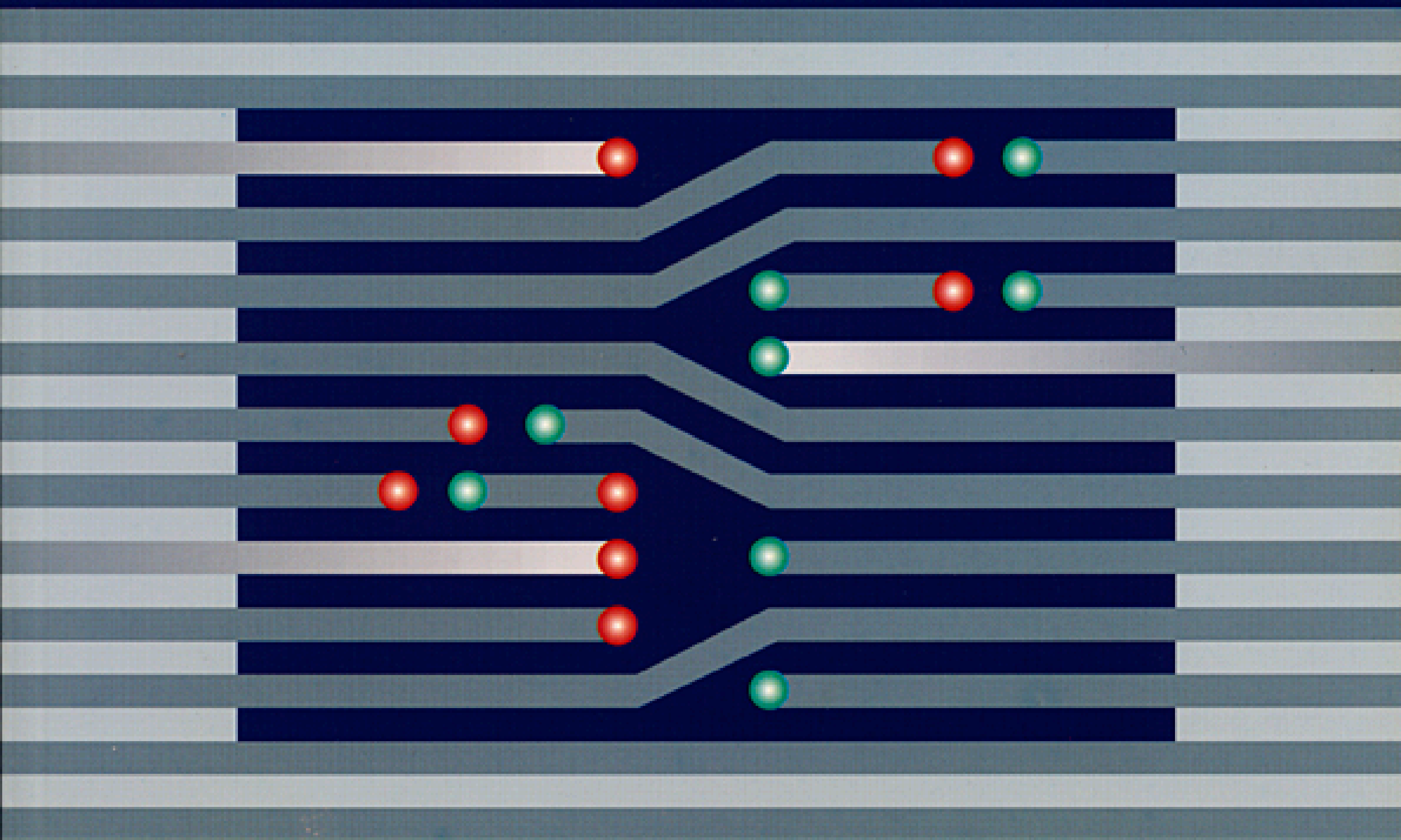
Telektronikk 4.94

Electromagnetic compatibility



Contents

Feature

Special

Status

Telenor

Telektronikk

Volume 90 No. 4 - 1994
ISSN 0085-7130

Editor:
Ola Espvik
Tel. + 47 63 80 98 83

Status section editor:
Tom Handegård
Tel. + 47 63 80 98 00

Editorial assistant:
Gunhild Luke
Tel. + 47 63 80 91 52

Editorial office:
Telektronikk
Telenor AS
Telenor Research
P.O. Box 83
N-2007 Kjeller, Norway

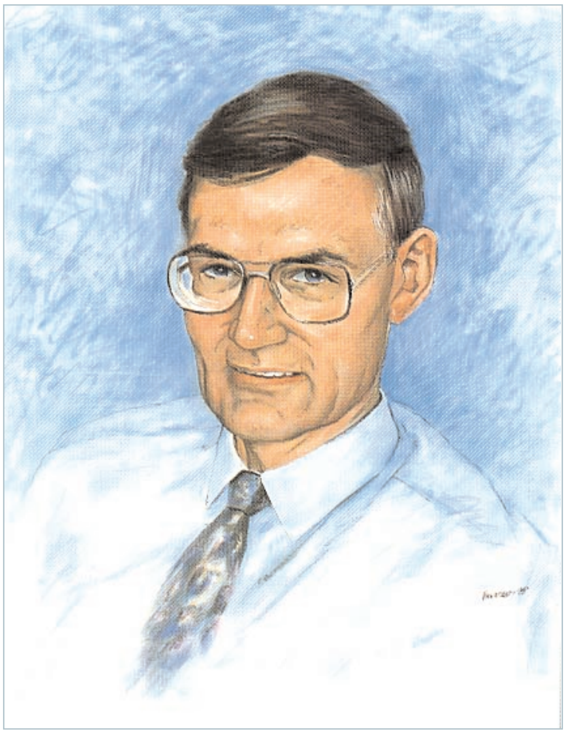
Graphic design:
Design Consult AS

Layout and illustrations:
Gunhild Luke, Britt Kjus, Åse Aardal
Telenor Research

Guest editorial

BY EGIL HAUGER
SENIOR ENGINEER, TELENOR RESEARCH

In 1895, when Marconi started wireless communication in Italy, he may have observed what we today define as EMC – electromagnetic compatibility problems.



Electromagnetic interference (EMI) a hundred years ago came mainly from natural sources such as lightning. But the electrification of the railway in the 1920s and the introduction of electric tramways in the big cities caused electromagnetic compatibility problems to the new communication medium – wireless broadcasting.

So, as long as man has been using the electromagnetic spectrum as a medium for communication, there have been EMC problems. What then highlights EMC problems today? The answer may be found in the statistics of the number of PCs, the number of mobile telephones, the number of pagers, and the amount of equipment using the radio spectrum for communication. The IT revolution has had a tremendous impact on the total pollution of electromagnetic energy in the radio communication part of the electromagnetic spectrum.

The IT revolution together with the digitization of the telecom network will bring the source of radio noise, the interferer, to every home in the industrialized countries in the years to come. Looking back 10 years when the limits of interference were laid down, the electromagnetic environment was quite different from today. The protection distance of 10 m was realistic in those days, but what is the reality 10 years later, and what can be seen on the horizon from a radio engineer's point of view.

The scenario for the year 2000 and beyond, and which for many people is a reality today, is the personal radio equipment. Technology has now enabled us to communicate with anyone, anywhere, anytime. But what about the communication link which we must all share? Is there enough space for everyone to travel along the same link at the same time with no restrictions on required bandwidth and with no restrictions on use wherever we want to use our radio communication link?

Today, there are areas where the use of mobile telephones or electronic equipment with active transmitters is prohibited, such as in hospitals and aircraft. These restrictions are for safety reasons and are not caused by the use of the radio spectrum as such.

If we go into the actual EMC situation today, where do we have the most pressing conflict between users of different radio equipment and other electronic devices? One of the areas which have been focused on in the media, is the possibility of dramatic situations arising if for example the airbag of a car is inflated because of incoming calls to the mobile phone when driving along the motorway, or a person's heart implant stops on the sound of the first signal from the mobile phone in his pocket. These dramatic situations have all been taken care of by the manufacturers of safety equipment, so there is no need to worry.

As mentioned above, the use of personal radio equipment constitutes a more significant EMC problem. If you are working at your PC, the reason you do not receive the beep-beep from the pager carried in your belt may be EMC problems. The same thing may apply when your colleague in the next office uses his GSM telephone and it disturbs your conversation in the old fashion fixed line telephone. The regulations authorities actually assume a protection distance of 10 m from an electromagnetic interferer to a radio receiver. The 10/1 distance ratio increases the interference by 20 dB!

What about the legal aspect of EMC and regulations authorities? In Europe, the EU has been the driving force since the 89/336/EU EMC directive "Council Directive on the approximation of the laws of the Member States relating to electromagnetic compatibility". This directive was to be implemented within the EU/EFTA area (18 countries) by January 1, 1992, with a transition period until December 31, 1992. That was the situation in 1989. In the period from 1989 to 1992 there were heavy protests from the European IT industry claiming that the time schedule for fulfilling the objectives was unrealistic and pointing at the lack of measurements facilities and administrative precautions taken to ensure the implementation of the directive. The result was, as we all know, that the date of implementation was put off until December 31, 1995. From that date no electronic equipment whatsoever is allowed to be placed on the European market without fulfilling the EMC directive. This includes restrictions on electromagnetic radiation as well as immunity. The discussion on the limits of both interference and immunity level will continue, but from a radio engineer's point of view, a great step has been taken to ensure proper operation of all kinds of electronic equipment.

Looking ahead, we see a lot of work still to be done in the area of EMC. The use of the electromagnetic spectrum for communications purposes will increase dramatically in the years to come, and the potential for conflicts between different users will increase accordingly. We must all be aware of the fact that the electromagnetic spectrum is a common resource for us all; its use cannot be based upon the rights of the strongest, but must be regulated in a manner which benefits everybody.

Egil Hauger

Electromagnetic interference - physical mechanisms

BY GEIR LØVNES

1 Introduction

All unwanted currents or voltages that may arise in a communication system are called electrical noise. All kinds of electronic equipment produce such noise.

The term Electromagnetic Interference – EMI – describes the situation that electrical noise destroys the functionality of systems. This noise may come from other parts of the system, from other systems or from natural phenomena (like lightning).

Depending upon the noise source, there are two main kinds of EMI:

- Intrasystem – The noise sources are within the disturbed system
- Intersystem – The noise sources are other systems.

Three basic elements must be present to create an EMI situation:

- Emitter (noise source)
- Susceptor (noise receiver, victim)
- Transfer or coupling medium (radiated or conducted).

When noise is transferred through cables, it may be common mode or differential mode. Common mode noise is currents flowing in the same direction in all wires. Differential mode noise is current flowing in opposite directions in two wires.

The opposite situation to Electromagnetic Interference is Electromagnetic Compatibility – EMC. An EMC situation occurs when all systems work as intended, both isolated and in their natural environments.

It must be noted that there will always be some interference between systems and internally within each system. However, the *term* EMI denotes the situation that the interference is so severe that the functionality of one or several systems is affected in an unwanted manner, while the *term* EMC denotes the situation that the interference is so low that all systems work properly. Hence, EMI and EMC may be regarded as two contrary situations.

The source of interference from electronic equipment is the circuits generating the signals. These signals may be emitted because cables and printed circuit board (PCB) traces work as antennas. The power of this emission depends on geometry and dimensions of the cables and PCB traces, and on the current flowing through them. Besides, the casing around the equipment may work as a resonant cavity.

Inductive or capacitive coupling may cause the current on one wire to affect other wires. Signals from one loop may affect other loops via common impedance couplings. Power and signal cables may in this way intercept unwanted signals from other wires inside the equipment. Such couplings contribute to both radiated and conducted interference.

A transmitter antenna also works as a receiving antenna. Therefore, all wires and PCB traces will also intercept interference. And noise may penetrate as conducted interference in power and signal cables.

In this paper we will look into the mechanisms that turn electronic equipment into emitters and susceptors. The same mechanisms largely apply to analogue equipment as to digital. The most important difference in receiving noise is the circuit's susceptibility to interference, while an important difference in emission is the frequency spectrum which the circuits produce. High speed digital circuits produce wideband noise and can give a substantial contribution to interference into other systems.

Most of the formulas are deduced in the text. This may impair the readability, but we have chosen to do it in this way in order to give the reader an understanding of the prerequisites behind the formulas – that is to say under which conditions the formulas apply.

2 Electromagnetic radiation and induction

2.1 Radiation from wires and loops

In section 2.1.1 we will see that any wire carrying a time variable current may radiate. Dimensions and geometry will decide whether the wires have good or poor antenna properties. In section 2.1.2 we will give some examples of wiring which have "good" – but unwanted – antenna properties.

2.1.1 Infinitesimal radiation sources

We will start by setting up the complete equations for the electrical and magnetic field vectors around an infinitesimal dipole and current loop.

We assume the dipole's length is l, and that it carries a sinusoidal current I. As shown in Figure 2.1, r is the distance from the dipole to our observation point. The requirement to infinitesimality is met when $l \ll r$, $l \ll \lambda$ and I does not vary along the length l.

Ignoring the transmission delay, we have for an infinitesimal dipole:

$$E_\theta = -\frac{I \cdot l \cdot k^3}{4 \cdot \pi \cdot \omega \cdot \varepsilon_0} \cdot \left[\frac{1}{(j \cdot k \cdot r)} + \frac{1}{(j \cdot k \cdot r)^2} + \frac{1}{(j \cdot k \cdot r)^3}\right] \cdot \sin\theta \quad [V / m] \qquad (2.1)$$

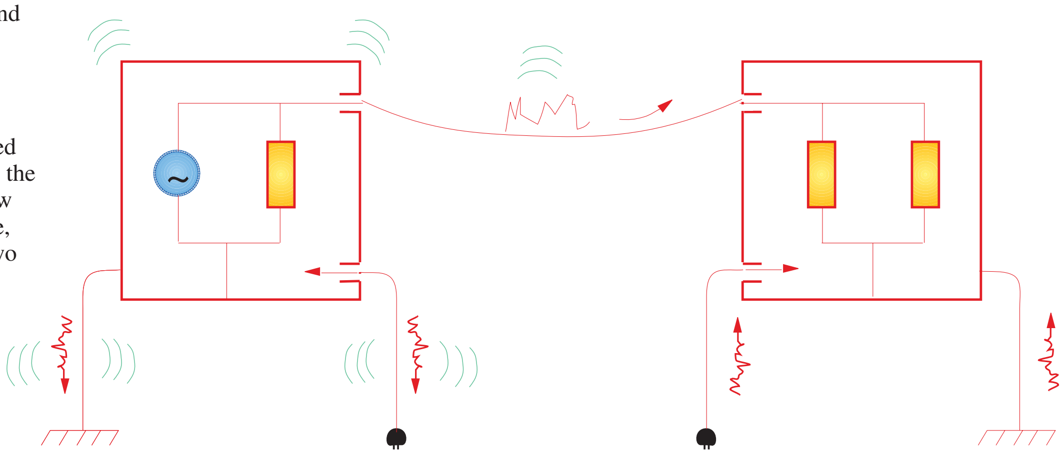


Figure 1.1 Electronic equipment as emitter and susceptor of interference. Interference may be radiated or conducted

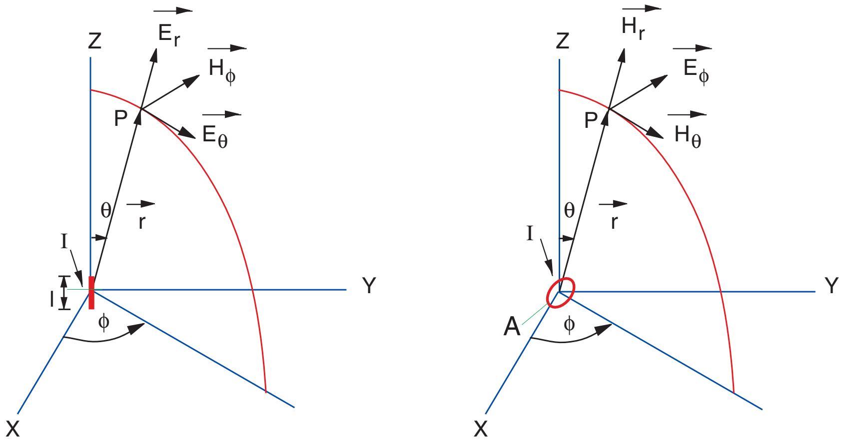


Figure 2.1 Infinitesimal dipole and current loop in spherical coordinate system

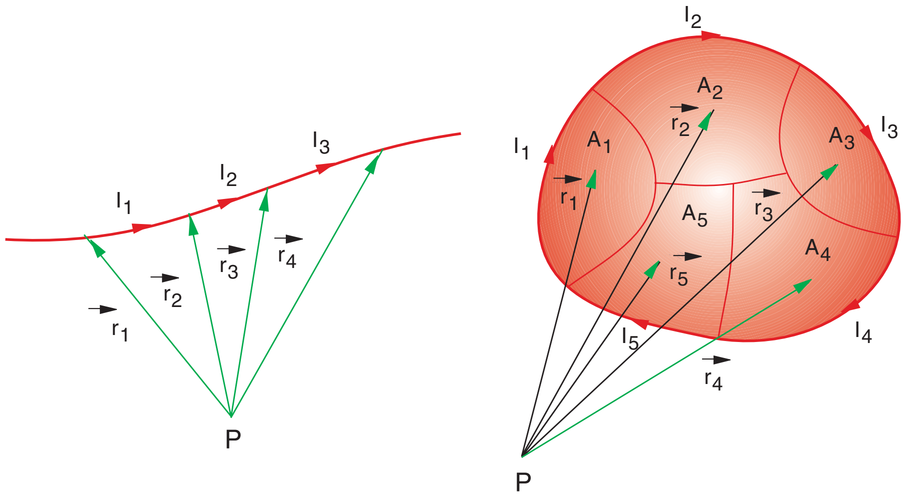


Figure 2.2 Total field from conductors is found by superposition of conductor segments or loop segments

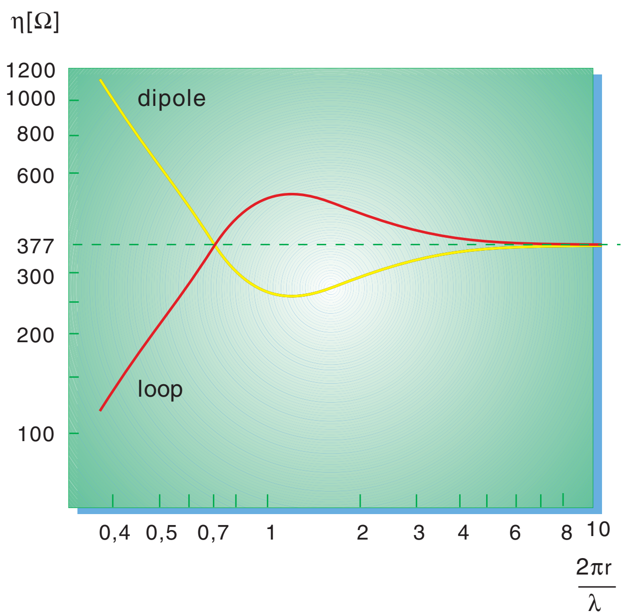


Figure 2.3 Wave impedance of infinitesimal dipole and current loop. (A small current loop is also called a magnetic dipole)

$$E_r = -\frac{I \cdot l \cdot k^3}{2 \cdot \pi \cdot \omega \cdot \varepsilon_0} \cdot \left[\frac{1}{(j \cdot k \cdot r)^2} + \frac{1}{(j \cdot k \cdot r)^3}\right] \cdot \cos\theta \quad [V / m] \tag{2.2}$$

$$H_\phi = -\frac{I \cdot l \cdot k^2}{4 \cdot \pi} \cdot \left[\frac{1}{(j \cdot k \cdot r)} + \frac{1}{(j \cdot k \cdot r)^2}\right] \cdot \sin\theta \quad [A / m] \tag{2.3}$$

$k = 2\pi / \lambda$

We will now look at a current loop with area A carrying a sinusoidal current I. The current does not vary around the loop. When the diameter of the loop is very small in relation to r and λ, and we ignore the transmission delay, we have the following equations for the field components:

$$H_\phi = -\frac{j \cdot I \cdot A \cdot k^3}{4 \cdot \pi} \cdot \left[\frac{1}{(j \cdot k \cdot r)} + \frac{1}{(j \cdot k \cdot r)^2} + \frac{1}{(j \cdot k \cdot r)^3}\right] \cdot \sin\theta \quad [A / m] \tag{2.4}$$

$$H_r = -\frac{j \cdot I \cdot A \cdot k^3}{2 \cdot \pi} \cdot \left[\frac{1}{(j \cdot k \cdot r)^2} + \frac{1}{(j \cdot k \cdot r)^3}\right] \cdot \cos\theta \quad [A / m] \tag{2.5}$$

$$E_\phi = \frac{j \cdot I \cdot A \cdot k^4}{4 \cdot \pi \cdot \omega \cdot \varepsilon_0} \cdot \left[\frac{1}{(j \cdot k \cdot r)} + \frac{1}{(j \cdot k \cdot r)^2}\right] \cdot \sin\theta \quad [V / m] \tag{2.6}$$

$k = 2\pi/\lambda$

When r is very small, the r^{-3} terms dominate. These terms describe what is called the static fields. The r^{-2} terms are induction fields, while the r^{-1} terms, which dominate when r is large, are called radiation terms. When $r = \lambda/2\pi$, the three terms r^{-1}, r^{-2} and r^{-3} are of equal magnitude.

We can draw some important conclusions:

Firstly: Emission from any conductor or loop may be calculated on the basis of these equations by dividing the conductor or loop into many infinitesimal segments and find the total field by superposition (Figure 2.2). Besides, we may find approximate expressions for the fields from conductors and loops which are "quite" small by using the equations for the infinitesimal instances directly.

This means that any wire carrying a time variable current will work as a transmitting antenna. However, the antenna properties may become poor because field contribution from various segments may be added in counter phase, so the far field may be strongly reduced.

In the same way any wire will work as a receiving antenna. The principle of reciprocity states that the radiation pattern of an antenna is the same whether the antenna works as transmitter or receiver. (The Radiation pattern is a three-dimensional graph describing the relative far-zone field strength versus direction at a fixed distance from the antenna.)

We observe that field strength from the infinitesimal sources is proportional to the current I.

When $r \to \infty$, E_r and H_r will approach zero, while the ratio between E and H components perpendicular to the propagation direction, will approach $\sqrt{\mu/\varepsilon_0} = 120\ \pi = 377\ [\Omega]$
The ratio between E and H is called wave impedance, η. The value 120π of the wave impedance is called free space impedance and is denoted by η_0. So, in the far field, wave impedance is equal to free space impedance and the field from a dipole and a current loop looks the same; we are unable to determine whether it was generated by a dipole or a current loop.

Very close to the dipole or the current loop, terms which are proportional to r^{-3} in equations (2.1) – (2.6) will dominate. This means that close to a dipole, $E/H > 377\ \Omega$, while close to a current loop, $E/H < 377\ \Omega$. When the field's E component dominates over its H component, or vice versa, we say that we are in the near field of the radiation sources. A current loop is then seen as a source of magnetic fields; a dipole is seen as a source of electrical fields. This is important when choosing shielding materials.

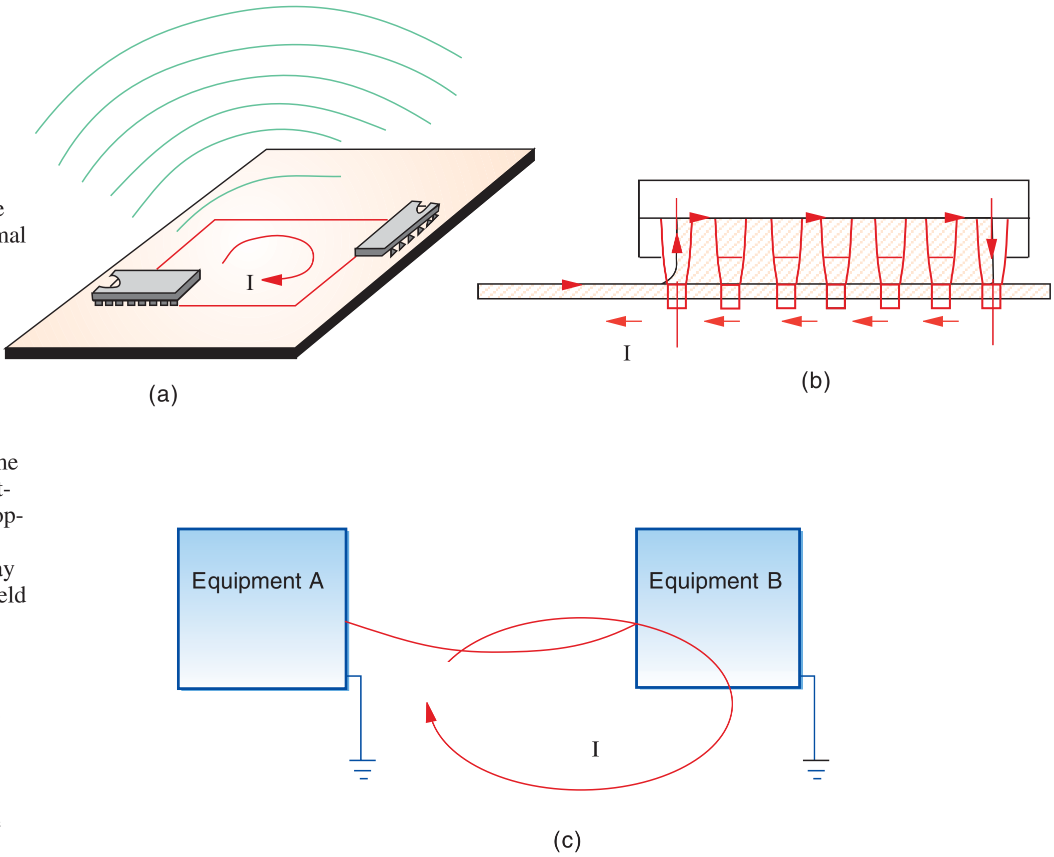


Figure 2.4 Example of current loops working as antennas

An antenna consisting of one or more current loops is usually called a magnetic antenna, while an antenna consisting of dipoles is an electrical antenna.

Figure 2.3 shows the wave impedance of the infinitesimal radiation elements.

An important issue is where to draw the border line between near and far field. As already mentioned, the three links in (2.1)–(2.3) and (2.4)–(2.6) are of equal size when $r = \dfrac{\lambda}{2\pi}$. For a point source we thus say that we are in the far field when $r > \dfrac{\lambda}{2\pi}$.

When the largest extent d of the source is of the same size as the wavelength, this far field condition is used [1]:

$$r > \frac{2d^2}{\lambda}$$

For a half-wave dipole we then have:

$$r > \frac{\lambda}{2}$$

For very directive antennas [2]:

$$r > \frac{4d^2}{\lambda} \quad \text{for the far field}$$

$$r < \frac{d^2}{\lambda} \quad \text{for the near field}$$

2.1.2 Examples of cabling acting as antennas

Any cabling[1] constituting a current loop will work as a magnetic antenna. Examples of loops which may create radiation are given in Figure 2.4.

Electrical antennas are usually created by power and signal cables. Unintended voltage drops in a circuit may, cause common mode currents to be set up in connecting cables. Such voltage drops are usually caused by potential differences in the ground plane (Figure 2.5). Even if a cable is terminated to a socket or other equipment, from the circuit it may look like a monopole. This is because the cable must be regarded as a transmission line when the cable length

---

[1] By cable is meant any pliable structure containing one or more isolated wires surrounded by a rubber or plastic covering. One single insulated wire will thus also be a cable.

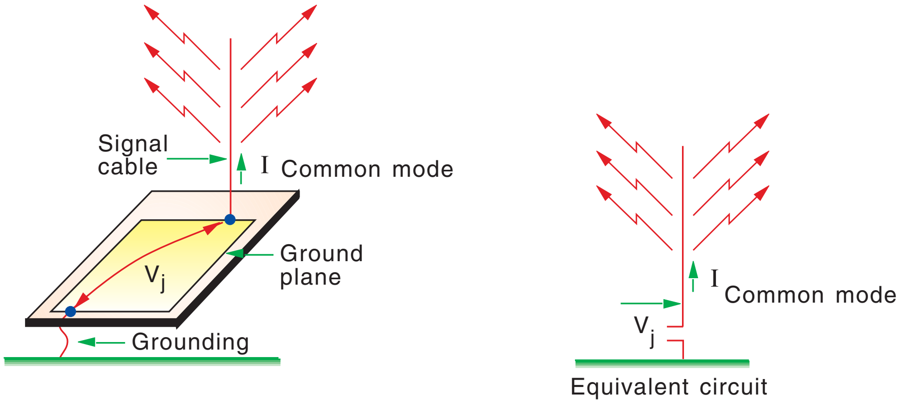
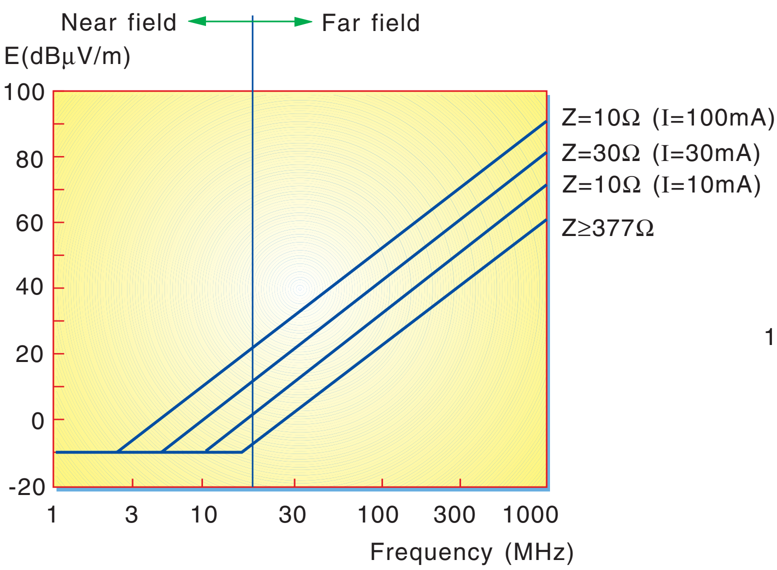


Figure 2.5 Radiation caused by potential differences in the ground plane



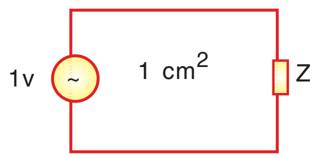


Figure 2.6 The field 3 metres from a 1 cm² current loop excited by a 1 V sinusoidal [4]. For other values corrections must be made by the formula: $E = E_{curves} + 20 \cdot lg(V) + 20 \cdot lg(A)$ (A = area in cm²)

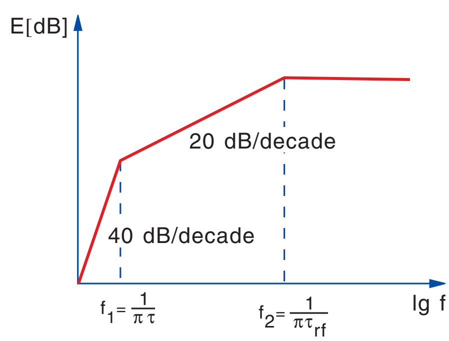


Figure 2.7 The far field spectrum from a current loop with periodic square pulses with pulse length τ. τ_{rf} is explained in Figures 7.3 and 7.4

is of the same order of magnitude as the wavelength. We will look closer at transmission lines in section 5.

A major part of the radiation from equipment with small dimensions compared to the wavelength is caused by common mode currents on cables between the equipment units [3]. Such radiation is often denoted common mode radiation.

2.1.3 Practical equations

Near field around current loop

From the equations (2.4) and (2.5) we have that the static field around a small current loop will never be larger than

$$H = \frac{I \cdot A}{2 \cdot \pi \cdot r^3} \quad [A/m] \tag{2.7}$$

This expression then gives the maximum magnetic near field from a small current loop when the frequency is so low that $r \ll \lambda/2\pi$.

Field close to a straight wire

At the distance r from a straight wire, the magnetic field is:

$$H = \frac{I}{2 \cdot \pi \cdot r} \quad [A/m] \tag{2.8}$$

when r is small in relation to the length of the conductor.

The formula comes from Maxwell's equation:

$$\oint \vec{H} \cdot d\vec{s} = I + \frac{\partial}{\partial t} \int\int \vec{D} \cdot \vec{n} \cdot dA \tag{2.9}$$

I is the conductor current and we presume harmonic current variation. We may with good approximation ignore the second part of the equation. The effect of the approximation is that the current I cannot vary along the conductor.

When integrating along a circle around the conductor at a distance of r, we have:

$$H_\phi = \frac{I}{2 \cdot \pi \cdot r} \quad [A/m] \tag{2.10}$$

Based on another of Maxwell's equations,

$$\oint \vec{B} \cdot \vec{n} \cdot dA = 0 \tag{2.11}$$

it can be shown that the other components of the H field is zero.

Far field from wire loop

The far field from a wire loop with area A is found from formula (2.6). The field component is

$$E_\phi = \frac{I \cdot A \cdot k^3}{4 \cdot \pi \cdot \omega \cdot \varepsilon_0} \cdot \frac{1}{k \cdot r} \cdot \sin\theta \quad [V/m] \tag{2.12}$$

We are not interested in the direction of the E field, and when rearranging the expression, we have this formula for the far field from a wire loop:

$$E = \pi \cdot \eta_0 \cdot \frac{I \cdot A}{\lambda^2 \cdot r} \cdot \sin\theta \quad [V/m] \tag{2.13}$$

We observe maximum radiation in the loop plane.

The E field increases by 40 dB/decade ($E \propto f^2$). This applies to sinusoidal signals. By combining the typical spectrum from digital circuits (section 6), we have the resulting frequency spectrum for the far field from a current loop carrying digital signals in Figure 2.7.

It is often more interesting to observe the field as function of voltage rather than current. If I in equation (2.13) is replaced by V/Z, we have

$$E = \pi \cdot \eta_0 \cdot \frac{V \cdot A}{\lambda^2 \cdot Z \cdot r} \cdot \sin\theta \quad [V/m] \qquad (2.14)$$

Z is the load, and if this is greater than 377 Ω, the field is largely related to the voltage V; the circuit becomes similar to a voltage driven dipole [4]. When $Z > 377$ Ω, the generated field will be the same size as at $Z = 377$ Ω.

Formulas (2.13) and (2.14) are derived from (2.6). Even if we in (2.6) assumed the diameter of the loop to be very small in relation to λ, we will get a reasonable result even when using (2.13) and (2.14) for larger loops. But none of the values we use in A (length, width, or diameter) must exceed λ/4 at any frequency. At these frequencies the values used must be reduced to λ/4 [4]. Above this limit the load impedance Z must also be replaced by characteristic impedance for the PCB traces.

In many cases, particularly when MOS circuits are used, the load impedance will be very capacitive. Assuming the load impedance to be purely capacitive, we can replace Z by $1/\omega C$ in (2.14) and have a spectrum as shown in Figure 2.8, that is a 20 dB/decade increase compared to Figure 2.7.

Far field from monopole

The far field from a short monopole over a ground plane is given by:

$$E = \mu_0 \cdot \frac{f \cdot I \cdot l}{r} \cdot \sin\theta \quad [V/m] \qquad (2.15)$$

l = length of the monopole

The basis for this formula is (2.1). When l is the length of the monopole, and we only observe the radiation term, (2.1) gives:

$$E_0 = -\frac{\mu_0}{2} \cdot \frac{I \cdot l \cdot f}{j \cdot r} \cdot \sin\theta \quad [V/m] \qquad (2.16)$$

The monopole is placed over an ideally conducting ground plane, therefore we must double the length of the monopole. We are interested in the absolute value of the E field, and have formula (2.15).

The formula presupposes uniform current distribution, which is a valid assumption when the length of the monopole is small compared to the wavelength. If the monopole is of the same order of magnitude as λ, we may apply the effective length of the monopole instead of l – see next section – and use (2.16) as an approximate expression for the E field.

The resulting spectrum with digital signals is shown in Figure 2.9.

2.2 Induction in wires and loops

As mentioned above, a transmitting antenna will also work as a receiving antenna. We will here consider the size of voltages which may be induced in a dipole/monopole and in a wire loop when these are placed in an electromagnetic field.

2.2.1 Voltage induced in a dipole/monopole

The relation between electrical field strength and voltage between two points a and b in space is:

$$V = \int_a^b \vec{E} \cdot d\vec{l} \quad [V] \qquad (2.17)$$

where the lines $d\vec{l}$ represent the integration path between a and b.

If a dipole has rectangular current distribution, i.e. the current is equal over the whole of the dipole length, the induced voltage will according to (2.17) be $V = E \cdot l_d$, where l_d is the dipole length.

In a real dipole the current decreases to zero at each end. Then the antenna does not receive as efficiently at the ends as in the centre. Assuming cosinus shaped current distribution over the dipole and the E field being parallel to the dipole element, we have for a half-wave dipole:

$$V = \int_{-\lambda/4}^{\lambda/4} E \cdot \cos\left(\frac{2\pi l}{\lambda}\right) \cdot dl = E \cdot \frac{\lambda}{\pi} \quad [V] \qquad (2.18)$$

This is the maximum voltage to be induced. If the polarisation of the antenna is different from that of the field, or the propagation direction of the field is not perpendicular to the antenna rods, the voltage will be lower.

