{(\beta r)} + \frac{1}{(\beta r)^2} - \frac{j}{(\beta r)^3}\right)e^{-j\beta r}\sin\theta$$

$$H_\phi = \frac{\beta^2 j\omega P}{4\pi}\left(\frac{j}{(\beta r)} + \frac{1}{(\beta r)^2}\right)e^{-j\beta r}\sin\theta$$

6.2 Magnetic Dipole

In spherical polar coordinates at a point $P(\rho,\theta,\phi)$ the non-zero field components of an oscillating small magnetic dipole of moment $M = IA$ are

$$H_r = \frac{\beta^2 j\omega\mu_0 M}{4\pi\eta}\left(\frac{2}{(\beta r)^2} - \frac{2j}{(\beta r)^3}\right)e^{-j\beta r}\cos\theta$$

$$H_\theta = \frac{\beta^2 j\omega\mu_0 M}{4\pi\eta}\left(\frac{j}{(\beta r)} + \frac{1}{(\beta r)^2} - \frac{j}{(\beta r)^3}\right)e^{-j\beta r}\sin\theta$$

$$E_\phi = \frac{\beta^2 j\omega\mu_0 M}{4\pi}\left(\frac{j}{(\beta r)} + \frac{1}{(\beta r)^2}\right)e^{-j\beta r}\sin\theta$$

6.3 Characteristics of near and far fields

The above expressions show that the distance $r = 1/\beta = \lambda/(2\pi)$ is of significance in determining the nature of the fields surrounding the dipoles. Within this distance the dominant fields may be recognised as being the same as the energy storage fields with which we are familiar for electrostatic or magnetostatic dipoles. This region is known as the near-field region, and the dominant fields therein are called the near fields, and simply store energy that periodically emerges from, and later disappears back into, the dipole. Outside of this distance, the dominant fields are those associated with energy propagation by electromagnetic waves away from the sources. This region is known as the far-field region, and the dominant fields therein are called the far fields, and transport energy continuously away from the dipoles.

7. EPC concepts

7.1 The Auto-ID Center

The Auto-ID Centre has introduced the concepts of identifying all objects in the supply chain, and many objects elsewhere, through an Electronic Product Code (EPC), kept on what they term as a Class 1 read-only RFID label, with associated data, coded in the Physical Markup Language (PML), being kept in a data base which is accessed via a Savant network using the Object Name Service (ONS).

For a full treatment of these concepts please refer to papers on the Auto-ID Center web site.

As illustrated below, there are four fields in the Electronic Product Codes, which are, in order: a *header*, defining the variety of EPC among a number of possible structures; a *domain manager number* which is effectively a manufacturer number; an *object class* which is equivalent to a product number; and a *serial number*.

EPC TYPE	HEADER SIZE	FIRST BITS	DOMAIN MANAGER	OBJECT CLASS	SERIAL NUMBER	TOTAL
256 bit	8					
96 bit	8	00	28	24	36	96
64 bit type I	2	01	21	17	24	64
64 bit type II	2	10	15	13	34	64
64 bit type III	2	11	26	13	23	64

7.2 EPGglobal

The above varieties of EPC have since been revised by EPCglobal to include inter alia a serialized GTIN. For up to date information on such codes, the EPCglobal web site should be consulted.

7.3 Other tag data

In addition to the EPC code, a Class 1 label carries a CRC for transmission verification, and a *password* that is used in the process of *label destruction*. As a guarantee of customer privacy, labels are always *electronically destroyed* before goods are released to the purchaser at their final point of sale.

8. Multiple label reading

In the modern supply chain with pervasive labelling of objects, the reading of the data from a number, sometimes are very large number, of labels simultaneously present in the field of the interrogator, is required.

This task is performed by the Auto-ID Centre through the definition of a number of *label classes*, of which more will be said later, and a range of suitable reading processes. These processes necessarily vary with the operating frequency as a result of variation of electromagnetic compatibility regulations within different parts of the electromagnetic spectrum. All of the classes are based hierarchically on the EPC concept.

The EPC data and its associated CRC, and sometimes password data, may be viewed as a serial data stream to be delivered or collected in a single data transmission, or viewed as levels in a tree-like data structure to be explored over a number of two-way data transmissions.

This dichotomy of views is useful in describing the two methods, namely *tree exploration* and a*daptive round data collection*, employed by the Auto-ID Centre in the reading of dense populations of electronic labels. These processes are discussed in turn below.

8.1 Tree exploration

An illustration of an EPC and its associated data in terms of a *binary tree* (hanging downward from a root in the sky, as such trees are wont to do), is provided in Figure 5. We are at this stage being deliberately vague about how the data is arranged, but it is clear that: (i) in any real application, the tree is huge in relation to the one illustrated, (ii) different length EPC codes lead to paths of different lengths, so the bottom of the tree is not of a uniform level, and (iii) for any population of labels in the field of view of an interrogator, only a small fraction of the tree is *populated*. A constant feature of the trees is that each *level* of the tree corresponds to a particular *position* in the data stream defined by: the EPC, the CRC and the password bits, taken in some order. In some arrangements, it is desirable to place the CRC adjacent to the root, and follow that by the EPC, header end first.

Some additional tree concepts will now be defined.

The *maximal tree* is the tree formed by the complete set of electronic product codes which has been defined for potential use. In the diagram of Figure 5, the maximal tree contains all the branches and nodes shown. The diagram has the benefit of simplicity, but is not representative of the maximal tree in two respects. Firstly, only five levels are shown, and secondly all products appear to be at the same level. Since 64, 96 and 256 bit EPCs have been defined, there are in fact product nodes after descents from the root to each of these levels. The reader of these notes is invited to form a mental picture of the true maximal tree.

The *global tree* is the tree formed by all of the electronic product codes which have been placed into service. It will be smaller than the maximal tree. Again the diagram is not representative of the global tree in the same respects as were discussed for the maximal tree.

The *local tree* is the tree formed by the electronic product codes of all the labels which are in the field of the interrogator at a particular time. It will be smaller than the global tree.

In the local tree we regard a node as *populated* if there are branches descending from it to a product or it is a bottom node corresponding to a product, otherwise we describe the node as *unpopulated*. We can classify a populated node as *singly populated* if only a single path leading to a product descends from it, or *multiply populated* if more than one path leading to a product descends from it. We have in the tree shown unpopulated nodes with a white interior and populated nodes with a shaded interior.

In many circumstances we place the root of the tree immediately above the most significant bit of the EPC version number. When this is done, we describe the resulting tree as a *version rooted tree*.

In some circumstances it is convenient to extend the EPC code leftwards by 16 bits to contain the CRC. In those circumstances, when we wish to illustrate the memory contents in terms of the tree, we place the root of the tree immediately above the most significant bits of the CRC. When this is done, we describe the resulting tree as a *CRC rooted tree*.

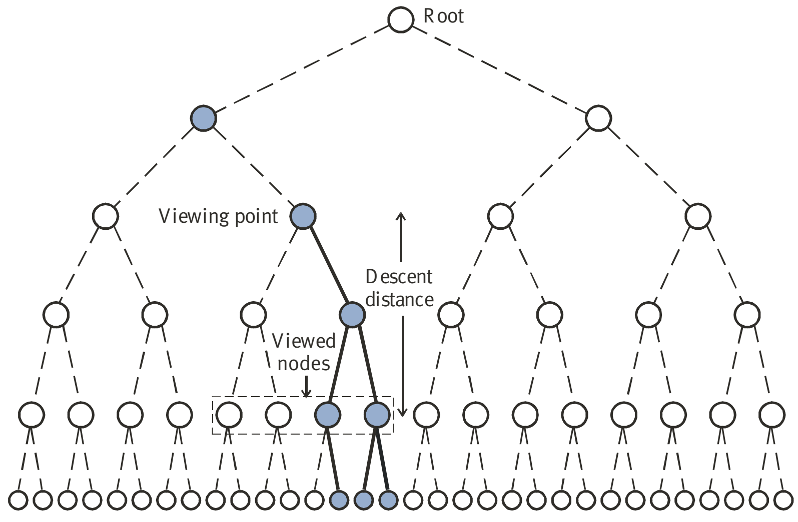


Fig. 5: Illustration of tree interpretation of EPC codes.

Next we define, in the context of tree scanning, the concepts of *viewing point*, *descent distance* and *viewed nodes*. These concepts are illustrated in Figure 5.

A *viewing point* is any node of the tree (or the root) from which we intend to take a view of a set of nodes descended from the viewing point by a number of levels equal to the *descent distance*. The set of nodes is defined as the *viewed nodes* associated with the viewing point. The descent distance may be any positive integer, but is defined in the context. In the example shown, the descent distance is 2. The definition and illustration of the viewed nodes is not likely to be disturbed by taking into account the variability of the bottom levels of realistic trees, as the separation of such trees into subtrees of different heights occurs well before we approach the bottom of any of them.

The *tree exploration* methods adopted by the Auto-ID Centre have the following features.

Initially they optionally perform a *label subset selection* through the definition of a particular path through a part of the tree. The label subset selection may correspond to a particular product selection, or may just be a mechanism to reduce, for a while, the number of labels responding. As illustrated in Figure 5, the label selection defines a *viewing point* within the tree, while the *internal structure* of the labels defines a *descent distance*, and hence the set of *viewed nodes* illustrated. All of these concepts are further explained below.

All labels whose position in the tree descend from the viewing point will respond to a suitable interrogator signal, after the completion of that signal. The responses are *separated in the time domain* according to the positions of the labels in relation to the *viewed nodes* illustrated in Figure 5. In Auto-ID Centre protocols, descent distances of 1 and 3 have been defined.

The responses provide information as to which part of the tree is unpopulated. Based on this information, the interrogator continues to explore specific paths lower into the tree until it is clear, either through the examination of the return signals from the labels or from the statistics of the various tree populations described earlier, that it is *likely that only a single label is responding*. At this stage the interrogator calls for a *complete response* from the label, so that the so far only partially defined, but hopefully unique in the local population, label becomes *fully identified*. That unique label is then *put to sleep* by means of an interrogator command, and the tree is explored further.

The power of this process is that the global tree does not have to be exhaustively searched. It is possible early on to gamble on having singulated a label, and to acquire its complete identity before the tree has been fully descended. Little is lost if the belief that a unique label had been singulated proves occasionally to be incorrect; the tree need simply be descended further to completely resolve such conflicts.

In the process of tree exploration, there is frequent interrogator signalling over a reasonably wide radio frequency band. While such signalling is possible within electromagnetic compatibility regulations at UHF, the regulations at HF preclude it, and the different approach described below is appropriate.

8.2 Adaptive round data collection protocol

The basic operating principle of this approach is what is known as the Slotted Terminating Adaptive Collection (STAC) protocol defined below. An illustration of the Slotted Terminating Adaptive Collection protocol is provided in Figure 6.

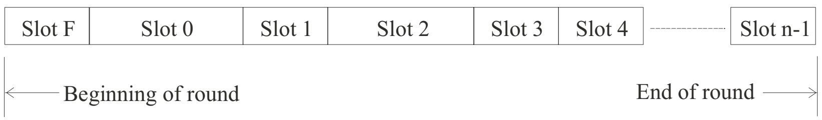


Fig. 6: Illustration of slots in a reply round.

In the Slotted Terminating Adaptive Collection protocol, labels reply with randomly selected positions or time intervals referred to as *slots* that have their beginning and end under interrogator control. An interrogator command signals both the end of the current and the beginning of the next slot. A number of slots form a *reply round*, and selected labels waiting to reply do so *once during each reply round*. The concepts are illustrated in Figure 6.

Labels which enter the energising or signalling field, when they have sufficient power for operation, wait before replying in a READY state for the reception of one of several commands. Such commands can be a *Write* command, by means of which a label may be programmed with its EPC, a *Destroy* command by means of which a label may be permanently disabled, or a *Begin Round* command, by means of which an appropriate subset of the labels may offer their data for collection.

The *Begin Round* command contains several parameters. Some of these define the *number of slots* in the forthcoming round. Others of these define a *selection* from among the number of ready labels that will participate in the round. Another parameter is a *hash value* that may regulate the way in which label memory contents are used to generate pseudo random reply positions within a round. It is intended that within each round, and between different rounds, label replies will occupy what are effectively random positions, so collisions between replies are neither frequent nor repeated.

As no label replies without first receiving a *Begin Round* command the protocol is consistent with what is known in the electronic labelling industry as the *Reader Talks First* (RTF) operating mode.

In the STAC protocol the issuing by the interrogator of a *Begin Round* command causes the definition, within each of the selected labels, of a *round size parameter*, *a selected or not selected flag*, *a proposed reply slot position*, and sets to zero a *counter of reply slot positions*. Such reply slot positions are arranged as shown in Figure 6, about which a number of points are made below.
In relation to Figure 6, we should note that:

- The figure is not to scale, but does correctly indicate that slots are not of equal sizes.
- There is a special slot (slot F) at the beginning of a round. This slot is followed by n further slots to complete the round.
- The number n of further slots is a power of two and is regulated by the interrogator.
- Slot F is of a special and fixed size, but the duration of other slots is regulated by interrogator signalling.
- The longer slots are of duration sufficient for a label reply.
- The shorter slots, (other that slot F) are short because they contained no reply and were closed early by the interrogator.
- Slot F comes to an end automatically without an interrogator signal.
- All slots make an allowance for necessary interrogator signalling.

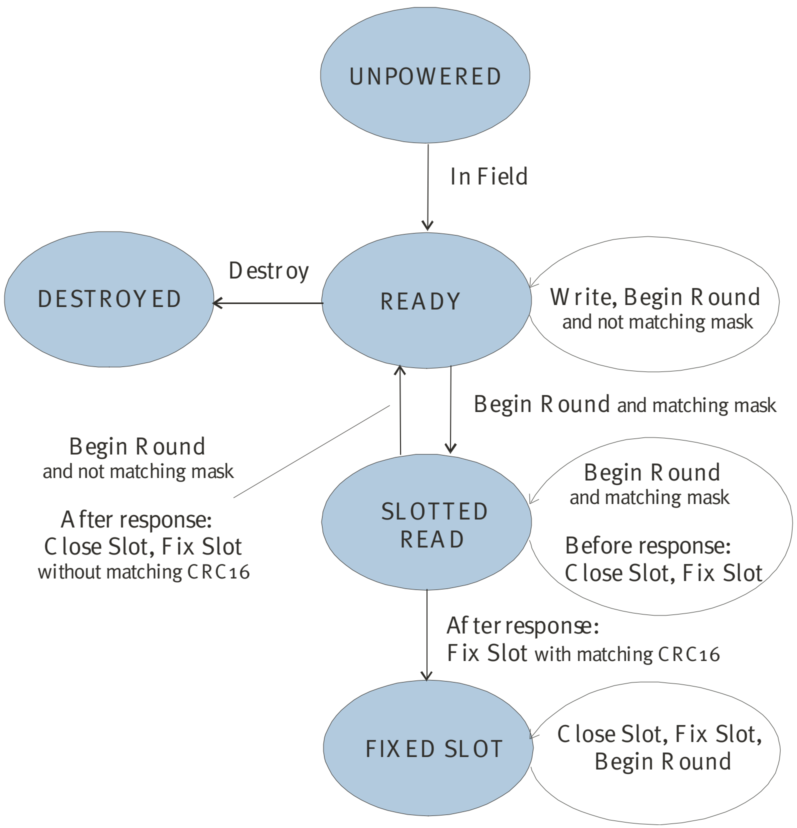
The issuing of a *Begin Round* command also causes a subset of the labels which were waiting in the READY state to enter the SLOTTED READ state. In this state the labels calculate for themselves a proposed reply slot and wait until their slot counter, which will advance each time the interrogator indicates the end of a slot and the beginning of a new one, reaches the *proposed reply slot position*, whereupon the label will reply during the slot.