From this we can determine the effective length, l_{eff}, as the length the dipole would have had if the current distribution were rectangular. l_{eff} will therefore comply with the equation

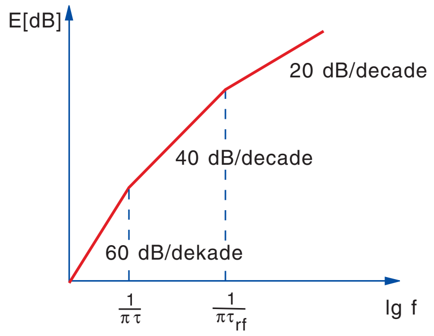


Figure 2.8 Same situation as in Figure 2.7, but the spectrum is assessed on the basis of voltage rather than current, and we have assumed a purely capacitive load

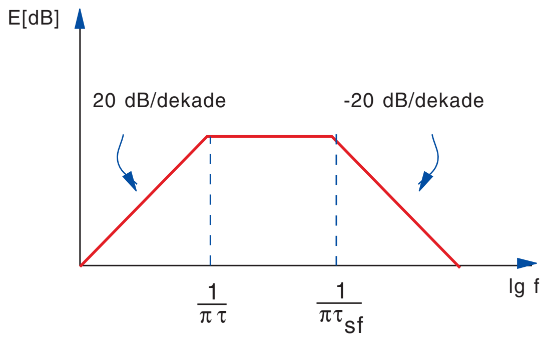


Figure 2.9 The far field spectrum from a monopole over a ground plane. The monopole is excited by square pulses as in Figure 2.7

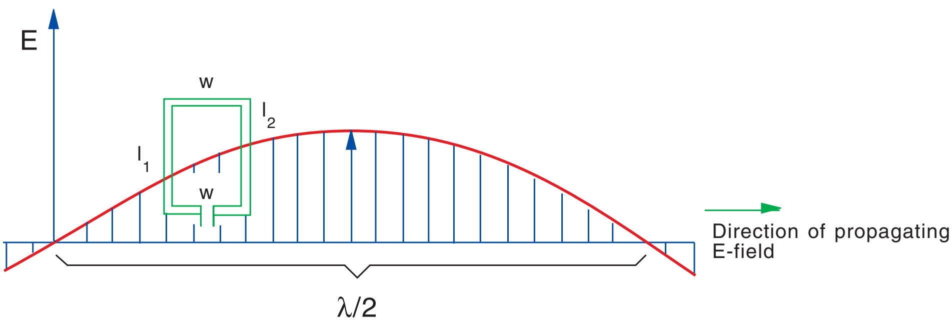


Figure 2.10 Situation when a plane wave moves through a frame antenna

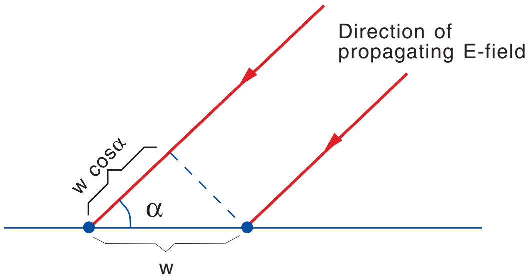


Figure 2.11 Angle of incidence in relation to the orientation of the frame antenna

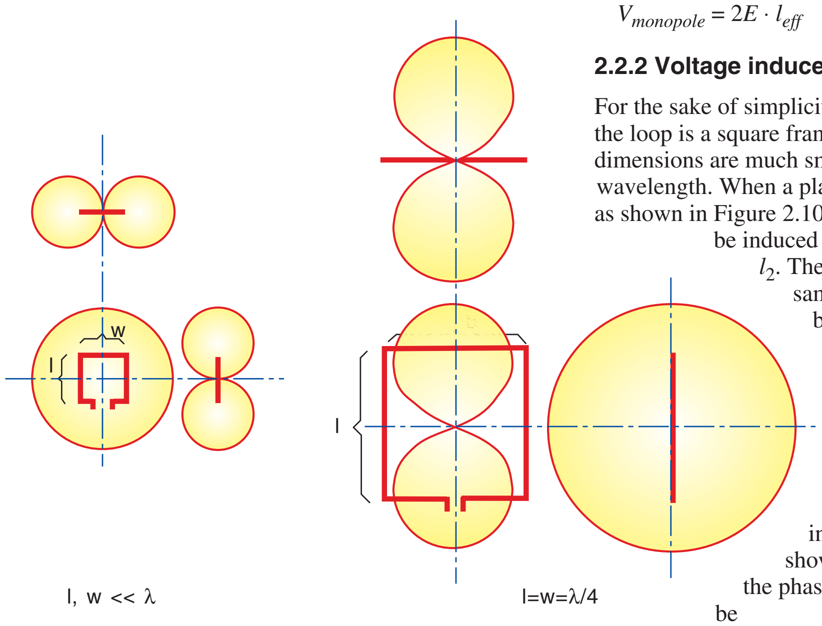


Figure 2.12 Radiation pattern for a small loop antenna and a square loop antenna with sides λ/4. For the small loop antenna the shape does not affect the radiation pattern as long as the loop is plane

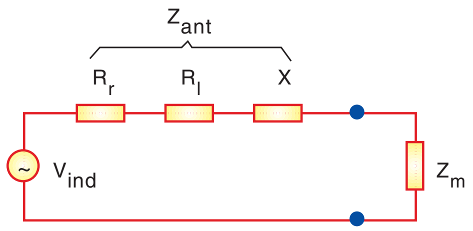


Figure 2.13 Equivalence diagram for antenna with receiver

$$V_{dipole} = E \cdot l_{eff} \quad [V] \qquad (2.19)$$

Thus, a half-wave dipole has effective length λ/π. For a short dipole the effective length will be half the dipole length because we may here assume triangular current distribution. An infinitesimal dipole has the same effective and real length because here the current distribution is defined as rectangular.

Likewise, a short monopole over a ground plane will have effective length equal to half the physical length, while a quarter wave monopole will have effective length λ/2π. But here we also have a field contribution by reflection in the ground plane, so the maximum voltage measured over the monopole will be

$$V_{monopole} = 2E \cdot l_{eff} \quad [V] \qquad (2.20)$$

2.2.2 Voltage induced in a loop

For the sake of simplicity we assume that the loop is a square frame and that the dimensions are much smaller than the wavelength. When a plane wave enters as shown in Figure 2.10, a voltage will be induced in the sides l_1 and l_2. The induction is the same in l_1 as in l_2, but because there is a certain distance between the sides, there will be a phase difference between the voltages. If the entering direction is as shown in Figure 2.11, the phase difference will be

$$\Delta k = \frac{2 \cdot \pi}{\lambda} \cdot w \cdot \cos\alpha \text{, and we have:}$$

$$V = 2 \cdot E \cdot l \cdot \sin\left(\frac{\Delta k}{2}\right) \cdot \cos\theta \quad [V] \; (2.21)$$

where we have multiplied by cosθ to make allowances for the E field not necessarily being parallel with l_1 and l_2.

Because the loop dimension is small compared to the wavelength and $\sin(x) \approx x$ when x is a small angle, we may use $\sin(\cdot) = (\cdot)$ in (2.21). Letting $l \cdot w = A$, we have:

$$V = E \cdot A \cdot \frac{2 \cdot \pi}{\lambda} \cdot \cos\alpha \cdot \cos\theta \quad [V] \; (2.22)$$

The formula applies independent of the shape of the loop, as long as it is plane.

Induced voltage from a time varying magnetic field is:

$$V(t) = -A \cdot \frac{dB}{dt} \cdot \cos\phi \quad [V] \qquad (2.23)$$

A = area of the loop

B = magnetic flux density

ϕ = angle between normal in the loop plane and the magnetic field.

If the magnetic field is sinusoidal, we have:

$$V = 2 \cdot \pi \cdot f \cdot A \cdot B \cdot \cos\phi \quad [V] \qquad (2.24)$$

In the far field $E/H = 377\ \Omega$, and the equations (2.22) and (2.24) give the same result. Equations (2.21), (2.22) and (2.24) presuppose a uniform current distribution on the loop.

If the dimensions of the antenna are so big compared to the wavelength that the current around the loop is no longer in phase, the radiation pattern will be greatly affected. For example, a square loop antenna with sides λ/4 will have the same radiation pattern as a half-wave dipole, see Figure 2.12.

When the dimensions are such that we do not have uniform current distribution, we may put α = θ = ϕ = 0 into formulas (2.22) and (2.24) and use the formulas as an estimate of maximum induced voltage. The induced voltage increases as long as the largest dimension in the loop is smaller than λ/2. Above this value the voltage will not increase further, but will oscillate with peak values for each half wavelength. As an estimate for maximum induced voltage we can use the value at λ/2 [4].

2.2.3 Transfer of power

An equivalence diagram for an antenna with receiver is shown in Figure 2.13. The antenna's impedance, Z_{ant}, consists of three parts: R_r which is the antenna's radiation resistance, R_l is ohmic loss in the antenna, while X is the reactive part of the input impedance. If the antenna is in resonance, $X = 0$. Radiation resistance is an imagined resistance consuming the same power as the antenna emits when used as a transmitting antenna. Z_m is the input impedance of the receiver.

V_{ind} is the voltage induced when the antenna is situated in an electromagnetic field. The voltage is distributed on both the antenna and receiver impedance. In order to obtain maximum power transfer, $Z_{ant} = Z_m{}^*$ must apply.

The radiation resistance of a loop antenna is [5]:

$$R_r = 320 \cdot \pi^4 \cdot \left(\frac{A}{\lambda^2}\right)^2 \quad [\Omega] \qquad (2.25)$$

Small current loops have very low radiation resistance. If for example $A = l \cdot l$ and $l = 0.1\ \lambda$, radiation resistance is 3 Ω. Transferred power from interference received with this current loop is small when the load resistance » 3 Ω.

Radiation resistance of a short dipole is also low – for example 8 Ω at dipole length 0.1 λ. Large electrical antennas have higher values. A quarter wave monopole has the value 36.6 Ω. A wire acting as a dipole antenna, not having symmetric feed, may have radiation resistances in the kΩ area.

On a printed circuit board most of the current loops are small compared to the wavelength of the interfering fields. Therefore, the current loops act as low impedance magnetic antennas. On printed circuit boards low impedant circuits will therefore be most susceptible to magnetic fields because this gives the best impedance match to the interference. Similarly, high impedant circuits will be most susceptible to electrical fields because the interference is received by the input cables which are often the same size as the wavelength, and therefore have higher radiation impedances.

However, we should be aware of the fact that the consequences of the interference are not necessarily determined by power transfer to the interfering element. For example, a current loop connected to a MOS circuit may interfere greatly even if the impedance match is poor, because MOS circuits are interfered by interference *voltages.*

3 Crosstalk

In section 2 we observed how cabling works as transmitting or receiving antenna independently. This section will deal with capacitive and inductive transfer between (at least) two conductors in the electrical and magnetic near field.

Unwanted inductive or capacitive transferred power from a wire or pair of wires to another wire or pair of wires is called crosstalk. The term wire may here mean anything from the traces inside an IC to a pair of wires in an Atlantic Sea cable.

Crosstalk is defined as:

$$Crosstalk = 20 \cdot lg \frac{V_{susceptor}}{V_{emitter}} \quad [dB] \quad (3.1)$$

In Figure 3.1 are shown two parallel wires with earth as return. The current in *E* produces a m,agnetic field affecting

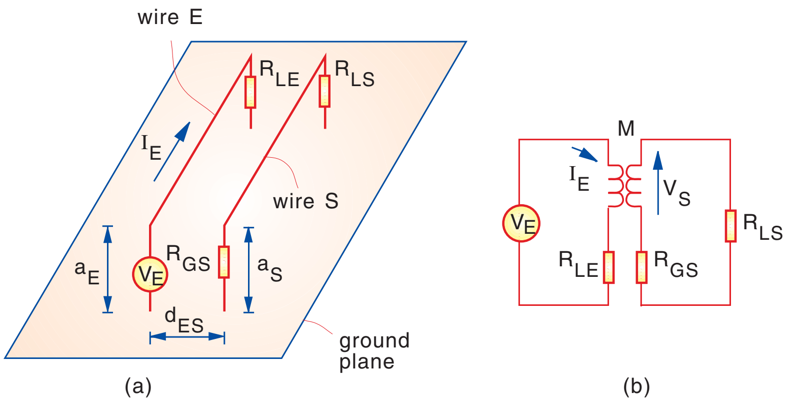


Figure 3.1 Inductive transfer between two parallel wires

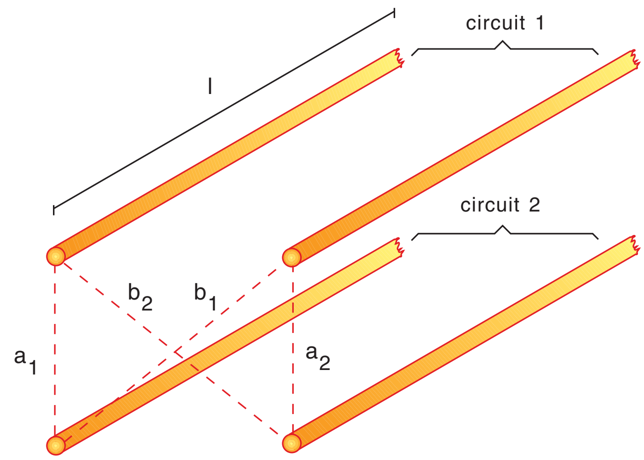


Figure 3.2 Transfer between two conductor pairs

conductor *S*. The inductive transfer may be expressed as:

$$V_S = M \cdot \frac{\partial I_E}{\partial t} \quad [V] \qquad (3.2)$$

where V_E is shown in Figure 3.1 and *M* is mutual inductance between wires.

With harmonic time variation we have:

$$V_S = M \cdot \omega \cdot I_S = M \cdot \omega \cdot \frac{V_E}{R_{LE}} \quad [V] \, (3.3)$$

If we have a configuration as shown in Figure 3.1.a, we have the following expression for the mutual inductance *M* [2]:

$$M = \frac{\mu_0}{4\pi} \cdot \ln \left| \frac{(a_E + a_S)^2 + d_{ES}^2}{(a_E - a_S)^2 + d_{ES}^2} \right| \quad [H] (3.4)$$

where all lengths are in metres and given by Figure 3.1.

If we have two wire pairs as shown in Figure 3.2, we have the following expression for *M* [2]:

$$M = 4.61 \cdot 10^{-7} \cdot l \cdot lg \frac{a_1 \cdot a_2}{b_1 \cdot b_2} \quad [H] (3.5)$$

Returning to Figure 3.1 we observe that there may also be capacitive transfer between the wires. We redraw the circuit and get the equivalence circuit as shown in Figure 3.3.

The voltage across R_{GS} and R_{LS} given by the capacitive transfer is, with harmonic time variation:

$$V_S = \frac{R \cdot V_E}{R + \frac{1}{j\omega C}} \quad [V] \qquad (3.6)$$

where

$$R = \frac{R_{GS} \cdot R_{LS}}{R_{GS} + R_{LS}} \quad [\Omega] \qquad (3.7)$$

C is the mutual capacitance between the wires. Ignoring the effect of the ground plane, *C* is given by [2]:

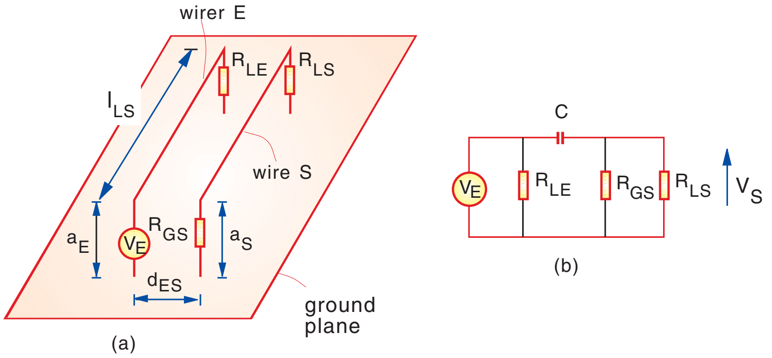


Figure 3.3 Capacitive transfer between two parallel wires

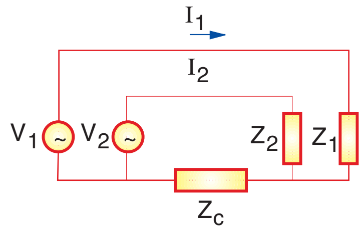


Figure 4.1 Common impedance coupling across Z_c

$$C = \frac{3.5 \cdot 10^{-10} \cdot l_{ES}}{2\pi \cdot \ln\left(\frac{d_{ES}^2}{g_E g_S}\right)} \quad [F] \qquad (3.8)$$

g_E = radius of wire E

g_S = radius of wire S

l_{ES} is mutual length of E and S.

All lengths are in metres and given by Figure 3.3.

If $1/\omega C \gg R$ we have:

$$V_S = R \cdot V_E \cdot j\omega C \quad [V] \qquad (3.9)$$

From (3.3) we observe that inductive crosstalk increases when impedance in the interfering circuit decreases. (3.9) shows that capacitive crosstalk increases when impedance in the interfered circuit increases.

Inductive crosstalk is strongest

- in transformers
- when wires are parallel
- usually at lower frequencies.

Capacitive crosstalk is strongest

- in multiconductor twisted cables
- at high frequencies
- at high impedance levels.

From (3.3) and (3.9) we observe that both capacitive and inductive crosstalk increases by 20 dB/decade when the frequency increases.

We get a rough estimate of crosstalk by using the formulas for capacitive and inductive coupling and choosing the one giving the highest result.

If wire lengths are greater than λ/4, we cannot use the formulas in this section. In this case we must consider the conductors as transmission lines and find mutual inductance and capacitance *per length unit*.

4 Common impedance coupling

This type of coupling is made by two loops in a circuit sharing an impedance. When the current I_1 in one loop produces a drop in voltage over the common impedance, the current in the other loop will be interfered by this (see Figure 4.1).

As for inductive and capacitive coupling, the common impedance crosstalk will also increase with frequency. When the common impedance is a ground plane, the increase is 10 dB/decade, when the common impedance is a conductor, the increase is 20 dB/decade [1].

5 Impedance mismatch

In case of poor impedance match we may have mutliple reflections on the transmission line between generator and load. This may cause high frequency oscillations to arise.

If the wire length from generator to load is considerably less than λ, where λ is the shortest wavelength of the transferred signal, we find currents and voltages easily by using Ohm's law. If this is not the case we must consider the wire as a transmission line and use transmission line theory in order to calculate voltages and currents.

A transmission line model of transmission from generator to load is shown in Figure 5.1. We see that if the length l in the transmission line model in Figure 5.1.a becomes very small ($l \ll \lambda$), we have a situation as shown in Figure 5.1.b, which is pure circuit technique.

In Figure 5.1.a there will be a voltage wave V_{GL} on the transmission line from generator to load. If there is impedance mismatch between load and line, a voltage wave V_{LG} is reflected from load to generator. The ratio between these two voltage waves is called reflection coefficient, and at load this is given by:

$$\rho = \frac{V_{LG}}{V_{GL}} = \frac{Z_1 - Z_0}{Z_1 + Z_0} \qquad (5.1)$$

The ratio between the currents is equal to (voltage) reflection coefficient with the sign reversed:

$$\frac{I_{LG}}{I_{GL}} = -\rho \qquad (5.2)$$

The line's characteristic impedance is

$$Z_0 = \sqrt{\frac{R + j \cdot \omega \cdot L}{G + j \cdot \omega \cdot C}} \quad [\Omega] \qquad (5.3)$$

R, G, C, and *L* are distributed ohmic resistance in the wires, distributed leakage conductance between the wires, distributed capacitance between the wires and distributed serial inductance in the wires, respectively – all per length unit of the wire.

Assuming $R = G = 0$, the line is loss free. We note that Z_0 is real in the case of loss free lines. Furthermore, we note from (5.1) that we have a positive voltage

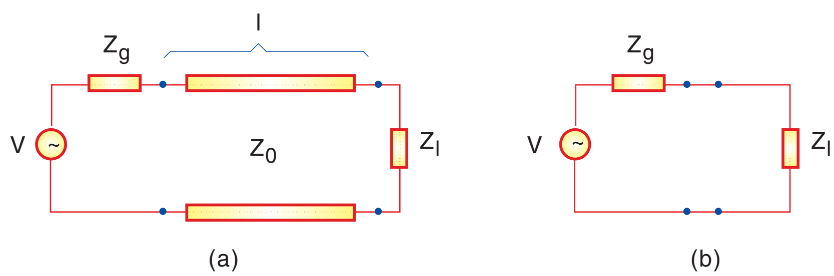


Figure 5.1 (a) Transmission line model, (b) circuit technique model

reflection when $Z_l > Z_0$, i.e. the voltage drop across the load is above the excitation level. When $Z_l < Z_0$ we have negative voltage reflection and the voltage drop across the load is below the excitation level.

Supposing, in Figure 5.1.a, there is impedance match either by generator ($Z_g = Z_0$) or by load ($Z_l = Z_0$), the network will enter a stable end state as soon as the pulse front on the line hits the matched load. However, transferred power may be low in the case of great mismatch. If, on the other hand, there is mismatch at generator as well as at load, we may have great voltage oscillations. The various situations are shown in Figure 5.2 where a step voltage is produced by the generator. The dotted line indicates the voltage across the line at generator, while the solid line is voltage across the load.

We particularly note the case $Z_g < Z_0 < Z_l$, where the multiple reflections produce strong signal distortion. These high frequency oscillations may be harmful, both for the circuit itself and because they may radiate strongly.

From circuit technique we know that maximum power is transferred from generator to load by complex conjugated impedance match ($Z_g = Z_l$*). This is also the case in the transmission line model. Maximum power is transferred from generator to line when $Z_0 = Z_g$*. But as we have seen, such a match may give reflection and signal distortion. This may be avoided by using a loss free or near loss free line and make sure that $Z_g = Z_0 = Z_l$ (all real values).

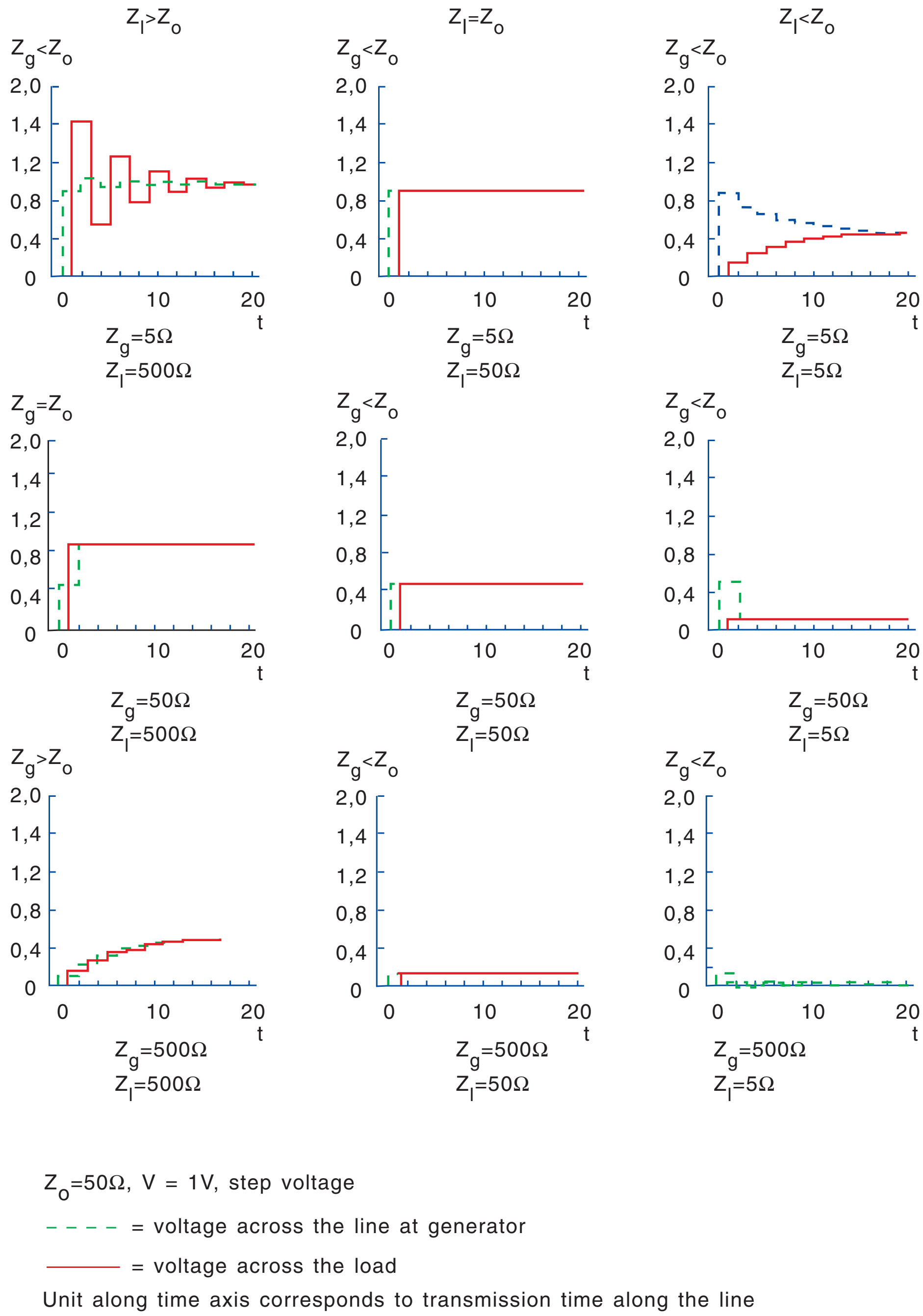


Figure 5.2 Comparison of various cases of impedance match and mismatch

6 Interference due to non-linear elements

If an amplitude modulated RF noise signal reaches the input of a transistor, this diode junction may act as a rectifier, hence demodulating the signal. The demodulated signal is then amplified by the transistor.

This effect explains why pulsed radars or TDMA (Time Division Multiple Access) radio signals – which from an interference point of view may be regarded as amplitude modulated signals – can create noise in other electronic equipment. If e.g. the amplitude modulation is in the audio range, this can create problems in all kinds of audio equipment.

Another kind of interference in non-linear elements is intermodulation. When two or more signals are transferred simultaneously through a non-linear channel, at the output of the channel we will observe not only the original frequency components, but also sums and differences of these frequencies and their harmonics. In addition, we will observe that the output of the original signals will not be proportional to the input levels.

7 Frequency spectrum from digital circuits

Digital equipment is an important source of electromagnetic noise due to its wide noise spectrum. We will thus in this last section look at the frequency spectrum from digital circuits.

The transition between time and frequency domains is given by the Fourier transformation.

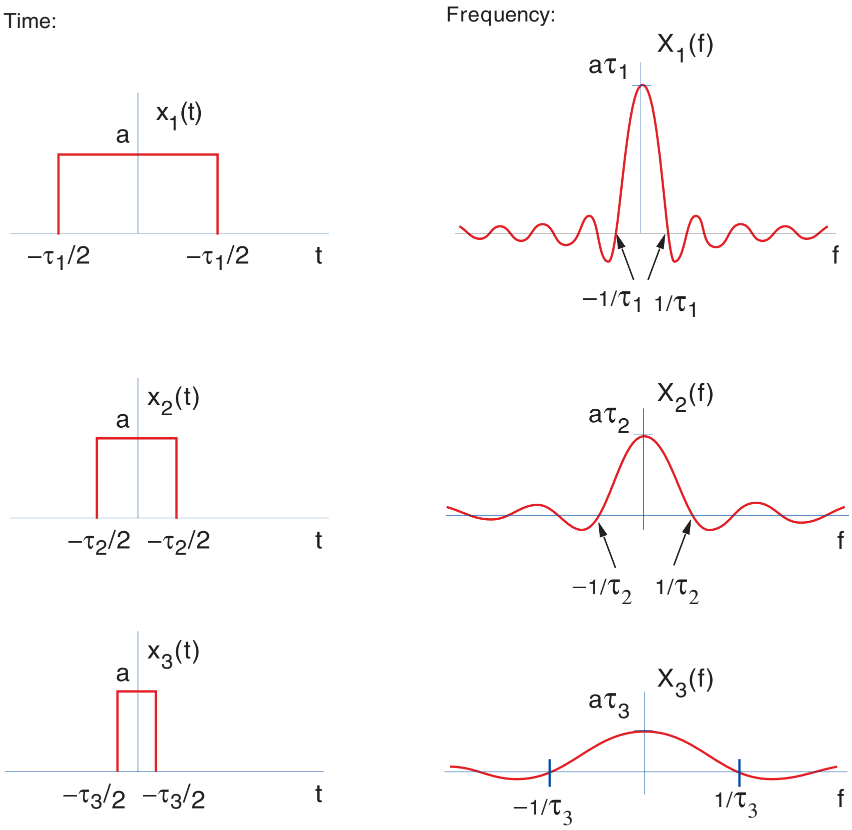


Figure 7.1 The Fourier transformed of square pulses of various lengths

The Fourier transformed of a square pulse is a sinc function,

$$\sin c(f) = \frac{\sin(\pi \cdot f)}{\pi \cdot f} \quad \text{– see Figure 7.1.}$$

A square pulse in time domain will therefore give a sinc shaped frequency spectrum.

Supposing that in time domain we have periodic square pulses, one pulse every T seconds, we will in the frequency domain have discrete frequency components separated by $1/T$ and a sinc shaped envelope. When the length of each square pulse is τ, the first zero point at the envelope will be $1/\tau$ (see Figure 7.2) The frequency component at $1/T$ is called the first harmonic, the component at $2/T$ is called the second harmonic, etc.

When the length of each square pulse is half the distance between the pulses, i.e. at 50 % duty cycle, the even harmonics will correspond to the zero points of the envelope. At 50 % duty cycle we will therefore only have uneven harmonics (1., 3., 5., etc.). (Duty cycle: $\delta = \tau/T$.)

There is no need to carry out Fourier transformations in order to quantitatively determine the envelopes of the frequency spectra. We will first look at periodic square pulses. Supposing that in the frequency plane we draw the frequency logarithmic and the amplitude in dB, we will see that the envelope of the maximum amplitudes decreases by 20 dB/ decade. If we also consider rise time and fall time of the square pulses, and define parameters as in Figure 7.3, we will have a frequency plane approximation as shown in Figure 7.4. Note that if the rise time is different from the fall time the lowest of the two is used. Figure 7.5 shows how the lines in Figure 7.4 are upper bars of the harmonics.

In Figures 7.6 and 7.7 are shown corresponding envelopes for an individual square pulse and periodic triangular pulses.

A bit sequence consisting of “ones” and “zeros” coded as +A volts and 0 volts may in the frequency plane be drawn in the same way as periodic square pulses. The points f_1 and f_2 are determined by the pulse width and the rise time or fall time of an individual “one”, while the distance between the harmonics is determined by the bit rate. In addition, we will here have frequency components lower than f_0, because we may have several consecutive “ones” or “zeros”. For the same reason we will also have components between $n \cdot f_0$ and $(n + 1) \cdot f_0$.

8 Conclusions

In this paper we have looked at physical mechanisms behind electromagnetic interference.

The term Electromagnetic Interference – EMI – describes the situation that electrical noise destroys the functionality of systems. The opposite situation is called Electromagnetic Compatibility – EMC. An EMC situation is when all systems work as intended in their natural environments.

To create an EMI situation, a noise source and a receiver or victim must be present, and the noise must be transferred from the source to the victim. This may be done either as radiated or conducted noise.

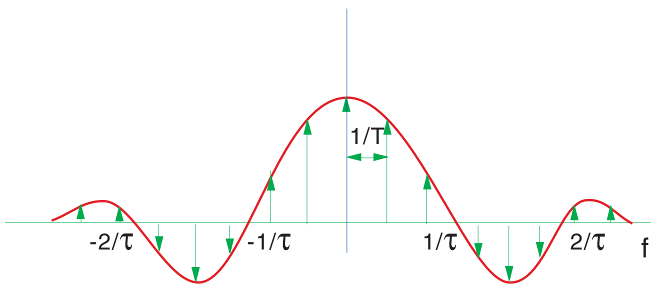


Figure 7.2 The Fourier transformed of periodic square pulses. Each square pulse is τ long, and the distance between them is 1/T

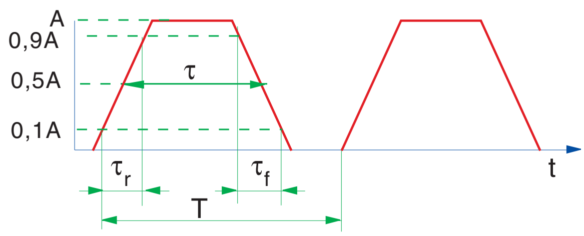
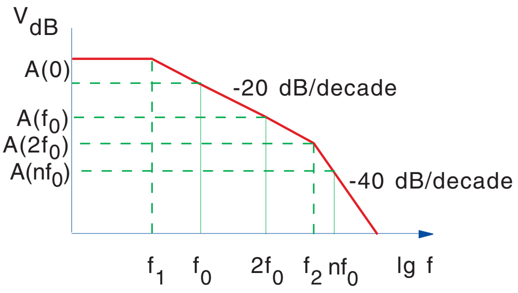


Figure 7.3 Periodic square pulses



$f_0 = 1/T$ fundamental frequency

$f_1 = \frac{1}{\pi\tau}$

$f_2 = \frac{1}{\pi\tau_{rf}}$

$\tau_{rf} = \min\{\tau_r, \tau_f\}$

$\delta \approx \tau f_0$

$A(0) = 20 \lg (2A\delta)$

$A(f_0) = A(0) - 20 \lg (f_0/f_1)$

Figure 7.4 Envelope of the frequency spectrum for the periodic square pulses in Figure 7.3

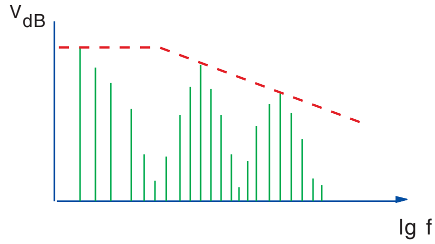


Figure 7.5 Example of correct frequency spectrum of the periodic square pulses in Figure 7.3

9 References

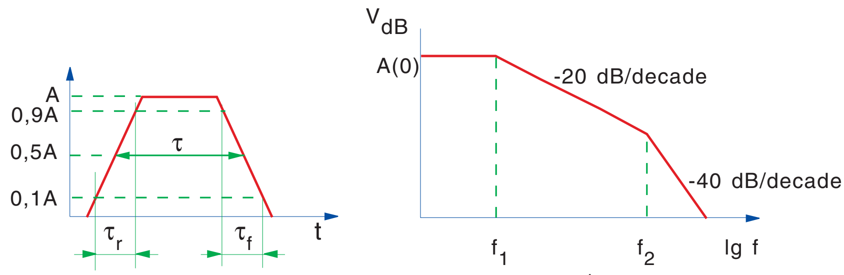
1 Krebs, P, Nissen, H G. *HMC-håndbog del 3: EMC-hensyn ved printudlægning.* Hørsholm, ElektronikCentralen, 1989.

2 Keiser, B. *Principles of electromagnetic compatibility.* Norwood, MA, Artech House, 1987. ISBN 0-89006-206-4.

3 Hubing, T H, Kaufman, J F. Modeling the electromagnetic radiation from electrically small table-top products. *IEEE transactions on Electromagnetic Compatibility,* 31, 74–84, 1989.

4 Mardiguian, M. *Interference control in computers and microprocessor based equipment.* Gainesville, VA, Don White Consultants, 1984. ISBN 0-932263-23-2.

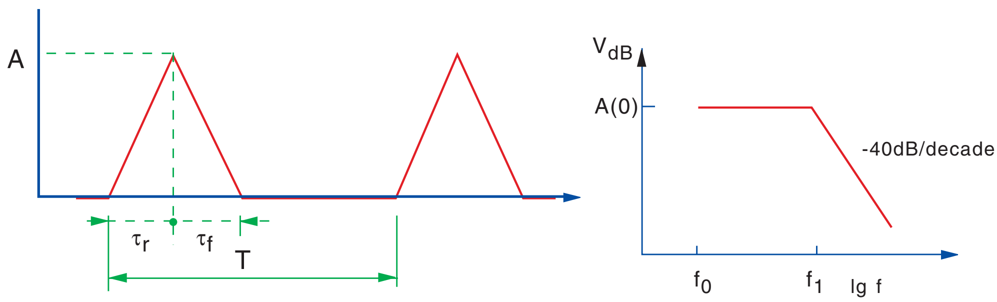
5 Andreassen, S O. *Antenner.* Oslo, Teledirektoratet, 1987.



$f_1 = \frac{1}{\pi\tau}$ $f_2 = \frac{1}{\pi\tau_{rf}}$

$A(0) = 20\lg(2A\tau)$

Figure 7.6 The envelope of the frequency spectrum of an individual square pulse. We will here have a continuous spectrum below the envelope



$f_0 = \frac{1}{T}$ $f_1 = \frac{1}{\pi\tau_{rf}}$

$A(0) = 20\lg(2A\delta)$

Figure 7.7 Periodic triangular pulses in the time and frequency planes

Radio interference in the band 30 – 1000 MHz
– the international limits of interference seen from a radio-engineer's point of view

BY EGIL HAUGER

The radio interference limits stated by EN55022 (European Norm) have its background in CISPR 22 (Comité International Spécial des Perturbations Radioéléctriques).

Before discussing the actual limits in this frequency range, let us have a look into the background for the chosen figure and what type of information the limits are based upon.

In a document from CISPR/B (Secretariat) September 30, 1983 the interference limits in the frequency range 30 – 1000 MHz are presented. This is a second draft recommendation for data and electronic office equipment (DP/OE) and takes account of the detailed consideration of the comments on Document CISPR/B (Central Office) 9. It was prepared by CISPR/B/WG2.

In section 2.1 we have the definitions of *Data Processing and Electronic Office Equipment (DP/OE)*:

> "Electrical/electronic units or systems which predominantly generate a multiplicity of periodic binary pulsed electrical/electronic wave forms and are designed to perform data processing functions such as electronic word processing, electronic computation, data transformation, recording, filing, sorting, storage, retrieval and transfer, and reproduction of data/or images *but excluding equipment designated solely for telecommunications purposes*". The emphasizing done by myself.

In CISPR Publication 22 First edition 1985, *Limits and methods of measurements of radio interference characteristics of information technology equipment*, we have mainly the same definition of the source of interference but under the abbreviation ITE – Information Technology Equipment. The interesting difference is that the exclusion of equipment designated solely for telecommunications purposes is removed.

What had happened from September 1983 to the official publication in 1985? The limits of interference are the same, "equipment designated solely for telecommunications purposes" seems to have no influence upon the chosen figures for recommended limits of interference. The only change is that in the 1985 Publication there is a section 4.2, *Limits of telecommunications line interference voltage*, which is "under consideration".

The interesting point is now what the interference limits are based upon?

We have a report in Norwegian which gives some background material. Unfortunately, we do not have the title of the report, but it is probably written by a delegate working in the area of radio interference, Radiostøykontrollen (the Radio Interference Service).

Section 5.5 in this document gives the background for limits and methods of measurements, *Bakgrunn for grenseverdier og målemetoder*.

In para *5.5.1 General,* we have:

"CISPR/B has concluded as follows:

- The potential of interference from DP/OE is different for professional use than for the common use at home.
- DP/OE commonly for professional use, e.g. within production, medicine, education, official business etc., may be used both in the domestic area and in the commercial and industrial area.
- In professional use we may have different factors of propagation compared to domestic use.
- These differences in propagation factors require different recommendations for interference limits.
- Electromagnetic energy may switch from the source to nearby radio system by conducting, inductive or capacitive coupling and radiation.
- In the frequency band below 30 MHz the dominating factor of coupling is by the conducting networks and the inductive and capacitive coupling from the same.
- In the frequency band above 30 MHz radiation is the dominating factor, but in some installations the inductive and capacitive coupling from cables connecting host units to peripherals may be of importance.
- In the frequency band from 30 to 300 MHz the source of radiation from smaller equipment giving interference will be the mains or other cables connected, mainly the cables nearest the equipment".