If any labels in the SLOTTED READ state have at the conclusion of Slot F a proposed reply slot position of zero they will reply in slot position 0. The reply conditions within that slot can then be separated into three categories: no label reply present; one label reply present; and two or more label replies present.

The first case above is known to the interrogator through its waiting for the time in which a label reply should have commenced, and by its observing from an examination of the amplitude of signals in the receiver that no reply has in fact commenced.

If the interrogator detects that no label reply is present the interrogator may issue a *Close Slot Sequence* which signals to all labels in the SLOTTED READ state, to increment their current slot number.

In the second and third cases above it will be clear to the interrogator that one or more labels are replying, and the interrogator will continue to keep the slot open for a time sufficient for the reply or replies to conclude and be evaluated.



Remark: in case the label leaves the field and is not powered anymore, the label moves from all stages except the DESTROYED state into the unpowered state.

Fig. 7: Illustration of the slotted terminating adaptive collection protocol.

The evaluation may take several forms. One of these is based on a special feature of reply coding and makes the detection of collisions, when they occur, highly probable.

Another of these is the checking of a CRC present in the reply, against an expected value that may be calculated from information present in the interrogator to label and label to interrogator signalling.

If there is evidence from these checks that the label data has not been correctly collected, the slot is closed by a *Close Slot Sequence* as already described for the case of an empty slot and the label returns to the READY state.

If however, it appears that the label information has been correctly collected, the slot is closed in another way, in particular by the issuing by the interrogator of a *Fix Slot* command, The effect of this command, if it is correctly detected by the label, and a parameter of the command matches that expected by the label, is to move the label from the SLOTTED READ state to the FIXED SLOT state, in which state the label continues to reply once per round, but always in the fixed slot F. When a label replies in Slot F, its reply is truncated, so as always to fit into slot F, despite the small size of that slot. All of these operations are illustrated in Figure 7.

8.3 Label selection

In this proposal, the labels have a *selection feature* in which groups of labels may be selected e.g. by header, domain manager, object class, or serial number. The interrogator therefore has a selection command which includes the selection data as *optional parameters of the command*.

While a selection command and its data are being processed within the label, the selection data is checked against memory contents on a bit by bit basis, Most Significant Bit (MSB) first. When the signalling of the number of selection bits stated within the command stops, labels in which the EPC matches each so far transmitted bit are *selected*. Labels which have a mis-match in one or more transmitted bits are *not selected*. Labels which are *selected* move from the READY state to the SLOTTED READ state shown in Figure 7.

When no effort is made to select labels, as a result of the Begin Round command not having a selection specification, all labels in a READY state when a Begin Round command is given are *automatically selected* and move from the READY state to the SLOTTED READ state shown in Figure 7.

8.4 Truncated reply

A selected label which subsequently offers a reply will, as shown in Figure 8, reply with the remaining contents of its memory, rounded up to an integral number of bytes, omitting all or most that part which participated in the selection process, and a CRC that is calculated over the entire EPC data.

This CRC is stored within the label, and introduced therein when the label is first programmed. The reply stops at the end (LSB) of the CRC. It does not extend into the part of memory containing the destroy code.

The reply signal also contains *reply start of frame* and *reply end of frame* signals.

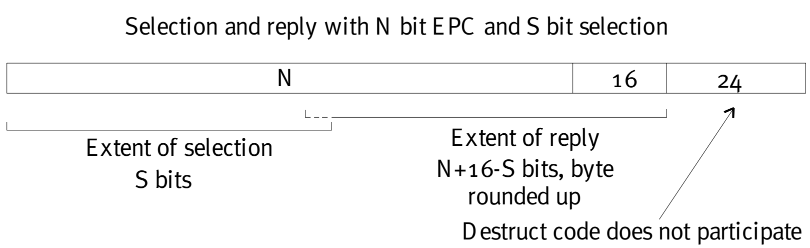


Fig. 8: Illustration of selection and reply.

An advantage of the above selection implementation is that label reply time is reduced as the number of bits transmitted is less than a full reply. The CRC provides protection against communication errors, either in the transmission by the interrogator of selection data, or by the label of the reply.

It is noted, that during the label programming the destroy code is stored in a dedicated portion of the memory. This destroy code portion of memory does not participate in normal selection or reply.

8.5 State retention

As long as it is energised, the label keeps a record in a volatile memory of whether it should be in a READY, SLOTTED READ or FIXED SLOT state. It can, as described above, be moved between these states by interrogator command.

8.6 Optional label kernel

In normal labels, when energising power is lost, the information on whether a label was in a READY, SLOTTED READ or FIXED SLOT state may be lost.

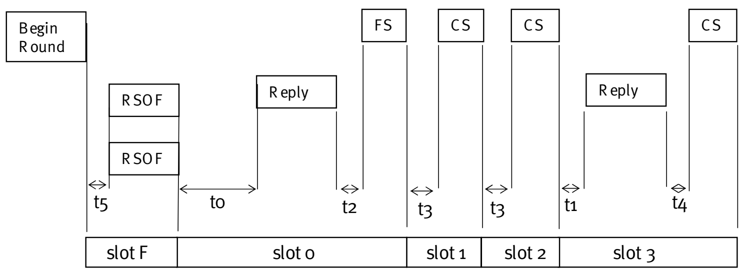
In a multiple read application, it can be desirable that such data be preserved for a short time during operations of label movement or necessary field re-orientation. For this purpose, as an option in this specification, a label may be fitted with controlled data retention memory, which will preserve the FIXED SLOT status when power is lost for a brief specified period, and move the label status to the READY state if energising power is restored after a greater period. That period is normally in the range between 100ms and 100s.

So that labels may, when it is desired, be guaranteed movement to a READY state even when the FIXED SLOT state has been preserved, the *Complete Reset* command has been provided as an option for those labels that have the capacity to remember to reenter the FIXED SLOT state.

This feature causes a slight modification to the protocol diagram of Figure 7.

8.7 Reply round timing

The Figure 9 following shows the timing within a reply round. Note that the diagram is not to scale, but does show the elements that make up a slot.



FS .... Fix Slot Command
CS .... Close Slot Sequence
RSOF ... Reply start of frame

Figure 9: Timing in a reply round.

	Minimum	Maximum
to	113.27μs – 2.36μs	113.27μs + 2.36μs
t1	302.06μs – 2.36μs	302.06 μs + 2.36 μs
t2	302.06μs – 2.36μs	302.06 μs + 2.36 μs
t3	t1nominal + tRSOF + 37.76μs	Infinite
t4	302.06μs – 2.36μs	Infinite
t5	151.03μs – 2.36μs	151.03μs + 2.36μs

Table of Reply Round Timing.

8.8 Generation 2 Protocols

Motivated by the desire of its members to have only a single UHF protocol, and to have one that is carefully shaped to provide good operation is the diverse world regulatory environments, EPCglobal has tasked its Hardware Action Group to produce a Generation 2 UHF protocol with those characteristics.

The emerging protocol has the features:

- A variety of forward link signalling modes is provided to cater for diverse regulatory environments.
- Operation is by a variety of slotted terminating adaptive round known as a query response protocol.
- Selection of groups of tags for participation in a round by various aspects of tag data content is possible.
- Provision for forward link encrypted write communication with a tag after it has been singulated.
- Tag responses may be concealed or unconcealed under a password protected ommand.

A very much simplified illustration of the emerging protocol is shown in Figure 10. Full details are available from EPCglobal publications.

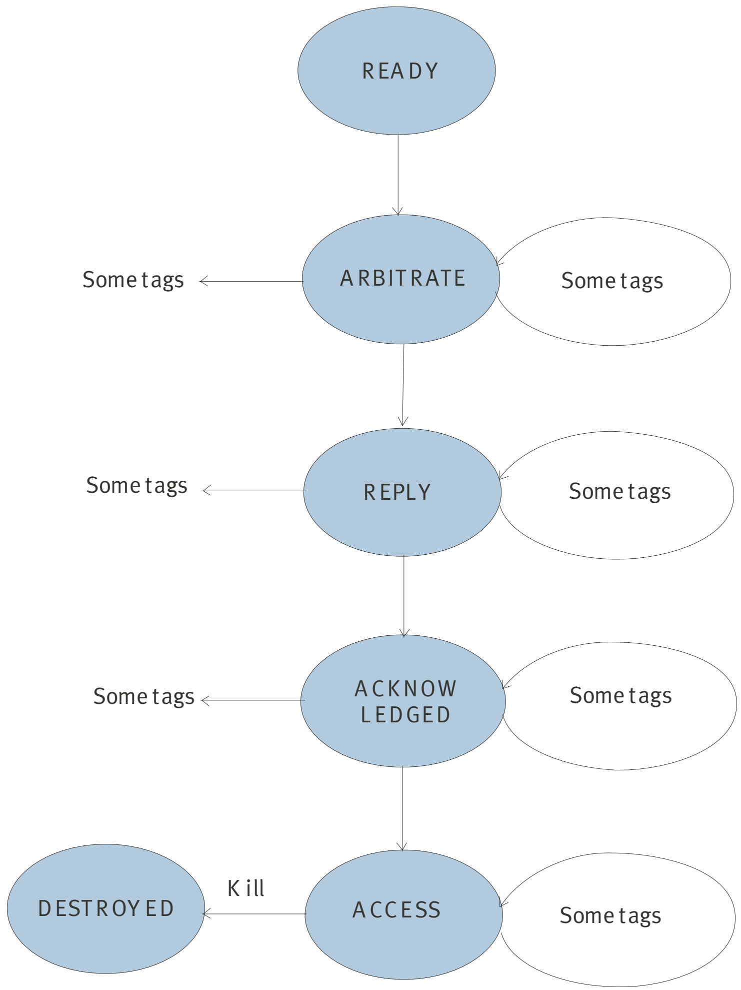


Fig. 10: Simplified Generation 2 Protocol State Diagram

9. Higher level labels

Although the simple read-only EPC data carrying label, known by the Auto-ID Centre as the Class 1 label, is expected to be the backbone of a RFID applications into the future, it may not cater for the future needs of (i) diverse functionalities, (ii) larger range, (iii) greater sensitivity, and (iv) additional autonomous relaying networks.

9.1 Class hierarchy

To cater for these needs, the Auto-ID Centre has defined a label class hierarchy below.

Class 1: Simple EPC read-only labels.
Class 2: EPC labels with additional functionality, such as read-write ability, provision of data security, and theft detection ability.
Class 3: Class 2 labels with battery support for longer range.
Class 4: Autonomous relaying RFID labels.

9.2 Inheritance properties

All of these classes exhibit an *inheritance of characteristics* of the lower classes, and in particular are based on the EPC concept.

An apparent conflict between the desire to maintain a simple linear hierarchy with a multifaceted nature of the additional functionality required within the Class 2 labels has been resolved by allowing, within that class, functionality and signalling characteristics, illustrated in Figure 11, to be defined within a database addressed by a label (as opposed to a product) EPC, and by the definition within the overall Class 2 specification of just those aspects of signalling needed to support this concept.

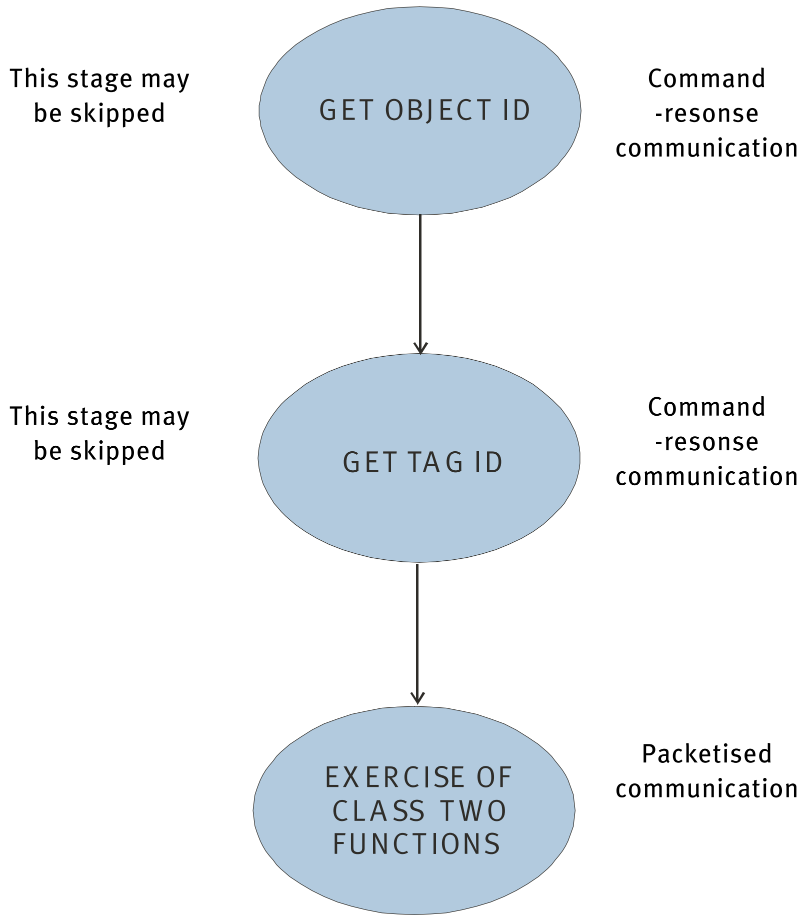


Fig. 11: Illustration of Class 2 operation.

9.3 Summary

The major concepts within this hierarchy are then:

- The EPC concept is used in unifying the classes.
- Differences Class 2 and Class 3 are small: just wake up, and battery, and levels of sleep.
- Thus benefits of Class 2 will easily propagate upward to Class 3.
- The dilemma of diverse functionalities was resolved by writing the standard so that manufactures can define and extend functionalities without revising the standard. The functionality is defined on a database, and pointed to by a tag-id. It is expected that the database will be small enough to be cached in readers.

9.4 Future developments

It is uncertain whether the concepts of Section 9.2 will be retained with the standards management process of EPCglobal.

10. Electromagnetic compatibility constraints

In understanding the effect of electromagnetic compatibility constraints that are normally applied to a RFID systems, we first note that all of the regulations, whether at UHF or at HF, are enforced in the far field. However, as shown in the equations of Section 6.1, near and far fields scale differently with distance, and in particular the near field energy density per unit volume decreases as the *inverse sixth power* of distance from the antenna. The result is that close to the antenna, substantial energy densities may be obtained, but these diminish very quickly as distance increases.

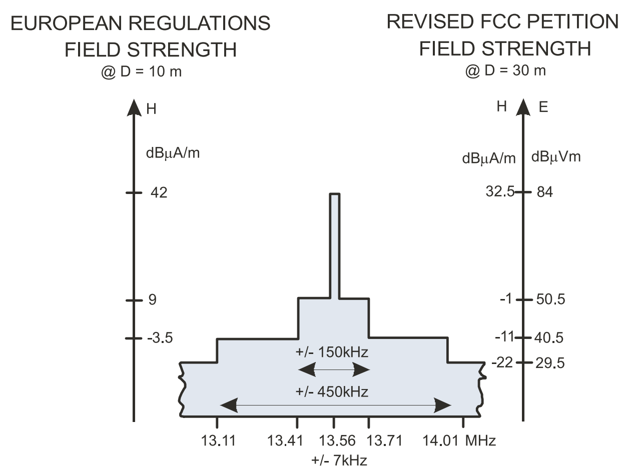


Fig. 12: HF electromagnetic compatibility regulations

Thus although the HF electromagnetic compatibility regulations, shown in Figure 12 allow only minimal (5mW) radiation, well inside the radian sphere distance of $r = 2\pi/\lambda$ it is practicable at HF to obtain a sufficient energy storage field for the operation of a RFID label. However, as efforts to increase reading distance are made, the previously mentioned inverse sixth power of the reactive power density is generally sufficient to reduce the label energising signal to a level below that for practical operation before the boundary of the far field, where its less severe inverse square dependence of energy density, is reached. Thus under current regulations operation of HF systems is almost entirely confined to the near field and short distances.

In an efforts to improve the range of HF RFID systems, efforts have been made, through the design of clever quadrupole antenna systems, to minimise far field radiation while enhancing close-in near fields. While these efforts do achieve some improvement, maintaining the necessary balance between antenna elements which are intended to produce far field cancellation is difficult to achieve in practice.

This focus on minimizing far-field radiation does tend to divert attention from an alternative productive approach to gaining long interrogation range. This approach is that of using *greater label antenna sizes*. To the extent that this approach is successful, ranges will increase to approach the mid field distance. Once this happens, it is reasonable to assume that many antenna fields will, in the ratio of mid-field to far-field amplitude, replicate the value that can be estimated from the dipolar field expressions given in Section 6.1. Once this is done, we conclude that we cannot really have a strong mid field without a related strong value for the far-field, i.e. without some radiation, and this conclusion does not significantly change with interrogator antenna size. In consequence, we can see that the rational approach to achieving long interrogation range includes the step of arguing for a greater allowed radiation from HF interrogator antennas, at least in respect of their carrier level, if not in the signalling sideband level.

It may be noted that the central portion of the spectrum shown in Figure 5, which regulates the operation of RFID equipment in the HF region, is an *industrial scientific and medical* (ISM) band, in which there is no limit to the radiation which is allowed for equipment which receives the classification industrial, scientific or medical. It may be argued that a substantial increase in HF radiation over the narrow central portion of the band shown in Figure 10, could be allowed with substantial technological benefits.

In Figure 10, the sidebands come from interrogator signalling in “reader talks first” systems. Reasonably frequent interrogator signalling is required in systems that manage large label populations. The result can be *significant interrogator signalling sideband levels* in such systems.

Interrogator signalling is generally by means of dips in interrogator signal amplitude. Dips of 18%, with a 1 in 4 duty cycle, and a symbol set which requires one dip for a binary one and no dip for a binary zero, have recently been adopted by the Auto-ID Centre.

Calculation of the sideband signalling spectrum resulting from these parameters is easily performed from the properties of Fourier transforms of single pulses and Fourier series for periodic functions. The result was illustrated in Figure 13 below for the above parameters and for random 1,0 signalling. Conformance to the regulations of Figure 12 is evident. However, if regulations are revised to permit an increased un-modulated carrier level but provide for no increase in sideband signalling levels, techniques to reduce the relative proportion of sideband signalling levels will have to be found. Fortunately, such techniques are available, and have been built into Auto-ID Center specifications for HF labels as an option.

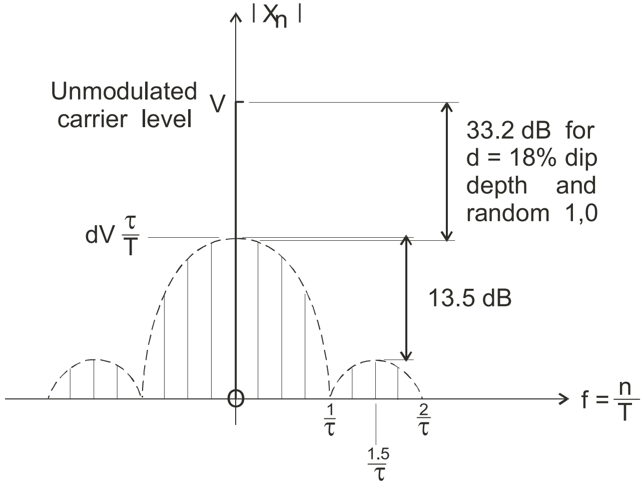


Fig. 13. Illustration of electromagnetic compatibility calculations at HF

11. Field creation structures

11.1 Near field creation

We first illustrate in Figure 14 and Figure 15 below structures which may be useful in the HF region.

Such structures can generate either *electric* or *magnetic* fields, either in the *near-field*, or the *mid-field*, that being the field at the boundary between the near field and the farfield.

As a consequence of the value (22m) of the electromagnetic wavelength at 13.56 MHz, the structures are always in practice electrically small. The first can be considered as creating in the near-field mainly electric field, but as a consequence of the expressions given earlier for dipole fields, which apply to some extent to this situation, the structures will also create some lesser value of near magnetic field.

In the far-field these structures will create electric and magnetic fields in equal proportion, in the sense that for radiated fields in the far-field $|\mathbf{E}| = \eta|\mathbf{H}|$.

The first structure to be illustrated is the monopole wedge illustrated in Figure 12. This structure is potentially useful for creating electric field over large volumes, and has found application in impedance studies of wedge and bow tie antennas, but has not experienced practical use in RFID applications.

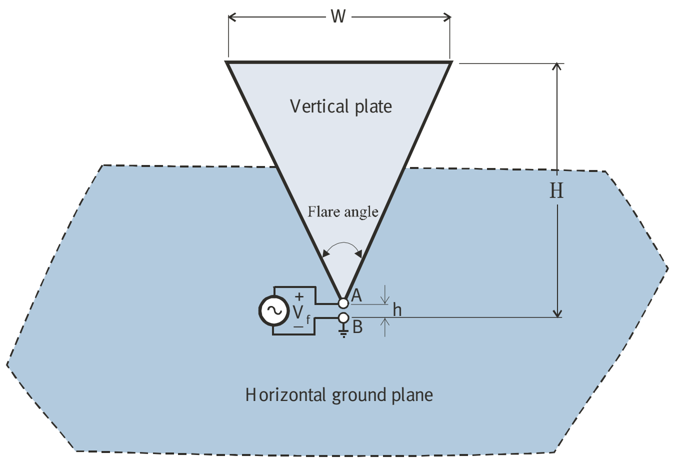


Fig. 14: A monopole plate antenna above a ground plane

The *driving point impedance properties* of the structure are given in Section 13.7.

Fig. 15. A small loop antenna

As HF labels often couple to magnetic fields because they are less easily stopped by conducting materials than are electric fields, there is in fact a more common interest in the creation of strong near-field *magnetic* fields. The usual structure by means of which this is achieved is a small current carrying loop such as is illustrated in Figure 13. This loop has tuning and matching elements at the top end. A strip line transmission line (not visible in the picture) on the underside of the right hand half conveys the driving signals from the connecting point at the centre of the bottom to the driving point terminals at the centre of the top.

As the algebra of Sections 6.1 and 6.2 shows, magnetic dipoles are superior to electric dipoles in the creation of magnetic fields in the near-field.

As is clear from the diagram, the antenna is made from metallic strip. In the mathematical analysis of such a structure, formulae for coils made from round wires are often used, with the diameter of the wire being set to *half the strip width*.

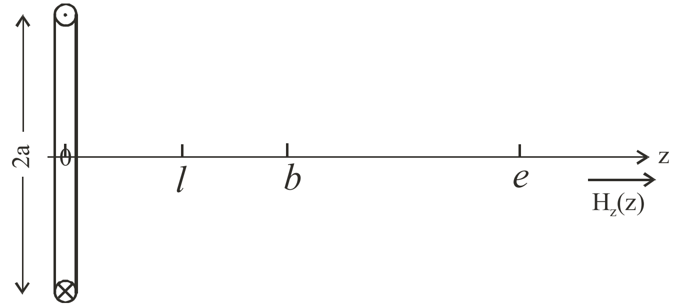


Fig. 16: Geometry for calculating loop and antenna near fields

The task of exciting a label in the near-field is substantially one of creation of a stored energy density per unit volume in the volume served by the antenna. As shown in Figure 17, below, the smaller the loop, the more the field is reasonably confined to the volume near the antenna, and the less the range of the antenna, but less total stored energy that needs to be provided.

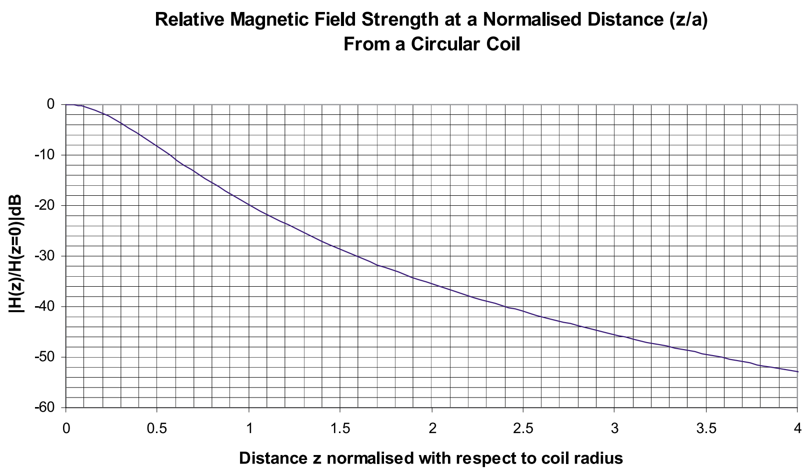


Fig. 17: Variation of loop field with distance along axis

But there is another feature of small loops that is of interest. If we take our task as that of maximizing the magnetic field at the label position *l* in the near field for a given electromagnetic compatibility enforcement distance *e* in the far field it can be shown that the optimum loop size is one where the radius *a* tends to zero.

Of course this is not a practical solution, as it has neglected important practical consequences and practical constraints. One of them is that as the antenna size is made small, an increasing concentration of very high and unused (because the label is further away) energy density per unit volume is created closer to the loop center than is the label, with the results that the total stored energy becomes very large, the quality factor of the antenna becomes unreasonably high, and also the real power density required to drive it becomes too large.

For this reason, attention passes to the *construction of larger loops* such as shown in the Figure 18. In such structures, an effort has been made to counteract the tendency of the current distribution on long wire antennas to become sinusoidal by breaking the loop resonating capacitance into several series elements which are distributed around the loop. Such structures have been used by ourselves to allow substantially increased range from high frequency labels.

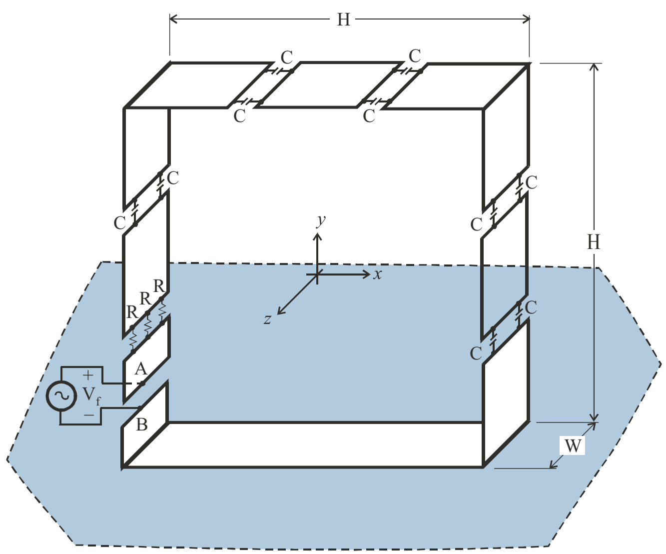


Fig. 18: A large loop antenna

11.2 Far field creation

The almost universal choice for the creation of far-fields, particularly in the UHF region, is the patch antenna illustrated below.

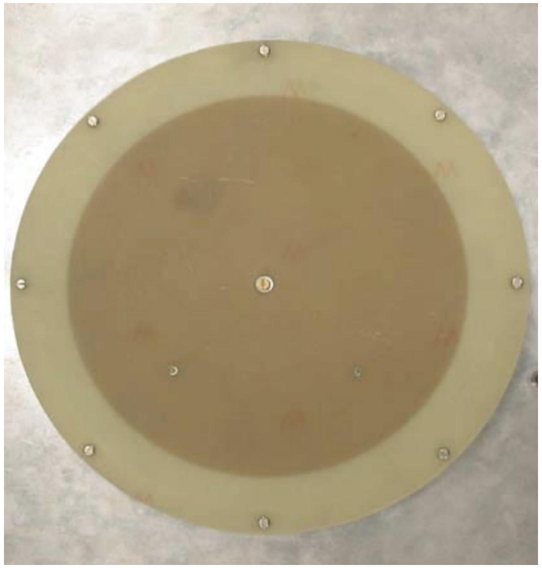


Fig. 19. A patch antenna

12. Antenna size effects

RFID label antennas may be electrically large or small depending upon frequency band. Reader antennas may also be electrically large or small depending upon frequency band. We will consider in this section some of the issues that arise.

12.1 Measures of label excitation

Appropriate measures of tag excitation in the near field are energy storage or oscillating *reactive power density per unit volume*, and in the far field *energy propagating per unit time per unit area*. At HF, we will model tag antennas using coupling volume concepts to be defined Section 13. In that modelling we can recognise both *electric field sensitive* and *magnetic field sensitive* designs.

12.2 Magnetic field sensitive antennas

A common example of a magnetic field sensitive HF label is shown in Figure 20 below. The label is 42 mm wide by 47mm high.

The label is designed to have sufficiently many turns to provide the resonating inductance for the microcircuit input capacitance, as well as a flux collecting area in the interior which is as large as practicable consistent with the size requirement for the label.

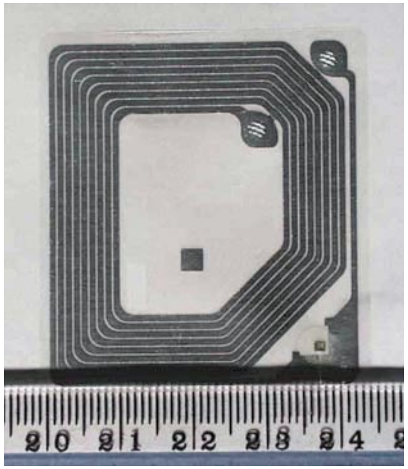


Fig. 20: A magnetic field sensitive HF label antenna

Advantages of working in the near field at HF rather than at LF are that the number of turns required to resonate the microcircuit capacitance is small enough for the low resolution of the lithography to be used in antenna construction, and no additional external resonating capacitance is required.

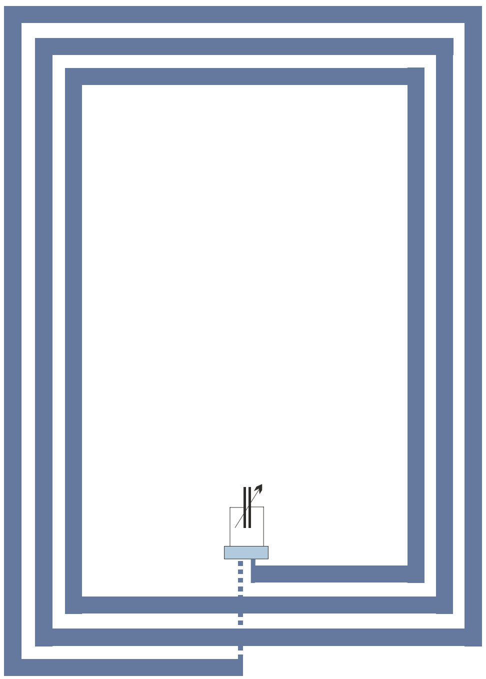


Fig. 21. A larger loop antenna for an HF label

When a larger space is available for the label, larger coil area should be used. As shown in Figure 21, fewer turns are then needed to obtain the required tuning inductance. It will be shown later that the figure of merit (the coupling volume) for a planar coil operating in the near field varies as the third power of size. Thus the antenna of Figure 21 is about 18 times more sensitive than that of Figure 20. Unfortunately, this increased sensitivity does not translate to a corresponding increase in range, as small coil interrogator antennas have an *inverse sixth power* decrease in energy density per unit volume as distance from the interrogator increases.