In the document from September 1983 we have the definition of Class A and B equipment:

> "Class A equipment is data processing and electronic office equipment which satisfies the Class A interference limits but does not satisfy the Class B limits. In some countries, such equipment may be subject to restrictions on its sale and/or use.
>
> Note: The limits for Class A equipment are derived for typical commercial establishments for which a 30 m protection distance is used. The Class A limits may be too liberal for domestic establishments and some residential areas.
>
> Class B equipment is data processing and electronic office equipment which satisfies the Class B interference limits. Such equipment should not be subject to restrictions on its sale and is generally not subject to restrictions on its use.
>
> Note: The limits for Class B equipment are derived for typical domestic establishments for which a 10 m protection distance is used."

The most interesting point here is to see that also for Class B equipment the CISPR committee recommended a protection distance of 10 m from the source of interference to the victim. In 1983 this may have been a reasonable distance between electronic equipment, but the question is: what is the reality today and what about the electromagnetic environment for the 1990s and beyond 2000?

When going into the background material we see that the maximum radiated field strength by a receiver placed at least 30 m (E30) from a class A DP/OE is given by:

$$E30 = ES - S/N + AB$$

where

- $E30$ = radiated limits of interference from professional DP/OE at least 30 m from a receiver in the chosen broadcasting band
- S/N = signal/noise ratio for the chosen broadcasting service
- AB = attenuation factor for the building
- ES = usable field strength for the receiver.

The calculated field strength for the chosen band is given in Table 1.

Investigations published by CBEMA (Computer and Business Equipment Manufacturers Association) shows that 89 % of the receiver antennas within 100 m from a class A DP/OE are located at least 30 m away from the installation.

Table 1 Calculated field strength for the chosen band

Frequency range MHz				
Primary service	54–88 TV	80–108 FM	108–136 Aero	470–1000 TV
ES dB (µV/m)	68	60	15	74
S/N dB	45	30	0	45
AB dB	8	8	21	8
E30 dB (µV/m)	31	38	36	37

Remarks:

- ES in TV band is minimum grade A "Statistical expected" usable signal defined by FCC.
- ES in Aeronautic band is chosen to 118 MHz where air–ground service starts because the interference potential by ground stations is higher.
- Experiments and literature studies show that buildings nearly almost give attenuation of signals. A factor of 8 dB attenuation for a single wall was chosen except in the aeronautic band.
- Empirical results published by CBEMA show a building attenuation from 21 to 39 dB in 27 American airports. A factor of 21 dB is used for all aeronautic band.
- The S/N ratio is found empirically by use of a CISPR quasi-peak receiver as reported by CBEMA.

Furthermore, is was stated that the quasi-peak receiver specified by CISPR publication 16 part 1 will give the same protection for the TV receiver both against wideband interference as well as narrowband signals. Therefore, the limits will be the same for both type of interference in the range 30 – 1000 MHz by use of a CISPR quasi-peak receiver.

For the Class B DP/OE a protection distance of 10 m was chosen. This distance should give sufficient protection when installing DP/OE in domestic areas.

The protection distance was based upon the following supposition:

- The receiver should be protected from the neighbour DP/OE
- There will be relatively few receivers closer than 10 m
- In case of less than 10 m distance, other factors of contributions may reduce the possibility of interference. These may include, but are not restricted to:

  1 Orientation of source and receiver

  2 More than 8 dB attenuation in the interjacent walls (also more walls)

  3 The frequencies with the highest radiation do not fall within the receiver selectivity.

The measuring distance was chosen to be the protection distance of 10 m. Since all conditions except protection and measuring distances are the same, the limits evaluated for class A equipment in the 30 m distance will give the same protection for the receiver also for class B equipment at a distance of 10 m.

Table 2 gives the interference limits in the frequency band 30 – 1000 MHz for class A and B equipment.

The aerial should preferably be located at the horizontal distance from the test unit as specified above. If the field strength measurement at 30/10 m cannot be recorded because of high ambient noise levels or for other reasons, measurements may be made at a closer distance. If so, an inverse proportionality factor of 20 dB/decade shall be used to normalize the measured data to the specified distance for determining compliance.

Comparing the E30 values in Table 1 to the interference limits in Table 2 we see a high degree of conformity. We can conclude that the figures in Table 1 are the basis for the EN55022 interference limits.

Radio communication and interference limits

The ES values in Table 1 give the "Statistical expected" usable signal defined by FCC.

In CCIR Rec. 412-4 and 417-3 1986, CCIR

"Unanimously recommends:

1 that when planning a television service in Bands I, III, IV or V, the median field strength for which protection against interference is planned should never be lower than:

Band	I	III	IV	V
dB(µV/m)	48	55	65(1)	70(1)

(1) The values for Bands IV and V should be increased by 2 dB for the 625-line(OIRT) system".

CCIR also assumes that, in the absence of interference from other television transmissions and man-made noise, the minimum field strength at the receiver antenna that will give a satisfactory grade of picture, taking into account receiver noise, cosmic noise, antenna gain and feeder loss is 47 dB(µV/m) in Band I, 53 dB(µV/m) in Band III, 62 dB(µV/m) in Band IV and 67 dB(µV/m) in Band V.

For the FM sound broadcasting at VHF, CCIR "Unanimously recommends" that the following planning standards should be used for frequency-modulation sound broadcasting in band 8 (VHF):

1 Minimum usable field strength

In the presence of interference from industrial and domestic equipment a satisfactory service requires a median field strength (measured 10 m above ground level) of at least:

1.1 for the monophonic service:
48 dB(µV/m) in rural areas
60 dB(µV/m) in urban areas
70 dB(µV/m) in large cities

1.2 for the stereophonic service
54 dB(µV/m) in rural areas
66 dB(µV/m) in urban areas
74 dB(µV/m) in large cities.

When comparing these values with CISPR emission levels, the different services require different radio-frequent protection ratios. CCIR has, in recommendations 412-4 and 655 (1986), given protection ratios for FM broadcasting and television services. The given figures are only valid for interference from FM or TV transmitters and other types of radio interference are not considered. In [1]

Table 2 Interference limits in the frequency band 30 – 1000 MHz for class A and B equipment

Class	Frequency range MHz	Quasi-peak limits dB (µV/m)	Test distance m
A	30 – 230	30	30
A	230 – 1000	37	30
B	30 – 230	30	10
B	230 – 1000	37	10

some investigations have been done concerning interference from different types of radio sources onto FM sound broadcasting and television services. The typical interference spectrum from a modern IT-equipment or tele-terminal (ISDN) is multiples of clock frequencies up to some hundreds of MHz and with a white noise spectrum 10 – 20 dB below the harmonics. This white spectrum is generated in the processor chips with its data lines and seen from the outside this acts as a random pulse generator with white noise spectrum.

If we consider the FM stereophonic broadcasting the NTR report measures the protection ratio to 42 dB when the interference is white Gaussian noise. The quality criteria are in accordance with CCIR rec. 641 with a S/N-ratio of 50 dB. For comparison CCIR recommends a protection ratio of 45 – 51 dB for stereophonic broadcasting when a standard FM signal is used as interferer. From the figure of field strength and protection ratio, we can now calculate the maximum interference level which can be accepted in the vicinity of the receiver antenna for an acceptable reception quality.

In the case of FM stereophonic service we have the following figure for maximum interference:

Large cities:
  74 dB(µV/m)
- 42 dB (protection ratio)
= 32 dB(µV/m)

Urban areas:
  66 dB(µV/m)
- 42 dB (protection ratio)
= 23 dB(µV/m)

Rural areas:
  54 dB(µV/m)
- 42 dB (protection ratio)
= 12 dB(µV/m).

These figures can now be compared with CISPR 22 Class B as mentioned above: 30 dB(µV/m) in 10 m distance from interferer. The figure for large cities is quite compatible with the calculated one, but when we go outside the city we may have problems. Especially in the rural areas where the wanted signal is weak, the interference from IT equipment may have severe consequences. One may ask why we have different figures from CCIR for the broadcasting services in cities and rural areas. The obvious explanation is the higher interference level. This interference has a lot of sources; heavy industrial plants with high powered machinery, high density of cars causing a high ignition interference and a higher utilization of the radio spectrum for communication purposes. But when we come to the use of IT equipment, our private, in-house environment is the same regardless of it being a flat in a large city or a villa in the rural areas. So, if we are talking of interference from IT equipment, there should be no difference depending on the customers place of residence. The figures from CCIR seem to take no consideration of the wide use of IT equipment even in non-commercial areas. Going back ten years, this is not so surprising, but we have to draw the consequences of the IT revolution.

The same exercise as for the FM broadcasting can be done for television services. As shown above we have the figure 48/55 dBµV/m in band I/III "for which protection against interference is planned". In the same CCIR recommendation we have the figure of radio-frequency protection ratio. This is based upon TV-signals as interferer and figures for CW signal or white noise are not given. As mentioned earlier, some comparisons between white noise and FM and TV as interferers have been done in [1]. When using a 120 kHz BW and QP detector the actual protection ratio is measured at 42.3 dB. This has to be compared with the continuous interference from 33 – 52 dB where TV signal is used. The variation of protection ratios depends on frequency offset in multiples of 1/12 line-frequency.

The maximum interference can now be calculated for band I/III

	48/55	dB(µV/m)
-	42	dB (protection ratio)
=	6/13	dB(µV/m)

Comparing with the interference of 30 dB(µV/m) 10 m away from IT equipment we obviously have a problem. The CCIR recommendation refers to field strength at a height of 10 m above ground level but no figure of antenna gain is given. Using an outdoor antenna we may have a front/back ratio of 10 dB, so the interference level can be increased to 16/23 dB(µV/m). The missing 14/7 dB for band I/III can now be achieved by increasing the distance from IT equipment to antenna to 50/22 m, turning the equipment in a direction where the radiation lobe has a minimum or by adding 14/7 dB to the wall attenuation. In practice, a combination of all these adjustments will be the best solution.

In both these example, the FM and TV broadcasting has the potential problem of interference from IT equipment which is in line with CISPR 22B.

Other types of radio communication may have different requirements to interference immunity. We will go through some examples illustrating the great variations of external interference in which the radio system is intended to operate.

An essential factor in all calculations involving field strength is the antenna factor. In the following we will use the equation:

$$AF(\text{dB}) = 20\log(f,\text{MHz}) - 29.8 - G(\text{gain, dBi})$$

In some installation we may have some dB gain, but for interference calculations where no information is given regarding antenna position and interference source we use isotropic antenna with $G = 0$ dB.

When referring to interference field strength we have to be aware of the difference in bandwidth for the different radio systems. In the examples given below we will use the equivalent bandwidth of 120 kHz which is in line with CISPR 22. From this follows that we are talking about *broadband interference* and not single frequencies generated by harmonics.

- Radiotelephony, 80 MHz, analogue:

Sensitivity		0.0	dB(µV)
AF	+	8.3	dB
C/I	-	8.0	dB
Imax		0.3	dB(µV/m)/10 kHz
Imax (120 kHz)	=	11.3	dB(µV/m)

- FM car-radio mono, 100 MHz:

Sensitivity		48	dB(µV/m)
C/I	-	36	dB
Imax	=	12	dB(µV/m)/165 kHz
Imax (120 kHz)	=	10.6	dB(µV/m)

- Paging, 170 MHz

Sensitivity		25.0	dB(µV/m)
C/I	-	10.0	dB
Imax	=	15.0	dB(µV/m)/10 kHz
Imax (120 kHz)	=	25.8	dB(µV/m)

- Mobile telephony analogue, 460 MHz

Sensitivity		4.0	dB(µV)
AF	+	23.4	dB
C/I	-	18	dB
Imax	=	9.4	dB(µV/m)/10 kHz
Imax (120 kHz)	=	20.2	dB(µV/m)

- Mobile telephony analogue, 900 MHz
  Sensitivity 4.0 dB(µV)
  AF + 29.3 dB
  C/I - 18.0 dB
  Imax = 15.3 dB(µV/m)/10 kHz
  Imax
  (120 kHz) = 26.1 dB(µV/m)
- Mobile telephony digital, 900 MHz
  Sensitivity 6.0dB(µV)
  AF + 29.3 dB
  C/I - 9.0 dB
  Imax = 26.3 dB(µV/m)/230 kHz
  Imax
  (120 kHz) = 23.5 dB(µV/m)

The figures for sensitivity are of vital importance for these calculations. For radio telephony in the 80 MHz band the sensitivity is decreased by 6 dB from the receiver's performance under static conditions due to fading under normal use, and we accept a C/I ratio of 8 dB. This will give a reasonably good speech quality, but not up to telephone standard. For mobile telephony in 460 and 900 MHz band (NMT), we have used the receiver sensitivity under fading condition. For digital systems in 900 MHz (GSM) we have used the reference sensitivity + 3 dB due to the higher BER with fading. The paging performance data are taken from the ETSI recommendation. For FM and TV broadcasting the CCIR recommendation is used. For the mobile system the equivalent bandwidth is the 3 dB bandwidth with standard signalling measured on the running system.

The maximum interference seen from a radio engineer's point of view can be summarised in Table 3.

One obvious conclusion can be drawn from Table 3: the higher the immunity, the higher the frequency. This is because the antenna factor is increasing by 6 dB/octave, or in other words, the receiver sensitivity is decreasing by 6 dB/octave when we refer to field strength with equal antenna gain. So, when we keep the interference field strength constant in a given frequency band, the interference power delivered to the receiver is dropping by the same 6 dB/octave when we use an antenna with a constant gain. This frequency dependency may be turned the other way round; the higher the radio frequency, the higher the wanted signal must be in terms of field strength. This is a general problem with all radio communication and is one of the reasons for the reduction of radio link distances with higher radio frequency.

The other interesting figure which can be seen from Table 3 is the comparison between analogue (NMT) and digital (GSM) when talking about interference immunity. The lower immunity to GSM, in spite of the 9 dB lower C/I ratio, is of course due to the fact that GSM requires a bandwidth of 230 kHz while NMT only occupies 10 kHz.

From a practical point of view, we see that the most exposed radio system is the analogue system in the lowest frequency band. In Norway we use the band I for TV transmission, and in rural areas where the TV-signal is on the margin, the introduction of IT equipment for common use may well cause problems for TV broadcasting. This frequency band is also the band with the highest interference level. When measuring interference from IT equipment we generally have a falling interference level above some

Table 3 Maximum interference to radio systems

Radio system	Max. interference dB(µV/m) 120 kHz	CISPR22B dB(µV/m) 10 m
TV band I	6	30
Radiotelephony 80 MHz	11.3	30
FM car radio, mono	10.6	30
FM stereo	12 – 32	30
TV band III	13	30
Paging 170 MHz	25.8	30
Mobile telephony analogue 460 MHz	20.2	37
Mobile telephony analogue 900 MHz	26.1	37
Mobile telephony digital 900 MHz	23.5	37

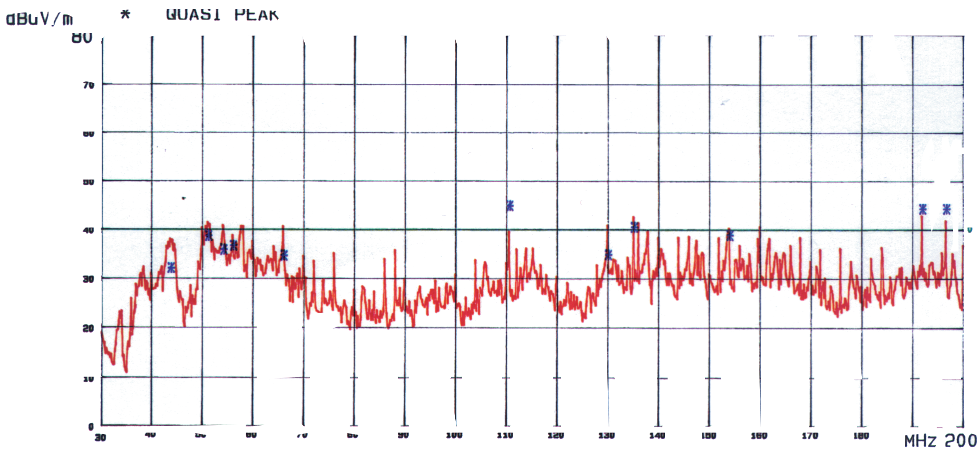


Figure 1 Interference ISDN terminal 30 – 200 MHz

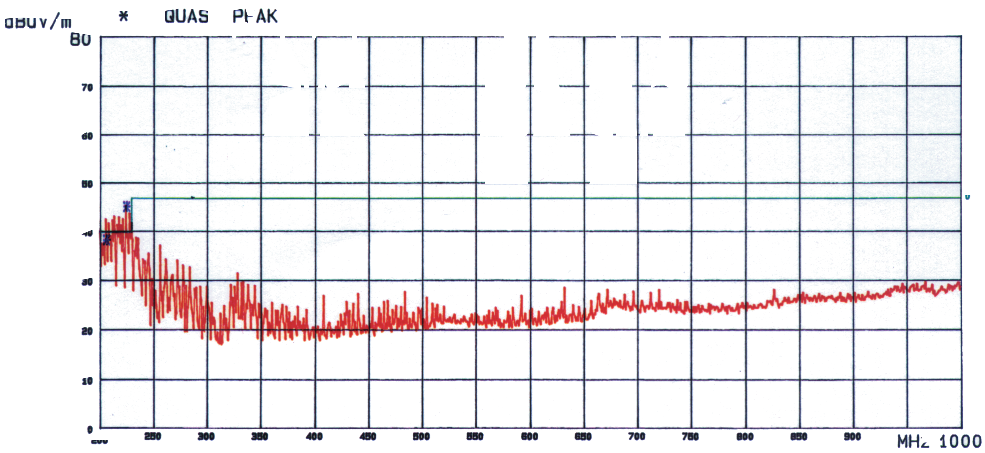


Figure 2 Interference ISDN terminal 200 – 1000 MHz

hundreds of MHz. This is of course related to the clock frequency and the processor speed, but with today's IT equipment with clock frequencies up to 50 MHz the highest interference level is in the frequency band 30 – 200 MHz. Above 200 MHz we may have high levels from harmonics of internal oscillators, but the "white" processor noise is normally reduced by more than 20 dB.

To illustrate a typical radio interference spectrum, we have in Figures 1 and 2 given the frequency plot of an ISDN terminal from 30 to 1000 MHz. According to CISPR the resolution bandwidth is 120 kHz and the green line is the CISPR 22B limit. For the quasi peak measurements the test receiving antenna and the terminal are orientated for maximizing the field strength.

In the range below 70 MHz we see the "white" processor noise. Above 70 MHz the spectrum is dominating by the harmonics of the internal oscillators with a level around 10 dB higher than the processor noise. This is very typical for this type of equipment. When we pass 230 MHz we see that the interference is dropping by 20 dB.

The probability of disturbances from IT equipment onto TV transmissions is of course influenced of many factors. One import factor is the fact that the radio interference must be within the receiver band, we must have "in-band" interference. The broadband processor noise may be 10 – 20 dB below the harmonics peak values. Taking 15 dB as a normal peak/broadband factor, we still have a too high interference level for the radio systems in the lowest frequency band. For the paging system we are just on the limit when the distance from source to victim is 3 m.

In many of the radio systems mentioned here, the 10 m distance source – victim is unrealistic. For FM and TV this can be achieved by using outdoor antennas, but for the mobile terminal in car installation or body worn equipment like pager or portable mobile telephone the distance in an ordinary office environment is in the range of less than 3 m. That means a 10 dB higher interference level from the disturbing IT equipment.

As seen from Table 3, the most exposed radio system is in the lower frequency band. As mentioned earlier, this band is also very attractive from a radio transmission point of view. New systems, especially the digital ones, need more spectrum and are usually located around or above 1 GHz. In this band the man-made noise is negligible in most environments, but here we have the drawback with radio link distance.

So if you do not receive the message on your pager when staying in a modern office environment, the reason may well be the interference from all the IT equipment and not the missing radio planning from the telecom operator. CISPR never considered today's penetration of "Data Processing and Electronic Office Equipment, DP/EO" when the limits of interference were stated.

What about ETSI and their contribution to the compatibility between IT and radio equipment?

In prETS 300 339 June 1993 *"Radio Equipment and Systems (Res); Generic Electro-Magnetic Compatibility (EMC) for radio equipment",* we have criteria for emission and immunity. The *"ETS is based upon the Generic Standards EN 50081-1 [3], EN 50082-1 [4] and other standards, where appropriate, to meet the essential requirements of the Council Directives 89/336/EEC [1] and 92/31 [2]. The ETS applies to domestic, commercial light industry and vehicular environments only and therefore may not give sufficient protection in more hostile environments",* and *"This ETS incorporates normative annexes which specify test methods relevant to a number of tests in this ETS. These test methods have been taken from the latest available information on the development of EMC Basic Standards in the International Electrotechnical Commission (IEC) and Comité Européen de Normalisation Éléctronique (Cenelec)".*

In all the international recommendations and standards the emission limits are based upon CISPR 22 which, as seen above, has its background material from a world totally different from today's reality when talking about electromagnetic environment. The sentence *"The limits for class B equipment are derived for typical domestic establishments for which a 10 m protection distance is used"* is among the key points. Today's tendency of increasing use of the radio spectrum for personal communication needs highlights the electromagnetic pollution, and an increased effort of "cleaning up" the radio spectrum is needed.

The mobile radio transmitter as an interference source to other electronic equipment

We have been talking about the IT equipment as a potential source of electromagnetic pollution, but this is only one side of the compatibility problem in the electromagnetic world.

All radio transmitters, like the mobile telephone, are intended to radiate electromagnetic energy. This radiation is what we can call "wanted" in contrast to the electromagnetic energy radiated from non-radio IT equipment whose communication is on copper cable or optic fibre. If we use Radio-LAN for communication we again have an ordinary radio transmitter included in the IT equipment.

The radio transmission is regulated in CCIR and co-ordination between different radio systems are part of the radio frequency management work. If this co-ordination is done as intended and the guidelines for transmitter power, antenna gain, antenna position etc., are respected, the radio systems should not give any interference to each other. The problem is the potential interference to the non-radio electronic equipment which can be used in the vicinity of the radio transmitter. This problem is not a new one. In the neighbourhood of high powered AM transmitters we have for several years had problems with interference to electronic equipment, particularly the telephone sets including audio amplifiers.

In the 27 MHz band, the private radio band, the amplitude modulation was prohibited in the early 1980s due to the interference to other electronic equipment.

The problem with AM is the way in which the demodulation can take place. All kinds of amplifiers have some non-linearity and the non-linearity itself demodulates the carrier giving the interference as a result. It is easily shown that when you express the in/out relation in an amplifier as seen in Figure 3 in terms of power series

$$U_{out} = a_0 + a_1 U_{in} + a_2 U_{in}^2 + a_3 U_{in}^3 + \dots + a_n U_{in}^n \qquad (1)$$

and insert the AM equation

$$U_{in} = U_o \cos(\omega_o t)(1 + \mathrm{m}\cos(\omega_m t)) \qquad (2)$$

where ω_o represents the carrier frequency, m the modulation index and ω_m

the modulation frequency, one term in the output signal is

$$a_2 m U_o^2 \cos(\omega_m t) \qquad (3)$$

This is the demodulated signal where we have the quadratic function from input to output.

In [2] we have done interference measurements on 12 analogue telephone sets. The reason for these measurements was the audible 217 Hz interference from the GSM telephone. As shown in Figure 4 the TDMA structure in GSM generates most of its energy in the lowest audio band where an ordinary telephone is most sensitive. In Figure 5 we see the 217 Hz interference and its harmonics together with the reference 1000 Hz tone with level -10 dBm. The measurements are done with 0.8 W GSM peak power 3 m away from the PSTN telephone. In this instrument set-up with perfect reflecting ground the electric field exposure to the test object is 3.7/3.9 V/m for vertical and horizontal field.

When measuring the interference as a function of the GSM RF effect, we see the quadratic function mentioned above: 1 dB increase in RF level gives 2 dB increase in audio interference.

This compatibility problem is due to the missing immunity standard for PSTN terminals. The GSM system operates as intended, the TDMA structure gives the amplitude pulse modulation and from an EMC point of view this problem should have been foreseen.

Hearing aids have the same problem as telephone sets. A person wearing the hearing aid cannot be a customer of the GSM telecom operator, he may use the car-mounted mobile telephone set, but a handheld GSM telephone may give a field strength exceeding 30 V/m to the hearing aid. The coming immunity recommendation for teleterminals will be 3 V/m. To achieve 30 V/m a completely new design for hearing aids is required.

Here we see again the compatibility problem with the penetration of personal radios and electronic equipment where the designer has paid insufficient attention to the RF layout or filtering.

The problem can only be solved by a total EMC concept where all the aspects of the electromagnetic environment are taken care of before launching new radio systems.

References

1 Hauger, E, Eriksen, J. *ISDN-terminaler som potensielle støykilder for kringkasting.* Kjeller, Norwegian Telecom Research, 1991. (Report TF R 50/91)

2 Hauger, E. *Interference from the TDMA structure in digital mobile communication to PSTN.* Kjeller, Norwegian Telecom Research, 1992. (Report TF R 40/92)

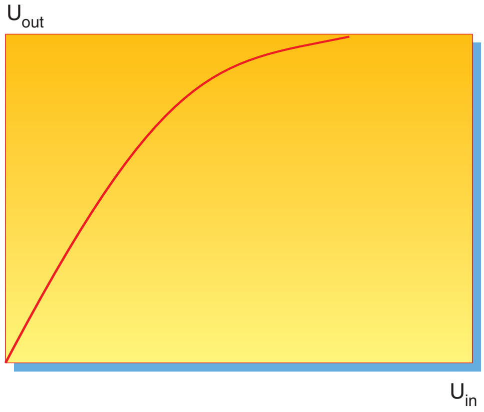


Figure 3 Amplifier in/out characteristics

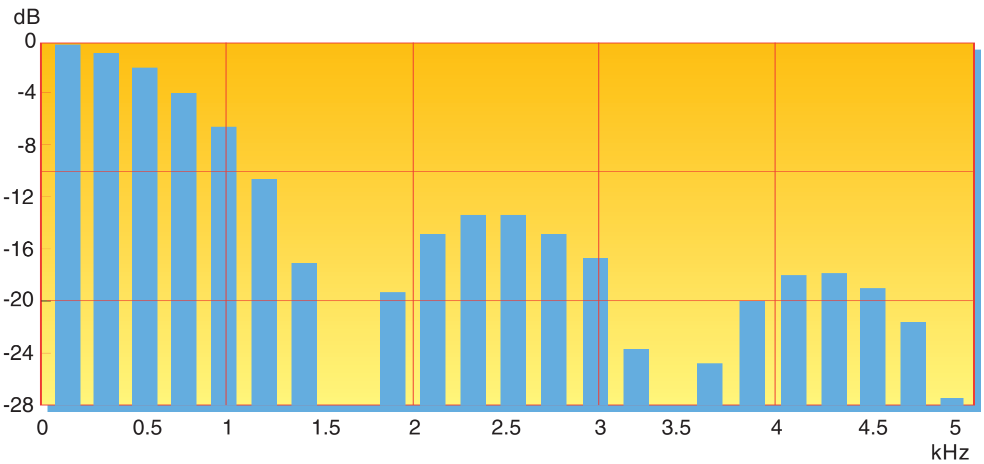


Figure 4 GSM TDMA frequency spectrum

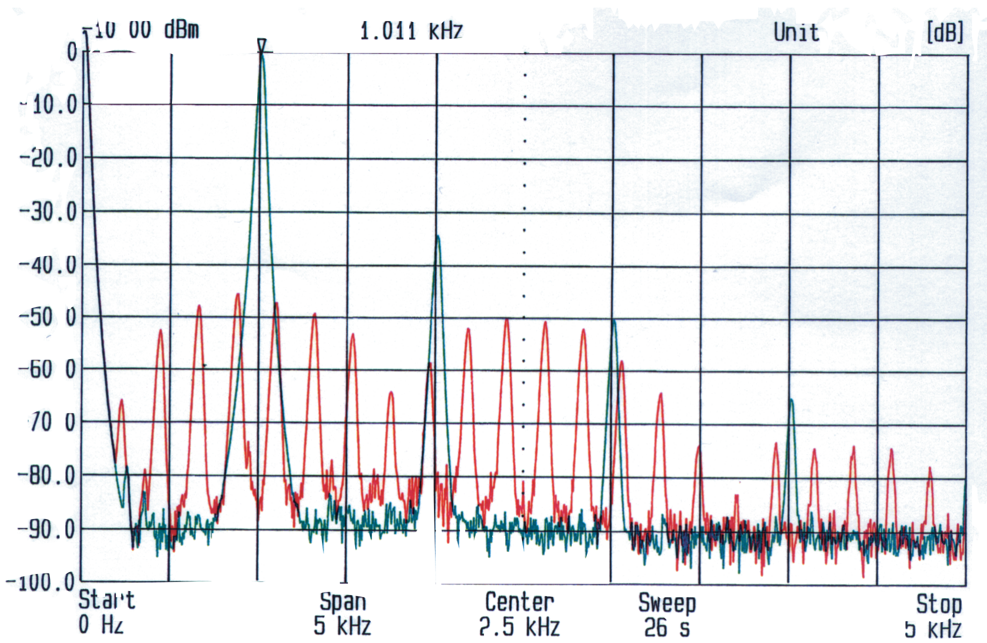


Figure 5 GSM TDMA interference in analogue telephone set

Design of electronic equipment with EMC in mind

BY AGNAR GRØDAL

Introduction

The design process of electronic equipment has traditionally concentrated on making a reliable product according to its functional specification, and at a competitive price.

Tests have shown that design along these traditional guidelines often fails the EMC Directive.

The multitude and complexity of electronic equipment in modern society have also led to numerous reliability problems and to accidents caused by unintentional interactions between electronic equipment of different types.

The EMC Directive has almost simultaneously come with a change in component and production technology, namely the surface mount technology. This technology gives important technical benefits, but it may also give rise to problems with respect to EMC.

Design for EMC thus imposes new restrictions to the design process, and more and better competence is required in order to succeed.

During the last few years we have been through a period where the design approach has been to concentrate mainly on meeting the functional specification, then to test the prototype and hope for the best with regard to EMC. This trial and error method inevitably leads to costly time consuming testing, fixing, new testing, and so on.

Lack of EMC experience is the main reason why this design method has been used. Experience over the last few years has shown that we today have a better understanding of the factors that will influence EMC. With this knowledge it is now possible to design economical competitive electronic equipment that has a fair chance of passing the relevant EMC test criteria, even at first attempt.

It may seem that too much time and money is spent on discussing EMC standards and levels of emissions, relative to the time spent on discussing how to meet these standards and levels of emissions by using adequate design principles. It is true that there is more broad knowledge available on standards and limits than there is in general, effective EMC design principles. Both disciplines, however, must be dealt with.

Specification

The starting point of a design process should always be to write the product specification. The product shall meet mainly two specifications:

- The functional specification, and
- The EMC specification.

It is important to incorporate the EMC specification at this early stage, because both specifications will influence the logical design, the choice of components, the number of layers in the printed circuit boards, and the enclosure.

Early knowledge of the EMC specification will also guard against a possible costly overdesign.

The EMC specification will be given in terms of norms and limits relevant to the product type. It is important to realize that these norms and limits reflect minimum performances for the equipment to pass legal requirements according to the EMC Directive.

Practice indicates that the EMC specification for some product types has to be more stringent than what is required to pass the legal minimum limits. The increasing use of mobile radio telephone systems has clearly shown that the immunity level of 10 V/m, which is the level most systems are designed for, may be far too low for certain types of electronic equipment. This fact will impose new and even more stringent restrictions to the design process.

The specification shall be tailored to CE-marking of the product. CE-marking demands that all relevant requirements shall be met, not only the EMC requirements according to the EMC Directive. A common additional requirement is that the safety norm EN 60950 shall be met. For equipment connected to the public telecom network, the Directive on Telecommunications Terminal Equipment 91/263/EEC applies. This directive poses requirements on EMC and safety, but also demands that the equipment shall be produced according to a quality standard, normally ISO 9001.

The specification should preferably be written in co-operation with the client. Alterations to the specification later in the design process can have severely costly and time consuming consequences, and may require:

- Extra filter components
- Components for overvoltage protection
- Shielding
- Screened connectors and screened cables
- Extra “real estate” on the printed circuit board.

A typical EMC specification will, as an example, dictate EN 55022 A or B for emission, and the IEC-801 (EN 61000-4) series for immunity. It is important that the relevant immunity severity level is indicated.

Components

The EMC challenge has come almost simultaneously with a radical shift in component technology. The reason for this is mainly new requirements from the surface mount (SMD) technology. The integrated circuit chip technology has not, however, altered as much as their encapsulation.

Passive components, mainly non-electrolytic capacitors, have undergone the most radical change. The reason for this is the massive change from hole-mounted plastic film capacitors to ceramic capacitors. Some ceramic capacitor types have parasitic properties that make their capacitance change radically with temperature and applied voltage. Wet electrolytic capacitors for surface mounting represent another problem, as these may be harmed by the soldering process.

The change in component technology has unfortunately led to a decline in the circuit designers’ understanding of which component types should be used where in the circuit design. In many companies there seems to be a shortage of qualified component experts. The component work has in some cases been reduced to mere registering and purchase. Using primarily CECC approved components is a good help in such cases.

The following list exemplifies new types of components used in modern logic designs:

- SMD components
- Components with small pitch (distance between terminals)
- Components with low output impedance, producing strong switching current transients
- Components with high edge rates
- Components with high output level, switching rail to rail

- Components using high clock rate, typically up to 50 MHz
- Microprocessors with wide buses
- Microprocessors with high bus activity.

Some of these components have qualities that are advantageous for realizing logical circuits with advanced system specifications. But the same qualities may also be disadvantageous with respect to EMC. CMOS integrated circuits from the “Advanced” series is a typical example. These circuits are capable of driving high capacitive loads at very high clock rates, and are often used in fast microprocessor bus circuits, but they are very noisy.

The change some years ago from Low Power TTL circuits to High Speed CMOS circuits was a great success. Practice has shown, however, that these are noisier than the corresponding LSTTL circuits, but the EMC threat was not an issue at the time of shifting from LSTTL to HCMOS.

The widespread use of the newer Advanced CMOS circuits has caused so many and serious EMC problems that many designers would rather use older TTL circuits, were it not for the TTL power dissipation problem.

Help has come in the form of the BICMOS circuit families, circuits with low power CMOS “kernels” and with TTL low noise and high drive capability. The shift to high speed 3 V families will also be a success with respect to EMC.

Some newer component types have been designed specially to achieve better EMC qualities. Examples are:

- Components with short terminals, giving lower parasitic inductivity
- Components with controlled output pulse edge rates (no sharp pulse corners)
- Components with better V_{CC} and ground pinout (centre pinout instead of corner pinout)
- Components with low V_{CC} voltage (3.3 V)
- Microprocessors with separate V_{CC} pinout for kernel and bus drivers.

Some general advice for selecting component types:

- Use fast, low output impedance components only where necessary for the function of the circuit
- Avoid using Advanced CMOS components, especially byte-wide types
- BICMOS types are preferable when high speed and high drive capability is needed
- Low voltage types are favourable
- Use special components for clock phase distribution
- Use special components for driving backplanes.

Be aware of the prices, components with favourable EMC attributes may be expensive!

General sources of EMC problems

The causes of the EMC problems can *simplified* be derived from Maxwell’s equations as:

Emission
$= K_1$ x Area x Current x Frequency^2

Susceptibility
$= K_2$ x Area x Field strength x Frequency

See Figure 1.

Observe that the term “Frequency” appears squared in the equation for the emission.

The term “Frequency” in the equations includes both fundamental frequency and its harmonics. A 1 MHz square-wave signal may thus emit noise up to tens or even hundreds of MHz. The task of reducing the emissions part of the EMC problem can be simplified as:

- Reducing the Frequency
- Reducing the Current
- Reducing the Area.

It is in fact, at least theoretically, that simple.

Solutions

Whereas the theory for reducing the EMC problems may seem simple, practice may be more complicated. The problems may in principle be solved by either:

- Designing the electronics without paying attention to EMC design rules, but instead to encapsulate the emissions within the products’ enclosure by using screens, tight metal boxes, filters on the cable outlets, and by using screened connectors and cables, or by
- Designing the electronics so that the generation of emissions is sufficiently limited to pass the relevant tests; hopefully without the use of costly mechanical screening.

The latter method is by far the cheapest and the most effective, even if some additional screening and filtering have to be used.

The use of extensive screening may, however, work. This may even be cost effective when trying to prolong the sale of an older design, but the method has its limitations. The enclosure will normally be more expensive, it may introduce cooling problems, and worst of all; the strong internal emissions may sneak out along the cables.

Designing electronics with little radiation requires wide competence within electronic design and manufacturing, and it is of utmost importance that involved departments and individuals co-operate.

The component price for a low-emitting design will in most cases not be significantly higher than for a comparable higher emitting design. The difference in design time for the circuit diagram will normally also be insignificant.

The printed circuit layout on a low-emitting board will normally take longer to design, and the layout designer must be EMC competent. The board will normally also require more layers than a conventional two-layer board. Low-emitting boards will thus be more expensive.

Some years ago, before the “discovery” of EMC, the electronics designer and the circuit board designer would perform their tasks without even knowing one another. Today, they have to co-operate closely to succeed with EMC. That also applies for the departments or persons responsible for the EMC-, component-, and design review activities.

All relevant expertise needed to succeed is seldom found within one company. Help must often be hired from external consulting firms specializing on EMC issues.

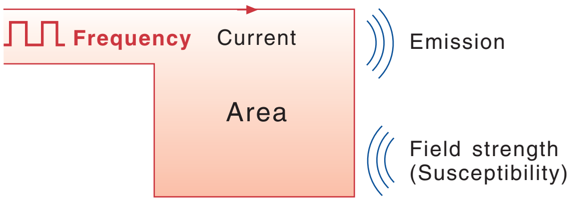


Figure 1 Simplified illustration of Maxwell’s law

The circuit board designer may have automatic routing programs at his disposal, which pay attention to EMC to some degree. Practice shows that most autorouting programs so far are of limited value, because they very often generate layouts with poor EMC performance.

Self compatibility

Newer component technology with lower output impedance, lower rise and fall times, higher signal levels, higher transient currents, and smaller distance between signal tracks on the printed circuit boards, have given rise to numerous problems with internal noise on the PCBs. As a consequence, circuits may disturb one another within the same PCB. A board that disturbs its own function is said to have low self compatibility.

Such problems are often marginal, and may come or go with a slight shift in the power feed voltage, with a shift in ambient temperature, with ageing, or with the use of “identical” components from a “second source” vendor.

Such circuit malfunctions may not be detected by simulation or testing. The faults may be detected or avoided by

- Applying good EMC oriented design rules
- Employing an EMC competent design reviewer
- Using more and better simulation tools
- Employing an EMC competent circuit board designer, paying particular attention to component placement and how the tracks are routed.

Immunity/susceptibility

The words immunity and susceptibility, despite having opposite meanings, are used to describe an electronic system’s behaviour when exposed to overvoltages or to an electromagnetic field.