Clearly, both of the designs illustrated above are unsuitable for being placed flat against metal, as the boundary conditions shown in Figure 3 will not allow a normal component of magnetic flux density at the metal surface. For this situation, the label antenna employing a solenoid with a magnetic core design shown in Figure 22, is employed.

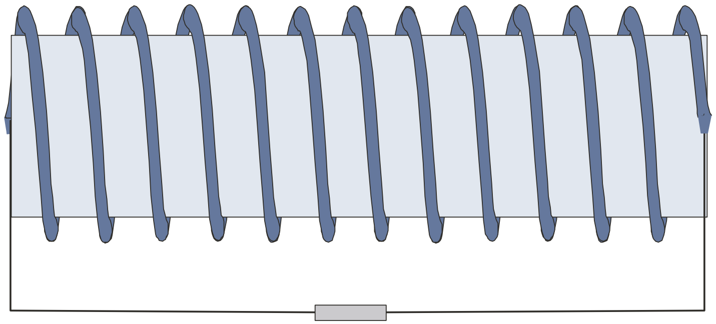


Fig. 22. Antenna for HF operation against metal

It will be shown in Section 13.3 that without the magnetic core the coupling volume of a long solenoid is just the physical volume, but when a magnetic core is inserted, the coupling volume increases by a factor equal to the *effective permeability* defined in that later section.

This behaviour may be contrasted with that of electric field labels, in which in Section 13.6 we will find that the inclusion of dielectric material into the interior of the label is not helpful.

12.3 Electric field sensitive antennas

Two varieties of *electric field sensitive antennas* are shown in Figures 23 and 24 below. Figure 23 shows a small bow tie antenna that is intended to be sensitive to electric fields in the horizontal direction.

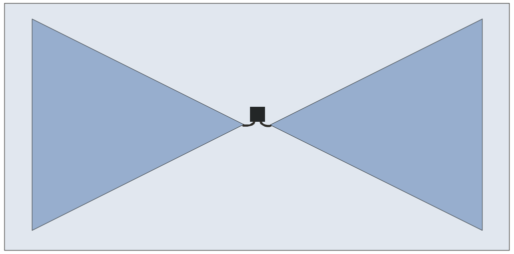


Fig. 23: An electric field sensitive label

Figure 24 shows an electric field sensitive antenna that is suitable for placement against a horizontal metal plate.

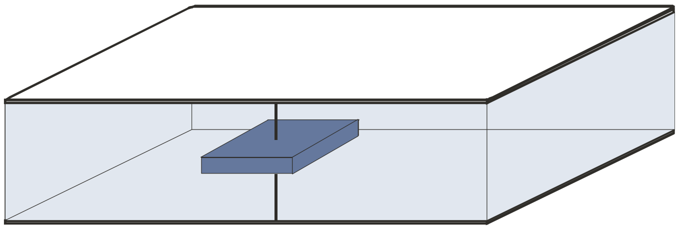


Fig. 24: A parallel plate electric field sensitive label

Analysis of both of these structures is provided in a later section. The analysis has a common feature that the figure of merit for these antennas, when placed in the energy storage electric field is a *coupling volume*, in the second case equal to the physical volume of the structure, and in the first case a volume derived from the label dimensions, even though the antenna itself has no physical volume.

Both of the antennas will also have an *effective area*, but this area should not be confused with the effective area concept of a radiating antenna or a far field antenna. The effective area for a near field electric field sensitive antenna describes the extent to which the antenna can e*xtract current from the displacement current density of the driving electric field*.

12.4 Label antenna equivalent circuits

Equivalent circuits for small *magnetic field sensitive antennas* and electric field sensitive antennas are shown in Figures 25 and 26 below. The range of validity of these equivalent circuits is where the reactance properties of the antenna may be described by a single parameter, L or C. When the antenna is larger, as will be reported in a later section, reactance properties might well be described by an appropriate mixture of L and C.

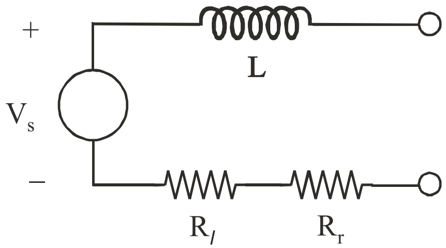


Fig. 25: Equivalent circuit for a small magnetic field sensitive antenna

In Figure 25, the source voltage is the voltage induced in the flux collecting area of the coil by magnetic fields other than those fields which resulted from current flowing within the coil itself. Those induced voltages are represented by the voltage drop in the inductor *L*. The parameters R_f and R_r are loss and radiation resistances respectively.

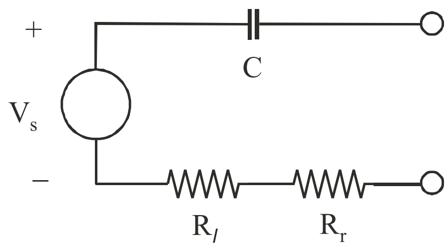


Fig. 26: Equivalent circuit for a small electric field sensitive antenna

In Figure 26, the source voltage is the voltage developed across the self capacitance of the antenna as a result of the current injected into it, by the displacement current density of the electric field in which the antenna is immersed. The parameters R_f and R_r are loss and radiation resistances respectively.

Calculation of the parameters of the magnetic field sensitive antenna is straightforward, the relevant formulae being obvious or contained in the Appendix. For the electric field antennas, determination of the relevant parameters is sometimes not quite so simple, as electrostatic field solutions for the relevant geometries are not readily available. Therefore empirical results or numerical modelling are more commonly employed for useful shapes. Results will be reported in a later section for some significant structures.

12.5 Electromagnetic field antennas

We consider an antenna as an *electromagnetic field antenna* on a couple of different bases. Firstly, if the antenna is capable of responding to both electric and magnetic fields we would consider it to be an electromagnetic field antenna. It is almost invariably true that unless the antenna is very small, it does have this property. Proper analysis requires that it be analysed using the full set of Maxwell's equations, rather than the subset or simplified versions that pertain to electrostatic or magnetostatic problems. A good example of this phenomenon is provided by the electromagnetic field sensitive antenna shown in Figure 27, in which there is no obvious effort to couple to either electric or magnetic field alone.

Fig. 27: An electromagnetic field sensitive antenna

Such electromagnetic antennas are generally useful for operation in the far-field, because far-field interrogation systems have shorter wavelengths, and antennas of acceptable size can no longer be considered to be electrically very small, but are merely small.

Despite this distinction, there are electromagnetic label reading environments in the UHF region in which, through reflections, either the electric or magnetic field is emphasised at the expense of the other. For such situations it is normally useful to take into account the nature of the driving fields in antenna design, and to shape the design so that it is recognisably attuned to one or other of those fields.

13. Coupling relations

13.1 Near and far fields

In the analysis of the performance of RFID systems it is important to consider whether the labels are placed in the far (propagating) or near (energy storage) fields of the interrogator antenna. When that antenna is of small gain, the distance which divides the near and far fields is the size or the *radian sphere* of radius r = λ2/π where λ is the free space electromagnetic wavelength at the operating frequency.

13.2 Field Measures

For a linearly polarised magnetic field described by a peak value phasor **H** we may develop the two measures of the exciting field

$$\text{Radial component of Poynting vector } S_r = \frac{\eta\,|H|^2}{2} \text{ in Wm}^{-2}$$

$$\text{Volume density of reactive power } W_V = \frac{\omega\mu_0|H|^2}{2} \text{ in VAm}^{-3}$$

For a linearly polarised magnetic field described by a peak value phasor **E** we may develop the two measures of the exciting field

$$\text{Radial component of Poynting vector } S_r = \frac{|E|^2}{2\eta} \text{ in Wm}^{-2}$$

$$\text{Volume density of reactive power } W_V = \frac{\omega\varepsilon_0|E|^2}{2} \text{ in VAm}^{-3}$$

In both cases the latter expression is ω times the peak value of stored magnetic or electric energy per unit volume. We can easily show that if β is the propagation constant at the frequency used

$$W_v = \beta S_r\,.$$

The last expression assumes that we are in the far field, i.e. there are no near-field reactive energy storage fields to augment W_v without contributing to S_r.

13.3 Near field operation – magnetic field

For near-field operation, the fields which excite the label are basically energy storage fields for which consideration of power flow is inappropriate, and for which pre- Maxwell versions of electrodynamics can give correct calculations. Both the interrogation field creation means and the label field detection means can be considered as weakly coupled inductors of self-inductances L_1 and L_2 and mutual inductance M. When both those coils are tuned to resonance with respective quality factors Q_1 and Q_2, it can be shown that the power P_2 dissipated in the losses of the label coil is related to the power P_1 dissipated in the losses of the "transmitter" coil by

$$\frac{P_2}{P_1} = k^2 Q_1 Q_2 \text{ where } k = \frac{M}{\sqrt{\mathrm{L}_1 L_2}}$$

This equation is useful to show the role of the quality factor, Q, of the resonances in both the label and interrogator coils in promoting power transfer, but it is not useful in separately optimising the properties of those two widely dissimilar elements.

For that purpose we can focus first on the *energy storage* measure of exciting field, which is the *reactive power per unit volume* in the field created by the interrogator at the label position. In terms of that measure, we can define, [4] a figure of merit of a label antenna as the ratio

$$V_c = \frac{[\text{Reactive power flowing in the untuned label coil when it is short circuited}]}{[\text{Volume density of reactive power created by the interrogator at the label position}]}$$

which clearly has the dimensions of volume, and is for this reason called the *coupling volume* of the label antenna. For the performance of the interrogator antenna, we can define the companion concept of *dispersal volume* V_d given by

$$V_d = \frac{[\text{Reactive power flowing in the inductor of the interrogator field creation coil}]}{[\text{Volume density of reactive power created by the interrogator at the label position}]}$$

When both antennas are tuned it is possible to show

$$\frac{P_2}{P_1} = \frac{V_c}{V_d} Q_1 Q_2$$

The benefit of this formulation is that the coupling volume is a property of the label parameters alone, and the dispersal volume is a property of the interrogator antenna parameters alone (and of the label position), and separate optimisation becomes possible, whereas k^2 is a complex function of the entire system geometry.

Coupling volumes for various label antenna are readily determined. For an air cored solenoidal antenna it is approximately the volume of that antenna, and when a magnetic core is in place that volume becomes multiplied by the relative effective permeability

$$\mu_{er} = \frac{\mu_{ir}}{1 + N(\mu_{ir} - 1)}$$

where μ_{ir} is the relative intrinsic permeability of the core material and N is the demagnetising factor in the direction of the interrogator field.

For a planar coil, which in its idealised state has no physical volume, the coupling volume is given by

$$V_c = \frac{\mu_0 A^2}{L}$$

where A is the flux-collecting area (incorporating by summation an area for each turn) of the coil, and L is the self inductance.

13.4 Near field – electric field

In RFID systems, the coupling can be via the magnetic field or the electric field. However it is almost always by way of the magnetic field. Nevertheless it is possible to couple to the electric field whether in the far field or the near field.

The energy transfer is provided by the electric flux terminating on the antenna surface and inducing a charge on the antenna. The induced charge will oscillate as the field around oscillates. The induced charges will thus produce a current.

The issues of understanding and optimising coupling to the electric field are important. The coupling volume theory developed here for the electric field case assists in that effort and allows comparisons between different antenna structures for their effectiveness and efficiency in terms of their actual physical volume and the coupling volume. It is also of interest in making comparisons with the magnetic field case.

When we can focus on the *energy storage* measure of exciting field, which is the reactive *power per unit volume* in the field created by the interrogator at the label position, we can define a figure of merit of a label antenna as the ratio

$$V_c = \frac{[\text{Reactive power flowing in the untuned label capacitor when it is open circuited}]}{[\text{Volume density of reactive power created by the interrogator at the label position}]}$$

AUTO-ID LABS

which clearly has the dimensions of volume, and is for this reason also called the *coupling volume* of the label antenna.

For the performance of the interrogator antenna, we can define the companion concept of dispersal volume V_d given by

$$V_d = \frac{[\text{Reactive power flowing in the capacitance of the interrogator field creation electrodes}]}{[\text{Volume density of reactive power created by the interrogator at the label position}]}$$

When both antennas are tuned it is possible to show

$$\frac{P_2}{P_1} = \frac{V_c}{V_d} Q_1 Q_2$$

The benefit of this formulation is that the coupling volume is a property of the label parameters alone, and the dispersal volume is a property of the interrogator antenna parameters alone (and of the label position), and separate optimisation becomes possible.

Coupling volumes for various label antenna are readily determined, as will be done in the next section.

13.5 Coupling volume of a general shape

In order to derive the coupling volume of antenna structures it is important to define a number of concepts. For a given antenna we can define an *electric flux collecting area* as the area in space required by the antenna to generate from the displacement current of the exciting field an oscillating current I when the antenna is placed in an oscillating electric field of flux density D. This area may or may not be equivalent to the physical area. This electric flux collecting area will be denoted by the symbol A_f will be referred to as effective area.

The formula for the current I flowing from an antenna placed in an electric field with a flux density D and oscillating angular frequency ω, using the effective area A_f of the antenna structure is

$$I = j\omega D A_f$$

The reactive power flowing in the antenna with a self-capacitance C when the antenna is open circuit is given by

$$\text{Reactive Power} = \frac{|I|^2}{2\omega C}$$

The reactive power W_V per unit volume in the field is obtained by

$$W_V = \frac{\varepsilon_0 |E|^2 \omega}{2}$$

Thus the coupling volume of an antenna structure with a self capacitance of C can be obtained as

$$V_c = \frac{\varepsilon_0 A_f^2}{C}$$

This formula provides the coupling volume of any antenna with a self capacitance of C and an effective area of A_f. It should be pointed out that A_f can easily be determined experimentally by measuring the current flowing from an antenna placed in an oscillating electric field of known strength.

$$A_f = \frac{I}{\varepsilon_0 E \omega}$$

13.6 Coupling volume of a rectangular capacitor

The derivation discussed below will prove that the coupling volume of an air dielectric parallel plate capacitor is equal to the physical volume of the capacitor.

Consider a parallel plate capacitor antenna, as shown in Figure 28, consisting of two plates, each having an area A and immersed in an external electric field E. Then the charge Q distributed on the plate is given by

$$Q = \varepsilon_0 AE$$

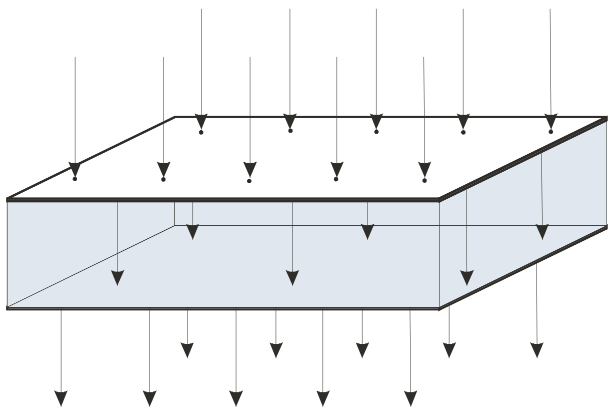


Figure 28: Field configuration for a parallel plate electric field antenna

The capacitance C of a parallel plate capacitor is given by

$$C = \frac{\varepsilon_0 \varepsilon_r A}{d}$$

Now the flux density between the plates is the same as the flux density outside, but the field strength between the plates is $1/\varepsilon_r$ times the field outside. So the voltage V between the plates is

$$V = \frac{Ed}{\varepsilon_r}$$

The energy stored in the capacitor is obtained as

$$\frac{1}{2}\frac{\varepsilon_0 AE^2 d}{\varepsilon_r}$$

As the energy stored per unit volume in an electric field is $\frac{1}{2}\varepsilon_0 E^2$, we have $V_c = \frac{Ad}{\varepsilon_r}$.