Insufficient immunity may lead to permanent destruction of one or more components, which requires manual repair for the equipment to resume normal functioning. Less serious failures may lead to a manual reset or restart of the equipment. Immunity can be designed into the system by using the same EMC rules as used to limit emissions. In addition to that, it may be necessary to add special, board space consuming, overvoltage protection components. It is important to be aware of the fact that the overvoltage protection circuitry of the equipment shall and will be tested according to levels specified by regulations. Real life conditions may expose the equipment to even higher levels.

It is also important not to overdesign the protection, especially on boards connected to a multitude of external signal lines. Practice has shown that a moderate protection circuitry may be sufficient, even when calculations point to the use of more costly and larger components.

Emission, principles of design

Alternative equivalent functioning circuit boards, but using different design principles or using different components, may have a different noise emission spectrum. The difference in noise level may be considerable. Price and power dissipation may differ a lot, too.

It is of prime importance to choose design principles that have the lowest possible price, the lowest possible power dissipation, sufficiently low emissions, and sufficiently high immunity. One should not aim for a better EMC performance than what is needed, as that will normally have consequences for price and power dissipation. Some alternative design principles with different EMC qualities will be mentioned:

- Do not use a higher microprocessor clock frequency than what is necessary. This is perhaps the most common design error. Designers tend to use the same clock frequency as they did in their previous designs, without checking what is really needed. Because the noise depends on the square of the frequency, the noise will be reduced by 75 % – at least theoretically (and simplified) – if the frequency is halved.
- If possible, use a microcontroller instead of a microprocessor. The microcontroller has the bus system built inside the processor chip. The bus system radiating area, which is usually one of the noisiest parts of the design, will then be greatly reduced.
- Do not use a clock controlled design if other design principles can be utilized. This is perhaps a naive suggestion, as most of today’s designs use microprocessors.
- Do not use multiplex controlled light emitting diode matrices. These are often used in displays, and have large, often unscreened areas carrying relatively high current transients, causing emitted noise. As a rule, liquid crystal displays are less noisy.
- Do not use a switch-mode power supply if a linear power supply can be cost effective. Switch mode power supplies are clock controlled, always noisy, and must in most cases have extra filtering components. This is an example where a clock controlled design may be replaced by a non clock controlled design.
- The use of optical fibre cables are always favourable with respect to EMC. They do not carry currents and will thus not emit noise. They will also be immune to overvoltages and to external electromagnetic fields.

Emission, design details

Many corrective circuit details will affect the EMC performance of an electronic design. Some of them may be very influential, and some only marginal. Some of them may be helpful over a broad band of noise frequencies, and some only at discrete frequencies. Some examples are:

- The capacitive circuit loads should be minimized. Noise from signal tracks on the printed circuit board is generated mainly because of the charging and discharging of the parasitic capacitive loads as the signal is pulsing.
- Signal tracks having high capacitive loads should be as short as possible to make the noise area as small as possible.
- Capacitive loading should normally not be higher than what the component data specify. Component delay data are given as a function of their capacitive load. If the load is higher, the delay will increase, and the delay data will no longer be valid. This may result in circuit malfunction.
- Noise generating currents in the board signal tracks may be reduced by inserting small series resistors at the signal driver outputs. Observe that these resistors will result in a slight increase of the signal delays.
- Series resistors may be useful on the most noisy bus signals; the read and write signals, and on the address bus’ least significant bit position A_0. This bus line has the highest frequency, because the address bus is incremented in the normal binary code. The signals mentioned are often noisy because they are routed to many components,

and thus have a large noise area and a high capacitive loading. Observe that series resistors may distort the signal wave form, especially if the parasitic capacitive loads are distributed along the signal track.

- All connections from drivers to receivers shall have correct voltage level matching. This is often ignored in modern designs where level incompatible bipolar and CMOS circuits are used. The result is poor self compatibility, which often may pass both simulation and test, whereafter circuit malfunction may appear after the product is sold.
- Correct impedance matching between driver and receiver will guarantee undistorted pulse shape, and thus correct functioning of the circuit. Impedance matching is often obtained using pull up and pull down resistors at the receiving end of the connection. This will normally increase the signal current, and may thus result in a higher noise level. A distorted signal can be acceptable if it has stabilized before being clocked into further circuitry. In such cases there is no need for costly, possibly noise generating, impedance matching.
- Avoid bus contention. Tristate bus logic, where the output of two or more circuits are connected together, may generate heavy noise if the two circuits are delivering signals to the common output at the same time. The situation may occur as a short-circuit glitch when control of the common output change from one driver to the other, and if one driver delivers a “0” and the other a “1”. The situation can easily be avoided utilizing proper design principles.
- Outputs from octal circuits – eight bit bus drivers – should not be used to control edge triggered components. The reason is the “ground bounce” phenomenon, which causes outputs to have transients that may trigger edge triggered components.
- Counters, such as the 74HC161, may cause malfunction if their carry-output is used to control edge-triggered components. This is misuse of the counting component because of the decoding transients on its carry-output.
- A circuit that controls an inductive load must be protected from the strong transient pulse that will be generated by the stored inductive energy in the load when its control current is shut off. The protection may consist of a shunt diode connected over the load. The transient pulse may generate a strong noise field. This can be limited by mounting the diode close to the load, to make the transient current loop as small as possible.
- Noisy transient pulses may be limited using components with TTL output levels.
- Design the overvoltage protection as inexpensive as possible. The testing, not the dimensioning, should dictate its complexity.
- SMD capacitors have low parasitic series inductivity ΔL. They are therefore effective as high frequency noise reducing decoupling capacitors.
- The use of discrete SMD transistors instead of transistor array – a multitude of transistors in an IC – will often result in a cheaper, smaller and less noisy coupling area.
- Avoid the use of IC sockets. They have inductive pins that may introduce a higher emission, higher cost, and blocking of the air cooling.
- Limit the use of filtering components. Use as few, inexpensive, and as small components as possible. It is normally not easy to determine the need for filtering components at the outset. One may therefore possibly use too many, and afterwards remove those not needed according to the test result.
- Filtering components are often needed close to the cable connectors. It is very important that they have short leads, preferably SMD, and that they are connected to a “clean” ground reference.
- The power supply voltage should be filtered on each card to avoid HF noise to contaminate the backplane.
- All cabling to the equipment, except the mains cable, should in most cases use shielded cables.

Placement of components

A sound placement of the components on a board is vital to obtain a low-noise printed circuit layout. But rules for low-noise component placement may conflict with placement rules to obtain self-compatibility, and with placement rules to obtain even cooling of the board. The following is a listing of EMC-effective component placement examples:

- For the bus noise area to be as small as possible, the component placement should be accommodated to a short and straight bus.
- The placement and orientation of the components should be accommodated to a short and straight signal track layout.
- Low output impedance components should have short output layout tracks. This is especially valid for buffer-circuits of the “Advanced” types that are capable of driving more than 100 mA into the receivers’ parasitic capacitances.
- The decoupling capacitors should be located close to the ground connection of the ICs. Each and every IC should have its own decoupling capacitor. Noise transients generated by switching of the components can be substantial, but is easily controlled by adequate decoupling.
- The analogue part of the circuitry should be separated from the digital part. Intermixing analogue and digital components is a common cause of functional noise in the analogue system, and of common-mode emissions.
- Interface circuits should be located close to its connectors. Noise from the interface part of the logic often results in radiated emissions from the cables, which can be hard to eliminate.
- Cable connector filtering capacitors should be located very close to the connector. The filtering performance will be substantially reduced if the distance between capacitor and its corresponding connector pin is increased only a few millimetres from its optimal location.
- The components for overvoltage protection should be located close to the line connector. This will increase the chance of correct functioning even under exposure to overvoltages.
- The components should be placed so that even cooling over the board area is obtained. The cooling problem resulting from the introduction of surface mounting technology and the use of higher frequencies, has led to the need of forced air cooling in many new designs.
- The location and orientation of the components should be adapted to the soldering process. This may be a problem when components are to be

mounted on both sides of the board; this mainly because wave soldering, which is often used on the secondary side of the board, may accept only resistors, capacitors, and transistors on that side.

Printed circuit layout design

The routing on a board consists of the signal tracks and of the power and ground system. It will be based on the location of the components, and its task is to ensure the boards' self-compatibility, so that it will function reliably. It shall also be designed to secure the EMC system specification. Its main goals should be:

- All nodes of the schematic diagram must be realized in the printed circuit.
- All signal tracks should have return tracks that follow the signal tracks as closely as possible. A signal track and its corresponding return path will form a noise radiating loop which has to be minimized. The return path may be a dedicated board track, or a less well defined current path in the ground plane.
- The signal tracks should have low cross-coupling. This demands that long adjacent signal tracks should be avoided. The introduction of the SMD technology with closer component connection pins (fine pitch), denser component packing on the board, finer tracks, and with higher system frequency, have increased the cross-coupling problems.
- Some signal tracks must be impedance-matched to its transmitting and receiving components. This can fairly easily be accomplished by an advanced CAD system, and normally results in wider or narrower tracks than those used for connections without impedance matching.
- The tracks must have sufficient insulation distances. This may be a problem when the board carries higher voltages than normal, such as the mains voltage.
- Multilayer boards should have as few vias as possible. Vias increase the board cost, the risk of scrapping the board, and they reduce the quality of the power and ground planes. They also introduce added inductivity to the tracks.
- Digital signal tracks should run only over digital power and digital ground, and analogue tracks only over analogue power and analogue ground. Not observing this rule will increase the risk of malfunction and of common-mode radiation.

The main goal for the PCB designer is to design a correct functioning board which also meets the EMC specification at first attempt. However, redesign is necessary more often due to changes in the board specification, than as a result of insufficient self-compatibility or because the EMC specification is not met.

It is interesting to observe, though, that the client usually accepts a layout that results in a board that does not function properly because of imperfect signal wave shapes, or the board does not meet the EMC specification, even if the layout designer should be to blame. The PCB designer is normally not made responsible for delays, and in addition he is even usually trusted to make a new try. When do we have a PCB designer who would guarantee his design, both with respect to a functioning board and to the EMC specification?

In his strive for success, the PCB designer may be tempted to use more board layers than necessary. The board, however, should only have as many layers as the component density and the signal and power track density demands. Only when the EMC restrictions dictate more layers, this should be done. This is because the board price increases as the number of layers goes up. Four layers have twice the cost of two layers, and for each successive pair of layers the cost goes up by some 40 %.

It is not possible before start of a design to figure out how many layers a certain board will need for the complete system to pass the EMC test. If the system passes the test, it is seldom checked if one or more boards could have had fewer layers.

In most cases at least four layers will be needed. Practice has shown, though, that an EMC layout expert who co-operates with the circuit designer, may in some cases succeed with only two layers. For low volume production, the board price has less significance justifying a higher layer count (e.g. six layers). For medium volume one may try to succeed with four layers, and for large volume, or if the EMC specifications are minor, one could try to succeed with only two layers. With tight time schedules it may sometimes be advantageous to first make a multilayer board, as this will have a higher chance of success getting the product as fast as possible on the market. Later, when the product is selling, one may try to cut costs using fewer layers.

Automatic routing of printed circuit layout tracks, "autorouting", has for some years been advantageous because of fast results and low cost. Autorouting has, however, traditionally not taken enough care of EMC considerations, and practice has shown that most designs based on such programs have resulted in systems that do not meet the relevant EMC specifications.

Autorouting programs that do take EMC into account have to some degree been available for some time. They are expensive, and to achieve acceptable results with respect to EMC, they need the same component data, and other relevant EMC data as is used by "manual" PCB design. The result may be acceptable, but not always cost effective.

It is not advisable to use autorouting programs that do not take EMC considerations properly into account. A typical result is more via-holes, longer tracks, and less control with critical signals with respect to EMC.

The result may be:

- The high via-count will degrade the uniformity of the power and ground planes, generating ground impedances and force the return currents to take uncontrolled return paths.
- Higher printed circuit board production cost
- Lower printed circuit board production yield
- Poor self compatibility of the logic circuit
- Uncontrolled and higher radiation
- Low overvoltage immunity
- Risk of more PCB redesigns
- Risk of costly repeated EMC measurements
- Risk of introducing delays in the development phase
- Risk of higher total development cost.

The above listing may seem somewhat exaggerated and indeed it is. Manual designs are sometimes even worse, but the listing illustrates the involved risks when using autorouting.

A mix between manual routing and autorouting, and between manual placement and "autoplacement" may give good results. Programs for autoplacement and autorouting are continually being refined.

Some general layout rules

The printed circuit tracks shall form smallest possible current loop areas, see Figure 2.

When discussing current loop areas, one tends to forget that the *total noise loop area* covers the complete area, from the power input through the signal track and back to the power output, see Figure 3.

The noise areas in boards having "complete" power and ground planes (they are never quite complete, due to lots of through holes), will always be substantially smaller than in a comparable two-layer board. This is because the return currents will obey Lenz' law and run as close to the signal currents as possible, thus reducing the area and the noise as much as possible (for once, the nature of electricity is on our side), see Figure 4.

The noise level from a four-layer board having power and ground planes and two signal layers, is some 20 dB lower than from a comparable two-layer board.

All tracks on a printed circuit board, and all return paths will contribute to the total noise spectrum and noise levels. When designing a low-emitting board it is not possible to look at all noise generating paths. One has to give priority to those considered to contribute most to the noise. A PCB designer may not be skilled enough to know which tracks that will contribute most to the noise, and when making up his priority list he has to co-operate with the logic circuit designer. A priority list may look as follows:

- Connections having high pulse frequency. The track carrying the highest system frequency is usually the most noisy one because the noise emissions are proportional to the frequency squared. The clock frequency is usually the signal that has the highest frequency. It is of utmost importance that the clock signal has a ground return path that runs as close as possible to the clock signal track. On multilayer boards this will normally be ensured by the ground plane. On two-layer boards, the clock signal return path should be routed separately on one or both sides along the entire clock signal track.
- Connections having high capacitive load. All signal receivers have a parasitic capacitance to ground. Signal voltage pulsing will generate noise current transients that charges and discharges these capacitances. Signal lines having many destinations will have multiple capacitive loads, and corresponding higher noise current transients. Typical connections of this type will be clock signals, and read and write signals.
- Connections having low impedance drivers. Most digital circuits have an output impedance at 5V V_{CC} that make them capable of driving signal noise transients of up to some 25 mA. Fast buffer circuits from the "Advanced" series, or from the BIC-MOS series have output impedances well below 100 ohms at 5V V_{CC}, and are thus capable of driving signal noise transients of up to 150 mA or even higher. All connections from these circuits should be short and straight, and have a nearby ground return path.
- All bus connections contribute significantly to the total noise picture. The address bus normally represents less problems than the data bus, and is also the easiest to cure. The address bus has usually only one source, and has also the lowest parasitic capacitive load, whereas the data bus normally has many alternating sources and a higher parasitic capacitive load. The data bus has therefore a stronger and more untidy noise spectrum. All buses should be short and straight, and not located near the board edges. The ground return should be good, preferably a ground plane with as few via-holes as possible.

Decoupling of the integrated circuits with capacitors was originally used to ensure the correct functioning of the logic circuits, and the capacitors were spread over the board according to more or less well defined rules. The change from TTL to HCMOS requires a more stringent practising of those rules, mainly because malfunctioning of the logic circuits occur more and more frequently. The reason for this is the higher voltage pulse levels, and the faster edge rates generated from HCMOS compared to TTL.

Decoupling rules should now be one capacitor for every integrated circuit, and

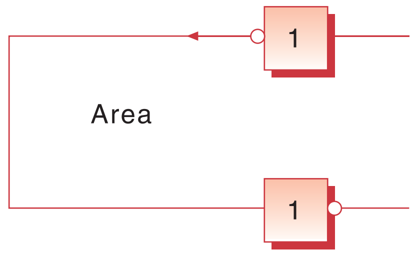


Figure 2 Signal loop noise area

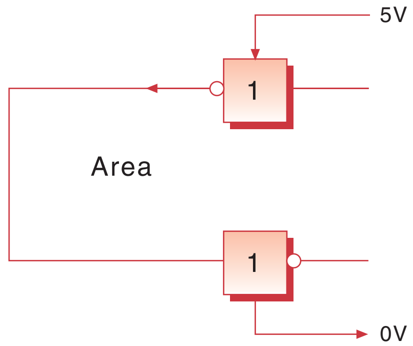


Figure 3 Two-layer board. Large noise area

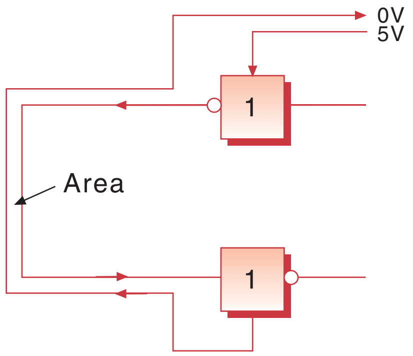


Figure 4 Four-layer board. Small noise area

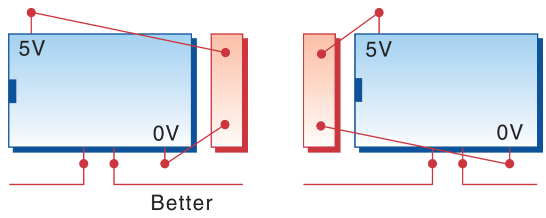


Figure 5 Location of decoupling capacitor

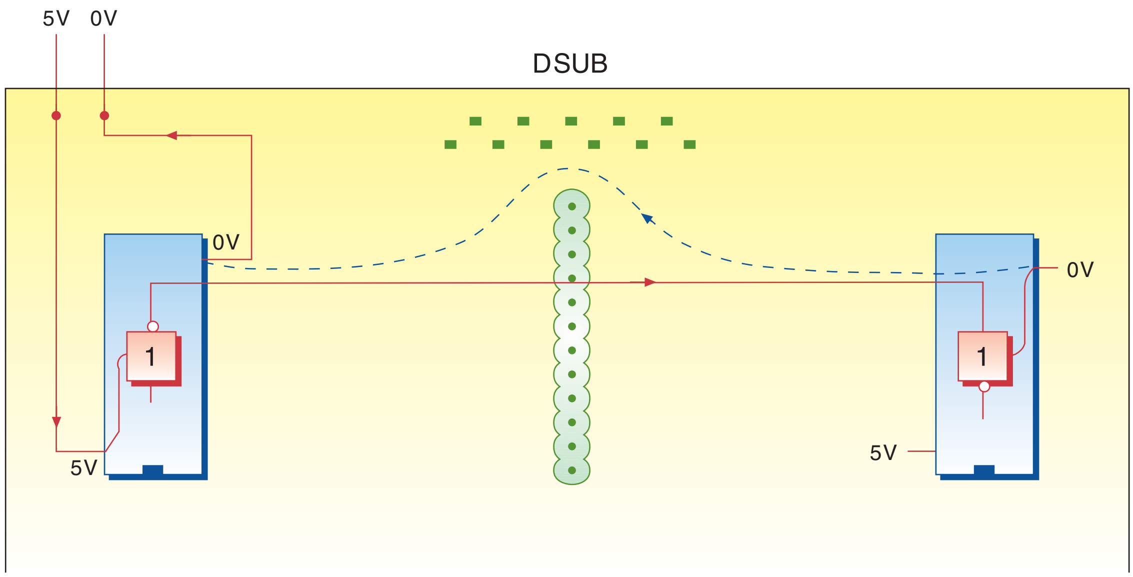


Figure 6 Ground return current contaminating connector area

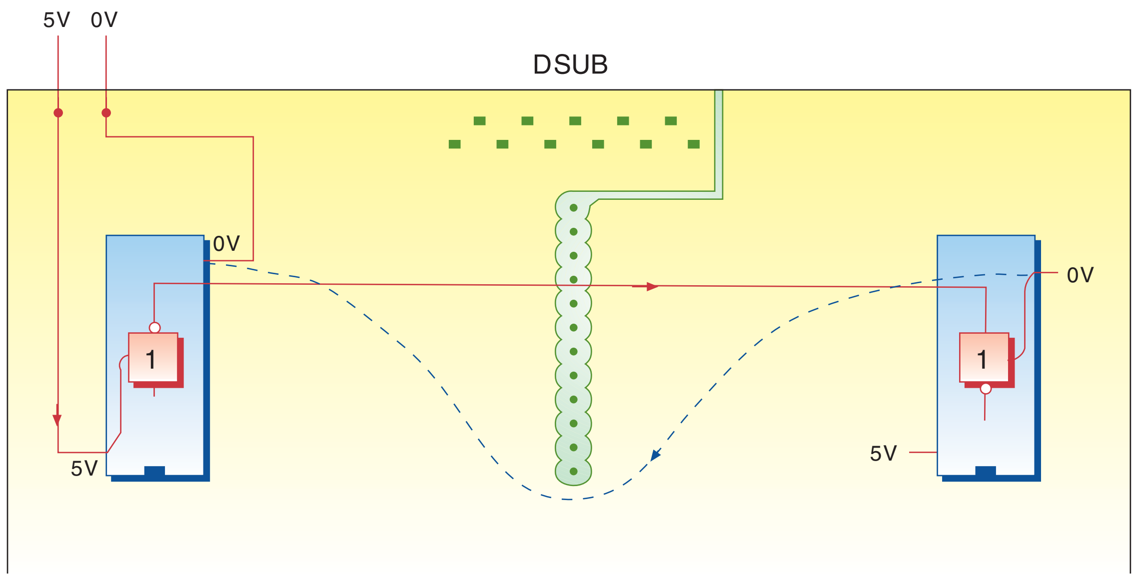


Figure 7 Diverting ground return current

for large ICs one should use several capacitors. The reason for this is the EMC problem, and that insufficient decoupling will readily be observed on the emissions plot.

The type of decoupling capacitor and its location with respect to the IC is also very important. Capacitors having long leads, or being located a couple of centimetres from the IC are of limited or no use. It is also important to locate the capacitor close to the ground connection of the IC, not close to the power connection, see Figure 5.

The decoupling transient current running from the capacitor to the 0 V termination of the IC will have a short path if the capacitor is located close to the 0 V termination of the IC. This current will thus contaminate the 0 V plane or the 0 V distribution layout as little as possible. This will reduce the mixing of the transient ground current with signal return currents running in the close vicinity. This is an example of how to keep the 0 V distribution or the 0 V plane as clean as possible, which is believed to be very important in the strive to reduce the generation of common mode emissions.

The use of electrolytic decoupling capacitors (LF-decoupling) was originally introduced to stabilize the V_{CC} of the board, and to thus secure the correct functioning of the board.

The high density of integrated circuits on modern printed circuit boards, and the use of one decoupling capacitor for every IC, often results in a total HF decoupling per board of several microfarads. This value is often sufficient LF decoupling, and because an electrolytic capacitor has no EMI-reducing function, it may be omitted. It may have a function, though,

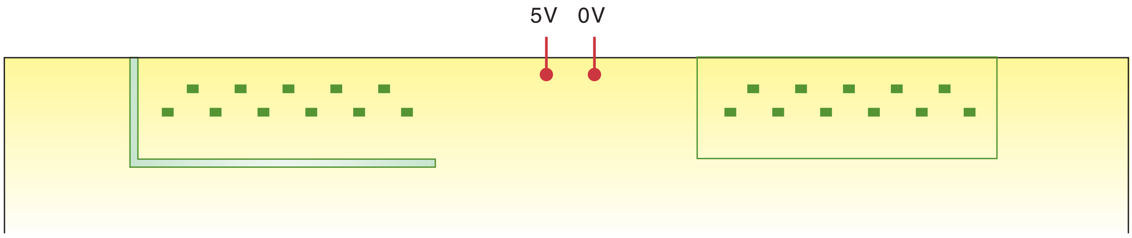


Figure 8 Hindering ground currents in DSUB connector area

as a stabilizer of the V_{CC} if an overvoltage should occur. Practice has shown that a microprocessor based system may fail due to electrostatic discharge if the electrolytic LF capacitor is omitted.

It is on two-layer boards, as well as on multilayer boards, important to know where the signal return currents will flow. The return current paths will be difficult to localize in ground planes. They will spread out to some more or less predictable degree, although they will obey Lenz' law and run as close as possible to the signal tracks.

If the natural return path in the ground plane is broken by a slit, as formed by a row of consecutive through holes, the return current may be forced to run around the obstruction and thus produce

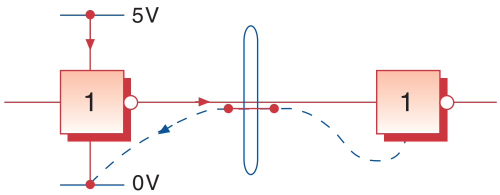


Figure 9 Ground return bridge over slit in ground plane

a noise emitting area, or the current may be forced into a noise sensitive area where it may generate circuit malfunction, or where it may generate common mode emissions. A typical example is a situation where a return current is forced to run in the ground plane close to a DSUB connector, where it will generate common mode emissions that will be conducted by the cable connected to the DSUB connector, see Figure 6.

Intentional use of slits in a ground plane may be used to divert ground return currents off from a noise sensitive area, see Figure 7.

Unwanted ground return currents in the ground plane close to a DSUB connector can also be avoided by a slit in the ground plane around the connector pins, or simply by removing the ground plane in that area, see Figure 8.

Slits in the ground plane made up by a row of consecutive circular insulation areas, as in Figures 6 and 7, can be avoided if the circular insulation areas are made small enough to let the current pass through. Another method is to bridge the slit by a ground current strap close to the signal current track, see Figure 9.

Electromagnetic compatibility between power networks and telecommunication networks

BY KNUT J RAMLETH

1 Introduction

Transmission of electric power as well as telecommunication signals on galvanic cables are necessary functions in a modern society. Transmission may take place on the same cable pair or on separate cable installations. This article will discuss the problems of compatibility between separate cable installations. We will further elucidate some problems on noise and dangerous voltages which the power network may cause in the termination points of the telecommunication cables.

Power networks, electrified railways and telecommunication networks are competitors in the selection of cable routes. Legally, the three parties are equal in that the Principle of Priority ("First come, first serve") forms the basis for financial settlements. A case of the telecommunication network acting as a noise source to power networks or electrified railway operation happens rarely or not at all, and in questions of compatibility between the three parties the telecommunication side is therefore always the "victim".

2 General

Any problem of compatibility comprises the three main elements

- source (emitter)
- receiver (victim)
- coupling path.

Provided both source and receiver are working and operating as intended, it seems likely that the coupling between source and receiver should be reduced in case of compatibility problems. Coupling and physical distance are usually two sides of the same issue, so an increase in distance will reduce the coupling. On the other hand, there are the general environmental requirements to keep the number of cable routes and encroachments on nature at a minimum. Economic considerations also favour common cable trenches, common poles and common routes. Besides, the telecommunication network is necessary for operating both the power network and the railway, and finally, the telecommunication operator and the power supplier have largely the same customers and users. In other words, a close neighbourhood between cable networks and between apparatus and equipment cannot be avoided in practical life.

The lack of compatibility between power supply and telecommunications affects both disturbance in the telecommunication network and safety of personnel and equipment connected to telecom services. Safety is regulated by Norwegian law through "Forskrifter for elektriske forsyningsanlegg" (Regulations for electric power supply plants) and "Forskrifter for elektriske bygningsinstallasjoner m.m." (Regulations for electric installations in buildings, etc.). Problems connected to noise have not earlier been as strictly regulated, but have now been highlighted through the introduction of the European EMC Directive.

3 Structure and characteristics of the power network

3.1 Alternating current networks

A distribution network for electric power may in principle be constructed as outlined in Figure 1.

A three phase alternating current can be transferred on three separate conductors as shown in Figure 2.

In a balanced three phase system the vector sum of the three phase currents is zero, and the current in the three phases seen together will therefore have no magnetic influence outside the near zone.

The example shown in Figure 3 represents an ideal case. In practice, there will always be degrees of unbalance in the vector diagram because of varying loads (current outlets) on the three phases; the worst case being when there is an earth-fault on one of the phases. The current in an unbalanced three phase network can mathematically be decomposed into

- a symmetric three phase system with normal phase sequence, the plus-sequence system
- a symmetric three phase system with opposite phase sequence, the minus-sequence system
- a third three phase system where all three currents are in phase, the zero-sequence system.

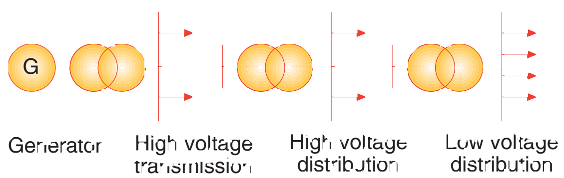


Figure 1 Transmission of electric power

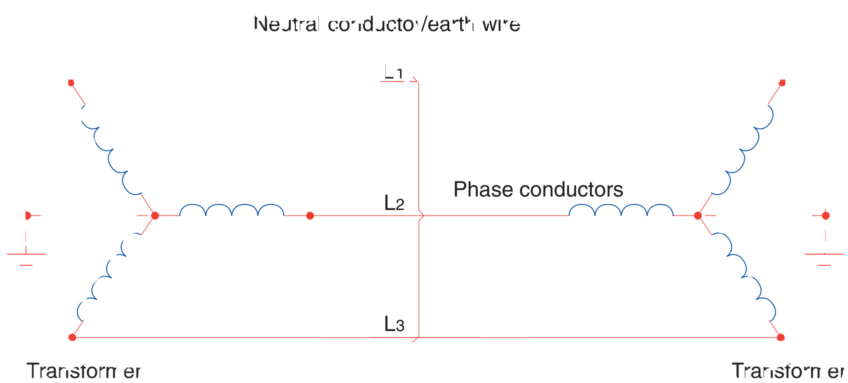


Figure 2 Three phase system

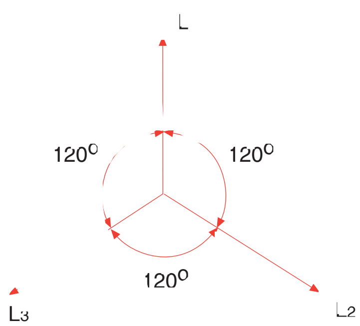


Figure 3 Phase currents in a balanced three phase system

The three decomposed systems are illustrated in Figure 4.

The individual plus- and minus-systems are balanced three phase systems which do not cause magnetic disturbance to the surroundings. The zero-system, however, represents all the unbalance because the zero-sequence currents must return to the source (generator/transformer) through the earth and/or a possible fourth conductor (top line). In the case of an earth-fault on one of the phase conductors the system earthing of the power network will therefore greatly influence the effect on the telecom network in the area. The part of the fault current which returns to the source through the top line is in opposite phase to the inducing fault current and thereby has a reducing effect on the induced voltage in the telecom network. The resistance in the top line will be decisive for its “reduction factor”. A well conducting top line has a good reduction effect; aluminium conductors in the top line will in other words be preferable to steel conductors. The reason why steel lines are still used is the relatively poor mechanical tensile properties of aluminium.

In systems with the neutral point connected to earth, ref. Figure 5, the fault current (3 x Io) may be quite large, but as previously mentioned, the inductive effect on nearby telecom cable may be attenuated when the high voltage towers are connected with continuous and well conducting top lines.

A large earth current induces high voltages in nearby telecom cables. However, the large fault current makes it technically possible to achieve a quick and safe automatic disconnection of the fault, and the effect on the nearby telecom network will therefore be brief (less than 1 sec.).

In a transmission system with isolated neutral the earth-fault current will normally be small. The whole of the fault current must return to “the healthy phases” through the conductor capacities, ref. Figure 6. This will normally limit the fault current to a few tens of Amperes.

A small fault current in the power network will in turn induce small 50 Hz voltages in nearby telecom networks. However, the combination of a small fault current and existing relay technique in the power network will render it impossible to achieve an automatic and quick disconnection of the fault. The duration of this type of earth fault may therefore be hours and days.

3.2 Direct current networks

The capacitive loss of power in long cables in alternating current network is considerable. As an alternative transmission system it would therefore be profitable to use direct current by introducing rectifiers/inverters at the cable termination points. This applies e.g. to the exchange of power between Norway and Denmark running through a 120 km long sea cable between Kristiansand and

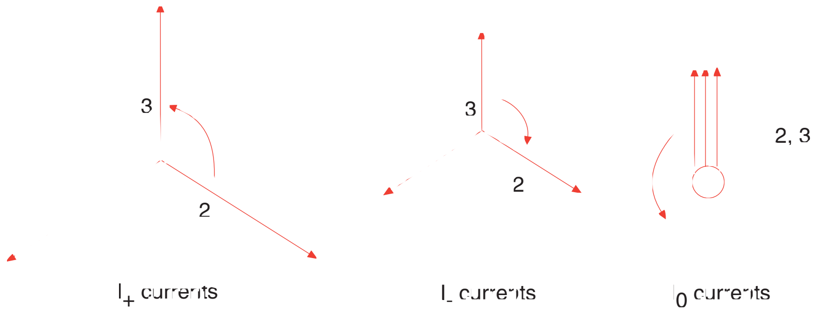


Figure 4 Decomposition of currents in an unbalanced three phase system

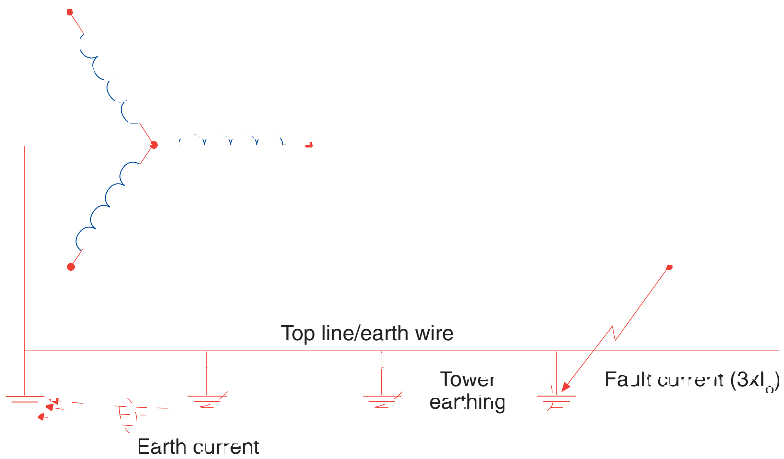


Figure 5 Fault currents generated from an earth-fault in a three phase network with directly earthed neutral

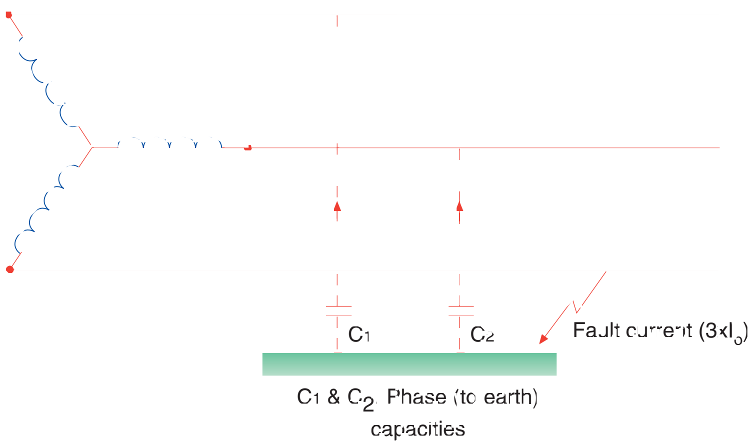


Figure 6 Fault currents generated from an earth-fault in a three phase network with isolated neutral

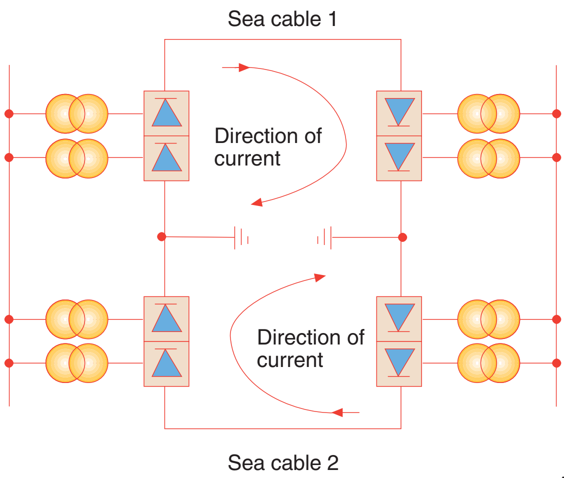


Figure 7 Co-operation of 2 DC cables

Jutland. Sea water may in principle act as return conductor, but as the loss of energy in sea water is considerable – in this case in the megawatt area – a coupling as shown in Figure 7 is used.

Polarities are chosen so that the return currents for the two DC cables are inverted in the sea area between the sea electrodes. If the cables have equal current load the current in the sea is zero, and the transmission losses are thereby minimised.

The rectifier/inverter installations in the transmission system are a source of major noise problems to the telecom network. The thyristor-based rectifier/inverter installations generate considerable amounts of harmonic components in the audible frequency band, and the introduction of effective smoothing filters after rectifying is limited by the fairly large operational currents. For environmental reasons, the coupling installation with transformers and thyristors on the Norwegian side was located some distance inland; however, at the cost of the whole aerial route with pole cables and electrode cables between the plant and the landing spot near Kristiansand now being a powerful electromagnetic polluter. Noise conditions are particularly bad during monopole operation with all the return current going via the sea electrode.

Another negative side effect of the monopole operation is that the return current does not only take the direct route between sea water and electrode: Parts of the return current act as earth current at a distance of several kilometres from the electrode and may cause electrolytic corrosion of metal sheathed telecom cables in the ground.

4 Noise sources in the power network

Rotating AC generators and power transmission systems with transformers and cables do not themselves produce noise in nearby telecom networks. It is power consumption and coupling techniques which "pollute the sinus curve" that cause the power network to act as a coupling element between noise source and the telecom network. As mentioned above, noise sources are found in large rectifier plants for power transmission, but also in similar plants for heavy industry. Static frequency converter plants also produce noise in connected networks.

Other examples of pollution sources in the power network are power regulators based on semi-conductors, coupling mechanisms like thermostats, relays, etc., power switches for switching on and off large power flows, as well as occasional faults involving poor or varying degree of contact and arcing. Corona discharge on conductors and coupling equipment may sometimes produce interference in the radio frequency band.

5 The power network as a source of dangerous voltage

A balanced three phase system will only have effect on telecom networks in the close proximity. As long as the power transmission takes place with full balance between the currents in the three phases, only capacitive and galvanic couplings will convey dangerous voltage from the power network to the telecom network. In cases where the power network and the telecom network are very close, the difference in distance between the individual phases in the power network and nearby telecom network may be so great that we also get a magnetically transferred voltage in the telecom network.

In the case of unbalance between the three phase currents in the power network, zero-sequence currents induced voltage in the nearby telecom network will occur. Particularly high inducing currents are seen in connection with earth faults on one of the phase conductors. The inducing fault current may then amount to tens of kA and cause dangerous voltages in telecom networks several kilometres away from the power network. On the spot where the fault current goes to earth, ground potentials of several thousand Volts may occur posing a threat to telecom networks in the area.