Examination of the above formula reveals a number of results. Coupling volume has a dependence on the value of the relative permittivity ε_r, and a dielectric if present will reduce the coupling volume. This last result stands in contrast to the case for a rectangular magnetic coil, where relative permeability μ_{er} multiplies the coupling volume, not divides it. However, for an air dielectric capacitor, the coupling volume is equal to the physical volume.

It can be observed that the coupling volume of a general structure reduces to the coupling volume of a parallel plate capacitor when

$$A_f = A \text{ (area of a plate) and } C = \frac{\varepsilon_0 \varepsilon_r A}{d}$$

13.7 Properties of the bow tie antenna

Calculating the self-capacitance, as depicted by the field lines of Figure 29, of the monopole bow tie antenna by seeking analytical solutions to Laplace's equation presents a difficult problem.

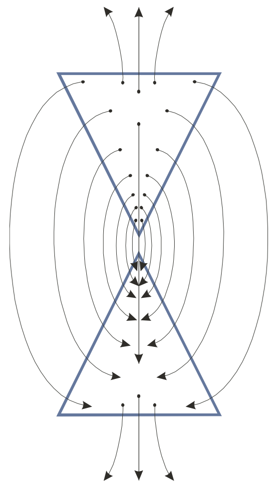


Fig. 29: Field configuration for self capacitance calculation of bow tie antenna

Nevertheless a numerical solution is both tractable and much simpler under the present circumstances and has been performed. The method of moments provides a suitable numerical approximation to the self capacitance of a bow tie antenna.

In addition, a detailed study undertaken of results first published by Woodward, [5] provided the following empirical model for an equivalent circuit shown in Figure 30 of a monopole bow tie antenna. The model derived from Woodward's results for the reactance $X(\omega)$ of a bow tie antenna provides the values of a capacitor, whose value is that of the self-capacitance of the bow tie, and an inductor placed in the series circuit shown.

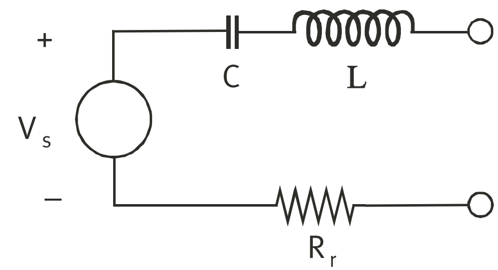


Fig. 30: Three parameter equivalent circuit for a bow tie antenna

$$jX = \frac{1}{j\omega C} + j\omega L$$

This model is only suitable for electrically small antennas obeying the strict limit of

$$h << \frac{\lambda}{6}$$

where h is the height of the bow tie antenna and λ is the wavelength measured in meters.

The model parameters obtained will vary for different flare angles of the bow tie antenna. However the capacitance can be expected to scale up with increasing height for any constant flare angle. The following results give the capacitor and inductor model values for a monopole bow tie antenna with a flare angle of 90°.

$$C = K_C \varepsilon_0 h \quad \text{with} \quad K_C = 7.6$$

$$L = K_L \mu_0 h \quad \text{with} \quad K_L = 0.2135$$

where h is the height of the bow tie antenna.

The significant finding is that, as expected, the low frequency impedance of a bow tie antenna is mainly capacitive and this value can thus be obtained by calculating the selfcapacitance of the bow tie antenna.

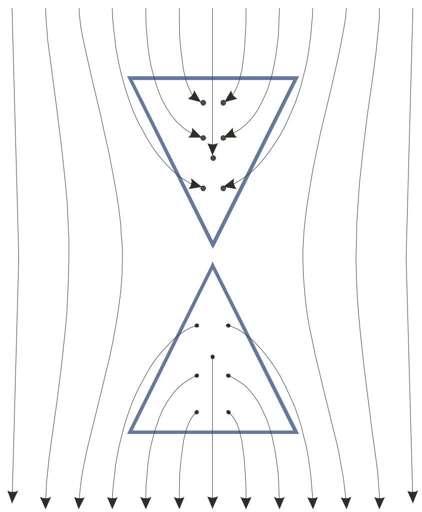


Fig. 31: Field configuration for calculation of effective area of bow tie antenna

Woodward's results and experimental results of radiation resistance of bow tie antennas have been used to produce a value for the radiation resistance R_r identified in Figure 30 for different antenna heights and flare angles. The following generic formula can be applied to evaluating the radiation resistance R_r of that antenna.

$$R_r = K_R(\beta h)^2 \quad \Omega$$

where the constant K_R, measured in Ω, is a specific value that depends on the flare angle of the bow tie antenna. The table below provides a set of values for K_R derived from Woodward's results [5].

Flare angle	K_R in Ω
5	15.0
10	17.7
30	22.6
40	24.9
50	26.4
90	30.0

Table of flare angles and radiation resistance constants for a bow tie antenna of height h

The radiation resistance obtained has significance in two ways. It allows the amount of radiated power to be calculated for a transmitting antenna, and also provides for a label antenna a means of calculating, using the reciprocity theorem, the effective electric flux collecting area, as depicted in Figure 31, of the antenna.

All of these results agreed with the results of our own direct measurements and our own numerical analysis performed using the method of moments.

13.9 Far field relations

For calculation of the power coupled in the far field to a label with a lossless receiving antenna the usual approach is to derive the available source power from the label antenna from the formula

$$P_r = S_r A_e = \frac{g\lambda^2}{4\pi} S_r$$

where A_e and g_r are the effective area and gain of the label antenna, related as shown above by reciprocity considerations. If the label has losses, or is not conjugately matched, the power delivered to the external load is derived from simple modifications of the above formula. It should be noted that the effective area for the far field is a concept unrelated to either the magnetic flux collection area of a loop or the electric flux collection area of a plate.

The power flow per unit area is given by

$$\text{Power flow per unit area} = \frac{g_t P_t}{4\pi r^2}$$

wherein g_t is the gain of the transmitter antenna and P_t the power which it transmits, and r is the distance from the transmitter antenna to the label position. In using this formula, we are implicitly assuming that the label has been placed in the direction of strongest radiation from the interrogator antenna.

The power P_r which may be extracted under optimum conditions of tuning and matching by a lossless label antenna placed at the above position is given by

$$P_r = A_{er} \times \text{Power flow per unit area}$$

wherein A_{er} is a property of the label known as its *effective area*. It is unrelated to the physical area of the antenna, (which if it is just a piece of thin wire, does not have a physical area), but has the desirable property that we may imagine the label antenna collects all of the radiated power which flows through that effective area which may be thought of as surrounding the label antenna.

The *Lorenz reciprocity theorem* of electrodynamics may be used to show that the effective area of a receiving antenna is related to the gain g_r it would have in a transmitting role by the equations

$$A_{er} = \frac{g_r \lambda^2}{4\pi}$$

$$\frac{P_r}{P_t} = g_r g_t \left(\frac{\lambda}{4\pi r}\right)^2$$

and

$$\frac{P_r}{P_t} = \frac{A_{et} A_{er}}{\lambda^2 r^2}$$

14. Relations Between Formulations

14.1 General remarks

As the Poynting vector-effective area formulation and the coupling volume-dispersal volume formulation are so apparently dissimilar, it is useful to show that they are equivalent, but useful in different contexts where different approximations may be made. It will be shown that the basic difference between the formulations is whether they emphasise the *radiation resistance* or the *internal losses* of the label antenna.

14.2 Power matching considerations

However we regard our label, we expect to have, in its series equivalent circuit, as well as a reactance, a radiation resistance R_r, a loss resistance R_c, and a load resistance R_L, and possibly a matching reactance. Considerations of maximum power transfer to the load will always require that at optimum we set

$$R_L = R_r + R_c$$

14.3 Choice of formulations

Because labels are mostly small in relation to a wave length, we will frequently (but not always) have $R_r \ll R_C$. When the label becomes particularly small, we will focus attention on Rc and neglect R_r entirely. Also because labels are mostly small we will have, in their tuning and matching, quality factors which are large.

Our analysis will be performed for the magnetic field sensitive antenna consisting of a small loop. We will leave it to the reader to produce a similar analysis for the electric field sensing antenna.

14.4 Analysis of a small loop

In the case of a small loop of area A and self inductance L, with loss and radiation resistances as described above, excited by a magnetic field, the radiation resistance is

$$R_r = \left(\frac{\eta}{6\pi}\right)(\beta)^4 A^2$$

Assuming the loop has the gain g_d of a small dipole it has an effective area A_e of

$$A_e = g_d \frac{\lambda^2}{4\pi}$$

and neglecting losses the available source power P_a when the loop is in a field of Poynting vector S_r is

$$P_a = g_d \frac{\lambda^2}{4\pi} S_r$$

This is the power which a *lossless* antenna would deliver to a load $R_L = R_r$. For comparison with coupling volume theory, which we apply when R_r is negligible with respect to the losses R_c, it is appropriate, in view of the definition in Section 16.3 of coupling volume, to calculate the power delivered to the losses R_c without any external load yet having been added. This power P_c is given by

$$P_c = \left(\frac{4R_r}{R_c}\right) P_a$$

If we now substitute for R_r and P_a from equations earlier in this section and S_r from Section 16.2 and use the value 1.5 for g_d, we obtain

$$P_c = \frac{\eta^2 \beta^2 |H|^2 A^2}{2R_c}$$

To manipulate this into a more familiar form, we replace R_c by $\omega L/Q$ and $\eta\beta$ by $\omega\mu_o$ and obtain

$$P_c = \frac{(\omega\mu_0 |H|^2)}{2}\left(\frac{\mu_0 A^2}{L}\right) Q$$

In this relation we find the familiar factors W_v in the first bracket and the coupling volume V_C in the second bracket. Thus we can rewrite the above equation as

$$P_c = QW_vV_c$$

which is the standard form of the result from coupling volume theory for coils coupling to the magnetic field.

What we have done is to show that the effective area and coupling volume formulations of loop antenna behaviour are entirely equivalent.

14.5 Available power from a small loop

When an optimum load resistor $R_L = R_C$ is added to a small loop, the power which can be delivered to that load is one quarter of that given by the equation immediately above, if we continue to interpret Q as defined by the loop losses and its inductance. If however we redefine Q to be the new and lowered Q, determined by the sum of the loop losses and the damping of the external load, then the power we can deliver to a matched load is now half that given by that equation.

14.6 Creation of an exciting field

We next consider issues concerned with the generation of label exciting fields. We can distinguish in the provision of exciting field a number of different situations such as: (a) a limited and enclosed volume; (b) a limited but not enclosed volume in which it is practicable to construct an antenna with dimensions of the order of the scanned space, and (c) interrogation over a distance large in relation to the dimensions of any practicable antenna.

We will consider the last of these situations below.

14.7 Interrogation at a large distance

By choice of frequency we may be able to place the large distance in the near field or in the far field. In the near field we are not concerned with radiation, and normally try to minimise it. We also minimise the losses and obtain a total reactive power U_t for a power P_t delivered to a field creation structure of quality factor Q_t given by

$$U_t = Q_tP_t$$

The energy density per unit volume at the label position is obtained by introducing the dispersal volume V_d, so

$$W_v = \frac{Q_tP_t}{V_d}$$

A good example to take is that of a circular coil of diameter D from which we wish to create a label exciting field at a large distance R. Since we are considering here the large distance case, analysis of the field properties of such a loop leads, to a good approximation when $R >> D$, to

$$V_d = F\left(\frac{4R^2}{D}\right)^3$$

where R is the distance from the interrogator to the label, and F is a factor of the order of 2.

Substituting this value into the second equation of Section 14.7 gives

$$W_v = \frac{Q_t P_t D^3}{F(2R)^6}$$

which is a result we will compare with another about to be derived. Now let us consider a different frequency for which we have the label in the far field of the transmitter antenna, which in this case will be an efficient radiator. If we have a transmitter gain g_t, the power density per unit area in the Poynting vector is

$$S_r = \frac{g_t P_t}{4\pi R^2}$$

and in view of the last equation of Section 13.2 we have

$$W_v = \frac{\beta g_t P_t}{4\pi R^2}$$

Replacement of β by $2\pi/\lambda$ gives

$$W_v = \frac{g_t P_t}{2\lambda R^2}$$

Comparing this result with that obtained earlier in this section, we see that we have lost the advantage of the high Q_t but gained the advantage that λ in the denominator of the equation just obtained is significantly smaller than the corresponding R in the fourth equation of this section. This remark applies when $R = D$; at greater values of R the situation gets rapidly worse. Furthermore, the factor 2^6 in the denominator of the fourth equation of this section tends to eliminate the benefit of any practical value of Q_t, and we also have in the equation immediately above the advantage (probably modest) of the transmitter antenna gain g_t.

Clearly the case for propagating communication becomes overwhelming when $R > D$. However, when we consider fields at the centre of a portal, it may be shown that $V_d = F D^3$ and our factor of 2^6 has disappeared. This is just showing us how relatively strong is the field at the centre of a portal in relation to the field a distance D away.

14.8 Operation under EMC constraints

The preceding analysis has explored the efficiency of power transfer between a transmitter and a label and is effectively drawn from the point of view of minimising transmitter power. A more real constraint is provided by electromagnetic compatibility (EMC) licensing considerations that are directed not to transmitter power but equivalent isotropic radiated power at stated distances. What this results in, is a direct constraint on S_r at the label site, whether or not the interrogation distance is the same as the EMC constraint distance. Our attention then changes to the problem of receiving power from a given area power density, with a label of size limited by the application.

As long as the antenna matching is efficient we find the power transfer is enhanced by lowering the frequency so the effective area of the label, considered as a receiving antenna, is increased. Eventually, however, as the frequency is lowered too far, the radiation resistance drops below the loss resistance of the antenna, and does so sufficiently rapidly for the increase in effective area of the label antenna to be more than offset by the mismatch efficiency. This conclusion is sustained for both electric dipole and magnetic loop antennas, although the details of the loss mechanisms vary between the two cases.

As a result of this analysis we are led to conclude that: t*he optimum frequency for operation of an RFID system in the far field is the lowest frequency for which a reasonable match to the* *radiation resistance of the label antenna can be achieved, at the allowed size of label, without the label or matching element losses intruding.*

15 Conclusions

The fundamental principles governing electromagnetic coupling between an interrogator and its labels in an RFID system have been outlined, and a set of concepts suitable for describing coupling in the near and far fields, using electric field, magnetic field or electromagnetic field sensitive antennas, have been defined.

Algorithms for the reading of large label populations at UHF and at HF have been described. Some of the properties of electric field and magnetic field antennas have been derived.

Some important theorems about optimising antenna sizes and operating frequencies subject to electromagnetic compatibility regulations have been derived. Optimisation policies which can be pursued in the near field, the mid field or the far field have been identified.

A proposal for relaxation of electromagnetic compatibility regulations in a narrow band in the HF region is investigated, and some of the interesting consequences for HF operation at substantially improved distances have been identified. These results, while attractive, are unlikely, without substantial increase in label antenna sizes, to challenge the supremacy of far-field systems for long-range RFID system operation.

AUTO-ID LABS

References

[1] International Organisation for Standardisation, “SI units and recommendations for the use of their multiples and of certain other units”. International Standard ISO 1000 (1992).

[2] Cole, P.H., “A Study of Factors Affecting the Design of EPC Antennas & Readers for Supermarket Shelves”, http://www.autoidcenter.org/research/ADE-AUTOID-WH-001.pdf, June 2002.

[3] J. A. Osborn, “Demagnetising Factors of the General Ellipsoid”, Physical Review, vol 67, pp 351, (1945).

[4] K. Eshraghian, P.H. Cole and A. K. Roy, “Electromagnetic Coupling in Subharmonic Transponders”, Journal of Electrical and Electronics Engineering, Australia, 2, 28-35, March 1982.