Agreements between the telecom and power supply parties in Norway are founded on the limits stated in the recommendations in the CCITT's "Directives concerning the protection of telecommunication lines against harmful effects from electric power and electrified railway lines." These directives have been prepared in a co-operation between the International Telecommunication Union represented by the CCITT, the International Union of Railways and the International Conference for Co-operation in High Voltage Systems. Without further national discussions, the stated limits have therefore been acknowledged by NSB (Norwegian State Railway), the power suppliers and Norwegian Telecom as a guide for determining when protective measures have to be implemented.

The following limits are used:

- 60 Veff permanent longitudinal voltage in a telecom cable. The limit may be increased to 150 Veff in certain cases, provided parts to be handled are clearly marked by danger signs.
- 430 Veff longitudinal voltage lasting from 0.5 to 1 sec.
- 650 Veff longitudinal voltage lasting less than 0.5 sec.
- 1000 V peak during a contact to earth of one wire of a nearby DC power or electrified railway line.
- For capacitive coupling the permissible limit has been set at 10 mA measured low-ohmic between a telecom conductor and earth, or between a telecom conductor and metal parts that may be touched.

If the estimated or measured voltages and currents in the telecom network exceed the above limits, protective measures will be implemented in the telecom network or the power network, or in both networks. The costs of the protective measures are shared according to the Principle of Priority.

6 Coupling mechanisms between power and telecom networks

The following coupling mechanisms apply between the power network and the telecom network:

- Galvanic coupling
- Capacitive (electric) coupling
- Inductive (magnetic) coupling.

Galvanic coupling between two circuits requires a common impedance, *Zk,* in both circuits, ref. Figure 8.

The current in circuit 1 produces a voltage drop over the common impedance *Zk*

$$Vk = i \times Zk$$

Zk is the coupling impedance and *Vk* is the operating voltage in circuit 2. *Zk* may e.g. be represented by the resistance in a common earth electrode (ref. section 8), the impedance in an earth connection cable, or similar.

Capacitive coupling between two cable systems over a common earth plan is outlined in Figure 9.

Transferred voltage from circuit 1 to circuit 2 may be calculated with the formula

$$V2 = V1 \frac{Z2}{Zk + Z2}$$

Normally, $Zk \gg Z2$, and in the case of sinusoidal current, the numeric value of *V2* may be calculated with the formula

$$V2 \approx V1\ 2\ \pi f\ Ck\ Z2$$

with good approximation.

From the formula for *V2* we have that the transmitted voltage is proportional to frequency *(f)*, coupling capacity *(Ck)* and impedance (*Z2*) between circuit 2 and earth. The numeric value of *Ck* is approximate to a logarithmic function as regards the separation between the two networks. As can be seen from Figure 10, quite a high reduction of attenuation may be achieved by increasing the separation when the networks are very close. At medium and large distances, on the other hand, variations in separation have less importance.

The *inductive coupling* between two aerial cables at a certain height above ground may be calculated by using quite complex formulas. The complete formulas given in the above mentioned directives from the CCITT are poorly suited to manual processing. Mutual induction be-

Table 1 Coupling impedance per km in V/kA for 2,500 ohm metre and 50 Hz frequency

Distance	0 m (6 m)	10 m	20 m	30 m	40 m	50 m	60 m	70 m	80 m	90 m
0 m	(421)	389	346	320	303	289	278	268	260	253
100 m	246	240	235	230	226	221	217	214	210	207
200 m	204	201	198	195	193	190	188	186	183	181
300 m	179	177	175	173	172	170	168	167	165	163
400 m	162	160	159	158	156	155	154	152	151	150
500 m	149	147	146	145	144	143	142	141	140	139
600 m	138	137	136	135	134	133	132	131	130	130
700 m	129	128	127	126	126	125	124	123	122	122
800 m	121	120	120	119	118	117	117	116	115	115
900 m	114	114	113	112	112	111	110	110	109	109

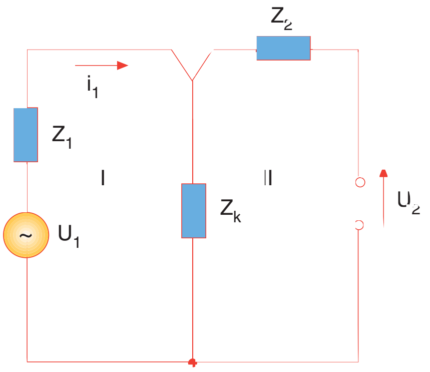


Figure 8 Galvanic coupling

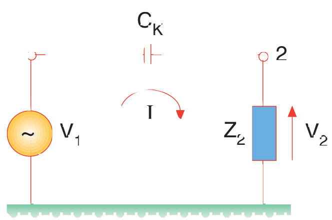


Figure 9 Capacitive coupling

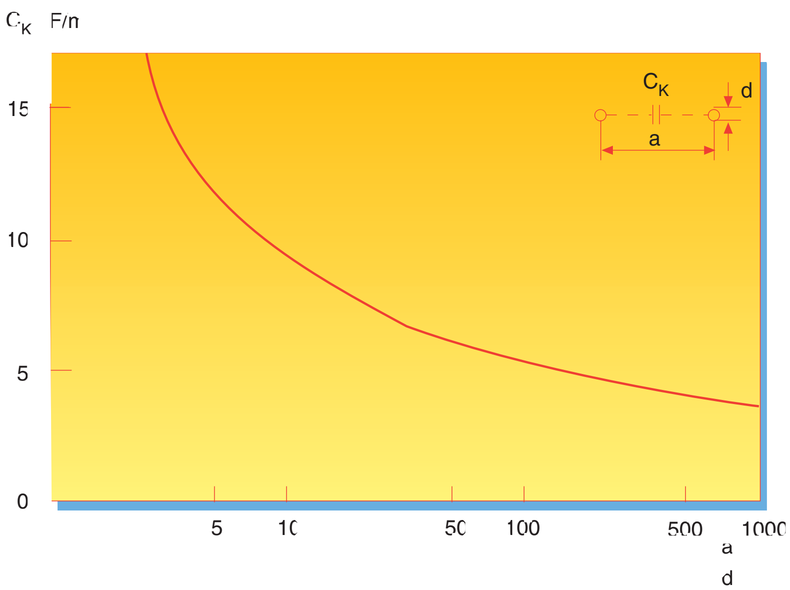


Figure 10 Capacitive coupling between two parallel conductors

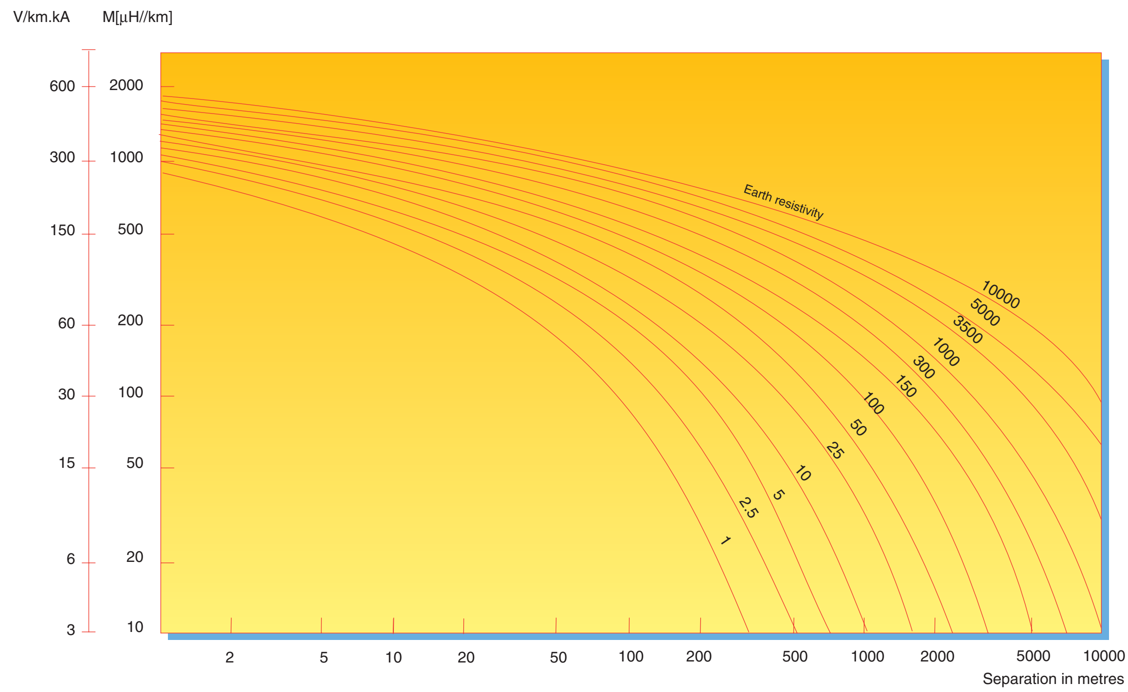


Figure 11 Mutual induction as a function of separation and ground resistivity for the 50 Hz frequency

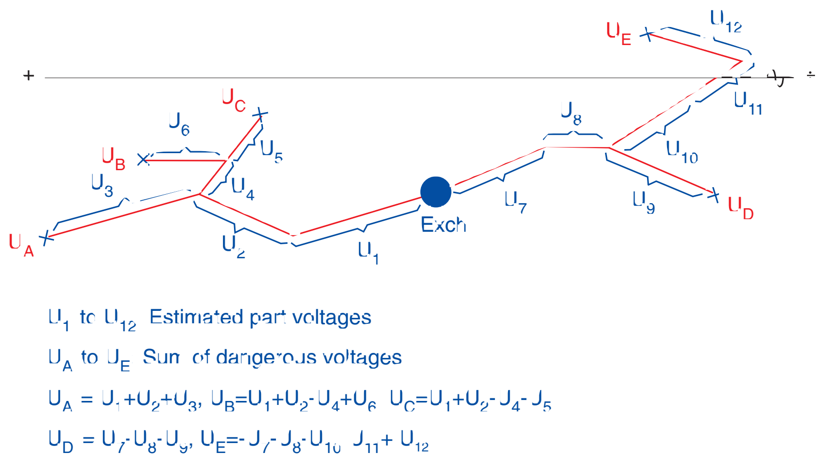


Figure 12 Inductive coupling. Summation of part voltages

tween a power cable and a telecom cable with earth being the return conductor for the inducing fault current, may be machine estimated by a complex function of frequency, earth resistivity, screening factor, separation, height above ground and length of exposure. An example of estimated mutual coupling *M* as a function of distance for the frequency 50 Hz and with earth resistivity as parameter, is shown in Figure 11.

For practical purposes we use pre-calculated tables of coupling impedance in Norwegian Telecom. The tables are based on typically Norwegian earth resistivities when calculating induced dangerous voltages in the telecom network. Table 1 is an example of such a calculation table.

The numeric values in the table presuppose parallelism between telecable and inducing line. In real life, the cable routes will be more or less at angles, and the stretches of proximity are then divided into shorter lengths where the two networks are approximately parallel. Each length is calculated separately, and the induced voltages somewhere in the network will appear as the sum of the induced voltages on the part lengths between that spot and the reference point.

For the purpose of induction calculation the two ends of the power cable are defined as plus and minus, respectively. The calculated part voltages add up to a positive or negative sum, depending on the direction of the part lengths in relation to the direction of the power cable. An example of summation of part voltages in an exchange area is shown in Figure 12. The exchange has been chosen as reference point.

Galvanic and capacitive couplings are typical proximity phenomena, while inductive coupling occurs from quite close, up to a distance of some kilometres.

7 Protective measures to reduce disturbing or dangerous effects in the telecom network

If calculations or measurements show that the power network may cause electric effects exceeding the stated limits in nearby telecom networks, protective measures will normally be implemented in the telecom network. Theoretically, protective measures might also be implemented on the power current side, but if the power network is already established, the most economical solution would normally be achieved by implementing protective measures on the telecom side.

7.1 Protective measures against noise

We distinguish between two principally different generations of noise on the power side. Rectifiers, power regulators, etc., represent unlinear loads which distort the sinus curve in the power network.

In addition to the 50 Hz component, voltage components will occur with a frequency as a multiple of 50 Hz, and amplitudes decreasing with increasing frequency. In such cases, nearby telecom networks will primarily be exposed to noise in analogue connections in the voice-frequency band. Measuring this kind of noise may be done with a psophometer, a broadband volt meter with a weighting filter in the input circuit. The filter has characteristics simulating the sensitivity of the human ear.

In connection with physical couplings (on/off switch, load adjustment, etc.) in an operating power network, various degrees of arcing will occur in manoeuvred switches and reversers. The arcing will in turn generate impulse noise in the form of singular bursts or transient bursts. Impulse noise mainly disturbs digital telecom connections.

Noise voltages/currents in the power network are transferred to the telecom network as longitudinal voltages (common mode). A reduction in the interference can be achieved by reducing the coupling between the two networks and/or reduce the effect of the transferred noise voltages. The coupling is reduced by using telecom cables with earthed screen and, if possible, increasing the physical distance between the networks. The use of coupling reduction in existing installations is quite limited, and if the noise source itself operates within permissible regulations and limits, the effect of the transmitted noise voltages should primarily be reduced in the telecom network.

In the efforts to reduce the effect of the transferred noise voltages in the telecom network, the balancing properties of the installations in the network are essential. Introducing isolating transformers between a telecom cable and its termination reduces the effect of unbalance in connected equipment; the same applies if longitudinal chokes are introduced in the cable pairs.

Cable- or insulation faults may cause a weakening of the balancing properties of the telecom cable and thereby increase the network's sensitivity to noise influence from the outside. The same applies if the cable screen is poorly earthed. Both examples represent cases which should be seen to before costly noise reduction measures are implemented in nearby power network or in a "noise source" which may already meet the stated requirements.

Low frequency longitudinal voltages are reduced by introducing a neutralising transformer in the telecom cable. The principle of such a transformer is that the current from a pilot wire is induced from a primary winding to secondary windings in series with each cable pair. The induced voltage is in opposite phase to the influence from the outside. A certain degree of noise cancellation is thereby achieved, but the solution requires vacant cable pairs or an insulated cable screen as pilot wire, and that low-ohmic termination earthing of the pilot wire can be achieved. Earthing of vacant pairs in a cable will often increase the unbalance of the other pairs in the cable, and this, together with difficult earthing conditions, renders this method of little interest in Norway.

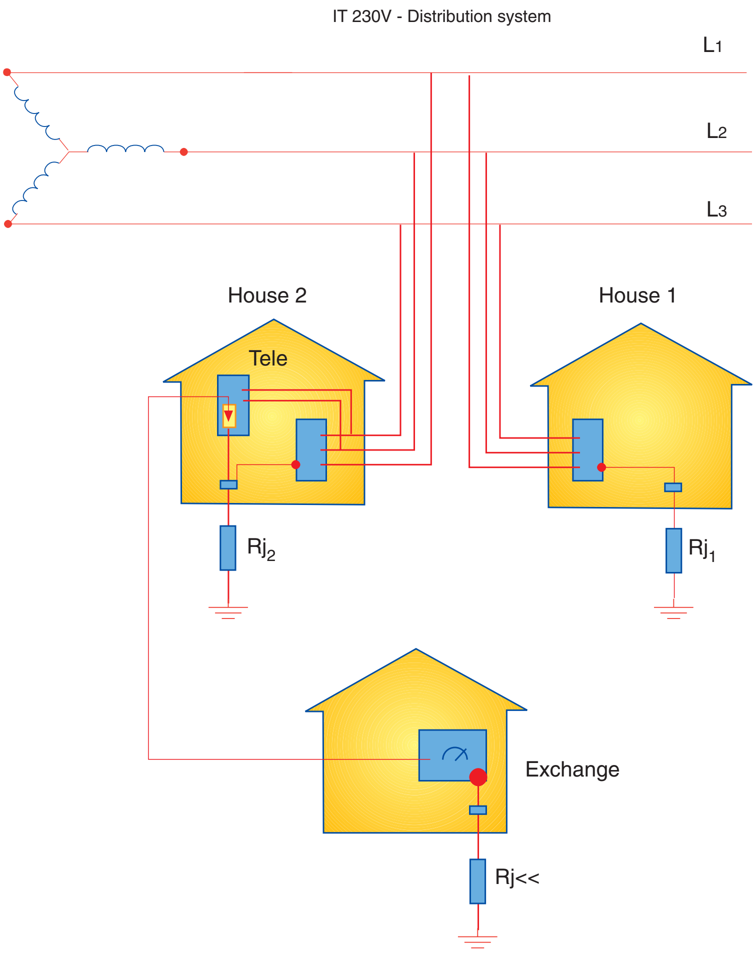


Figure 13 230 V IT distribution system

7.2 Protective measures against overvoltages

The most frequently used method of keeping permanent longitudinal voltage in the telecom network below the stated maximum limit is the use of distance attenuation. When using joint trenches and tunnels for power and telecom cables it is important to keep in mind that the separation between telecom network and power network should be kept large enough for a normal load fluctuation in the power network not to cause unacceptable voltages in the telecom network. The corresponding rule applies to aerial networks where an unprotected telecom cable may be exposed to both electric and magnetic coupling from a nearby power network. If unacceptable, lasting overvoltages occur in an established network, the telecom cable may be divided into shorter lengths with 1:1 isolating transformers. A major disadvantage of the method of division is that it causes blockages for the transmission of direct current in the telecom cable.

During earth-faults in the power network brief overvoltages will occur in the nearby telecom network. In real life, such earth-faults are impossible to avoid, and measures have to be taken in the telecom network to reduce the overvoltages to safe levels. The most frequently used protective methods in telecom networks are

- Metal sheathed, steel reinforced cable (cable with reduction factor)

- Surge arresters coupled between cable pair and earth
- Division of exposed telecom cable by introducing isolating transformers with a high dielectric strength
- Neutralising transformers.

The protective methods may be applied separately or in combination, depending on the level of overvoltages. The different protective methods have their advantages and disadvantages – discussing them is outside the scope of this article.

8 Low voltage network without an earth reference

Norway is just about the only country to use the internationally regulated IT system for 230 Volts power distribution within large areas. While all other countries use distribution systems with defined earth references (the TN or TT system), large parts of Norway has a low voltage network without any permanent earth reference. This leads to the peculiar Norwegian compatibility problems between telecom network and low voltage power supply with earth-faults. The problem has been taken care of for new installations in that the 1991 issue of "Forskrifter for elektriske bygningsinstallasjoner" (Regulations for electric installations in buildings) contains the requirement that new electrical installation shall have an earth-fault protection or insulation protection. However, the regulations do not apply to previously erected installations, and we should therefore expect to live with our particular earth-fault problems in old IT installations until well into the next century. The difference regarding earth-faults between IT and TN or TT systems is that in an IT system without earth-fault protection "standing earth faults" may occur, raising the local ground potential to more than 200 Volts. This is not possible in TN or TT systems.

Compatibility problems between the telecom network and an IT power network without automatic earth-fault protection may be illustrated by the example in Figure 13.

In the most widely used IT system in Norway there are three live phase conductors (*L*1, *L*2 and *L*3) with 230 Volts between the phases. In order to prevent equipment with insulation faults being dangerous to touch, *local* earth electrodes are established. The earth resistance of the individual installation is denoted *Rjx* in the figure.

Should an insulation fault (line to earth) occur on phase *L*3 in house 1, a capacitive current of a few hundred mA will flow through *Rj*1, via earth and the cable capacities in the low voltage distribution network, back to phases *L*1 and *L*2. The magnitude of this current depends on whether aerial or ground cables are used, but measurements carried out show that as a rule of thumb, the earth current, measured in mA, is estimated at a magnitude of the same number as the transformer size in kVA, i.e. in a distribution system with a 500 kVA transformer, the first earth-fault will produce a current of some 500 mA. If *Rj*1 is less than 100 ohms the ground potential of house 1 in relation to distant earth is below the danger limit of 50 Volts. Irrespective of the value of *Rj*1, the *local* danger of handling parts will be taken care of, provided there are no electrical breaks in the earthing installation in the house.

With a sustaining earth-fault on phase *L*3 in house 1, an additional situation may arise of an insulation fault on phase *L*1 or *L*2 in house 2, whereby there is a total system voltage of 230 Volts between the internal earthing installations in house 1 and house 2. Capacitive earth currents will now flow in both installations and in addition, we will have an earth current between two phases, only limited by *Rj*1 in series with *Rj*2. Ideally, the earth current ought to cause a fuse break, but because of the difficult earth conditions we have in Norway, the added sum of *Rj*1 and *Rj*2 is very rarely small enough for any current fuse to operate. On the contrary, the usual case in large parts of the country is that we will have an earth current of only a few Amperes between house 1 and house 2. When a voltage of 230 Volts is distributed on the two resistances *Rj*1 and *Rj*2, it is physically impossible to meet the requirement of the earth potential of the two individual installations to be less than 50 Volts, as stated in the regulations. Provided that the two earthing installations are otherwise in accordance with the regulations, there is, however, still no voltage which is dangerous to touch *locally.*

We will now assume a situation as described above with two earth faults without automatic switch-off, and that *Rj*1 is very small and *Rj*2 correspondingly large. The ground potential of house 1 will then be a few Volts, while the earthing installation in house 2 increases towards 230 Volts. In house 2 is shown a telecom installation with a gas discharge tube installed between the telecom cable and local earth. If this arrester has a spark-over voltage smaller than the ground potential for house 2, the arrester will operate and open for a 50 Hz current into the telecom network. The extent of damages that such a current may have on the telecom network may be indicated by the fact that a fire at Frogner exchange in Oslo in the mid-1980s happened as a result of a 50 Hz current from a telephone subscriber, and that the material damage in the exchange amounted to some NOK 100 mill. One should also bear in mind that a gas discharge tube with a continuous flow of current will be so hot that the arrester itself represents a potential fire hazard. Preventing a recurrence of the phenomenon which led to the Frogner fire needs requirements to make sure that gas discharge tubes at the subscriber end coupled between telecom cable and local earth, shall have a nominal DC spark-over voltage of at least 420 Volts.

9 Conclusion

A modern society needs both telecom services and electric power in order to function. In the foreseeable future, both telecommunications and power supplies will make use of galvanic cable pairs. As long as we operate on galvanic connections, electromagnetic compatibility problems may occur between telecom and power. The problems may be limited to noise and disturbance, but situations may also arise where life and property are at stake. All compatibility problems may be solved technically; the practical limitations are usually of an economic nature. Questions of electromagnetic compatibility are legally taken care of by the Norwegian regulations. Conditions of rights and responsibilities are based on the Principle of Priority, provided the limitations in established regulations are fulfilled.

EMP in telecommunication networks

BY BJØRN FOSSUM

1 What is EMP?

EMP (Electromagnetic Pulse) is a brief and intense electromagnetic wave (radio wave) which may consist of one or a few single oscillations. Its duration is usually less than 1 ms with a low frequency. EMP is transmitted as electromagnetic radiation and induces currents in metal constructions. Thereby, EMP may also be distributed along cables like current (conducted EMP).

EMP is generated by all couplings of electrical circuits. EMP is thus a frequently occurring phenomenon, and telecommunication equipment must therefore have a certain degree of immunity against EMP in order to "survive" in the daily electromagnetic environment.

EMP from nuclear explosions was discovered at an early stage by the fact that it complicated measurements or destroyed instrumentation. It was soon made clear that EMP could cause serious damage to electronic equipment over great distances. EMP from nuclear explosions has caused crackling in radios up to 6,000 km away and has been registered after five circuits around the Earth. This EMP is usually denoted NEMP (Nuclear EMP).

Approximately 0.01 % of the released energy from a nuclear explosion is converted into EMP, i.e. approximately $4 \cdot 10^8$ J per kiloton (kt). Radiated energy may exceed 10^{11} J (≈ 25,000 kWh). The majority of the energy lies in the frequency range 10 kHz – 100 MHz.

2 EMP from nuclear weapons

The most serious threat to the telecommunication network is denoted HEMP (High altitude EMP). HEMP is generated when nuclear weapons are detonated at great heights (> 40 km above ground level). At ground level the effect of the weapon is just an electrical phenomenon, as opposed to if the weapon had been used on the ground or in the air space near the ground.

HEMP is the type of EMP which is generated at ground level and in the atmosphere by detonations of nuclear weapons above the atmosphere. At the time of detonation gamma rays are released which at the speed of light are widely distributed because the attenuation in space is very low. When the gamma rays hit the upper layers of the atmosphere (at approx. 40 km height), they collide with the air molecules. In the collision an electron is emitted ("Compton electron"), which moves in the earth's magnetic field. This electron movement produces currents which in turn generate an electromagnetic field. This field influence will cover a wide area, ref. Figure 1.

The range (in kilometres) is the arc length from a point on Earth directly under the point of detonation to the tangent to earth. The range depends on the height of the detonation and may be calculated by the formula:

$$R_t = R_e \arccos\left(\frac{R_e}{R_e + h}\right)$$

where R_e = earth radius (6,378 km) and h = height of detonation (km).

The area of earth surface (km^2) affected by this HEMP may be calculated by the formula:

$$A = \frac{2\pi R_e^2 h}{R_e + h}$$

Range and affected area as a function of height of explosion is shown in Figure 2.

A detonation above the atmosphere over the North Sea at a height of 100 and 450 km, respectively, produces a coverage area as shown in Figure 3.

The energy density of EMP from nuclear explosions at great heights is expected to be up to 1 J/m^2 with a field strength of 50 kV/m at distances of hundreds of kilometres from the point of explosion. The magnetic field strength (in A/m) is given by the relation

$$B = \frac{E}{Z_0}$$

where Z_0 is characteristic impedance for vacuum, i.e. $120\,\pi$ (≈ 377) ohms.

A typical model pulse for EMP is shown in Figure 4. It has a rise time, t_r, of 10 ns and a half-life, t_h, of 200 ns. The amplitude, A, is 50 kV/m. It may be transferred to the frequency spectrum with the result as shown in Figure 4.

As seen from Figure 5 most of the energy lies below 100 MHz.

The model pulse in Figure 4 is the first part of a long electromagnetic pulse. Because of the high field strength and the wide frequency spectrum, this is what traditionally has been denoted HEMP. The subsequent part of pulse may last for several minutes but with quite a reduced field strength compared to the first 200 ns. The frequency components in this subsequent "tail" are near DC. The field may induce voltages of tens of

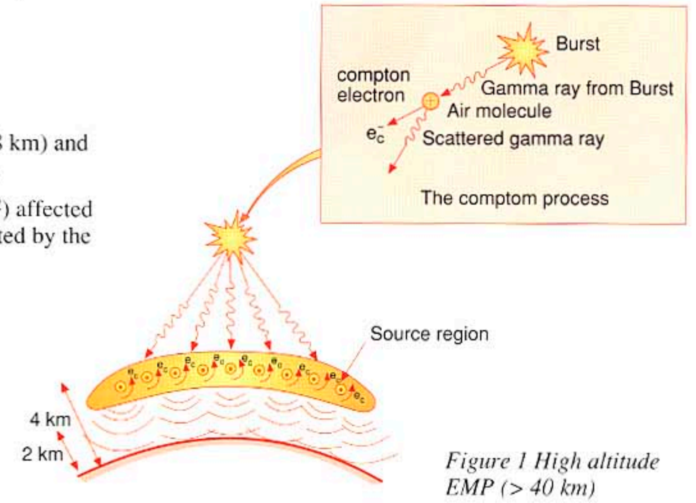


Figure 1 High altitude EMP (> 40 km)

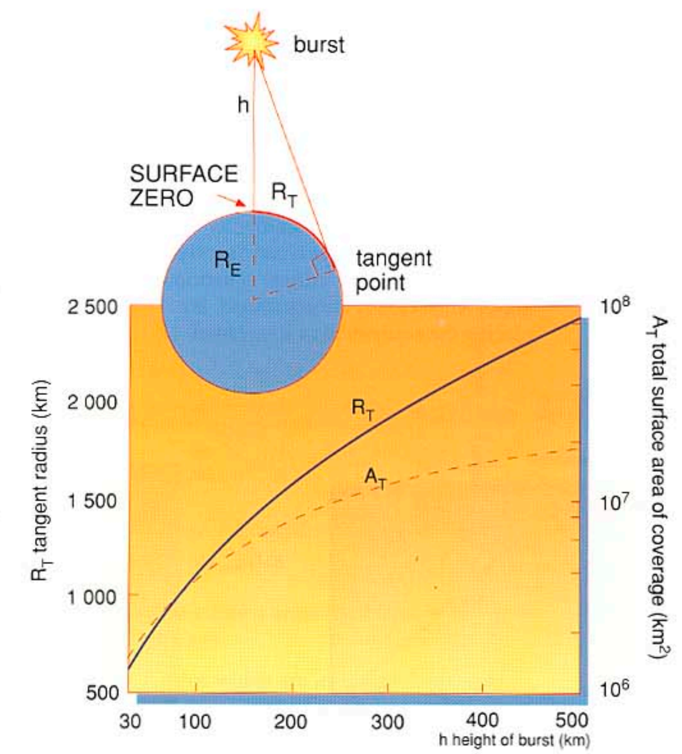


Figure 2 Range and affected area as a function of the height of explosion

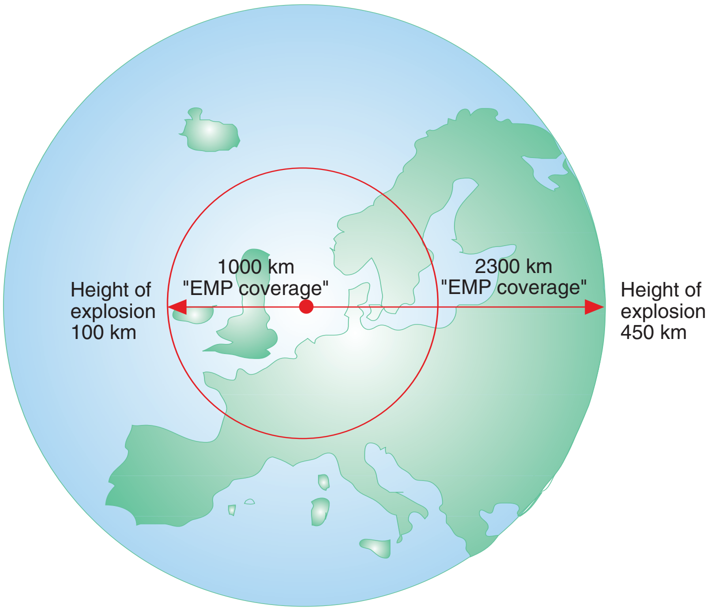


Figure 3 Coverage area of a detonation above the atmosphere over the North Sea

V/km. The phenomenon resembles induction caused by geomagnetic storms which may cause great long-lasting induced voltages on long communication cables.

3 Protection of telecommunication installations

The general method of protecting against radiated EMP is to protect the equipment with the help of a screen cage (Faraday's cage). The attenuation requirements are dependent on the equipment's resistibility against NEMP. We have placed different types of telecommunication equipment in an EMP simulator in order to check the intrinsic EMP immunity of the equipment and whether it can "survive" such a radiation. The critical situation arises when cables are connected. By placing the equipment in a screened room, we have several options as regards the cabling within the installation. Cabling inside a screened room may as a general rule be done like an ordinary telecommunication installation. If the EMP screen is open, e.g. wire mesh is used, screened cables have to be used when cabling close to the EMP screen. Use of and requirements to screened cables are otherwise as for an ordinary telecommunication installation.

All electrical connections with the outside world, like communication cables, power supply, ventilation ducts, etc., should thus be handled as specially. Long cable routes may act as antennas for EMP and actions must be introduced to reduce the harmful effects. Surge arresters, filters and limiters may be used; all mounted with the lowest possible impedance to the screen. All lead-in cables should have a screen which is connected to the room screen via a bushing sheet. The connection between the cable and the room screen is critical; it should be done with the minimum impedance and with contact to the whole circumference of the cable.

3.1 Building protection

Depending on the construction of the building and at which stage the EMP screen is to be installed, the building's own reinforcement may be used as protection, or an additional protection is installed, made of foil or wire mesh.

In EMP screens with single reinforcement consisting of iron reinforcements 8 mm thick and with a mesh size of 150 mm, the attenuation in the frequency range 10 kHz – 10 MHz is around 28 dB, while for wire mesh with a mesh size of 25 mm and wires 0.8 mm thick, the attenuation is around 40 dB. The attenuation applies to rooms with a minimum size of 4 m. Double reinforcement is very often used, producing greater attenuation depending on the distance between the screens.

For a room of height *h* and longest side *l*, the lowest resonance frequency is given by the formula

$$f_0 = 150\sqrt{\frac{1}{h^2} + \frac{1}{l^2}}$$

where the room's dimensions are given in m and the resonance frequency in MHz.

By use of reinforcement mesh protection a combination of bonding and binding is employed. In Figure 6 bonding points are indicated by circles, while binding points are marked by a slash. Regardless of the type of screen chosen, good electrical connection at the joints must be assured.

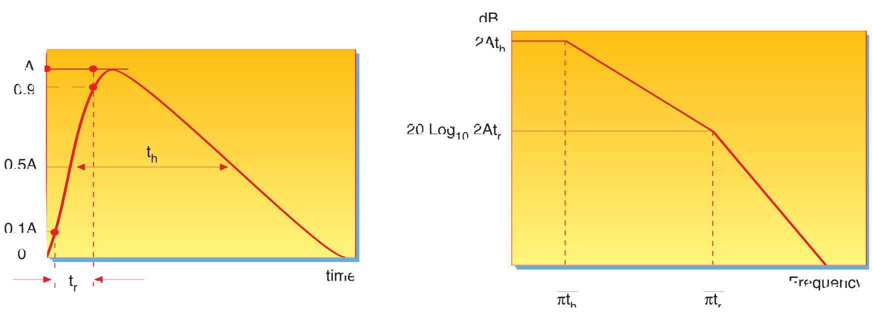


Figure 4 Typical model pulse for EMP

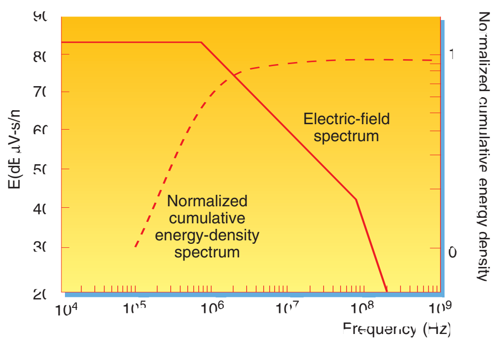


Figure 5 High altitude EMP spectrum and normalized energy density spectrum

It is important that the screen is not "punctured". Outer bolts e.g. should not be led through the screen. All outer metal constructions, like cable rack from the antenna tower, work as an antenna for EMP, and their chance of leading destructive currents into the installation should be prohibited.

3.2 EMP barrier

The ideal requirement of all cables being fed through a single entry combined with filters and surge protection placed in the barrier, usually has to be abandoned and replaced by compromises. A telecommunication installation may be an exchange with tens of thousands of connections. It is hard to visualise a single entry with limiters and filters for tens of thousands of line pairs placed in the barrier!

It is of great importance that the electrical connections between the screen room and the outside world are concentrated in one point; a so-called single entry. Having multiple separate entries in the screen may easily cause currents in the screen wall which may destroy the effect of the room screen.

The best discharge of sheath currents is achieved when a copper braided-wire shield is thread on the outside of the metal sheath, so that the whole circumference of the cable is covered. If the screen termination is done as pig-tail, its impedance may be considerable ($U = \omega LI$ or $L \cdot di/dt$). All bondings or screen connections should therefore be as short as possible and preferably done with a copper lace/ribbon in order to limit the inductivity. The inductivity, L, is approx. 1.5 µH/m for wire. In Figure 7 are shown examples of screen termination.

Telenor has standardised the shield connector, so that the whole of the cable's circumference is covered. The braided-wire sleeves are shown in Figure 8. They are provided in dimensions from 16 mm to 84 mm (inside measurements).

3.3 Physical connections

A telecommunication network may consist of both screened and unscreened cables and installed mainly as earth cables as regards protected installations. But particularly on the outskirts of a subscriber network there may be aerial cables.

For earth cables an accumulated current of 1 – 1.5 kA may be estimated (1). For unscreened cables the voltage between conductors and earth may reach 100 kV. Due to the thickness of the plastic insulation, damage is not expected on our standard telecommunication cables, but connected equipment may be harmed. Currents in an earth cable as a function of the length is shown in Figure 9.

Aerial cables may hold larger currents and as characteristic impedance for aerial cables is greater than for earth cables, the voltage will also be higher and a flash over to suspension equipment may occur if the cable is not protected.

For screened cables the voltage will be reduced. We have chosen screened lead-in cables where coupling impedance is reduced by increasing frequency. 0.2 mm aluminium sheath or 2 mm lead sheath is used. These types of cables are shown as graph A in Figure 10. As a comparison is shown the coupling impedance for cables with an aluminium foil cover (graph C) and with a braided copper screen (graph B). One may clearly see that for cable screens which are not HF tight, the coupling impedance increases considerably with high frequencies.

If unshielded cables are used in a network it is still a requirement that shielded cables should be employed for the last 300 m leading up to the EMP protected installation. Lead cables are preferred; not least because it is relatively easy to have sheath currents discharged to the barrier.

Outside the installation the shielded cables may be connected to long unshielded cables. By using shielded lead-in cables, we have in principle introduced a low-pass filter, even if the attenuation is not very high. Typical value for plastic insulated cables is 3 dB/100 m at 1 MHz, and the attenuation in dB is approximately proportional to the square root of the frequency.

In order to achieve the wanted effect of the cable screen it has to be earthed; the best solution being a continuous discharge to earth. Such discharge is achieved quite simply for lead sheathed cables without any outer insulation and for Al-sheathed cables by the outer plastic sheath being added coal, thus turning the plastic into a semi-conductor (ρ ≈ 20 Ωm).