[5] Brown, G.H. and Woodward Jr, O.M., “Experimentally determined Radiation characteristics of Conical and Triangular Antennas”, RCA Review, vol 13, pp 425-452, (1952).

Appendix I
Radiation Made Easy

1. Objective

We summarise here, in simpler terms than provided in the main paper, and without proofs, some of the import concepts and equations of electromagnetic radiation theory that are relevant to RFID.

The material is provided for the benefit of persons who are not yet well versed in electromagnetic theory.

2. Is there electromagnetic radiation?

It would seem so. You are invited to turn on your radio (or your mobile) to experience the effects of electromagnetic radiation.

3. Why there is radiation

In brief, Maxwell. Before Maxwell, the laws of electrodynamics including Gauss' law, Ampere's law of magnetostatics, and Faraday's law did not predict waves. They correctly gave what is known as the *near field*, i.e. the electrostatic field of an electric charge, or the magnetostatic field of a current loop. Maxwell realised those laws were inconsistent with the *law of conservation of charge*. He corrected Ampere's law of magnetostatics to become Ampere's law as corrected by Maxwell, so that consistency with the law of conservation of charge now occurred. He added a term which said that *vortices of magnetic field* can be *displacement current density* (time varying electric flux density) as well as conduction current density. The resulting corrected equations predicted *electromagnetic waves*. Hertz confirmed experimentally that they exist.

4. How to calculate radiation

We use the concept of *retarded potentials*. These are just the same form as the electrostatic scalar potential and the magnetostatic vector potential, but have the added concept that the *potential propagates* away from the source (or vortex) at the *speed of light*. Formulae can be found in the lecture notes, but these simple notes here do not need them.

5. Simple transmitting antenna concepts

We assume lossless antennas here. The result will be that the concepts of *gain* and *directivity* (soon to be defined) will be merged.

5.1 Power density per unit area

Assume and antenna is radiating a power P_r. The power-like measure of radiated field is the *power per unit area* (Poynting vector) S_r. An *omni directional radiator* has

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

It is the same for all directions. It is the average over all directions. Notice the inverse square law, as expected.

5.2 Antenna gain

If radiation is not omnidirectional, then at a given distance, some power density per unit area in the stronger directions will be stronger that the average, and other power density per unit area in the weaker directions will be weaker that the average. Hence the concept of *antenna gain* g_r for a lossless radiating antenna

$$g_r = \frac{\text{power density per unit area in strongest direction}}{\text{average power density per unit area over all directions}}$$

Combining the concepts above

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

Unless the antenna is very large, like a multi-element array, or a large dish, the gain turns out be close to but only a little more that one. Examples are a gain of 1.5 for a *small dipole* or a *small loop*, or a gain of 1.64 for a *half wave dipole*. Those values should be learned.

5.3 Input impedance

$$Z_i = R_i + jX_i$$

This is just the definition of the expected notation. Effectively we are defining the parameters for a *series equivalent circuit* of the antenna impedance.

5.4 Radiation resistance

The input reactance X_i we will not consider, but R_i for a lossless antenna is the *radiation resistance* R_r, such that if the radiation power is P_r and the input current peak value phasor is I, then

$$P_r = \frac{1}{2}|\mathrm{I}|^2 R_r$$

6. Simple receiving antenna concepts

The *output impedance* of a receiving antenna is the same as the *input impedance* of the same antenna when it is used in a transmitting role. We can say this is just the behaviour of any linear network.

The maximum power which a receiving antenna can deliver is the power delivered to a conjugate matched load. We can say this the maximum power transfer theorem for a linear network. This power is also known as the *available source power*.

The *available source power* of a receiving antenna is *proportional* to the *power density per unit area of the field* in which it sits. This result is a consequence of the *linear behaviour* of the fields, voltages and currents, and the fact that the powers are *quadratic in the field amplitudes*.

6.1 Effective area

This result leads to the *definition* of *effective area of a receiving antenna* A_e by taking the ratio of the available source power to the above power density per unit area.

$$A_e = \frac{\text{Available source power of the antenna in a field}}{\text{Power density per unit area of that field}}$$

So far, this is just common sense. Now we introduce two results which can be proven by the *reciprocity theorem*, which we have not seen, and will not see this year. The results are:

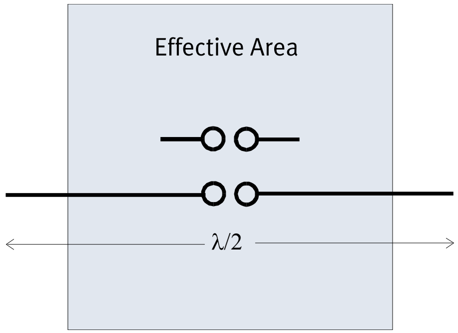
- The effective area of an antenna in its receiving role is proportional to its gain in its transmitting role. Thus $A_{er} \propto g_t$
- The constant of proportionality is $\dfrac{\lambda^2}{4\pi}$

Combining these results gives $A_{er} = \dfrac{g_r \lambda^2}{4\pi}$

Notice the dimensions (units) are expected.

6.2 How large is the effective area?

The effective area could be *larger or smaller* than a *square of side equal to the antenna length*. Consider the cases of a small dipole, and a half wave dipole. See the diagram below. Despite the difference in antenna size, the effective area is about the same, as it depends on gain and wavelength, and the gain does not vary much with size.



6.3 Power transmitted over a communication link.

If we have a communication link with lossless antennas, with transmitted power P_t, transmitter antenna gain g_t, receiver antenna gain g_r, optimum orientation of antennas and a separation between antennas of *r*, and an available received power of P_r, it is easy to combine all of the above results to obtain

$$\frac{P_r}{P_t} = g_t g_r \left(\frac{\lambda}{4\pi r}\right)^2$$

You should produce this result as an *exercise*, as well as the equivalent result below.

$$\frac{P_r}{P_t} = \frac{A_{et} A_{er}}{\lambda^2 r^2}$$

7. How to calculate far fields

Sometimes it is necessary to calculate for a radiating antenna the amplitude of either the *electric field* or the *magnetic field* at a distance *r* which is known to be in the *far field*.

For simplicity, will assume that the radiation is *linearly polarised* and present results for that case.

The answers may be obtained easily by *combining two separate results* for the radiated power density per unit area.

The first of these comes from the Poynting vector for a *uniform plane wave*, which may be expressed for a linearly polarised plane wave either in terms of the *peak value phasor* E representing the electric field or the *peak value phasor* H representing the magnetic field as

$$S_r = \frac{|\mathrm{E}|^2}{2\eta} = \frac{\eta\,|\mathrm{H}|^2}{2}$$

We assume that this result is very well known to attendees.

The second of the results has already been derived in Section 4 of this document, and is re-quoted as

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

The appropriate procedure then is to calculate the power density per unit area produced by the antenna from the second of these results, and then to rearrange the first of these results to obtain the magnitude of the linearly polarised electric or magnetic field which is found in the associated (approximately) uniform plane wave.

We have not done the rearrangement here, as any such rearrangement would be merely a different expression of the same basic truth, and we do not wish to encourage students to commit to memory more than a single version of a basic result. We believe that in practical cases the required rearrangement can be easily obtained after the basic result is first written down.

AUTO-ID LABS

Appendix II
Analysis of Power Transfer at UHF to Very Small Label Antennas

1. Objective

There is in the RFID community frequent discussion of labels with on-chip antennas. That discussion frequently omits important practical issues and limitations to performance that come from the use of very small antennas.

The objective here is to present an analysis in which the penalties paid for using such tiny antennas are evident.

2. Summary

An analysis is performed of the power transfer to UHF labels employing very small magnetic dipole label antennas. A magnetic dipole antenna of size 2mm by 2mm is assumed.

The method used is an appropriate mixture of r*adiating antenna theory* and coupling *volume theory*.

The results show that in the absence of any losses in the antenna, substantial power transfer can take place, but only over extremely narrow bandwidths. Of course these results do not relate to reality, because in practice losses are always present.

When the real losses in practically available materials are taken into account, or a greater operating bandwidth is desired, the picture becomes less rosy.

With a high-quality factor inductance, it is probably possible to obtain an adequate rectified voltage from an unloaded rectifier, but whether we can also get enough output power to run an RFID label is in doubt.

The conclusions depend upon assumptions which need to be made about the feasible rectifier capacitance, achievable quality factor in coils, and label circuit current drain, would need to be revisited if those assumptions, which are quoted in detail later, are found to be in error.

My conclusion is that the argument for using off chip label antennas, which can be made of adequate size, remains strong.

So size really does count!

3. Objective

The objective of these notes is to record a summary of an analysis on the properties of very small RFID label antennas for operation in the far field, particularly at a frequency of 915 MHz.

4. Method of analysis

In this analysis we will make use of both *coupling volume theory* and *radiating antenna theory*. Coupling volume theory is not widely used by workers in the field of RFID, but is reasonably outlined in the main section of these notes.

Radiating antenna theory, commonly used in radar calculations, is appropriate in the context in which labels are placed in the far field, and also when the label antenna size is large enough for the theoretically available source power from a lossless antenna to be actually extracted, or nearly so, within the constraints imposed by the facts:

- that there is some loss, and complete extraction of the available source power is not possible, and
- the radiation resistance of the antenna is accompanied by a reactance which results in the fact that good transfer of power to an external load can only be accomplished over a limited bandwidth.

At a frequency of 915 MHz, those constraints do not intrude too badly for antennas sizes up to about 30 mm by 15 mm.

Coupling volume theory was devised for the situation in which labels are placed in the near, i.e. energy storage, field of a transmitter antenna, and also in the situation in which the radiation resistance of the label antenna is small in relation to the losses in that antenna. When we are operating at 13.56 MHz, both of these conditions are normally satisfied, and coupling volume theory in its pure form is appropriate.

For the situation when labels are placed in the far field of an interrogator antenna, but the labels are so small that their own losses are large in relation to the radiation resistance of the label antenna, it is appropriate to use a hybrid version of radiating antenna theory and coupling volume theory. Radiating antenna theory is used to calculate the energy density at the label position, and coupling volume theory is used to work out what useful power the label antenna can extract from the field.

This is what is done in this Appendix.

5. Range of antennas

The full analysis in the main part of the paper covers both *small electric dipole* and *small magnetic dipole* antennas. Some interesting results on their similarities and differences have already emerged. However, for a number of reasons, this Appendix will cover only magnetic dipole label antennas in the form of *small planar coils*. The reasons for this restriction are twofold. Firstly, experimental results are in practice necessary to define the properties of the most useful shapes of electric dipole antenna, particularly the bow tie antenna. While those experiments have begun, the results are not yet available. Secondly, I have encountered some difficulties in making universal statements about the efficiency of appropriate matching networks between those antennas, the output impedance of which is largely capacitive, and label input impedances, which are themselves also largely capacitive.

These difficulties do not seem to appear in the matching of magnetic dipole antennas, whose output impedance is largely inductive, and label input impedances, which as just stated are substantially capacitive. It appears to be possible in many cases to arrange for resonance between these two elements to occur at the frequency of operation. It is well known in coupling volume theory that such resonance is required to make good use of the power transferred to the label antenna.

6. Single and multi-turn inductors

In power transfer calculations it is often possible to base results on a calculation based on a single turn inductor. This comes about because many properties of the inductor, including its inductance and its loss resistance, will vary as the square of the number of turns, provided the same amount of surface area is allocated to the provision of those turns as was allocated to the provision of a single turn. This is one of the interesting properties of coupling volume theory.

Of course in some aspects of the calculation, we become interested not merely in the power transferred, but also the particular voltages and currents at which the power transfer occurs. That is, we become interested in particular independence levels. In such cases, it becomes necessary to consider whatever number of turns may be required to achieve resonance with particular load impedances.

7. Dimensionless ratio analysis

In the spirit of Buckingham's pi theorem, we will work, as far as possible, with dimensionless ratios. Quality factors of various resonant circuits, which are dimensionless, and power transfer ratios, which are also dimensionless, will therefore frequently appear in the analysis.

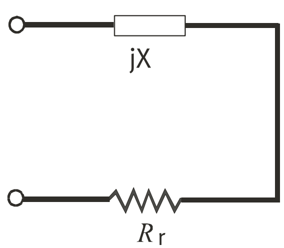
8. The role of junction capacitance

At some stage in the powering of a label, we do become interested in the amount of power transferred to that part of the label beyond the rectifier. If we are going to operate in the far field, we need real power to generate a reply, and to radiate that reply.

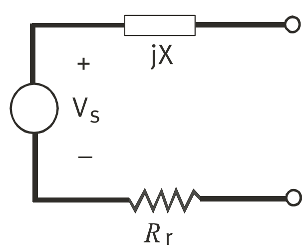
But before any of that can happen, we do need to produce some rectification. Practical rectifiers have a junction capacitance which must be periodically charged and discharged to a sufficient potential for rectification to occur. That process involves an exchange of reactive power between the junction capacitance and some complimentary energy storage element. This process has been dealt with at length in U.S. Patent 5,305,008. One of the benefits of magnetic dipole antennas is that the self inductance necessarily present in the antenna can be that energy storage element provided it has been suitably adjusted, and suitable adjustment is available by appropriately choosing the number of turns.

9. Antenna equivalent circuits

The figure below shows equivalent circuits for an electrically small antenna, operating at the left in its transmitting role, and at the right in its receiving role. It shows that in both cases there is a radiation resistance in series with an antenna reactance. The same resistance and reactance is found in both circuits.



Transmitter



Receiver

For an electric dipole antenna, the reactance is negative, and for a magnetic dipole antenna the reactance is positive.

From these elements we can construct a radiation quality factor

$$Q_r = \frac{|X|}{R_r}$$

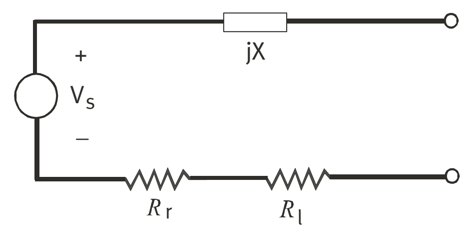
10. Scaling with antenna size

We can show that for both electric and magnetic dipole antennas the radiation quality factor scales as the *universe third power* of antenna size.

This means that as either antenna becomes very small, a relatively speaking increasingly large reactance stands between the radiation resistance and any external load to which we might wish to match it.

11. Effect of antenna losses

When losses are to be taken into account the antenna will also have a loss resistance R_l so its equivalent circuit becomes modified to that shown below.



Receiver

The optimum load impedance, previously $R_r - jX$, now becomes $R_r + R_l - jX$, and the power which can be delivered to that load impedance is reduced.

When the antenna becomes very small, the radiation resistance becomes very much less than the loss resistance. It is in this situation we have the option of applying coupling volume theory to the determination of how the circuit behaves.

In this theory, the source voltage in the above antenna circuit might as well be calculated from Faraday's law, the radiation resistance neglected, the self inductance calculated from the magnetostatic formula, and the loss resistance be determined by the reasonably familiar methods, those methods, however, taking into account that conduction will only occur within a skin depth of the metal surface.

12. Practical values and approximations

12.1 Operating frequency

We have already announced that our calculations will be based on a frequency of 915 MHz.

12.2 Antenna material

The antenna material will for a number of reasons be assumed to be copper with an electrical conductivity of 5.8×10^7 S/m.

12.3 Antenna size

We will assume our antenna to be a square of side 2mm. This will not prevent us from the making some approximations including where convenient replacement of the square by circle of equal area.

12.4 Antenna shape

Although the antenna coil will undoubtedly be realised as a planar structure with a flat strip conductor, we will be cheerful about replacing the flat strip by an equivalent round wire.