Cables which are insulated from earth should be earthed at various points some 30, 60 and 90 m away from the screen room. In addition, the cable screens should be discharged to the screen wall

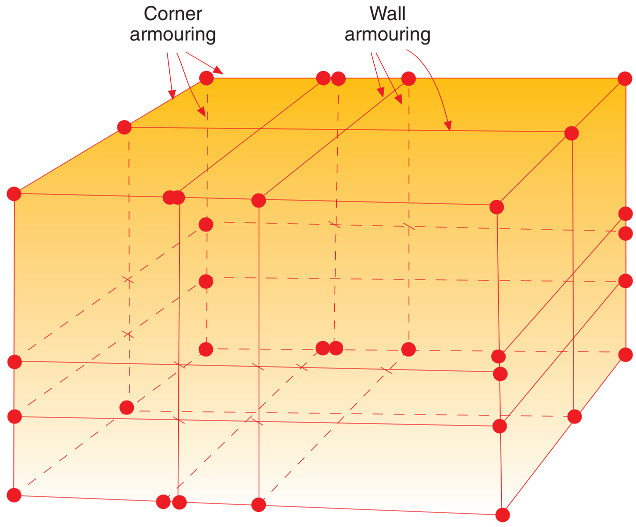


Figure 6 Reinforcement mesh screen

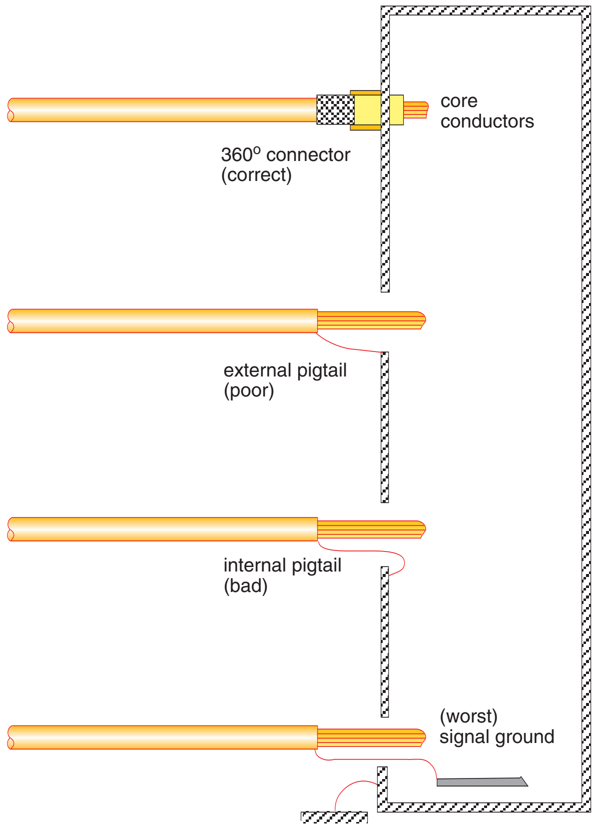


Figure 7 Examples of screen terminations

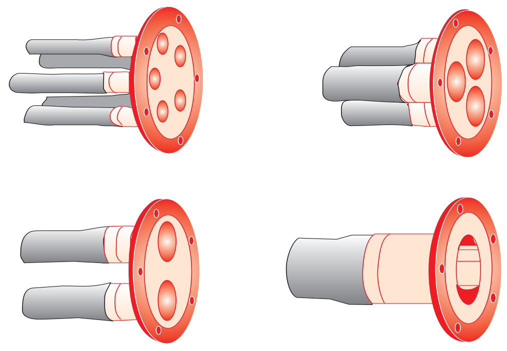


Figure 8 Examples of braided-wire sleeves

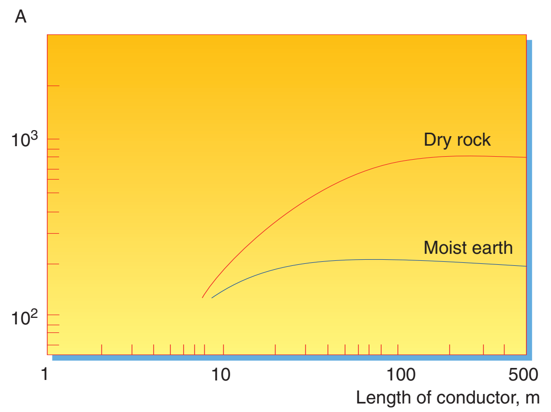


Figure 9 Current as a function of cable length

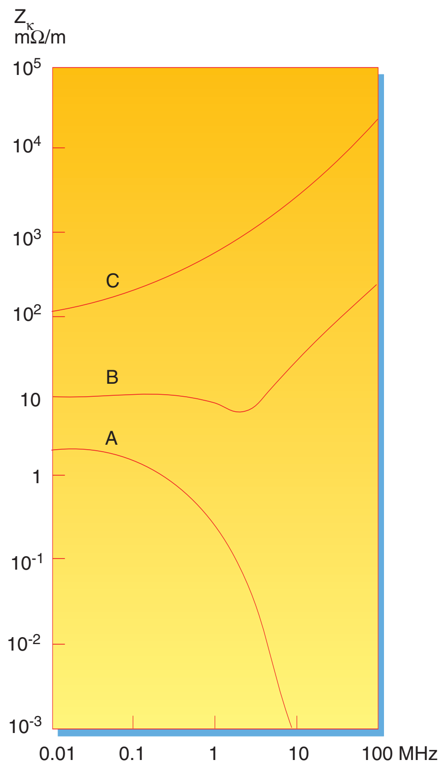


Figure 10 Shielded cables

by braided-wire sleeves in single entry. The outer sheath of the cable is important for the screening effect. Lead cables without any outer insulation are favourable from an EMP point of view, as they may be installed directly on earthed cable racks or put directly in contact with the ground. The attenuation of the sheath current because of insulation related to earth may perhaps be illustrated by the following numbers (1):

- Approx. 2 dB/100 m on cable with insulated sheath,
- Approx. 4.5 dB/100 m on cable with conducting sheath in dry rock, and
- Approx. 33 dB/100 m on cable with conducting sheath in moist ground.

Cable with conducting sheath laid directly in cable racks of metal and being continuously earthed, also show favourable values for sheath discharge.

3.3.1 Cable termination/protection

Large parts of the telecommunication cable network is unshielded, and the voltages fed into the installation can therefore be great. In order to reduce the conducted EMP to levels which the equipment may stand, all connections should be equipped with surge arresters. For paired cable conventional gas discharge tubes (GDT) are used, e.g. mounted in Krone/Siemens Trennleiste 71 or Krone LSA + cable termination. The reason why we may use conventional GDTs is the “filter” in the form of shielded cables in front of the barrier. For protection of antenna cables are used coaxial mounted GDTs. For protection of power supply and short connections for fire alarms and so on are used various types of varistors.

3.4 Optic cables

Fibre optic cables are well suited as connections between the protected installation and the outside world. This is particularly favourable if the cable is metal free. If the cable contains metal, the metal should be connected to the screen wall as for conventional paired cables. Regenerators or terminals may, however, be influenced by EMP.

4 Earth electrodes

Telenor uses mainly earth rods in order to avoid an increase in earth resistance caused by drought or frost. In connection with lightning protection it is well known that several short electrodes are more favourable than one long one. This particularly applies to NEMP because of the short rise time of the pulse. The effective length and pulse impedance for one earth rod may be calculated by the following formulas (2):

$$l_{eff} \approx 0.9\sqrt{t_r \rho}$$

$$R_{pulse} = \frac{1}{G l_{eff}}$$

$$G = \frac{2\pi}{\rho} \frac{1}{\ln \frac{l}{r}}$$

where

G is the conductance of the electrode (S/m)

l_{eff} is the effective length of the electrode (m)

l is the actual length of the electrode (m)

t_r is the rise time of the pulse (µs)

ρ is earth resistivity (ohmm)

r is the radius of the electrode (m).

5 References

1 *Elektromagnetisk puls, virkninger og vern.* FFI/Rapport-82/3003.

2 *Handbuch für Blitzschutz und Erdung.* Hasse/Wiesinger.

Telecom overhead cables as antennas for long wave radio signals

BY ARNE THOMASSEN

Introduction

It has been observed over quite a long period of time that telephone subscribers in some areas were annoyed by false telephone calls. At the same time the normal functioning of the telephone equipment was disturbed, especially for frequency division multiplex (FDM) subscribers. Investigations in 1987/1988 showed that the noise problems came from a navigational transmitter in Germany, and the problems will here be dealt with in 9 sections.

1 The actual transmitter

Among the navigational transmitters, a transmitter in Germany gave us the strongest received signals. The frequency 23.4 kHz was also well inside the Division Multiplex frequency band in the direction from the customer to the exchange (bandwidth 19 – 37 kHz). In addition, even the carrier was from time to time amplitude modulated with 10 Hz and interfering with our own 10 Hz signalling. The transmitter also came on and off from time to time, and the transmitted level seemed to vary.

2 Time propagation effects

Because of the special propagation effects the level of the received signal could be several times higher than normal about one hour after sunset and one hour before sunrise. Maximum values ("worst case") was then obtained, but even this varied with time, latitudes and longitudes (fading). It was assumed that it was advisable not to make routine measurements in these periods. A reference antenna was established in Arendal and maximum values were continuously recorded for several days, and a multiplication factor was then given to the measuring teams travelling along the coast controlling the FDM systems for one exchange at a time.

3 Geographical propagation effects

Also a geographical kind of propagation effect was observed. A low frequency signal has a tendency to bend (1, page 365) to the lowest conductivity. Therefore, the 23.4 kHz signal could follow the coastline up to Karmøy and Bergen. In addition, the signal on the south coast of Karmøy had a higher received signal than Stavanger because of the recovery effect (1, page 284). This is in fact also a result of the use of Millington's method (1, page 282).

The signal strength is much attenuated over land because of low conductivity. Lakes, fjords and valleys with rivers will give relatively good conditions for the signals, while especially mountains and woods may increase the attenuation. Table 1 and Figure 2 show some maximum field strengths in Southern Norway.

4 22 kilovolt overhead power lines as antennas

Long 22 kilovolt overhead power lines in the north-south direction may act as travelling wave antennas. The induced currents are added to another all the way along the power line and may reach large levels, and then it is induced into parallel telecommunication cables. If the power line is long enough the energy may be transmitted from the power line to the telecommunication cable, or from power line via power cable to telecommunication cable, ref. Table 2 and (2).

5 Travelling wave antennas

Overhead paired cables will act as travelling wave antennas and may even form quarter-wave antennas and common mode LC-resonators. Shielded cable sections may play a part in resonances, RC-attenuation and impedances to the ground. Signals may be reflected from power lines, paired cables, and big metal structures and be induced into the cables with different phases and with different levels. The result of all this is therefore that the condition is rather complicated and difficult to predict.

The most convenient way is therefore to measure the standing waves and then make some estimates. One thing that seems certain, however, is that the induced currents travel with the same phase and the same direction on all the single wires in the cable (common mode). A change in one of these cable pairs therefore also affects all the other cable pairs and alters the standing wave patterns on all the pairs. Individual variations will appear where the cable has taps to smaller cables.

Table 1 Maximum magnetic field strength in µA/m, Southern Norway

Locations	19 kHz	23.4 kHz
Arendal, ref	50	100
Dølemo	35	60
Vrådal	25	20
Kviteseid	30	20
Brunkeberg	25	20
Mandal	40	160
Vigeland	35	140
Flekkefjord	35	140
Jøssingfjord	25	110
Stavanger	25	100
Karmøy	45	140
Ølen	25	35
Røldal	20	15
Hønefoss	40	15
Hægeland	15	45
Byglandsfjord	40	60
Lillehammer	25	25
Vågsli	35	25
Haukeli	25	25
Ålvik	25	20
Br. Knarvik	25	110
Lindås	20	25
Vadheim	25	25
Florø	20	8
Ramsdal	10	15
Førde	25	10
Stryn	25	35
N.fjordeid	35	35
Ålesund	10	6
Namsos	4	20

Comments to the table:
All measurements are related to worst case. 23.4 kHz is the strongest frequency component most places.
19 kHz is varying somewhat different. A few values could be high because of reflection from metal structures, but they should give a good idea of the situation as examples.

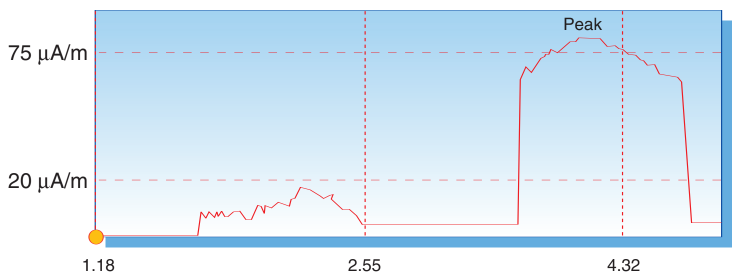


Figure 1 Peaking level a little before sunrise

5.1 Measurement at the filter location in Hægeland

It is evident that low levels between cable pairs and ground is normal in the exchange because of low impedance to the ground while levels far out on the cable can be high. This is due to standing waves, impedances and mismatch. At such critical places only a slight impedance imbalance or mismatch can cause trouble. Transformers and common mode noise filters, correctly placed, can reduce the problem. The best, however, is to avoid FDM systems when such conditions occur.

6 Cable pair impedances to ground

Low impedances to ground at the exchange is an important factor. At some places with lead cables into the station, signals to ground were completely extinguished while they were still high in the outer end of the attached overhead cable network. Unfortunately, it is in this outer end that bad impedance balance in line filters made common mode currents into signals between a- and b-wire. If there are several sections with shielded cable (see Table 3, line 1), the induced noise is greatly attenuated both because of less induction and because of lower impedance to the ground. Overhead cables with lengths of 100 metres between filter and customer did not cause any trouble. Cables up to 20 metres between filter and customer gave no resonance problems at relevant frequencies. Shielded overhead cable reduced the problems, but even distance between the grounding should be avoided because of the danger of multiple resonance.

7 Longitudinal balance and mismatch on cable pair

The last model of the FDM 1+1 system had a line filter which had a serious longitudinal balance error. Induced common mode noise on cable pairs were then transformed to differential mode and could then disturb the receiver at the exchange. The fault was in a common mode filter-coil so that the longitudinal unbalance came between input and output of the filter and the common mode currents on the line then caused problems. The filter was "floating" and had no connection to the ground. This caused mismatch in the filter. There was also a mismatch as well in the exchange as at the customer. Errors in the cable pairs

Table 2 Magnetic field strengths (23.4 kHz) at Kristiansand

Magnetic field in the open fields	100 µA/m
Magnetic field 5 metres from 230 V power line	200 µA/m
Magnetic field 5 metres from 22 KV power line	500 µA/m
Magnetic field 50 cm from 22 KV cable	5000 µA/m
All measurements in this table are related to worst case.	

Table 3 Results in 10 exchanges from Kristiansand to Karmøy

235	customers	had between	001 and 499	µV	(Safe values)
30	customers	had between	0.5 and 1.5	mV	(Shadow zone)
19	customers	had between	1.6 and 3.9	mV	(Danger zone)
6	customers	had between	4.0 and 6.0	mV	(Some failure)
2	customers	had	6.1 and 15	mV	(Much failure)
1	customer	had (Hægeland)	53	mV	(Much failure)

themselves, however, could also cause an unbalanced condition to the ground so that the noise voltage (measured up to 850 mVrms) would give severe problems. These filters were later removed from problem locations and replaced by modified filters.

7.1 Measurement showing mismatch and unbalance at filter

The measurement shows that the mismatch and unbalance are creating a signal between a-wire and b-wire at the Hf-customer (90 mV). The signal may then be attenuated on its way to the demodulator in the exchange (50 mV).

8 Susceptibility of the FDM 1+1 system

Customer-HF demodulators on the FDM 1+1 systems use carrier on/off for signalling. (28 kHz gave dialling tone. Short breaks in the 28 kHz gave the dialling signals). Of course, the exchange also receives tone signalling, but both systems existed simultaneously. When the noise voltages exceeded the receiver threshold of about 2 – 4 mV, this was treated as a call and the dialling tone started. Keying was detected when level varied or the unwanted radio signal was keyed. This caused automatic set-ups to other customers. The customers who had been phoned automatically complained to our service department. They claimed that someone was bothering them because nobody answered. The FDM customers

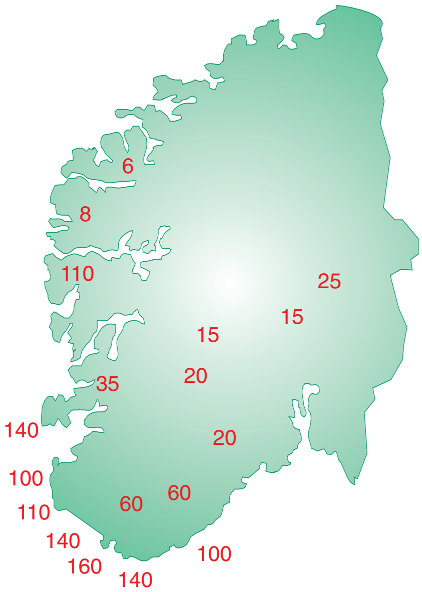


Figure 2 Maximum field strength in µA/m, Southern Norway

also made complaints because they often could not get the dialling tone. The noise had activated the dialling tone for such a long time that the exchange had deactivated it.

9 Routine measurements

293 FDM customers in Flekkerøy, Randesund, Hægeland, Kvinesdal, Storekvina, Hjelmestad, Ferkingstad, Koparvik, Kjernagel and Sveio in the

Southern and South-Western part of Norway, were measured. A lot of other places were later controlled (ref. Table 3). The lines were controlled with the help of a HP-3580 spectrum analyzer, with high impedance and in parallel with communication signals between a-wire and b-wire. Not all customers had problems, even at the worst locations. Only a few customers in some of the locations had overhead cables with problematic lengths which could cause resonances. The measurements were calibrated using the antenna in Arendal simultaneously.

Comments to Table 3:

Worst exchange was Hægeland in Setesdal with many unshielded overhead cables. Most of the FDM customers were in trouble there. After changing from old filter to a modified one, nearly all customers had a disturbing signal level below 0.5 mV, which was well below the level at which the equipment would operate. But the high levels between cable and ground are still there. Just a slight error in line impedance balance may therefore bring the problem forth again.

Randesund was a well operating exchange because of many shielded cables (the 500 KV-DC power line goes through the area on its way to Denmark). Only one out of a hundred customers had trouble there.

The exchanges at Åna-Sira and Flikka sometimes had so big problems that the entire exchange was blocked. As a first step the service-team had to disconnect some of the FDM equipment.

An example in Åna-Sira exchange shows the resonance problems. From the exchange there was at first 3,500 m shielded overhead cable, grounded with even intervals. Then came a 450 metre unshielded overhead cable up to the line filter. From the filter to the customer the distance was 300 metres. All worked well, but then a new customer was connected to this line. The only difference now was the fact that the distance between filter and customer was reduced to 200 metres. This caused the signal to exceed the problem level. This shows how difficult it is to predict these conditions. It is absolutely necessary to make measurements.

Koparvik exchange at Karmøy had no customers above 50 µV disturbing signal. The lead shielded cable caused these very low levels. Some overhead cable outside the lead cable here did not seem to cause problems.

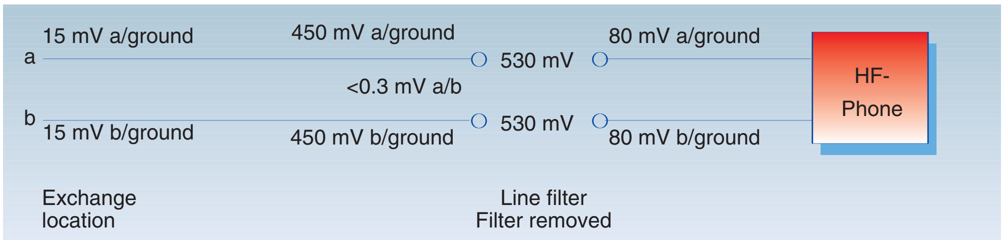


Figure 3 23.4 kHz levels between wires and ground

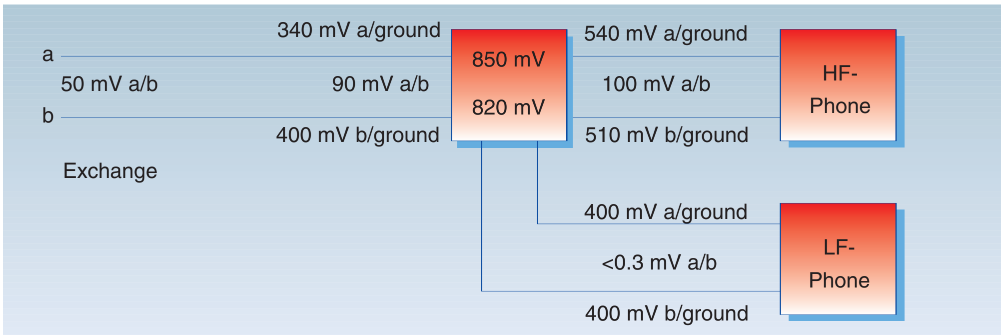


Figure 4 23.4 kHz levels with filter included in Hægeland

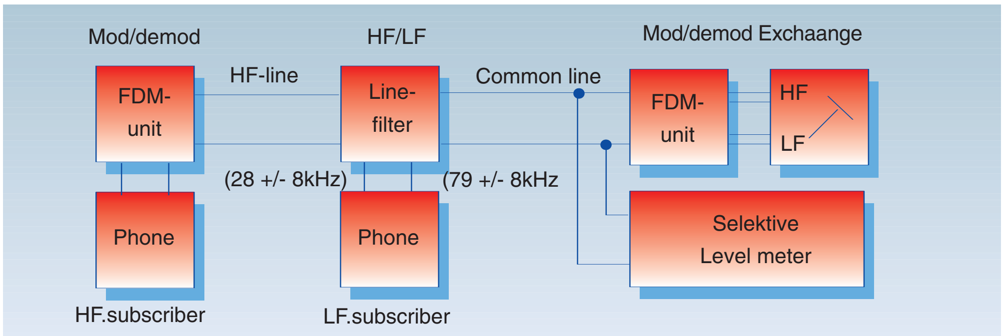


Figure 5 Method of routine measurement

Kjernagel exchange in Haugesund had overhead and underground cables without shielding. The levels here were comparatively low. 5 years later, however, some problems arose at this exchange. Something had presumably happened in the cable network.

Routine measurements should be done at all exchanges from the Swedish border up to Bergen because of all the uncertainties. Every measurement should be referred to the reference level in Arendal. Exchanges with many overhead cables along the coast should be measured first. Some places which are thought to be safe may not actually be so.

References

1 Stokke, K N. *Radiotransmisjon: kabler, bølgeutbredelse, antenner.* 4. utgave. Oslo, Universitetsforlaget, 1982.

2 Stokke, K N. Travelling wave antennas. *Telektronikk,* this issue.

EMC-norms and regulations – the situation in Norway

BY SVERRE TANNUM

The regulations on electrical devices and systems are based upon the low voltage directive and the EMC-directive of the European Union (EU). The substantial content of this directive was implemented in the Norwegian regulations as of January 1, 1991 (Forskrift om utførelse og kontroll av elektrisk utstyr som tilbys eller omsettes til bruk i lavspenningsanlegg). The EMC requirements are incorporated in § 11.

Traditionally, it has been demanded that a number of products shall be controlled in one way or another before marketing. We say that a major part of all electrical equipment belong to the regulated product area.

The situation now is that there is no longer an obligation for approval tests of electrical equipment before marketing in Norway, but the authorities have established a registration system which requires full technical documentation of the product. For all electrical apparatus and equipment which can be connected to the low voltage network a declaration is required from the producer identifying the product itself and the actual Norwegian regulations or norms which have to be fulfilled for this product.

The EMC directive covers all electrical and electronic equipment with very few exceptions.

A short time after the approval of the directive it was agreed that there should be a transition period from January 1, 1992 to January 1, 1996. In this period the producers can choose if they want to fulfil the national requirements of each country or the EMC directive with CE-marking, according to the Cenelec European norms (the new approach).

As mentioned, the Norwegian regulations were put into force as of January 1, 1991. In another regulation put into force on the same date it is stated that the self-declaration shall contain a reference to the relevant Norwegian regulations and shall confirm that the equipment complies with their requirements. This is valid for EMC when norms exist. In other words, Norwegian regulations require compliance with national Norwegian norms.

The situation now is that some 20 European EMC norms exist. They are already put into force as Norwegian norms and consequently, they have to be complied with.

Because Norway already has adopted the European EMC Norms as Norwegian norms, we have in principle no transitional period like many other European countries have.

The so-called generic norms for emission and immunity for domestic commercial and light industry equipment is very important in this respect.

To put the new European Norms (EN) into force at an early stage without waiting until 1996 could make the Norwegian export industry more competitive, but on the other hand, it would reduce the turnover for Norwegian import firms who get their products from countries without the same EMC requirements.

These problems are now being discussed and the question is whether the economic interests will overrule the importance of the electromagnetic environment until 1996.

If we establish the same transitional period as many other European countries have, these years should not be spent without preparing for the coming situation.

As a minimum, we should demand that all products which are developed now and in the coming years must comply with the requirements of the EMC directive. This view is expressed by the leader of NEK (The National Electrotechnical Committee of Norway), Mr. Bjørn I. Ødegård, and he recommends a positive attitude to the EMC quality requirements.

An overview of Norwegian electrotechnical norms concerning EMC

NEK-EN 50 065-1	(1991)	Signalling on low-voltage electrical installations in the frequency range 3 kHz to 148.5 kHz. Part 1: General requirements, frequency bands and electromagnetic disturbances.
NEK-EN 50 081-1	(1992)	Generic emission standard. Part 1: Residential, commercial and light industry.
NEK-EN 50 081-2	(1993)	Generic emission standard. Part 2: Industrial environment.
NEK-EN 50 082-1	(1992)	Electromagnetic compatibility – Generic immunity standard Part 1: Residential, commercial and light industry.
NEK-EN 55011	(1989)	Limits and methods of measurement of radio disturbance characteristics of industrial, scientific and medical (ISM) radio frequency equipment. Modified.
NEK-EN 55 013	(1989)	Limits and methods of measurement of radio disturbance characteristics of broadcast receivers and associated equipment. Amdt 1 (1983).
NEK-EN 55 014	(1988)	Limits and methods of measurement of radio interference characteristics of household electrical appliances, portable tools and similar electrical apparatus. Modified. Amdt 2 (1989) A2 to EN 55 014.
NEK-EN 55 015	(1988)	Limits and methods of measurement of radio interference characteristics of fluorescent lamps and luminaries. Modified. Amdt 1 (1989) A1 (1990) to EN 55 015.
NEK-EN 55 020	(1988)	Immunity from radio interference of broadcast receivers and associated equipment.
NEK-EN 55 022	(1988)	Limits and methods of measurement of radio interference characteristics of information technology equipment. Modified.
NEK-EN 60 555-1	(1987)	Elektrisk støybegrensning i forsyningsnett. Part 1: Definisjoner.
NEK-EN 60 555-2	(1987)	Part 2: Harmonics. Modified. Amdt 1 (1985).
NEK-EN 60 555-3	(1987)	Part 3: Voltage fluctuations. (+ Corrigendum 1990) A1 (1992) to EN 60 555-3.
NEK-EN 60 601-1-2	(1993)	Medical electrical equipment Part 1: General requirements for safety. 2. Collateral standard: EMC – Requirements and tests.
NEK-EN 61 000-4-1	(1994)	Part 4: Testing and measurement techniques Section 1: Overview of immunity tests.
NEK-EN 61 000-4-7	(1993)	Part 4: Testing and measurement techniques. Section 7: General guide on harmonics and interharmonics measurements and instrumentation for power supply systems and equipment connected thereto.
NEK-EN 61 000-4-8	(1993)	Part 4: Testing and measurement techniques. Section 8: Power frequency magnetic field immunity test.
NEK-EN 61 000-4-9	(1993)	Part 4: Testing and measurement techniques. Section 9: Pulse magnetic field immunity test.
NEK-EN 61 000-4-10	(1993)	Part 4: Testing and measurement techniques. Section 10: Damped oscillatory magnetic field immunity.
NEK-HD 481.1 S1	(1987)	Electromagnetic compatibility for industrial process measurement and control equipment. Part 1: General introduction.
NEK-HD 481.2 S1	(1987)	Part 2: Electrostatic discharge requirements.
NEK-HD 481.3 S1	(1987)	Part 3: Radiated electromagnetic field requirements.

EMC measurements

BY SVERRE TANNUM

Open Area Test Site (OATS)

The new Open Area Test Site (OATS) of Radiostøykontrollen (the Radio Interference Service) is located at Maridalen near Oslo. It is built for the normalized measuring distances 3–10 and 30 metres and complies with the international requirements for OATS.

The test site must be free from reflecting objects within the ellipse having a major diameter 2 d and a minor diameter √3 d (d = measuring distance).

The glass fibre weather protection covers the 3 – 10 metres measuring sites. 30 m measurements is made without weather protection of the receiving antenna.

The measuring site is located about 15 km outside the city boundary and the topography gives good protection from the ambient noise which is at least 30 dB lower than in the city.

Specifications:
Weather protection:
20 x 10 x 8 m (LxWxH)

Ground plane:
18.5 x 10 m

Turntable diameter:
4 m

Photo: Telenor AS

Figure 1 OATS in Maridalen

Photo: Telenor AS

Figure 2 OATS interior, 3 and 10 metres measuring sites with weather protection

Turntable load:
5,788 kg

Power supply:
30 A 400 V AC three phase

30 A 230 V AC three phase

16 A 230 V AC single phase

100 A 48 V DC

By special arrangement
160 A 230 V AC

Site attenuation:
Within the specified CISPR Limits ± 4 dB.

Electrical and electronic consumer equipment as well as telecommunication apparatus must comply with national or international EMC (Electro Magnetic Compatibility) requirements. EMC means the ability of a device, unit of equipment or system to function satisfactorily in its electromagnetic environment without introducing intolerable electromagnetic disturbances to anything in that environment.

Testing of EMC includes testing of both emission and immunity.

Radiostøykontrollen performs measurements according to the following norms:

EN 55011 corresponding to	CISPR Publication 11
	CISPR Publication 12
EN 55013 corresponding to	CISPR Publication 13
EN 55020 corresponding to	CISPR Publication 20
EN 55022 corresponding to	CISPR Publication 22
EN 50081-1	
EN 50081-2	
EN 50082-1	

and corresponding national norms.

A short description of some EMC measurements

Measurement of radiated interference

The most common method of measuring emission of radio noise and interference is to measure the field strength of the emitted signal from the equipment under test (EUT) which is placed upon a turntable rotating 360 degrees.

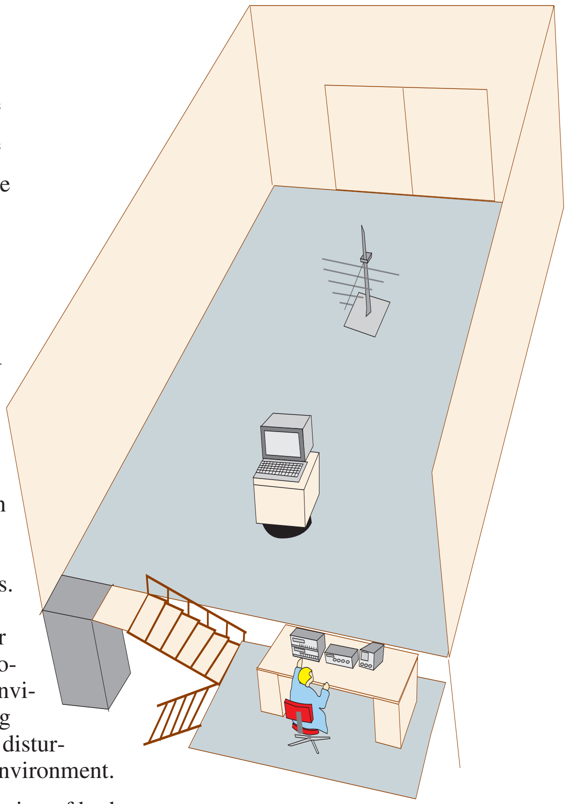


Figure 3 EMC test site, emission measurement

The field strength is measured via the receiving antenna by a measuring receiver. The antenna height is adjusted between 1 and 4 metres above the conducting ground plane for maximum reading.

Measurements of conducted interference

Conducted interference (normally below 30 MHz) is measured either directlt on the signal terminals or via a Line Stabilizing Network (LISN) or an Artificial Mains Network (AMN). The interference level is recorded by the measuring receiver.

The AMN and LISN are required to provide a defined impedance at high frequencies and they have filters to isolate the test circuit from unwanted radio frequency signals on the signal- and supply lines. The measurements of conducted interference are performed in a screened room to avoid influence from background noise.

“Clamp” measurements

An alternative measuring method to radiation measurement is the so-called “clamp” method.

It is generally considered that for frequencies above 30 MHz the disturbing energy produced by appliances and similar devices is propagated by radiation. Experience has shown that the disturbing energy is mostly radiated by the portion of the mains lead near the appliance. It is therefore agreed to define the disturbing capability of an appliance (EUT) as the power it could supply to the mains lead. This power is nearly equal to that supplied by the appliance to a suitable absorbing device (clamp) placed around this lead at the position where the absorbed power is at its maximum.

Immunity measurements

Immunity is the ability to maintain a specified performance when the equipment is subjected to interference signals. Unwanted signals can attack the victim equipment in different ways like conducted disturbances, radiated fields (wanted and unwanted) from other equipment or even from modules in the victim equipment itself. Concerning broadcasting receivers we usually specify the following:

- Input immunity (immunity from unwanted signals present at the antenna input terminal)
- Immunity from conducted voltages (unwanted signal currents present at the mains- and signal terminals of the equipment)
- Immunity from conducted currents (unwanted signal currents present in cables connected to the equipment)
- Immunity from radiated fields (electromagnetic fields present at the equipment)
- Screening effectiveness is the characteristic of a coaxial connector terminal to attenuate the transfer of external currents into internal voltages.

The EMC laboratory is equipped to perform these measurements.

Stripline

The immunity of broadcasting receivers from radiated fields can be tested in an open stripline device (TEM). The equipment under test (EUT) is placed in the open stripline with the wanted signal supplied at the input terminal. The unwanted signal in the range 0.15 MHz to 150 MHz is supplied by a generator via a matching network to the stripline which is loaded with a terminating impedance (150 ohms). The wanted signal is specified according to the operating mode of the receiver and the unwanted signal is amplitude modulated with 1000 Hz at 80 % depth.

The stripline is normally placed in a shielded room. The level of the unwanted field is monitored by a probe close to the EUT, and the level is increased until interference is detected in picture or sound quality.

The immunity test based upon a large number of single frequencies will be very time consuming. A scan-method is now evaluated to reduce test time and costs.

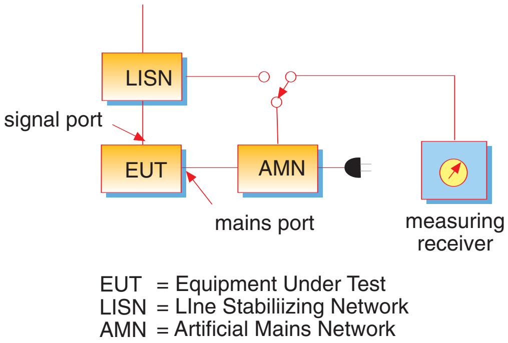


Figure 4 Measurements of conducted interference

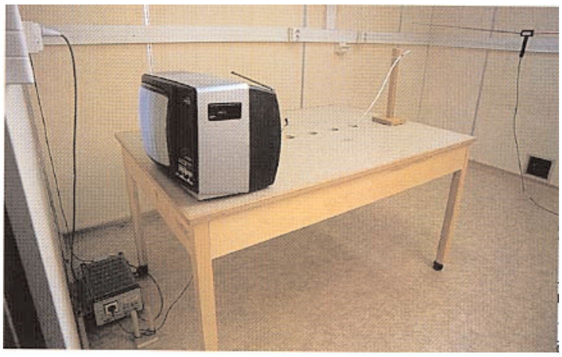


Photo: Telenor AS

Figure 5 Measurement of conducted interference, the artificial mains network (AMN) down to the left

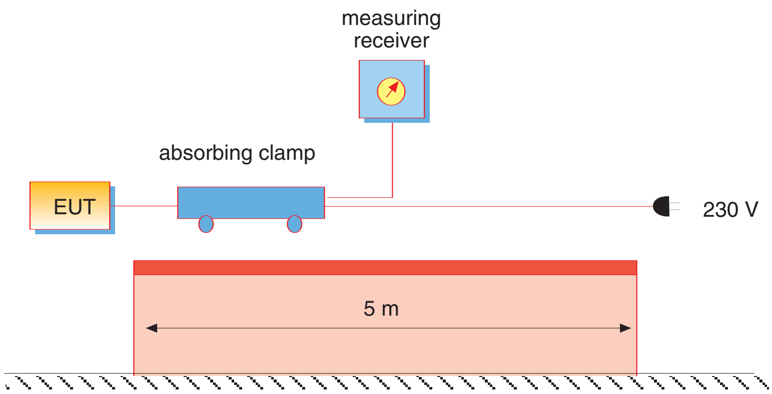


Figure 6 “Clamp” measurements

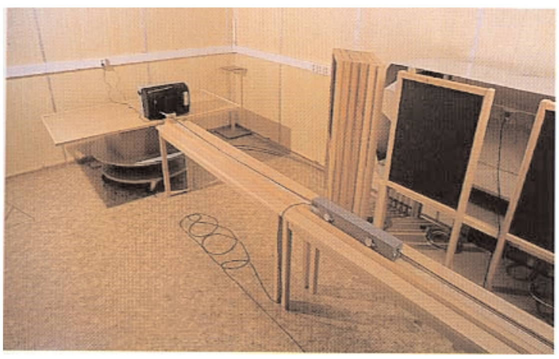


Photo: Telenor AS

Figure 7 Measurements by clamp in a screened room

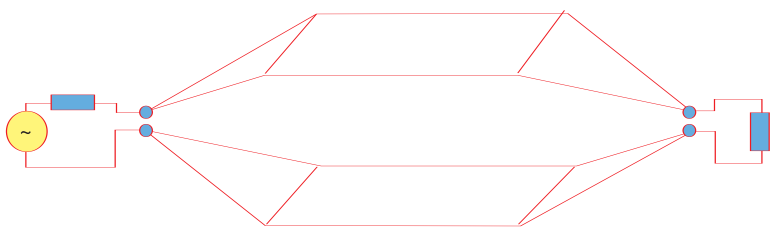


Figure 8 Open stripline – TEM device, basic configuration with matching network and terminating impedance. The equipment under test is placed between the two plates

Table 1 Measuring instruments and antennas

Emission measurements			
CISPR EMI Receiver:			
Spectrum Analyzer	100 Hz – 22 GHz	HP 8566B	Hewlett Packard
RF Preselector	20 Hz – 2 GHz	HP 85685A	Hewlett Packard
Quasi Peak Adapter		HP 85650A	Hewlett Packard
EMI Test Receiver	9 kHz – 30 MHz	ESVS30	Rohde & Schwarz
EMI Test Receiver	9 kHz – 30 MHz	ESH3	Rohde & Schwarz
EMI Test Receiver	20 – 1300 MHz	ESVP	Rohde & Schwarz
Spectrum Monitor		ESM	Rohde & Schwarz
EMI Software		EZM-K1	Rohde & Schwarz
XYT Recorder		ZSKT	Rohde & Schwarz
Spectrum Analyzer	9 kHz – 2.6 GHz	R3261	Advantest
Synthesized CW Generator	1 – 20 GHz	HP 83711A	Hewlett Packard
Signal Generator	0.1 – 990 MHz	HP8656A	Hewlett Packard
Colour TV Pattern Generator		PM 5415TN	Philips
Programmable RF Generator	100 kHz – 1 GHz	PM 5390S	Philips
Process Controller		PSA 17	Rohde & Schwarz
Royce Field Site Source	10 – 600 MHz	4610	EMCO
Impulse Generator		91263-1	ETN
Active Loop Antenna	10 kHz – 30 MHz	6502	EMCO
Dipole Antenna Set	28 MHz – 1 GHz	3121	EMCO
Biconical Antenna	20 – 300 MHz	HUF-Z2	Rohde & Schwarz
Logperiodic Antenna	200 – 1000 MHz	HUF-Z3	Rohde & Schwarz
Horn Antenna	1 – 18 GHz	3115	Rohde & Schwarz
Logperiodic Antenna	1 – 18 GHz	HL 025	Rohde & Schwarz
Pulse Limiter	10 dB	ESH3-Z2	Rohde & Schwarz
Two-line V-network, 2 x 10(16) A, 50 ohm // 50 uH + 5 ohm	9 kHz – 30 MHz	ESH3-Z5	Rohde & Schwarz
Artificial Mains Network, 2 x 10(16) A, 50 ohm // 50 uH + 5 ohm	9 kHz – 30 MHz	NNLA 8119	Schwarzbeck
Artificial Mains Network, 4 x 16(25) A 50 ohm // 50 uH + 5 ohm	9 kHz – 30 MHz	NSLK 8126	Schwarzbeck
Artificial Main Network, 2 x 6 A Delta, 50 ohm		NNBM 8116	Schwarzbeck
Step Attenuator		HP 355C	Hewlett Packard
Step Attenuator		HP 355D	Hewlett Packard
Manual Step Attenuator	0 – 11 dB	HP8494B	Hewlett Packard
Manual Step Attenuator	0 – 110 dB	HP8496B	Hewlett Packard
Photometer Lumacolor		J17	Tektronix
Motorized turntable	4 m, 5988 kg	1081-4.03	EMCO
Antenna positioning mast	1 – 6 m	1050	EMCO
Positioning Controller		1050-84	EMCO
Positioning Controller		101991B	EMCO
Plotter		R9833	Advantest
Plotter		HP7440A	Hewlett Packard

Cavity antenna

The dimensions of the EUTs are limited by the stripline. For physically large EUTs it is necessary to expose the whole volume of the shielded enclosure to the electromagnetic field.