12.5 Rectifier junction capacitance

A principal target of the analysis will be the determination of the feasibility of providing adequate reactive power to the junction capacitance of a small rectifier diode. We will assume that a diode depletion layer capacitance of 0.1pF can be achieved. This capacitance is believed to be within the range of achievement of modern technology.

12.6 Interrogator power

We will assume that our interrogator radiates a power of 1W through an antenna of gain 4 (6dB), as is permitted in the FCC frequency hopping regulations.

12.7 Interrogation distance

We will assume an interrogation distance of 1m, which is well into the far field, but is not at such a distance as to represent an unfair challenge to the feasibility of the system we are seeking to investigate.

13. Skin depth calculation

For calculation of the skin depth K we use the formula

$$\delta = \frac{1}{\alpha} = \sqrt{\frac{2}{\omega\mu\sigma}}$$

Using this formula, we have calculated the skin depth for copper at a frequency of 915MHz to be 2.18μm.

Some remarks may be made on this result. Firstly it is feasible to deposit material to at least this depth. Secondly, depositing more metal will not produce a benefit as the metal deeper than one skin depth will not contribute to the conduction.

It follows that all resistance calculations should be based on the assumption that inductors of all sizes are made from strip material of this thickness, even if curved into a circular form to resemble a round but hollow wire.

14. Surface resistivity

A simple step towards the calculation of the coil resistance involves the determination of the surface resistivity

$$R_s = \frac{1}{\delta\sigma} = \sqrt{\frac{\omega\mu}{2\sigma}}$$

For the selected material and frequency the surface resistivity turns out to be 7.909 mΩ per square.

15. Resistance

Using the surface resistivity obtained above and a diameter ratio of 5, we will calculate a coil loss resistance of $R_l = 39.55\text{m}\Omega$.

16. Scaling of antenna resistance

Provided we do not change the number of turns, an antenna in the form of a single circular coil of round wire in which the ratio of the diameter of the circle to the diameter of the wire is a constant will have a constant resistance, independent of size.

17. Scaling of inductance

Provided we do not change the number of turns, an antenna in the form described immediately above will have an inductance which scales as the first power of size.

18. Scaling with number of turns

If the single turn coil is replaced by coil of *N* turns, both the self inductance and the coil resistance will both scale as the second power of *N*, and the series induced voltage will scale as the first power of *N*. Most of the calculations below, such as the available source power, the short-circuit reactive power, the coupling volume, and the quality factor will be unaffected, and might as well be based on calculations for a single turn coil.

19. Properties of a single turn coil

19.1 Self-inductance

For the self-inductance we will use the familiar formula

$$L = \frac{\mu_o D}{2}\left[\log_e\left(\frac{8D}{d}\right) - 2\right]$$

which applies to a circular coil of diameter *D*, made from a round wire of diameter *d*, and apply it to such a coil which approximates, through having the same area, a single turn square coil of size 2mm by 2mm, and will use a diameter ratio *D/d* of 5. The result of our calculation is a coil self-inductance *L* of 2.122nH.

19.2 Coil reactance

At the planned operating frequency of 915MHz, the reactance *X* of this self-inductance is 12.20Ω.

19.3 Radiation resistance

Although we propose to use coupling volume theory, in which radiation resistance is assumed to be negligible, we will assure ourselves that this is so by performing a calculation. Using the formula

$$R_r = 20\pi^2 (\beta a)^4 \ \Omega$$

for a circular coil of radius a, we obtain at a frequency of 915MHz, a radiation resistance of

$$R_r = 43.10 \ \mu\Omega$$

Thus we do see that the radiation resistance is indeed negligible compared with the antenna coil loss resistance, and our selection of coupling of volume theory for the analysis of the label antenna is appropriate.

19.4 Radiation quality factor

From the coil reactance and the radiation resistance we can calculate the radiation quality factor Q_r of 283,063. This calculation has, however, no practical significance, as the coil losses substantially exceed the radiation resistance, as is clear from the calculation in the next paragraph.

19.5 Inductor quality factor

From the coil reactance and the just calculated resistance we can calculate an inductor quality factor which turns out to be 308.

20. Comment on this calculation

This calculated value has turned out to be quite high, as experiments with even larger coils tuned by the depletion layer capacitance of a semiconductor diode gave values more of the order of 40. Of course in such experiments there may have been significant other sources of loss in addition to the conduction losses in the copper coil.

Because the reactive power which we are able to supply to the junction capacitance of the rectifier depends strongly upon the overall quality factor of the tuned circuit formed by the inductor and that depletion layer capacitance, we will probably perform two calculations, one for a quality factor of 308, and another for a quality factor of 40.

21. Reactive power density per unit volume

Because the label is assumed to be in the far field, we will begin our calculation of the reactive power density per unit volume by a calculation of the real power flow per unit area using the radiation antenna theory formula

$$\text{Power flow per unit area} = \frac{g_t P_t}{4\pi r^2}$$

It is easily shown that the reactive power density per unit volume can be obtained from the real power flow per unit area, simply by multiplication by the propagation constant β. Thus we are able to obtain the formulae below.

$$\text{Reactive power density per unit volume} = \beta \times \text{power flow per unit area} = \frac{\beta\, g_t P_t}{4\pi r^2}$$

The result of all these calculations, under the assumptions of interrogator power, antenna gain, and interrogator to label distance defined previously, is

Reactive power density per unit volume = 6.1VAm^{-3}.

What we have just calculated is an appropriate, power-like, measure of the field strength available for energising the label.

22. Label coupling volume

We remind ourselves that the definition of label coupling volume V_c is

$$V_c = \frac{\text{Reactive power in the label inductor when short circuit}}{\text{Reactive power density per unit volume at label position}}$$

and that it may be shown that the coupling volume for a planar coil of N turns is given by

$$V_c = \frac{\mu_0 N^2 A^2}{L}$$

For the coil of self inductance calculated above, and assuming a single turn, with a flux collecting area of 1.6mm by 1.6mm, we find we have a coupling volume of

$$V_c = 3.881 \text{ x } 10^{-9} \text{ m}^3$$

We remember that the coupling volume will not change with the number of turns, so we will be free to vary the number of turns later to achieve a different inductance if that is desired to produce resonance with the depletion layer capacitance of the rectifier.

23. Reactive power in short circuited label

Forming the product of the coupling volume just calculated and the reactive power density per unit volume calculated earlier gives us the reactive power which would flow within the label coil when it is short-circuited. The result of the calculation is short circuit reactive power of 23.67nVA. We note that the coupling volume scales as the third power of size, so if more short circuit reactive power is required, we can make the label bigger.

24. Power delivered to a tuned label

To estimate the power which will be delivered to whatever losses exist within the label when it is tuned to resonance, we need multiply the reactive power in a short circuit label antenna by the quality factor of the resonance.

We have agreed earlier that we will perform this calculation for two values of quality factor.

If we assume from our previous quality factor calculation a Q of 308, we will conclude the power of 7.29μW can be delivered to the losses of the tuned circuit.

If we assume that various factors have limited our quality factor to 40, we would predict that we can deliver power of only 947nW to the losses of the tuned circuit.

Just where the power goes must be clarified. If all losses are in the coil, the power goes to those losses. If we then ask what power can we extract to a fixed load placed beyond those losses, then the answer is one-quarter of the coil losses just calculated. This corresponds to the Q = 308 case. If, however, the substantial quality factor reduction has occurred through the introduction of a real load to which we wish to deliver power, then the calculated power is going substantially to that real load. This could be the Q = 40 case.

25. Reactive power in tuned coil

To determine the reactive power which flows in the inductor, and also in the capacitance that it feeds, assuming those two things are resonant, we can either multiply the power calculated in the previous paragraph by the assumed quality factor, or we can multiply the reactive power in the short-circuited label by the square of the assumed quality factor. For the two values of quality factor for which we are doing calculations, we obtain reactive powers of 2.25mVA and 37.88μVA respectively.

26. Reactive power needed in depletion layer capacitance

If we assume that our rectifier requires a 1.5V r.m.s. voltage across it for a useful output, and we assume that the capacitance is the previously mentioned 0.1 pF, we see we need a reactive power of 1.294mVA.

27. Some conclusions

- if the quality factor of 308 can in fact be achieved we can generate a reactive power in excess of that needed to service the junction capacitance of the rectifier diode.
- That being the case we have room for other sources of loss, or for genuine output power.
- Such other losses or genuine output power could be from:
  - access resistance loss in the diode;
  - losses in the rectification processes, such as junction conduction and minority carrier persistence;
  - any power output required from the rectifier;
  - losses from the magnetic fields of the self inductance encountering nearby objects in the environment and inducing eddy currents on their surfaces;
  - dielectric losses from the electric field between the inductor terminals.
- the quality factor which is required is substantially above what has been observed in antenna structures of about 40 mm by 15mm.
- The power we have calculated as being able to be extracted from the antenna at either the quality factors of 40 or 308 is still very much less than the theoretically available source power which a completely lossless receiver antenna at this distance could provide. That power would be given by the farfield relation

$$\frac{P_r}{P_t} = g_r g_t \left(\frac{\lambda}{4\pi r}\right)^2$$

  which would predict a power transfer ratio of 4.8×10^{-3}, and therefore an output power of 4.8mW.
  Of course such a power transfer ratio is in practice unachievable in the face of losses in the small label antenna.

- When the label antenna is sharply tuned, it is not clear what manufacturing tolerances can be achieved in order to keep it on tune. Nor do we know how much the environment will detune the antenna. The latter problem will be less severe if the antenna is tiny, but the first problem will be present with all label sizes.
- Estimates have been sought from circuit designers of the power likely to be required to actually run a backscatter label circuit. A power of about 20μW has been suggested. That power is in excess of the 1.82μW (one quarter of 7.29μW) which could be extracted from the label even if the coil quality factor is as high as 308.
  For an anharmonic label, as opposed to a backscatter label, we cannot provide from recent experience an estimate of the power required, but consultation with expert designers confirms my opinion that it will be substantially in excess of that required for a backscatter label.

28. Effects of size increase

If we contemplate an increase in label size the radiation resistance will increase as the fourth power of the size increase, but will still in practical cases remain significantly less than the loss resistance in the circuit. The loss resistance, with constant skin depth, will remain the same. The amount of power which can actually be extracted will increase as the fourth power of the size increase.

This fact explains more than anything else why we have never before contemplated tiny antennas.

Probably the only beneficial effect which might be claimed for using tiny antennas is that they are easier to prevent from becoming environmentally detuned because their self-fields will extend only a small distance from the antenna, but we are not sure without detailed calculation whether or not this argument is valid. Certainly the energy storage field extends out to a distance equal to the radius of the radian sphere, which at the frequency of consideration has a radius of about 50mm. But because the fields of the coil are not quite the same as those of an infinitesimal dipole, we would have to look carefully at the way in which the stored energy varies with distance from the coil, and how much of it is within certain distances, and we are prepared to believe that some validity to the argument could be found.

29. Some general conclusions

It is perhaps surprising to find that if one assumes that only copper losses are present in a very small magnetic loop antenna, it should be possible to provide at least the reactive power required to service the junction capacitance of a rectifier.

We cannot offer any experimental evidence that the actual losses in the label circuit will be anywhere near as low as was calculated when only copper losses were assumed.

Certainly when one considers the problem that a reasonably low antenna quality factor may be required to provide for manufacturing tolerances which might vary the resonant frequency, or we make the assumption that significant power must be supplied to the circuits beyond the rectifier, it is difficult to continue to sustain the belief that very tiny antennas will suffice.

Provided that the basic assumption of the coupling volume theory, namely that the loss resistance within the label antenna tuned circuit is significantly greater than the radiation resistance of that antenna, is satisfied, we have complete confidence in the validity of the hybrid analysis using radiating antenna theory for the calculation of the field at the label position, and coupling volume theory for the determination of what label power can be extracted from that field.

In this connection we remind ourselves that the main paper has established an equivalence between radiating antenna theory and coupling volume theory and that the results of this appendix could have been obtained from radiating antenna theory alone.

Appendix III
Useful Formulae

In addition to the formulae quoted in the body of the paper, the following relations are often useful.

1. Physical constants

$\mu_0 = 4\pi \times 10^{-7}$ H/m by definition; $\varepsilon_0 = 8.854$ pF/m; k =1.380 650 x 10^{-23} J/K.

The conductivity σ of copper is 5.8×10^{7} S/m.

2. Inductance Calculations

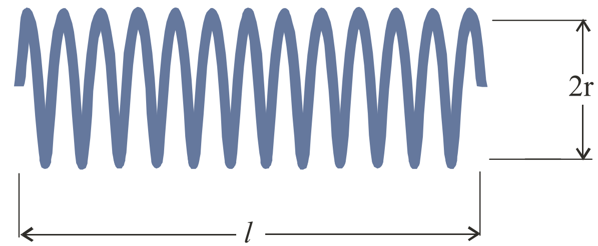
2.1 Planar Circular Coil

The self inductance of a single-turn circular coil of diameter *D* made from wire of diameter *d* is, when the currents flow on the surface, given by

$$L=\frac{\mu_o D}{2}\left[\log_e\left(\frac{8D}{d}\right)-2\right]$$

2.2 Solenoidal coil

An empirical but useful formula for the self inductance of a thin wire solenoidal coil of *N* turns wound over a length of *l* on a former of diameter of 2*r* as shown in the diagram below is



$$L=\frac{\mu_0 \pi r^2 N^2}{l+0.9r}$$

2.3 Twin Wire Line

The self inductance *L* of a twin-wire line in which the conductors have diameter *d* and separation *s* is given by

$$\begin{aligned} L &= \frac{\mu_o}{\pi}\,\text{arc cosh}\left(\frac{s}{d}\right) \\ &\approx \frac{\mu_o}{\pi}\log\left(\frac{2s}{d}\right) \qquad \textit{when } s >> d \end{aligned}$$

3. Axial Field of a Circular Coil

In the magnetostatic approximation, the field at a point at a distance *z* along the axis of a single turn circular coil of radius *a* carrying a current *I* is given by

$$H_z(0,0,z) = \frac{Ia^2}{2\left(a^2+z^2\right)^{\frac{3}{2}}}$$

4. Skin effect

Skin depth δ in a metal at an angular frequency ω is given by

$$\delta = \frac{1}{\alpha} = \sqrt{\frac{2}{\omega\mu\sigma}}$$

The surface resistivity R_s per square due to skin effect is

$$R_s = \frac{1}{\delta\sigma} = \sqrt{\frac{\omega\mu}{2\sigma}}$$

and the wave impedance η at the surface is

$$\eta = (1 + j)R_s$$

5. Radiation Resistances

5.1 Electric dipole

The radiation resistance of a short electric dipole of length L, operating at a frequency for which the free space propagation constant has magnitude β, is given in Ohms by

$$R_r = 20(\beta L)^2$$

5.2 Magnetic dipole

The radiation resistance of a small current loop of radius a, operating at a frequency for which the free space propagation constant has magnitude β, is given in Ohms by

$$R_r = 20\pi^2 (\beta a)^4$$

Small loops of other shapes but the same area have the same radiation resistance.

6. Fourier transforms

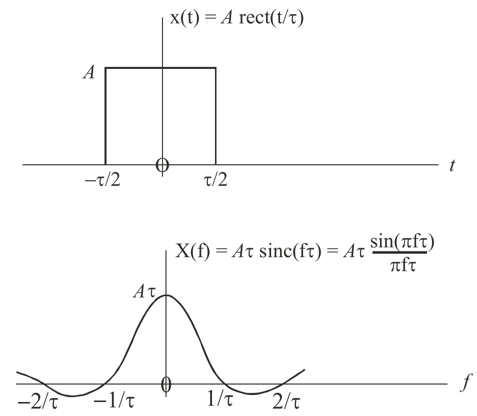
6.1. Useful reminders

We take the opportunity to remind ourselves that (i) calculating Fourier series for a quantity does not change its units, (ii) performing a Fourier transform does, and that (iii) delta functions of time have units of inverse time.