A cavity antenna placed on one of the walls will create an HF field, but because of reflections the homogeneity is poor. To compensate for this, sensors are placed close to the EUT. The output from the sensors regulates the power input to the antenna to maintain a constant field strength close to the EUT.

A fully anechoic room will give better possibilities to create a more homogeneous field over a large volume.

EMC measurements in situ

When the equipment is too large to be tested in the laboratory or test site, in situ testing has to be performed. This is often the case when testing large industrial and offshore installations or cable-TV networks.

Measuring equipment

In addition to the described OATS in Maridalen, Radiostøykontrollen has the following equipment:

Screened enclosures (not anechoic)

A
4 x 6 x 3 m DOOR 80 x 200 cm

B
2.4 x 3.8 x 2.4 m DOOR 80 x 200 cm

C
2.25 x 1.7 x 2.3 m DOOR 90 x 227 cm

D
A fully anechoic screened enclosure
8.96 x 6.08 x 4.56 m DOOR 250 x 250 cm
Absorbing material walls, ceiling and floor: Ferrite tiles
Turntable: Air floating, 4 m diameter
Load: 6,000 kg

Table 1 continued

Immunity measurements			
Signal Generator	10 kHz – 2.7 GHz	2031	Marconi
Signal Generator	10 kHz – 1000 MHz	2022	Marconi
Signal Generator	10 kHz – 1040 MHz	2019A	Marconi
Broadband Amplifier	125 kHz – 250 MHz 10 W	model 10L	Amplifier Research
Broadband Amplifier	0.01 MHz – 100 MHz 25 W	model 25A 100	Amplifier Research
Broadband Amplifier	1 MHz – 1000 10 W		Amplifier Research
Broadband Amplifier	10 kHz – 220 MHz 150 W		Amplifier Research
CISPR Stripline 80 cm	150 kHz – 150 MHz		
IFI-field Antenna EFG-3	10 kHz – 220 MHz		
Eaton biconical Antenna 96002C	20 MHz – 300 MHz		
AR log periodic Antenna AT-1000	150 MHz – 1000 MHz		
Vertical Antennas for HF, VHF and UHF AR Fieldprobe FP1000	10 kHz 1 GHz		
Cavity Antenna AT 2000	30 – 1000 MHz 3500 W		Amplifier Research
AR Isotropic field monitor model FM 1000	10 kHz – 1 GHz 300 V/m		
AR Levelling amplifier model 888	10 kHz GHz		
Audi Elektronik Spark generator		SS 4361503	
Absorbing Clamp	30 MHz – 1000 MHz	MDS-21	Rohde & Schwarz
Absorbing Clamp	30 MHz – 300 MHz	MDS	Rohde & Schwarz
Instruments for Industry E-field meter EFS-1 1 v/m-300 v/m	10 kHz – 200 MHz		
Broadband sampling voltmeter 1 mV – 3 V	10 kHz – 1200 MHz	3406A	Hewlett Packard
Sennheiser audio voltmeter		UPM-550	
Microphone system for audiomonitoring			
Hagener light intensity meter 52			
Function generators	10 kHz – 0.1 MHz		

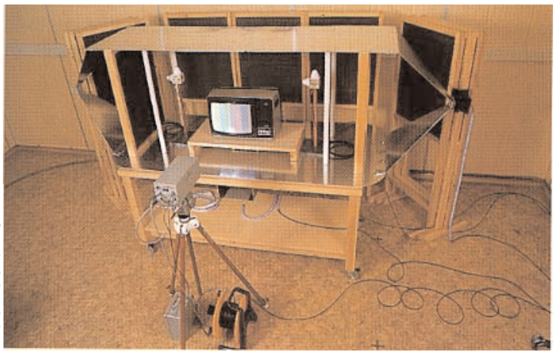


Photo: Telenor AS

Figure 9 Immunity test of TV receiver in the stripline

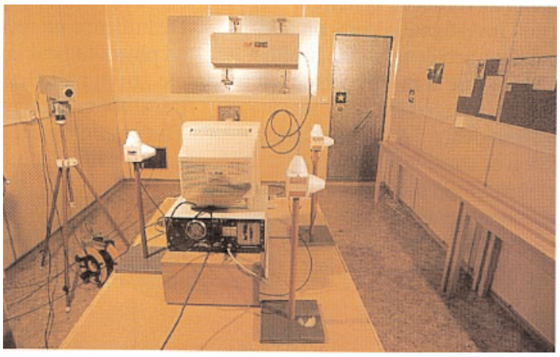


Photo: Telenor AS

Figure 10 Cavity antenna and field strength sensors in the screened enclosure

Electrical Medical Equipment Control and Vigilance in Norway

BY ERIK FØNSTELIEN

In the late sixties and the early seventies it became increasingly clear to biomedical engineers and safety technicians that hospitals and other health institutions represented an important area of hazards due to defective medical equipment and harms to patients by equipment users' errors. Worth mentioning is also some reports on adverse effects by the use of correctly functioning equipment (1). In 1970, Dr. A.F. Pacela of Beckman Instruments Inc. published a bibliography containing more than 300 articles on electrically induced accidents in hospitals (2). The cited articles, of course not being anywhere near the complete story, nevertheless describe thousands of single incidents, many with fatal outcomes. Until about 1985, device and component defects due to sub-standard construction and maintenance counted for some 80 % of all accidents reported (3,4). Today, the accelerating use of highly complicated, miniaturised, and processor-controlled equipment and the development of clinical practice and the steady movement of the limits of the possible, some 80 % of equipment-induced harms to patients, are ergonomic or users errors problems. Incidents with EMC are recorded and statistical analysis of the registered trends indicates that the problems will be increasingly important in the future.

The Norwegian Electrical Safety Directorate established the Special Inspection for Electrical Medical Equipment, STEM, in 1980. In 1992, STEM merged with the Electrical Safety Directorate as the Unit for Medical Devices, EMU. EMU has obligations and rights in accordance with given instructions based on law to:

- Do technical inspections on installations and equipment in all the hospitals once a year and in other health institutions when deemed necessary.
- Run a national vigilance system to receive and investigate reports on accidents and near-accidents in the health and social welfare area. All personnel categories are obliged to report incidents without delay, i.e. by telephone. EMU receives some 200 reports a year. 10 % are fatal incidents. It is estimated that non-reporters may represent commensurable numbers, primarily minor harms and near-accidents (5,6).
- Evaluate and discuss maintenance and repair and the proper management of equipment for diagnosis, treatment, and monitoring of patients.
- Monitor the development and attend to national and international standardisation work related to health and medical devices, notably both IEC and CENELEC working groups on EMC and medical devices.

EMU publishes a quarterly magazine on technical medical safety matters and tries, in various ways, to inform clinicians and hospital engineers on such topics within the limits set by financial and personnel resources.

Electromagnetic compatibility problems have always been a matter of concern by the use of electrical devices connected to the patient. Living tissue contains electrolytes and the geometry of the human body makes it susceptible to electrical and magnetic fields. Sensors, transducers, electrodes, and fixtures on and inside the body are leads to electric currents and liquid infusions are electrolytes as well. Some commonly used devices like defibrillators, pulse generators to restart a failing heart, and high frequency surgical diathermy, generate extremely high power densities, $>10^4$ W/cm^2, when in use, and tend to disturb everything in the neighbourhood. Most clinicians are aware of these unavoidable problems, and hospital procedures are accordingly modelled. Serious problems arise when electrical medical equipment is used in adverse environments and situations. Modern medicine tends to relocate more and more patient care and treatment outside the traditional institutions.

Typical reported EMC incidents are:

- Disturbance of neighbouring devices and navigation systems in ambulance aircraft on patient defibrillation.
- Disturbance on recorded and monitored physiological signals, notably ECG etc., from radio transmitters, power transformers, navigation systems, etc. The physiological signals are micro- and low milliVolt levelled.
- Portable telephones have become an important source of electromagnetic emission disturbing medical equipment. EMU has registered an increasing number of reports, at present 5–15 a year. The cause of the incidents is mainly that patients and visitors take their telephones into the hospital and thereby bring the transmitters close to vulnerable medical devices, the antennas sometimes touching electronic pumps, ECG-electrodes, etc. The equipment accidents are, with few exceptions, seen with infusion pumps and dialysis systems. Both equipment types regulate the flow or composition of potent fluids to the patient's body.
- Handheld phones are recorded as being able to inhibit fire alarms and surveillance systems. It has been observed that automatic clinical chemistry analysis machines have changed their readings when antennas are close to their circuitry.
- Electrical wheelchairs' controls are susceptible to fields from transmitters and similar electronic devices. The Radiostøykontrollen (the Radio Interference Service) has made both field and laboratory (unpublished) tests on some types of battery driven, electronically controlled wheelchairs – confirming that some types are susceptible to electromagnetic fields from, for example, cash registers and mobile telephones.

Many accidents are believed to be induced by conducted interference, but the facts and knowledge are scarce. Norwegian findings compare well with the ECRI, USA (6).

The investigation of incidents involving EMC are inevitably complicated, the reproducibility being very low due to missing information on the actual geometry and electronic systems. The Radiostøykontrollen has been most helpful to assist the EMU with competence and measurements, both in hospitals and in the laboratory to establish actual emission from mobile telephones and immunity of some widespread medical infusion pumps and a typical dialysis machine (7).

The Radiostøykontrollen has carried out tests on field emissions and conducted RF from diathermy devices to establish levels of signal noise both in the neighbourhood of the active electrode and on the power line (8). The levels and spectra registered indicate that diathermy is an important source of potential harm to almost any other type of instrumentation in hospitals, both battery operated and power line connected. These results and those earlier reported (7) are one of the background data sources for the development of the IEC 601-1-2 standard, the main standard on EMC and medical devices.

This article gives a short overview of the activities at EMU, the medical device unit at the Electrical Safety Directorate in Norway. The directorate has seen important co-operation with the Radiostøykontrollen, the EMC work at the EMU having been completely dependent on the laboratories and the competence at Radiostøykontrollen. The findings, incident reports and general information on medical device EMC are regularly published to keep Norwegian hospitals, clinical workers, and engineers continuously up-dated to enhance safety to patients.

References

1 Bruner, J M R. Hazards of electrical apparatus. *Anaesthesiology,* 28 (2), 396–425, 1967.

2 Pacela, A F. *Bibliography on electrical hazards, safety & standards in biomedical engineering.* Report CRR-21A. Beckman Inst. Inc. Fullerton C 92634.USA, 1970.

3 Fønstelien, E. Comments from a government. *Proceedings Workshop on the European standardisation of medical devices, Brussels, Dec. 1988,* Brussels, CEN, 317–322, 1989.

4 Fønstelien, E. Equipment failures, patient safety and governmental control. *Proceedings of the World congress on health technology standards, 1990,* Dublin, IEC/ISO, 509–512, 1991.

5 STEM-report (in Norwegian). *10 års erfaring med undersøkelse av uhell og feil på medisinsk teknisk utstyr i norske sykehus.* 1990.

6 Fønstelien, E, Nilsson, F (in Norwegian). Uhellsetterforskning ved Det særlige tilsyn med elektromedisinsk utstyr. *Tidsskrift for Norsk Lægeforening,* 109 (16), 1801–1803, 1989.

7 Fønstelien, E (in Norwegian). Feltimmunitet og mobiltelefoner. *Helstj. med. teknikk,* 9, 10–12, 1991.

8 Fønstelien, E (in Norwegian). Om nettbåren elektromagnetisk støy; med særlig referanse til forholdene under bruk av diatermi. *Helstj. med. teknikk,* 6, 24–26, 1992.

Company EMC organization

BY AGNAR GRØDAL

Introduction

New electronic products have to meet the EMC specification, and from 1996.01.01 the EMC Directive will apply. One should be prepared for meeting the EMC specification when writing the product specification and during the development phase, rather than hoping for the best during the EMC test phase. This as well as other EMC related activities will involve many departments and individuals in the company. To succeed, all these must co-operate.

The technical problems that must be solved are complicated, have many facets, and require competence in various mechanical and electronic design topics, often of high-frequency radiotechnical nature.

The investment in, and the use of EMC measuring instruments, also need special competence.

The field of EMC problems involves even legal aspects. It is necessary to know which EMC norms and other norms and regulations in the vast jungle of such related documents that apply to the actual product.

Documentation of the EMC activities is also a comprehensive, essential, and time consuming activity.

Solving the various EMC related problems as they evolve, and deciding who should be responsible in each and every case, will inevitably lead to frustrations and delays in the development process.

To succeed, it is therefore vital that the company has some kind of organization for the various EMC activities.

EMC responsible

To succeed with the various EMC activities, it is important that the company appoints one person responsible for all EMC activities. He should not necessarily be competent in the whole EMC field, nor should he himself perform all the relevant work.

He should, however, see to that this work is being done, and he should preferably also be EMC competent and perform central parts of the work.

An important part of his work should be to make the company aware of the EMC issue, as it most probably, at least to some degree, will be the concern of all employees.

The person in charge of the EMC activities should have a similar function as the person in charge of the quality activities. The EMC activities may in fact be defined as a part of the quality work, and the two job functions may, especially in smaller companies, be performed by the same person.

It is important that the EMC responsible person and his job function is backed by a mandate, to guarantee that the relevant EMC activities will be performed compulsorily for all products that are designed and manufactured by the company.

It is likewise important that all EMC activities are documented, and that the documents – especially the test documents – are filed and stored in a safe manner and over a time period as specified in the EMC Directive.

The work burden on the EMC responsible person may well result in a full-time job. He must thus be given adequate time and budget resources, as the job seldom can be performed as an extra job for an employee already occupied with other duties.

It is difficult to measure or to document in terms of money how profitable the work of the EMC responsible person will be. It is therefore essential that the company management has a correct understanding of the necessity of the EMC work, and that it supports it.

The person appointed as company EMC responsible may have the following duties:

- He should always be updated on EMC norms and regulations.
- He should know which literature that can be recommended.
- He should know where to find the best consultancy.
- He should know and be able to recommend the best test laboratory in each particular case.
- He should be responsible for the EMC training of the employees.
- He should gather and spread information on relevant EMC courses and seminars.
- He should check that the functional specification of a new product incorporates the correct EMC specification.
- He should check that the clients' EMC demands are included in the product specification.
- He should check that the choice of components and the hardware design principles have a fair chance of meeting the system EMC specification.
- He should check that the placement of components on the printed circuit boards and the layout of the signal and ground return tracks will have a fair chance of meeting the EMC specification.
- He should check that the mechanical design will give the product a fair chance of meeting the EMC specification.
- He should check that the EMC issues are discussed during all design reviews.
- He should check that all work performed by consultants has taken EMC into account.
- He should check that all purchased sub-units have the relevant EMC certificates.
- He should check that all EMC measures are adequate, but to the lowest possible cost.
- He should check that the product is tested according to the relevant EMC norms.
- He should be responsible for the cost analysis, evaluation of usefulness, and the purchase of company EMC test instruments.
- He should be responsible for the writing of the products' "suppliers declaration of conformity" document.
- He should be responsible for new EMC testing and for revision of the relevant EMC documentation when a product is modified.
- He should be responsible for the CE-marking of the products.
- He should be responsible for the archiving of all EMC documents for the products.

It is important to realize that issuing a "suppliers declaration of conformity" document, or CE-marking a product without having the relevant confirming documents is a criminal offence. The documents shall clearly show that the product meets the relevant specifications.

One should observe that an external EMC test laboratory normally does not issue any kind of "certificate", just a measuring protocol. The laboratory is therefore not responsible for issuing documents implying that the relevant specifications are met. The laboratory is, however, responsible for having performed the ordered tests correctly. The producer is responsible for the contents and the writing of the "suppliers declaration of conformity" document. He is also responsible for having performed the relevant tests.

CE-marking of a product implies not only that the EMC specification is met, but also that it meets other relevant Directives and norms. A typical example is that a product cannot be CE-marked without meeting the safety norm EN 50950.

The EMC documentation shall be at disposal on request from the relevant inspecting authority for ten years after the last sample of the product was manufactured.

EMC specification

A product's possibility of success in the market can be studied before the product is realized by reading its system specification. This document is, however, normally incomplete and erroneous.

The project manager and the EMC responsible person are responsible for checking that the EMC requirements are incorporated in the product specification before the start of the design phase. Neglecting this will inevitably have costly consequences.

The product specification has two main parts, the functional specification and the EMC specification. Both will affect the logical design and the types and number of components that will be used. The product specification will also incorporate the safety specification and the environmental specification relating to mechanical and climatic endurance such as vibration, bump, temperature and humidity.

Producer and client should agree upon the product specification before the start of the development process.

International standardization work is continually aiming to put different electronic equipment into different "product categories", each with a specific set of EMC requirements. When developing a new product one must first decide to which "product category" it belongs. Based on this decision the detailed requirements to Electromagnetic Interference (EMI) and to Electromagnetic Susceptibility (EMS) are given in the relevant norms. The aim should always be CE-marking of the product. This marking is mandatory from 1996.01.01. Some products on the market already have this marking.

Standardization of components

There has, over the last few years, been a radical increase in the number of new component types; this despite some "experts" believing that this evolution should have levelled out some years ago, as more and more logic hardware was supposed to be hidden in VLSI (Very Large Scale Integration) components.

The new components, both active and passive types, have changes in their internal technology as well as in their physical size.

Most new component types have better EMC qualities than the former similar types, but some have worse or even significantly worse EMC qualities. It is therefore important to choose component types with sufficiently good EMC qualities, but above all, to avoid the worst types.

The process of choosing the right component types has, however, the last few years tended to be more difficult than before. The reason for this is the increased and vast diversity of component types, and a seemingly decreasing interest in component technology among hardware designers.

A new problem is the decreasing possibility of getting components with equal function and quality from more than one vendor, the so-called "second source" policy. Each vendor prefers more and more often to give his particular versions of standard function components some specific attractive properties, which hopefully will increase the sale. These components may then not have second source components in the market. Even if components have functional second sources in the market, their EMC properties may be different. This can cause products that meet the EMC requirements to fail the same requirements if functional second source components are used.

Components with superior EMC properties are often higher priced; their prices are often unreasonable for use in low cost products. Their availability may also be a problem, and the lead times – the time between ordering and delivery – can often be counted in months rather than days or weeks.

The standardization of components must always take their EMC properties into account. The EMC responsible person may in these matters often be more competent than the local component responsible person.

Design review

Design review of a circuit diagram and its accompanying parts list has shown to be a very efficient means to disclose design errors, both functional errors and economic ones.

Design review is also very efficient for securing design principles, choice of components, placement and orientation of components on the circuit board, and design of the printed circuits so that the generation of radiation will be sufficiently low, and so that the product will be sufficiently immune to incoming radiation and overvoltages – all at the lowest possible cost.

Design review should be performed using documented guidelines, where EMC should be among the main topics. Design review is a part of the quality standard EN 29001, and the review procedure is documented in the standard IEC 1160, Formal Design Review.

All product designs, whether they are made by company employees or by an external consulting company, should undergo the same review procedure. The procedure should also cover the review of purchased OEM products, Original Equipment Manufacturer products, subunits that will be incorporated in the company's own product.

Succeeding with EMC during the design phase requires that the company has an appointed and competent design review group of which the EMC responsible person is a member.

It is very important that the leader of the review group has sufficient mandate, so that the design review process can be performed for absolutely all designs and revisions of designs, and that this work will be documented.

Printed circuit board

The printed circuit boards, their types and their number of layers are normally discussed for the first time late in the design phase, often after the logical diagrams are designed. However, as a rule, it is better to consider the number of layers as early as possible when writing the product specification. This because an unnecessary high number of layers will be costly, and because one may never know if a lower number of layers would have been sufficient.

High EMC requirements will as a rule require multilayer boards. Signal track routing on densely populated boards will also as a rule require multilayer boards.

Design of the signal tracks and of the ground return current system without considering EMC will as a rule result in boards that have a limited chance of meeting the EMC requirements. One may in that case be forced to revise the board design one or more times until tests eventually and hopefully will show that the EMC requirements are met.

It is quite essential for an acceptable result that the printed board designer is trained in EMC board design. It is also important that he co-operates closely with the designer of the circuit diagram. He should also be in close contact with manufacturers of printed circuit boards, and be able to evaluate their quality profile, their punctuality, and the total price of a printed circuit board, dependent on its number of layers and the pricing policy of the producer.

The PCB designer must continually have competence in evaluating the cost effectivity of the used CAD system, and be able to evaluate if a two-layer solution can be used without getting too serious EMC problems, and also be able to evaluate if autorouting of the board tracks can be used. One should observe, however, that most autorouting systems of today result in a rather poor EMC performance.

Well screened circuit boards having moderate EMC requirements may be based on a two-layer solution. High volume production series, where a higher cost printed circuit board design can be acceptable, may also be based on a two-layer solution.

Recent practice based on EMC knowledge has shown that some complicated boards with high EMC requirements may be based on only two layers, or even on just one layer.

Soldering process

The change from hole mounting technology to surface mounting technology has resulted in smaller board areas, and thus to lower emissions. This change will normally require investment in new soldering machines, changing from single wave machines to double wave machines and to infrared soldering machines.

Newer component types with shorter distance between terminals, circuit tracks with shorter insulating distance, and the mounting of components on both board sides may require even newer soldering principles and machines, which again may require new investments.

This escalation of investments in production machinery partly caused by the EMC requirements, may force smaller companies to let the mounting of components and the soldering process to be performed by external companies specialized in production.

Instruments

Companies developing and producing electronic products, no matter how small, should invest in EMC measurement instruments, both for measuring emission and for testing the products' immunity to overvoltages and static electricity. The type and number of instruments may vary, depending on the need and the possibility to borrow or rent.

A useful first set of instruments may consist of a spectrum analyzer including test probes, antennas, plus an electrostatic discharge (ESD) pistol. The set may cost about GBP 20,000.

A miniature test chamber, a TEM-cell (Transversal ElectroMagnetic cell), may also prove helpful. The smallest chambers are priced at about GBP 10,000.

EMC measurements based on the above mentioned instrumentation will give indicating measurements only, but will prove very helpful for checking EMC modifications on prototypes. Correct use of the ESD pistol may give absolute-value measurements, as done in a test laboratory.

To achieve absolute-value emission and susceptibility measurements, a reflection-free test room must be used. This is affordable for the largest companies only, but reasonably acceptable results can be obtained in a sufficiently large empty cellar room or other less valuable space.

Investments in test equipment may seem expensive, but one should not forget the necessary additional investments in testing competence and manpower.

Test laboratory

The writing of a "supplier's declaration of conformity" and the CE-marking demands that the product has been tested on an open area test site (OATS) as defined in EN 55022, or in a large anechoic test chamber with properly controlled and acceptable unwanted reflections and resonances.

The EMC responsible person should be well informed of all domestic and some foreign test laboratories, their strengths and weaknesses, their accreditation status, their accessibility and their pricing policy.

It is important that the EMC responsible person is well informed of the most used test norms, and of their practical execution. He should also be present during the testing to gather experience for future EMC work and decisions.

Domestic EMC competence centres

The EMC responsible person should be well informed of all domestic competence centres such as:

- Testing laboratories
- EMC consulting companies
- Publishers of normative documents
- Sales offices of domestic and foreign normative documents.

Consultants

Succeeding with EMC and other relevant disciplines necessary to develop a new electronic product requires a very broad competence, not only in technical matters, but also in a complicated and constantly growing jungle of norms and regulations.

Few small or medium sized companies can afford or hope to have internal competence to handle all these questions. A way out of the problems is to use specialized external EMC consultants. This may, however, sometimes give rise to other problems, as many companies have decided to solve all problems with the use of internal workforce only. They have usually no budget for the use of external consultants, and see this as extra unnecessary expenses.

This way of thinking may indeed prove to be expensive.

The EMC responsible person should be well informed of the consultants market, and be able to recommend a specific consultant for solving a specific EMC problem.

Approval of products

The "approval" of a product may have different definitions. "Approval" may sometimes not be related to EMC, and sometimes the word is even used to deceive potential customers. Foreign customers may have their own definition, sometimes more severe and sometimes less severe than the EMC Directive.

Equipment connected to the Public Telecommunication Network must have "Type Approval" which in Norway is issued by the Norwegian Telecommunications Regulatory Authority. Some products are "subject to compulsory certification," and some "subject to compulsory registration." Certifications and registrations are in Norway issued by the Norwegian Electricity Inspectorate. All type approved, certified, or registered products must comply with the EMC Directive for both emission and susceptibility. Companies producing equipment that has to comply with the Directive on Telecommunications Terminal Equipment (91/263/EEC), must also comply with the quality norm ISO 9001.

The EMC responsible person must know before starting the design phase which approval will be relevant for the new product, as this information may influence the design of the product.

Compliance with the EMC Directive may be documented by a "supplier's declaration of conformity" document, or by having the importer or manufacturer to generate a "Technical Construction File".

The "supplier's declaration of conformity" route is expected to be the most common method. The tests may in this case be performed by the producer, provided he have the facilities and technical competence to do the job. As only the largest companies can afford to have acceptable testing facilities, the testing will normally be performed by an external independent test house. The test house issues test result protocols, and based upon these documents the producer will issue a "supplier's declaration of conformity" document, providing the test protocols show that the relevant EMC requirements are met.

The "Technical Construction File" route specifies a description of the product, a description of the procedures used to warrant the relevant safety requirements, plus a technical report or certificate issued by a "Competent Body" stating that the relevant EMC requirements are met.

The use of a "Competent Body" may imply that no tests will be performed.

Both the "supplier's declaration of conformity" and the "Technical Construction File" route qualifies for the use of the CE mark.

The EMC responsible person will be responsible for the "supplier's declaration of conformity", or the administering of the "Technical Construction File" route, and of the CE-marking.

Competence and co-operation

It may seem as meeting the EMC Directive, both technically and administratively, may be complicated. Indeed it is, and it requires a multitude of competence and duties on behalf of the EMC responsible person, and of the other involved employees. The key word to success is competence, and above all, close co-operation between

- The client
- The writer of the technical and the EMC specification
- The EMC responsible person
- The quality manager
- The designer of the circuit diagram
- The component responsible person
- The circuit board designer
- The designer of the mechanics.

EMC is the concern of all employees.

The Radio Interference Service in Norway – a historic review

BY SVERRE TANNUM

The beginning of EMC was radio noise and in most countries this turned out to be a problem when broadcasting started around 1925. Radio noise and interference had to be defeated and on September 25, 1925, the Norwegian Telegraph Administration addressed a letter to the Department of Trade concerning a draft of international rules for interference free broadcasting.

The act (“Lov om tilsyn med elektriske anlegg”) from 1929 was modified preliminary for a period of 3 years ending in 1936. The proposal of amendment of the regulations was put into effect some time later.

Radiostøykontrollen (the Radio Interference Service) was formally established in 1939, and the aim was:

1 As far as possible to reduce the radio noise from electrical machinery and apparatus before marketing

2 To localize and suppress radio noise from existing equipment

3 To contribute to the best possible immunity of the receiving installations.

Radiostøykontrollen was organised under the Norwegian Telecommunication Authorities (Telegrafstyret).

The service was established in 6 control districts corresponding to the organization of the Norwegian Electrical Safety Directorate who is acting as the legal authority when injunctions have to be executed.

The early years of broadcasting

Despite the problems and the special situation during the second World War the organization of Radiostøykontrollen was continued and the first 10 persons were engaged 1940 – 1941. After the confiscation of radio receivers during the occupation, the engineers were employed on transmitter stations and laboratories.

In 1945, bicycles and instruments were put into service and the German supplies were searched for radio noise suppressor material, because the noise sources were numerous and strong after 5 years of war and insufficient maintenance.

Transportation was difficult and public communications poor, so the employees of Radiostøykontrollen were among the very first to get licences for buying cars. For transportation in the coastal districts a 34 feet boat was purchased after the war. The main purpose, apart from radio noise suppression, was to investigate the coverage of long- and medium wave transmitters.

There was a great need for assistance in the remote areas and the coverage from the Norwegian AM transmitters was often insufficient and the interference from more powerful transmitters in the rest of Europe was a problem. The boat was frequently used during the summer and the crew was made up of different radio noise inspectors. One of them, though, Mr. Kåre Kristensen from Tromsø, was a skipper, engineer, navigator, deckhand and cook, all in one person.

The instruments for radio noise localization were very simple in those years, mostly portable radios were used. The noise sources were mainly rotating machines, a lot of local DC power plants of different types. The distribution networks were worn and repaired by all available means – copper and iron in a mixture. Generators from 12 to 2 x 220 volts had not been maintained for many years. A lot of these installations were dangerous and some had to be closed down.

The municipal distribution of electricity developed, the maintenance was better and in the early 1970s this problem disappeared.

Some districts, like Sogndal and Åndalsnes, had got equipment for transmitting radio signals over the power distribution network. Problems occurred with 50 Hz interference caused by modulation of power frequency. The reason was usually corrosion which was difficult to locate, and this problem existed as long as the transmitters were operating.

The high voltage overhead power lines also created radio disturbance. It could be permanent noise from insulators which might be difficult to get rid of because of the high cost of rebuilding. Tracking down such noise sources with the equipment available at that time was very time consuming. However, surprisingly often it was a success. The managers of the local power networks were often impressed, but the ones who did not take account of it had unexpected breakdowns which were much more expensive than a programmed cut-out.

After the introduction of FM transmitters, the radio noise searching equipment became more advanced and localization of faults on overhead power lines was easier. The power lines were also rebuilt and the permanent noise disappeared after a while.

Interference from transmitters

Interference from citizens band (walkie-talkie) and illegal pirate FM transmitters became a problem. The radio amateurs usually operate within their specified power- and frequency limits but nevertheless, they can produce interference and noise because of unwanted detection of their signals in stereo sets and low frequency amplifiers even when they are turned off. This problem occurred when radio tubes were replaced by transistors, and even signals from far away broadcasting stations could be picked up by long loudspeaker cables and detected.

Protecting broadcasting reception from the signals from neighbouring radio amateurs can be very difficult and often requires good co-operation from all parts: the amateur, the listeners and the radio noise inspector.

Even broadcasting transmitters can be a problem for radio reception. Particularly the high power transmitters on the island of Kvitsøy off Stavanger created problems for the local population. Medium wave (1200 kW) and short wave (1000 kW) transmitters interfered with TV and radio reception, video tape recorders (VCR), video cameras and telephones. Radiostøykontrollen was engaged from 1979 to 1983 to solve the interference problems and the costs were estimated at about NOK 270,000.

The electric tramway

One of the most serious sources of radio noise in the first period of broadcasting was electric tramways. Broadcasting at that time was performed in the long- and medium wave bands and electric tramways and railways was a serious problem in Norway, as in other countries.

The tramway is normally driven by DC and the radio noise from the rectifiers was transmitted through the wire along the track and also generated in the motors and regulators. The most dominating reason for radio noise was the contact device (current collector) feeding the current to the machinery.

The generated radio noise is very much dependent on the magnitude of the DC current which is occasionally interrupted at the contact point. If the current is below the critical value, it generates sparks causing severe radio noise. The interruption of the DC current above the critical value generates an arch which does not disturb to the same extent. The critical value of the current is dependent upon the voltage and the material of the contact surface.

The noise problems were most severe with a moving car when the motors were not working. During the cold season the noise was reduced because of extra power consumption for heating the tram cars.

Reducing the noise was tried by increasing the current with the help of extra load resistors and by installing capacitors across the main circuit and by increasing the contact pressure. A certain reduction was accomplished, but the best results were achieved by use of broad and often double contact pieces. Different materials were tested and carbon proved to give the best result. The so-called Fischer phantograph and extra load resistors introduced a great improvement to the radio noise problem, but resulted in increased costs which had to be shared by the broadcasting service and the tramway company.

After the war, in the years 1947 – 1950, radio noise from the trolley buses in the city of Drammen caused complaints, and the solution was the same as in the 1930s: extra load resistors ensuring the current drain to always exceed 2 – 3 ampères.

The problem with radio noise from the tramway disappeared after 1955 when the FM network was introduced.

Radio noise from high voltage overhead power lines

The valuation cases of the 1960s

Norway's consumption of electrical energy is high, and we have a comprehensive distribution network. Long- and medium wave reception is often disturbed by radio noise from the high voltage overhead power lines. Especially the corona discharge in humid weather is the source of serious radio noise when receiving AM (long and medium wave) while FM reception is not disturbed by this type of radio noise.

Establishment of new lines caused a lot of valuation cases where Radiostøykontrollen acted as technical expertise in courts of law.

People subject to disturbance were offered a so-called FM module mounted on the radio which made FM reception possible, or alternatively a compensation of NOK 350. Measurements were made and reports were given by Mr. Falk-Pedersen to elucidate the problem.

It was proved that listeners closer than 100 – 200 m to the lines would have a degradation in their AM reception during dry weather and almost no possible reception during humid weather. This was also demonstrated by tape recordings in the Court of Law.

It was concluded that the degree of disturbance was dependent upon the type of program transmitted. Speech and entertainment music required a signal to noise ratio of 35 – 40 dB, while 20 dB was considered sufficient for pop music!

The Supreme Court decided, however, that because of considerable investments and intensive extension of the FM network, no compensation should be given for poor long- and medium wave reception. A lot of people were after all given FM modules by NVE (the State Electricity Board).

Instrumentation

The quality of the radio noise service depends very much on how experienced the personnel is. Normally it takes 2 – 3 years of practice before full competence is reached.


Photo: Telenor AS

Figure 1 Trolley bus

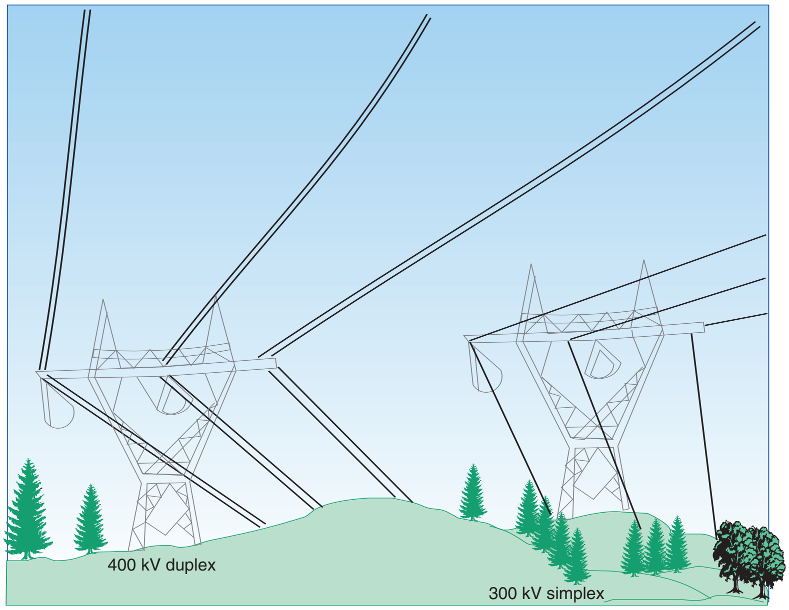


Figure 2


Photo: Telenor AS

Figure 3 Marconi radio noise instrument

Photo: Telenor AS

Figure 4 The T.K. noise instrument

Photo: Telenor AS

Figure 5 Kurér portable radio

Photo: Telenor AS

Figure 6 HUZ field strength meter

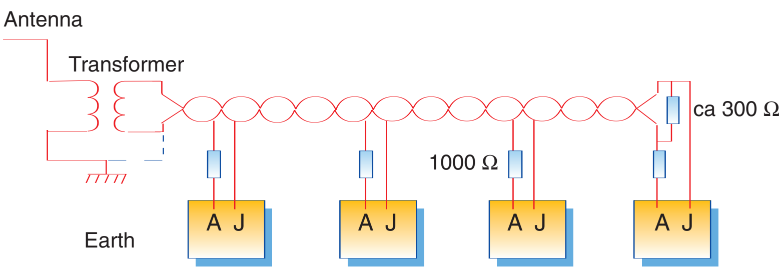


Figure 7 Diagram included in the guidelines for communal aerial systems

But good and proper instrumentation is also very important. Today, we have a selection of suitable instruments to choose from, which is a situation entirely different from the early years.

A report from 1938 in Hammerfest gives the following description of what was available in the district:

1 Marconi radio noise instrument (LW + MW)

2 Siemens radio noise instruments

1 Tandberg radio noise instrument.

These pieces of equipment were heavy and impractical. A modernised type called T.K. (Thor Kaltvedt) was used from 1940. The instruments were portable receivers, more or less modified.

From around 1950 the Kurér portable radio was used supplied with universal instrument and megger. The David Andersen portable receiver also had an instrument for signal strength indication.

The FM transmissions in the VHF band were started around 1955. This introduced a new technique because propagation of both wanted and unwanted signals was different.