6.2 The rectangular function

The Fourier *transform* of the rectangular function A rect (t/τ), of amplitude A and duration D centered on the origin of time axis, is the sinc function of amplitude $A\tau$, centered on the origin of the frequency axis.

Both the rectangular function and its Fourier transform are illustrated in the figure below.



As shown in the figure, the sinc function is defined by

$$\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$$

The Fourier transform has zeros at integer multiples of the inverse of the pulse length.

6.3 The isolated delta function

The Fourier transform of an isolated unit amplitude delta function $\delta(t)$ (which we note has units of inverse time) occurring at the origin is just the dimensionless constant 1.

6.4 A constant

The Fourier transform of dimensionless constant 1 is an isolated unit amplitude delta function $\delta(f)$ (which we note has units of time) occurring at the origin of the frequency axis.

6.5 A periodic function

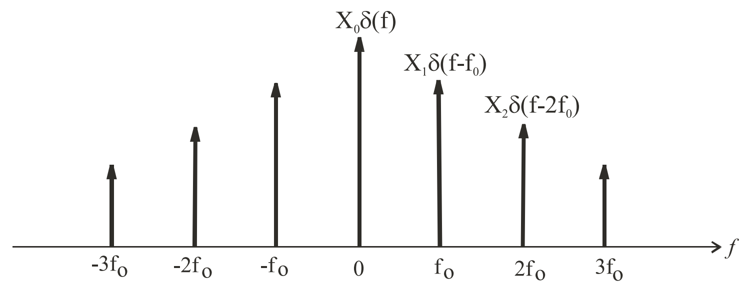
If $x(t)$ is a dimensionless periodic function of period T centered on the origin of the time axis and $f_o = 1/T$ is its fundamental frequency, and X_n are the (dimensionless) values of the Fourier series, i.e.

$$X_n = \frac{1}{T} \int_{T/2}^{-T/2} x(t) e^{-j2\pi f_0 t} dt$$

we can, with the aid of the delta function, define for the periodic function the Fourier transform

$$X(f) = \sum_{n=-\infty}^{\infty} X_n \delta(f - nf_0)$$

which we see has the correct units of inverse time. The transform is illustrated in the Figure below.



6.6 The comb function

When the periodic time function is just a regular sequence of unit amplitude delta functions, each of dimensions of inverse time (as is appropriate), occurring at intervals a multiple of T, then in the above results each element X_n of the Fourier series is equal to $1/T$ (and has units of inverse seconds). We will denote this series of delta functions by $\mathrm{comb}_T(t)$.

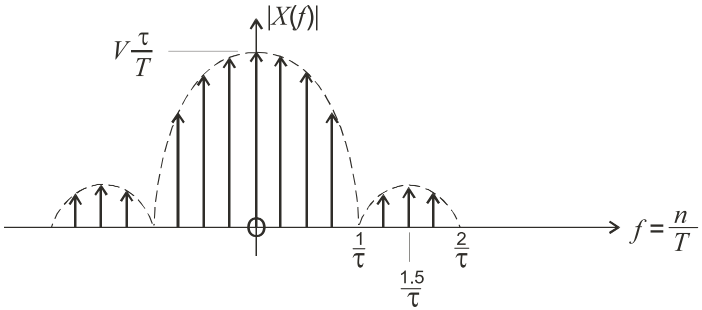
From the above results we find its Fourier transform is a dimensionless comb function in the frequency domain in which the delta functions have amplitude $(1/T)$ and frequency separation $F=1/T$. We denote this sequence by $(1/T)\ \mathrm{comb}_F(f)$. Taking into account the fact that the factor $(1/T)$ has dimensions of inverse time, and a comb of unit amplitude delta functions in the frequency domain has dimensions of inverse frequency, the resulting transform $(1/T)$ $\mathrm{comb}_F(f)$ has no units.

6.7 A repeated sequence of pulses

From the above results and the facts that

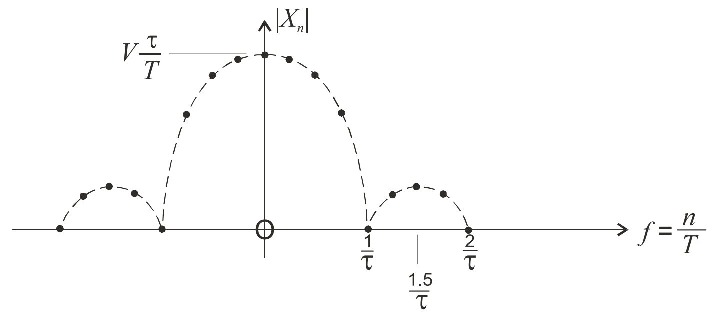
- a repeated sequence of rectangular pulses in the time domain can be regarded as the convolution in the time domain of a single pulse and a sequence of unit amplitude delta functions, and
- the fact that convolution in the time domain is equivalent to multiplication in the frequency domain

we conclude that the spectrum of the result is the product of the Fourier transform (which has units time) of the original isolated rectangular pulse and the unit-less transform (1/T) $\text{comb}_F(_f)$.



This result is illustrated in the figure above, which shows the Fourier *transform* (a sequence of delta functions) for a sequence of rectangular pulses of amplitude V volt and duration τ, repeated at interval T.

The result could alternatively have been obtained by a straightforward calculation of the Fourier *series* for the above described sequence of pulses, the result being illustrated in the figure below.



In the representation of the Fourier *transform*, which has units of V_s, we use delta functions of frequency, which have those units.

In the representation of the Fourier *series*, which have the units of volts, we use dots.

In both cases we have simplified the diagrams by assuming that T is a multiple of τ.

7. Frequency modulation

7.1 Fourier series

The Fourier series for a FM signal $x_c(t)$ of amplitude A_c, carrier frequency f_c, modulated with a modulating signal of amplitude A_m and modulating frequency f_m is given by

$$x_c(t) = A_c \sum_{n=-\infty}^{n=\infty} J_n(\beta)\cos(\omega_c + n\omega_m)t$$

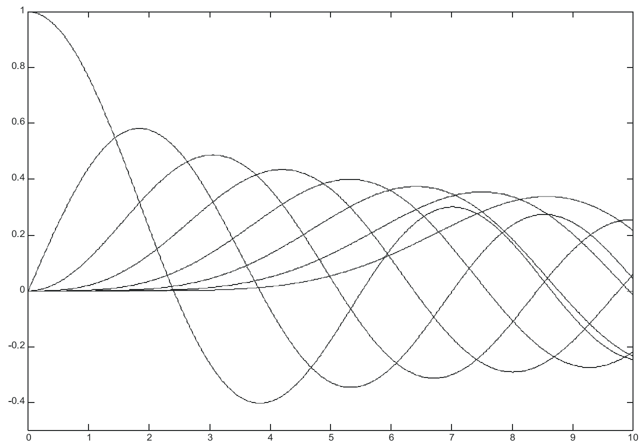
where $J_n(\beta)$ is the Bessel function of the first kind of order n and β is called the modulation index and is given by

$$\beta = \frac{f_\Delta A_m}{f_m}$$

where f_Δ is the maximum deviation in frequency produced by the modulation.

7.2 Bessel functions

The Bessel functions are illustrated in the figure below.



It may be observed that for large modulation index, which may be achieved by having a large ratio of f_Δ to f_m, that

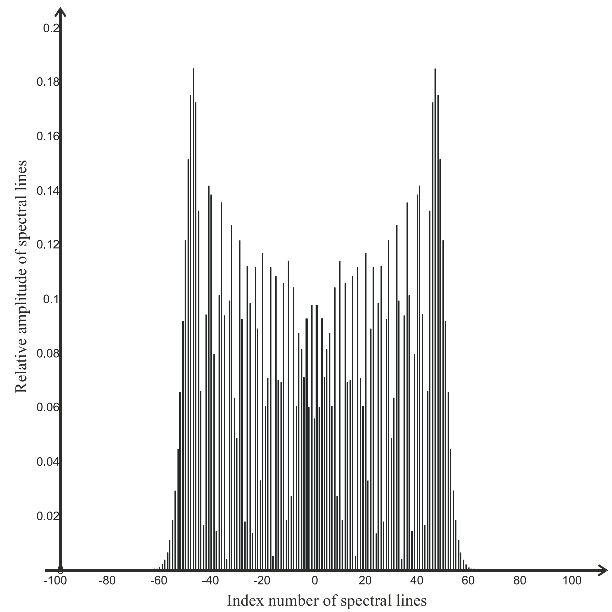
- all of the sidebands may be kept small, and
- sidebands above a certain order are negligible.

7.3 Application to spreading spectrum of regular pulse signaling

To illustrate what is possible we have plotted in the figure below the relative amplitude of spectral lines generated for a single spectral line when a maximum deviation of 5kHz is used with a sinusoidal modulating signal at a frequency of 100Hz.

7.4 Significant question

Would our original spectral lines be sufficiently far (10kHz) apart for this spreading to be useful? If our signalling were based on regular pulses 40μs apart, they would be separated in their spectrum by lines by 25kHz apart.



AUTO-ID LABS

Appendix IV
Exercises

1. Objective

This and the next Appendix are designed to assist readers new to the subject to better understand some of the important concepts through the performance of simple exercises.

The exercises in the next Appendix are somewhat simpler than those in this one.

2. Skin depth calculations

Determine the skin depth of copper at

(a) 13.56 MHz
(b) 915 MHz.

3. Loss resistances of small antennas

(a) For a single-turn square planar coil antenna of outer di-men-sions 40mm square made from copper strip material of width 5mm and operating at 13.56MHz, determine the series loss resistance R_l.

(b) For a single-turn square planar coil antenna of outer dimensions 160mm square made from copper strip material of width 5mm and operating at 13.56 MHz, determine the series loss resistance R_l.

(c) For a single-turn square planar coil antenna of outer dimensions2 mm square made from copper strip material of width 0.25mm and operating at 915 MHz, determine the series loss resistance R_l.

4. Radiation resistances of small antennas

(a) For the above described single-turn square planar coil antenna of outer dimensions 40 mm square made from copper strip material of width 5 mm and operating at 13.56 MHz, determine the series radiation resistance R_r.

(b) For the above described single-turn square planar coil atenna of outer dimensions 160 mm square made from copper strip material of width 5 mm and operating at 13.56 MHz, determine the series radiation resistance R_r.

(c) For the above described single-turn square planar coil antenna of outer dimensions 2 mm square made from copper strip material of width 0.25mm and operating at 915MHz, determine the series radiation resistance R_r.

5. Efficiency of small antennas

For each of the above small antennas determine the efficiency $R_r/(R_l + R_r)$.

6. Quality factors of small label antennas

(a) For each of the above small loop antennas operating at the stated frequencies, determine the quality factor and the 3dB band width.

(b) State whether the bandwidth is determined mainly by the radiation resistance or the loss resistance.

(c) Comment on the band width in relation to a desired communication bandwidth of 400kHz.

7. Power density in two forms

A small loop antenna radiates a power of 5mW. The antenna has the gain of 1.5 of a small magnetic dipole, and the radiation istrongest in the equatorial plane. Determine for this antenna:

(a) The power density Sr per unit area carried by electromagnetic waves away from this point that is at a distance of 10 m in the equatorial plane from the loop centre.
(b) The peak value of the magnetic field at this point.
(c) The reactive power density W_v per unit volume if we were in the near field, and the same value of magnetic field were present.

8. Range extension by increase in power

An HF interrogation system with a transmitter antenna in the form of a large circular loop of diameter 1m is operating at a range of 1m with a current in the loop that causes a radiated power of 200μW.

(a) Using the magnetostatic approximation for the magnetic field on the loop axis, and the curves supplied in this paper, determine the expected range if the radiation were increased to 500mW.
(b) Is this distance still in the near field?

9. Evaluation of signalling compliance at HF

(a) Perform the calculations that should confirm that when the near-field HF interrogation system described in Section 10 of the accompanying notes is operating at the limit of compliance in its carrier signal level, its sideband signals resulting from interrogation signal amplitude modulation are also compliant.
(b) By studying the effects of the frequency modulation of an HF interrogator signal by a 100Hz sinusoid producing a deviation of 5kHz in the interrogation carrier as described in Appendix II, determine by how much signalling sidebands of the interrogation signal of the form described above can be reduced.

10. Pulsed interrogation at UHF

A short range radio transmitter designed for sale in the USA radiates, from a linearly polarised patch antenna of gain 5dB, within the ISM band extending from 902 to 928MHz, a pulse amplitude modulated 915MHz carrier wave with the rectangular modulation envelope illustrated in Figure 1. In its modulation cycle, the carrier wave is on for a period of 10ms and off for a period of 90ms. The repetition interval T of the modulation cycle is thus 100ms.

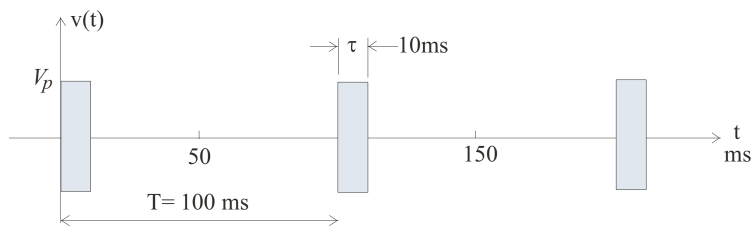


Fig. 1: Illustration of periodic interrogation pulses

One of the electromagnetic compatibility regulations within the said ISM band allows such pulsed radiation, but requires that the electric field strength of the radiated signal, as measured by an average detector, be limited to a value of 500mV/m r.m.s. at a distance of 3m.

(a) Illustrate the spectrum of the interrogator signal. You should take note of the fact that the carrier frequency is an harmonic (albeit a large one) of the frequency of the modulation envelope.

(b) Determine, for the modulation as described above, the ratio of the average amplitude of the electric field at the position of a receiving antenna which is indicated by an average detector with an averaging time long compared with the period of the modulation cycle to the amplitude of the interrogation electric field at the same position during the on-period of the carrier wave.
(c) Neglecting any effects from ground reflection, determine the power allowed to be radiated by the transmitter during its on period.
(d) Assume that the effect of the ground plane is to produce, at the position of the receiving antenna, in addition to a direct signal from the transmitter antenna, a single ground reflected signal, and that the transmitter antenna is angled so that both the direct signal and the reflected signal are radiated with a strength approximately equal to the on-axis signal for the transmitter antenna. Determine the effect of the ground plane on the amplitude of the signal reaching the receiving antenna.
(e) Hence determine the power output from the transmitter allowed during its on period when ground reflections are taken into account.

11. Comment

We should point out that this exercise was conceived in an era when, in calculating the regulatory compliance in some jurisdictions of a RFID system, the effect of ground reflections needed to be taken into account in the way outlined here.
However, electromagnetic compatibility regulations are being increasingly cast in terms of “conducted measurements” or “substitution measurements” in which the effects of ground reflections no longer serve to reduce the allowed power that can be radiated below that which would be allowed if the ground reflections did not occur.

Appendix V
Simple Exercises

1. Objective

This and the previous Appendix are designed to assist readers new to the subject to better understand some of the importance concepts through the performance of simple exercises. The exercises in this Appendix are somewhat simpler than those in the previous one.

2. Calculate the following

(a) The free space electromagnetic wavelength at 1 GHz.
(b) The propagation constant for electromagnetic waves at 1GHz.
(c) The size of the radian sphere at 1GHz.
(d) The radiated power density S_r in Wm^{-2} at a distance of 2m in the direction of strongest radiation for a linearly polarised reader interrogator antenna of gain 4 transmitting a power of 1W at 1GHz.
(e) The effective area of a label antenna of gain $\pi/2$ at 1GHz.
(f) The available source power from that antenna placed in the field of the first.