The HUZ field strength meter from Rohde & Schwarz covered the frequency range from 47 MHz to 225 MHz (Band I, II and III). This instrument had the possibility to measure spark generated noise and was equipped with a special noise detector (CISPR detector) in addition to standard mean value detector.

The instrument had built-in calibration signal and a tunable dipole antenna mounted on a telescopic tube and it weighed 4 kg. The HUZ generation lasted for 20 years and was succeeded by an Italian field strength meter, PRESTEL, used from 1975 until a few years ago, but now succeeded by another Italian instrument, UNAOHM.

For the lower frequency range (150 kHz – 3 MHz) the Siemens STTM was used from 1960, fitted with an optional ferite antenna and the possibility to measure conducted radio noise via a probe or a line Impedance Stabilizing Network (LISN).

It is difficult to find a replacement for this instrument and it is still in use. 30 years is a very ripe age for a mobile instrument and the reactions from today's technical environment are not always positive when veteran equipment like this is exposed.

The introduction of television in Norway (around 1958) meant a challenge to Radiostøykontrollen concerning signal measurements. The signal strength alone was not sufficient for quality judgement; reflections could damage the picture completely.

Small portable TV measuring receivers for combined battery and 220 V supply were used for the purpose.

Colour TV and Text-TV have further increased the demand for better measuring equipment, and during the 1980s portable colour measuring receivers, oscilloscopes and frequency analysers were supplied.

Special equipment for control of cable-TV networks (CATV), ultra sound equipment for localizing radio noise from overhead power lines, and equipment for registration of magnetic and electric fields have been introduced.

The EMC laboratory is also well equipped with measuring facilities, such as instruments screened rooms and open area measuring sites.

CATV (Cable-TV)

In 1937 the first description of a network for the supply of antenna signal to concentrated flats was given by Christian Nome.

Before the TV age, in February 1956, guidelines for communal aerial systems (fellesantenneanlegg) were edited.

These guidelines were based on research done by Radiostøykontrollen and contained the diagram shown in Figure 7.

The transmission of TV signals along the Swedish border, especially in the south-eastern part of Norway, started an activity in this sector in the 1960s. A document dated 26.08.1965 describes 6 networks around Oslo containing 20,000 subscribers.

It was evident that there was a need for guidelines and regulations to avoid interference and to ensure that the subscribers were given a decent signal quality and good reliability.

The technical specifications based upon recommendations from IEC are given by

Radiostøykontrollen in 1974 concerning signal levels, frequency allocations, attenuation between antenna outlets, amplitude response, frequency stability and intermodulation plus description of measuring methods.

In 1975 these specifications were transferred into Norwegian norms (No. NEN 59.75). However, they were not obligatory until 25.10.1980, when Norwegian Telecom edited regulations for cable-TV. At that time, Norwegian Telecom was the legal authority and the control was executed by Radiostøykontrollen.

In 1983, a total of 385,000 subscribers were connected to CATV in Norway and 142,000 had more than 3 channels. Two networks had more than 10,000 subscribers and 15 networks had between 3,000 and 10,000 subscribers.

In 1984 preliminary regulations were given, and the government confirmed the technical regulations as being permanent the same year.

The control today is performed by Radiostøykontrollen on behalf of The Norwegian Telecommunications Regulatory Authority, which is legally in charge.

Satellite reception

The large number of satellites now operating have extended the offer of programmes considerably. This has led to rebuilding and extension of the cable-TV networks.

In 1989 CATV networks had a capacity of 24 channels – the majority being satellite programmes, because local programme production is expensive.

The CATV can be used as a general broadband network for transmission of other services like alarm, monitoring, digital communication, etc.

The use of the network is mainly a political question to be decided in the future.

Approval test

An important contribution to the reduction of radio noise is testing of consumer equipment to verify if the requirements are fulfilled. The regulations on radio receivers based upon IEC/CEE norms were put into force on January 1, 1954.

The principal idea is that electrical equipment shall not disturb broadcasting reception (this is also valid for broadcasting receivers themselves).

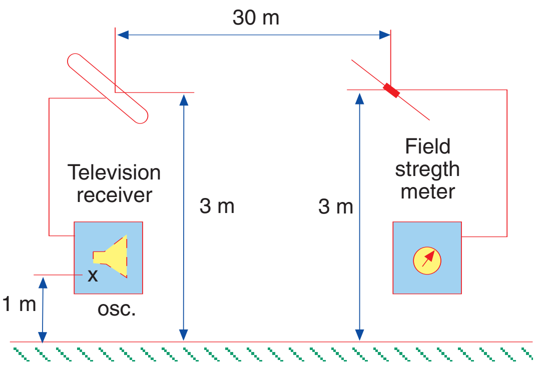


Figure 8

The measurements of conducted and radiated radio noise were performed by Radiostøykontrollen and the formal approval was the responsibility of NEMKO. Radiated noise was measured at a distance of 30 m (see Figure 8).

The national regulations were modified in 1983 in accordance with international norms which included measurements of conducted radio noise via a special (delta) network, and the radiated signals were measured at a distance of 3 m.

In the beginning, radio and TV receivers were the most numerous test objects, but in later years different types of other equipment are included, like video recorders, TV games, electromedical equipment, satellite receivers and also some telecommunication equipment and personal computers (PCs).

Since 1964, Radiostøykontrollen has also acted as a consultant engaged by the traffic authorities for type approval of vehicles and for road controls. International standardization has, however, reduced the problem with radio noise from cars and lorries to a minimum.

International collaboration on radio noise reduction

Since radio transmissions were started in Norway in 1923 it has become evident that radio noise from all kinds of electrical equipment has to be fought internationally. Exact limiting values have to be defined for all kinds of apparatus like household appliances, lighting equipment, broadcasting receivers, medical apparatus, computers, etc.

National differences in the regulations could lead to barriers of trade, so from this point of view international co-operation is also wanted. It started with a conference in Paris 1933 and this led to the foundation of the international radio noise committee CISPR. The first meeting was held in Paris in 1934 by IEC with participants from the different national committees and representatives from some international organizations.

There were only European members in CISPR before the second World War, and the most active countries were Great Britain, France, Belgium, Germany and the Netherlands. The CISPR work was resumed after the second World War in 1946. At a meeting in London it was decided that CISPR should be a committee under IEC with participation from other international organizations. The following organizations are now members of CISPR: The different national committees (NEK in Norway) and EBU, OIRT, ECMA, CIGR, UNIPEDE, UIC, UITP, UIE and IARU. There is also a close co-operation with CCIR and ICAO.

The work of CISPR is to establish norms concerning radio noise (EMC), defining measuring methods and limits.

Norway participates in CISPR through NEK (Norsk Elektroteknisk Komité). Radiostøykontrollen together with NEMKO have played an active role.

The Norwegian EMC norms are worked out to the best possible compliance with the CISPR publications.

The European Union has increased the requirement for co-ordination of norms and regulations. The European Committee on Electrotechnical Standardization

(CENELEC) has established a set of norms on radio noise/EMC (EN 55xxx) which are based upon CISPR publications.

Abbreviations

CCIR	International Radio Consultative Committee
CENELEC	European Committee for Electrotechnical Standardization
CATV	Cable Television
CIGR	International Conference on Large Electric Systems
CISPR	International Special Committee on Radio Interference
EBU	European Broadcasting Union
ECMA	European Computer Manufacturers Association
EMC	Electromagnetic Compatibility
IARU	International Amateur Radio Union
ICAO	International Civil Aviation Organization
IEC	International Electrotechnical Commission
NEK	Norsk Elektroteknisk Komité
OIRT	International Radio and Television Organization
UIC	International Union of Railways
UIE	International Union for Electroheat
UITP	International Union of Public Transport
UNIPEDE	International Union of Producers and Distributors of Electrical Power

Literature

Isaksen S et al. *50 år mot radiostøy. Radiostøykontrollen 1939 – 1989.* Teledirektoratet, Fagenhet for radiostøy, Oslo, September 1989.

Special

Travelling wave antennas

BY KNUT N STOKKE

Introduction

A travelling wave antenna is normally a single conductor line over the ground. In order to get the best efficiency, the line should be terminated to the ground by its characteristic impedance.

The travelling wave antenna effect

The most used antennas are based on standing waves. One definition of antennas is that *a standing wave is an antenna* [1, page 168–175]. This means that wherever there are standing waves (because of mismatch in cables, lines, etc.), we have radiation of electromagnetic energy, and reciprocally a mismatched cable or line will receive energy from an electromagnetic field.

The travelling wave antenna is based on another principles. As shown in Figure 1, a forward travelling wave is initiated because the antenna is terminated with a resistive load. As already mentioned, the most effective antenna is the one which is terminated with its characteristic impedance. If there is any mismatch, as often experienced in practice, this will create standing waves which may interfere with the travelling wave effect.

In Figure 1 is indicated the conditions when a positive and a negative charge move along a conductor. As the charges move along the conductor, we get an electric field as indicated by only two field lines. But at the same time a charge creates a magnetic field which moves out from the conductor with about the same velocity as the velocity of the charge along the conductor (dependent on the electric and the magnetic conditions in the surroundings). The resulting movement of the magnetic field is then in 45° out from the conductor, and a part of this magnetic field may be combined with a part of the electric field and cause radiation. Poyntings vector is at right angle to both the electric and the magnetic field where the field lines cross each other. For short conductor lines the travelling wave effect is rather weak, but when the antenna wire is several wavelengths, we may have some gain relative to a short monopole over ideal conducting plane. In fact, the travelling wave antenna is one of the few antenna types which may have some gain at very low frequencies.

The antenna pattern for a travelling wave antenna may be calculated by using the assumption that the antenna is an end-fire array of collinear Hertzian dipoles coupled in series. The equation for the antenna pattern will then be [2, page 329]

$$E = \frac{60 \cdot I_o}{r} \cdot \frac{\sin\theta}{1-\cos\theta} \cdot \sin\left(\frac{\pi l}{\lambda}(1-\cos\theta)\right) \qquad (1)$$

where I_o is the r.m.s. value of the current along the antenna wire, r is the distance to the measuring site, θ is the angle to the antenna wire, l is the length of the antenna wire, and λ is the wavelength.

Concerning the antenna pattern, we may consider the current along the wire to be constant.

The antenna pattern is measured at constant distance, and therefore the first part of (1) may also be considered to be constant, for instance 1. The equation for the relative antenna pattern may therefore be written as

$$E_r = \frac{\sin\theta}{1-\cos\theta} \cdot \sin\left(\frac{\pi l}{\lambda}(1-\cos\theta)\right) \qquad (2)$$

In Figure 2 are shown three examples for antenna lengths $l = \lambda/2$, $l = 3\lambda$ and $l = 8\lambda$. We have to be aware of the fact that the patterns are calculated for free space conditions, and they are therefore rotation diagrams around the antenna wire. In practice, where the antenna conductor is relatively near the ground with a conductivity and a permittivity, the losses and the ground reflections will cause changes in the patterns.

The gain of a travelling wave antenna with a length around 3λ is about 3.8 dB, and for an 8λ antenna the gain is about 6.8 dB; that is, gain relative to a short monopole over a perfectly conducting plane. With ground losses near the antenna wire the gain will be reduced, and such antennas are very much dependent on the ground characteristics. However, for the lowest frequencies the ground losses are small because of large penetration depths.

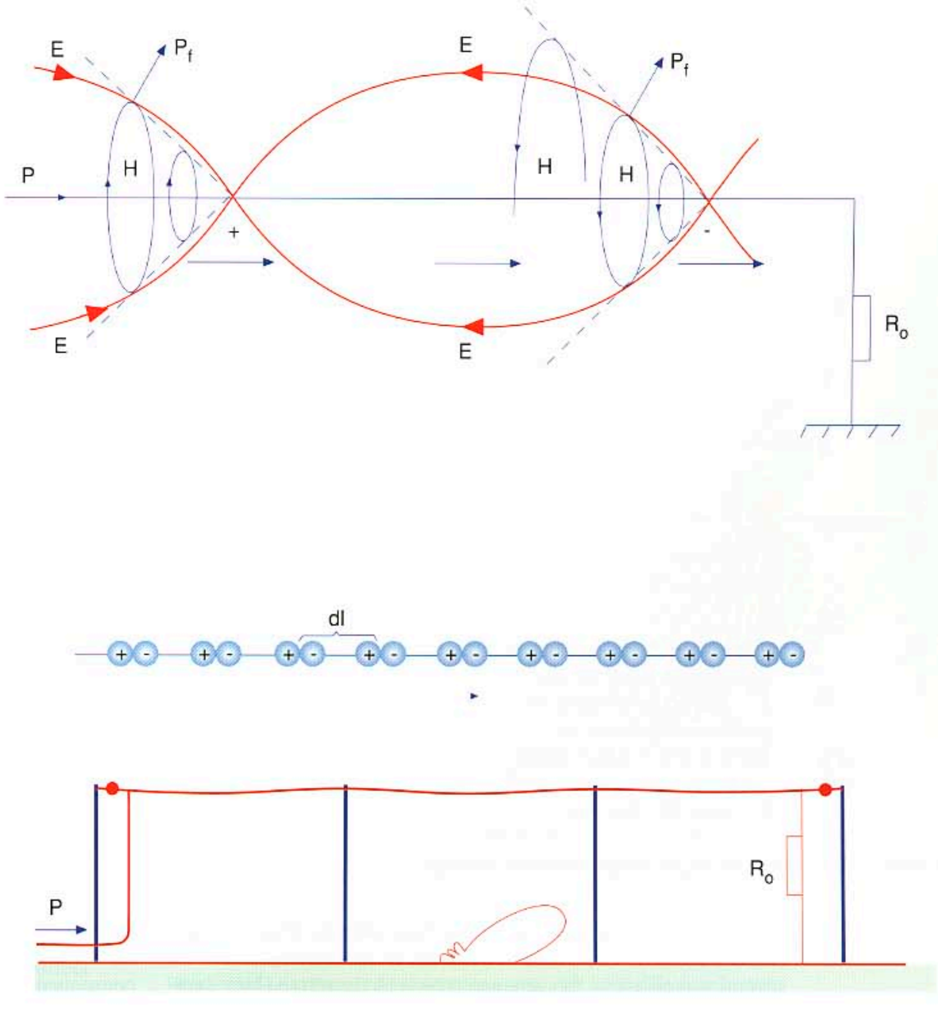
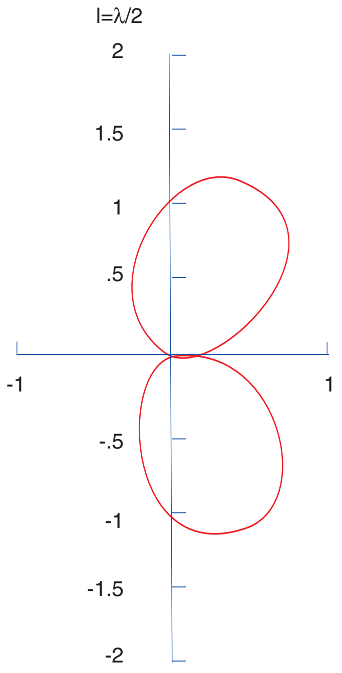
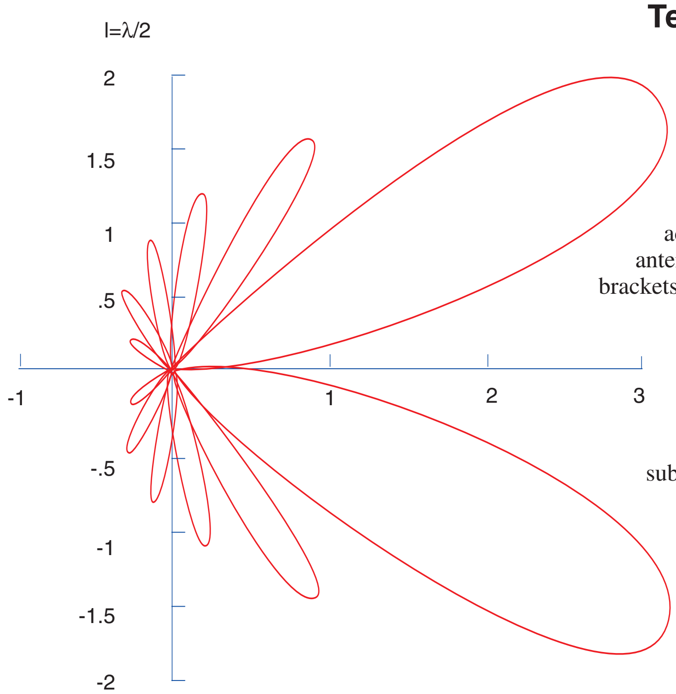


Figure 1 Travelling wave antennas





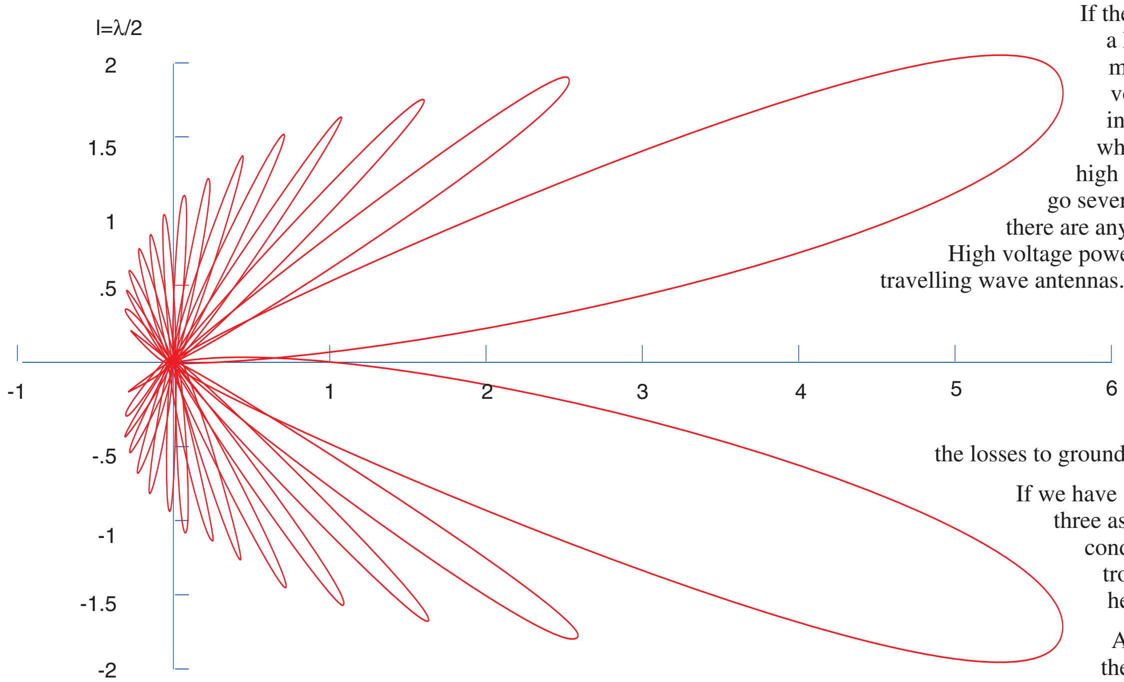


Figure 2 Antenna patterns for travelling wave antennas

We have a special type of travelling wave antenna when we use two conductor lines formed as a rhomb. If the conductors are several wavelengths, this antenna type is characterized by a very narrow main lobe in the antenna pattern.

Telephone cables and lines as travelling wave antennas

For example a telephone cable along a valley may act as a travelling wave antenna. The cable is hung on brackets of metal as sketched in Figure 3, and the capacitive coupling to the brackets and masts causes a leakage of energy to the ground. In addition, there are often several taps to subscribers along the cable, and this also reduces the efficiency of the cable as a travelling wave antenna. However, telephone cables are travelling wave antennas, and when receiving at very low frequencies, this effect may cause serious problems especially for frequency division multiplex (FDM) systems.

Power lines as travelling wave antennas

If the telephone cable is parallel to a high voltage power line, we may get another effect. The high voltage lines have very good insulators, that is, insulators which are very effective also for high frequencies. These lines often go several tenths of kilometres before there are any taps (to transformers, etc.). High voltage power lines are therefore very good travelling wave antennas.

The termination of such an antenna may be in a transformer, but we may also get a travelling wave effect because of the losses to ground along the line.

If we have several conductors, for instance three as shown in Figure 3, the three conductors get about the same electromagnetic energy, and may here be considered as one line.

As also indicated in Figure 3, the current in the power line because of the received electromagnetic field, will give a magnetic field which may induce a current in the telephone cable. We also have a capacitive coupling, and when there is a mismatch and consequently standing waves, we may have electromagnetic radiation from the power line to the telephone cable.

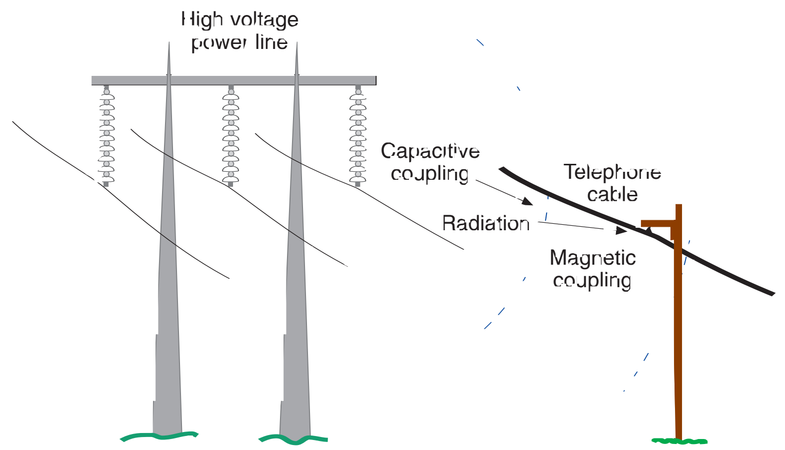


Figure 3 Telephone cable parallel to a high voltage power line

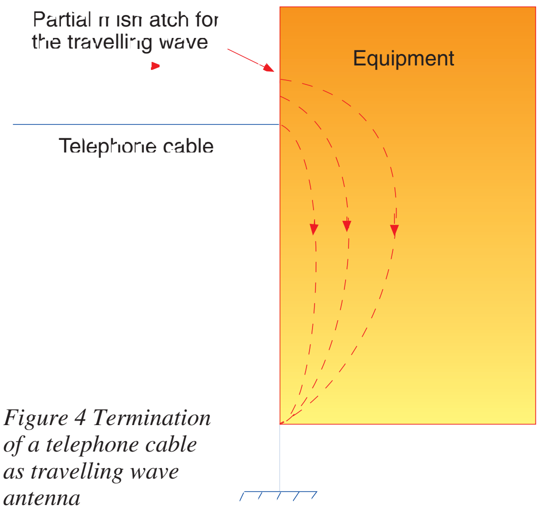


Figure 4 Termination of a telephone cable as travelling wave antenna

When underground telephone cable is used, the problem of radiation into the cable is highly reduced. However, the magnetic coupling is still there because ground normally gives very weak magnetic shielding.

Mismatch for electromagnetic waves

It is now of interest to take a look at the problems concerning mismatch for electromagnetic waves. It is very important to be aware of the difference in matching for a paired line in a telephone cable and the matching for the telephone cable considered as a travelling wave antenna.

The matching conditions for paired cables are well known. But concerning the travelling wave antenna, the matching conditions will be more complicated.

In Figure 4 are indicated the conditions when a shielded telephone cable is terminated at for instance a telephone exchange. The current caused by the travelling wave antenna effect goes via the equipment shielding and to the ground. The current may also go through the electronic devices and thereby cause problems in the equipment.

However, in practice there are seldom good matching for a travelling wave at the telephone exchange. The matching will be dependent on the dimensions of the equipment, the earthing conditions, etc. We therefore get a standing wave on the telephone cable, and this standing wave may cause severe interference problems especially for FDM systems.

Also, when there is no shielding in the cable, we will have the same problems. The problems may even be worse because the current then must go through the electronic devices and to the ground.

If, as sketched in Figure 5A, we have a mismatched coaxial cable with an input signal *P* we get standing waves in the cable. Inside the outer conductor we then have potential differences along the cable. The + and - signs indicate the direction for the potential differences. The result is that these potential differences will go through the conductor to the outer side of the outer conductor and cause a standing wave there. This effect depends on the thickness of the outer conductor.

If we have a standing wave on the outside of a properly terminated coaxial cable as indicated in Figure 5B, we get the reciprocal effect. The standing wave will produce a current in the coaxial cable even if it is properly terminated. Here we must take into consideration that the coaxial cable is an unbalanced device. If in addition we have a mismatch, the phase conditions will cause a standing wave in the cable.

Concerning paired lines, a mismatched symmetrical line will have standing waves as sketched in Figure 5C. These standing waves are then both between the conductors and on the two conductors considered as a single line with ground return.

If we have a standing wave imposed on the balanced line from outside, we may have a perfectly balanced line (referred to ground) and with perfect termination, as indicated in Figure 5D. Under these circumstances the standing wave will cause no signal between the two conductors. However, in practice we very seldom have good balancing, even if we have a twisted line. In addition, the influence of the surroundings causes changes in the impedance of a paired line, often resulting in difficult terminating problems.

In Figure 5E is shown a mismatched balanced line. A standing wave imposed on the line from outside will here result in a signal between the conductors because of the phase conditions at a mismatch. In this context there is no difference if we use twisted line.

In order to stop or reduce the influence of the travelling wave antenna effect, it is possible to have filters for these signals in every paired line. It is also possible to reduce the travelling wave by introducing leakage of energy via taps from the cable shielding to the ground. These taps should be done with a resistance which will consume a part of the electromagnetic energy. The problems may then be reduced, but often such attempts will move some of the problems to other parts of the network.

In principle, it is also possible to have taps to the ground for high voltage power lines for instance by using high voltage condensers. However, such condensers have large dimensions and are very expensive. In addition, the impedance of such a line varies with the power consumption, causing unstable conditions for matching.

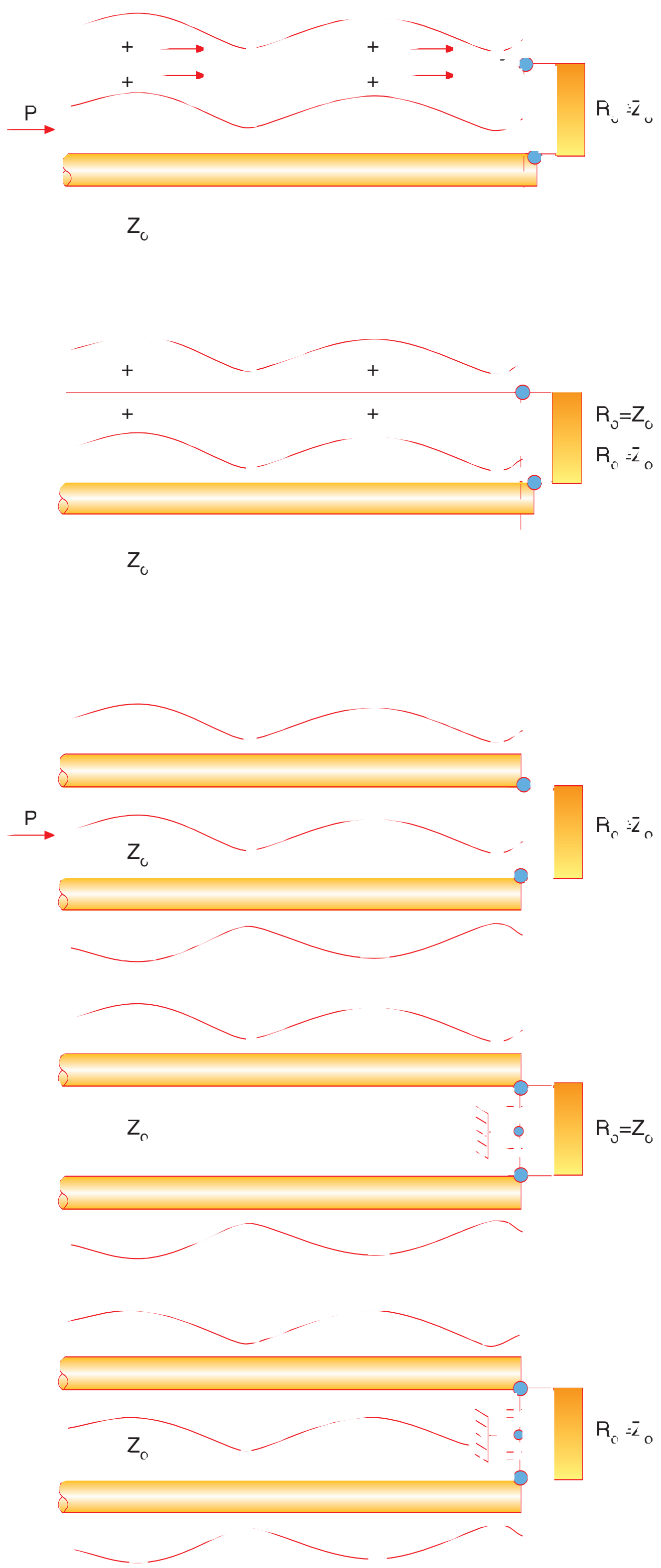


Figure 5 Standing waves in cables and lines

The conclusion is that when there are very strong received signals, some result may be obtained by using filters. But often the only way of solving the problems for the FDM system is to avoid using the disturbed channels.

Observed problems

The problems concerning the travelling wave effect have been observed especially in the southern part of Norway. In some areas the power lines along the valleys have a direction towards the European continent where there are several navigational transmitters in the 20 kHz to 30 kHz band. We here have to take into account the fact that the antenna pattern of a travelling wave antenna has maxima at certain angles to the antenna wire.

The received signal voltages have been rather high, up to more than 0.5 V, and the results have been false calls and even blocking of telephone exchanges.

When the antenna is directed towards the continent, we may have additional effect. Because of high atmospheric discharge activities such as lightning in continental areas, the natural noise in continental areas is higher than in coastal areas. The travelling wave antenna effect may therefore increase the noise in the network, and natural noise pulses may cause problems in the telephone systems.

References

1 Stokke, K N. *Radiotransmisjon: kabler, bølgeutbredelse, antenner.* 4. utgave. Oslo, Universitetsforlaget, 1982.

2 Glazier, E V D, Lamont, H R L. *The services' textbook of radio, volume 5: transmission and propagation.* Her Majesty's Stationary Office, 1958.

XDR/RPC without TCP/IP

Client-server communication with file exchange as transport mechanism

BY PER SIGMOND

1 Introduction

External Data Representation – XDR (1) and Remote Procedure Call – RPC (2) are the names of two protocols used for client-server communication. XDR provides abstract type definition and concrete transfer encoding syntax. RPC provides a client-server message format, and is defined partly with XDR type definition. XDR/RPC serves as a basis for the definition and implementation of various client-server protocols such as Network File System (NFS) and Network Information System (NIS). The *rpcgen* code-generation tool makes it easy to specify and implement customized client-server applications. XDR/RPC is now regarded as an industry standard.

XDR/RPC is normally closely related to the TCP/IP (3)(4) protocol stack, using UDP (5) or TCP as the transport protocol. In the last few years Transport Independent RPC (TI-RPC) has evolved to make this relation less tight, but for the PC (DOS) most solutions use TCP/IP. This means that PC-users with no IP (or SLIP (6) or PPP (7)) network connection lack the ability to use client-server communication based on the industry standard XDR/RPC.

This article presents a simple, transparent solution to this problem using file exchange as a substitute for the TCP transport protocol. Provided that the PC can exchange files with the server machine, the user can do client-server communication with his PC as the client. All existing TCP/IP-based RPC-servers are accessible via a special "transport bridge" at the host machine. Thanks to the modularity of the client-server model no changes are necessary in the client-code; only the client-stub needs a little tweaking.

2 Client-server model

Figure 1 shows the general client-server model as described by Tanenbaum (8). The main idea is that the client and the server is sheltered from the communication protocol activity by the client- and server-stubs. The stubs handle all client-server specific protocols (like XDR/RPC) while the transport entity provides transmission of the byte-stream across the network.

In the model there is a clear distinction between the transport entities and the stubs, suggesting that different transport protocols can be used. In traditional XDR/RPC one can choose between UDP and TCP.

3 XDR/RPC client-server session

A traditional client-server session using XDR/RPC is shown in Figure 2. A session normally has at least two transactions. The first one starts with a request from the client to the server-machine's portmap-server. The reason for this transaction is that the client needs to know what TCP port or UDP port a specific RPC-server listens on. Knowing the full transport address of the server, the client now transmits the "real" request to the RPC-server optionally followed by other requests.

4 Substituting the transport protocol

In order to allow client-server communication without using TCP or UDP, we have to substitute the transport functionality by some kind of file exchange or binary bulk-transfer. In the case of serial line connections the "modem community" use several file transfer protocols which have proven to be reliable and efficient, and there is no reason for not choosing one of them for a sample implementation. Among the candidates we find the [xyz]modem protocols and kermit. In a local network environment the file exchange can be done by simple file-sharing. But even if we substitute the transport protocol, we want to preserve connectivity to existing RPC servers that use TCP or UDP for transport. The solution is a "transport bridge".

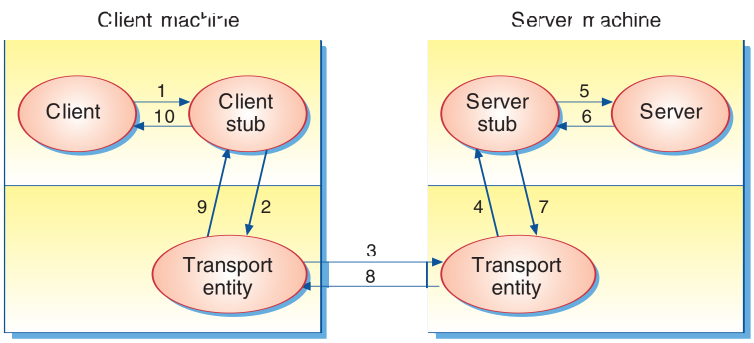


Figure 1 The client-server model

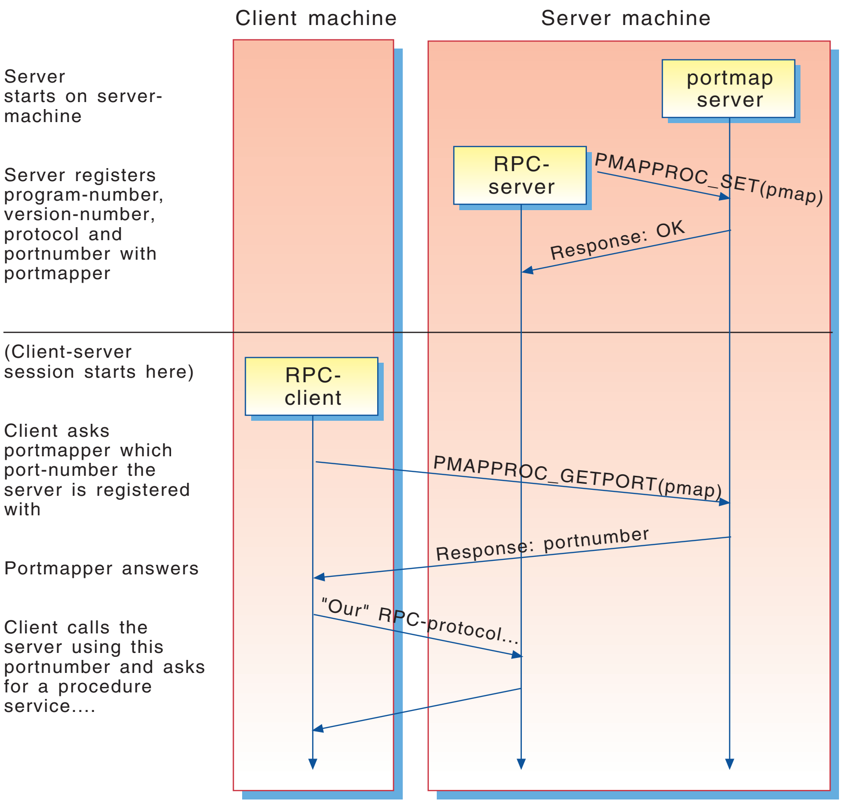


Figure 2 XDR/RPC client-server session

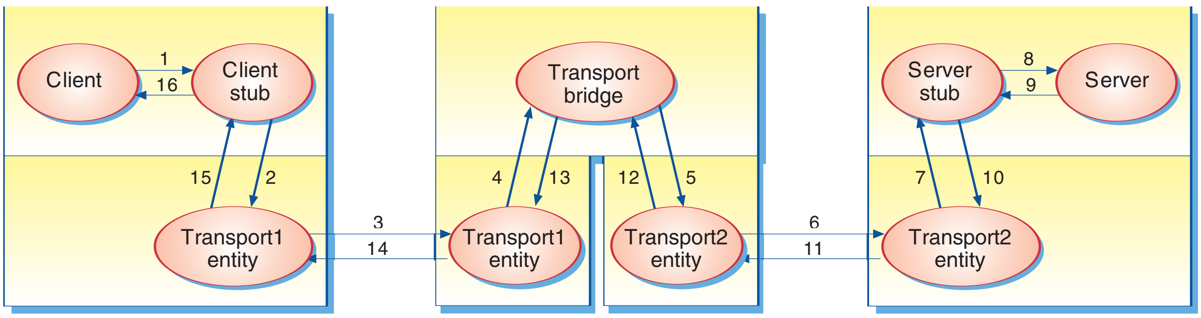


Figure 3 Transport bridge in relation to the client-server model

5 The transport bridge

The task of the transport bridge is to transparently transform the stub-data from one transport protocol to the other. From the clients-stubs point of view the transport bridge must look like a server stub and vice versa.

The bridge is not supposed to change the format of the stub-data, so the XDR format applied to both transports must be the same. Because of the stream-nature of the file exchange transport, it seems natural to use TCP as the transport protocol across the IP network. The XDR-bulk therefore has to be encoded according to the Record Marking Standard for byte stream transport protocols as defined in RFC1057 (2).

Figure 3 shows how the transport bridge relates to the client-server model. The “transport1” entities perform the file exchange, while the “transport2” entities are TCP-peers.

6 XDR/RPC client-server session using the transport bridge

An XDR/RPC client-server session using the transport bridge is a bit different than in the TCP/IP case. From Figure 4 we can see that the client is not issuing a portmap call. The only reason for the client to call the portmapper is to find the TCP or UDP port. Given that the client does not use TCP or UDP, it does not need the port number. In addition, every RPC-request carries full information about RPC program number and version number. So the transport bridge can extract these numbers from the RPC request and do the portmap call by itself.

7 A sample implementation

7.1 The PC-library (client-stub)

In an implementation we have to make some choices. The file exchange is done via serial line by the zmodem protocol, and the PC XDR/RPC code is my own port of the SUNRPC-4.0 for the “Waterloo TCP” package called “WATRPC”. I have chosen not to incorporate modem initialization and dialling-scripts into the XDR/RPC code. This connection set-up (and close) must be

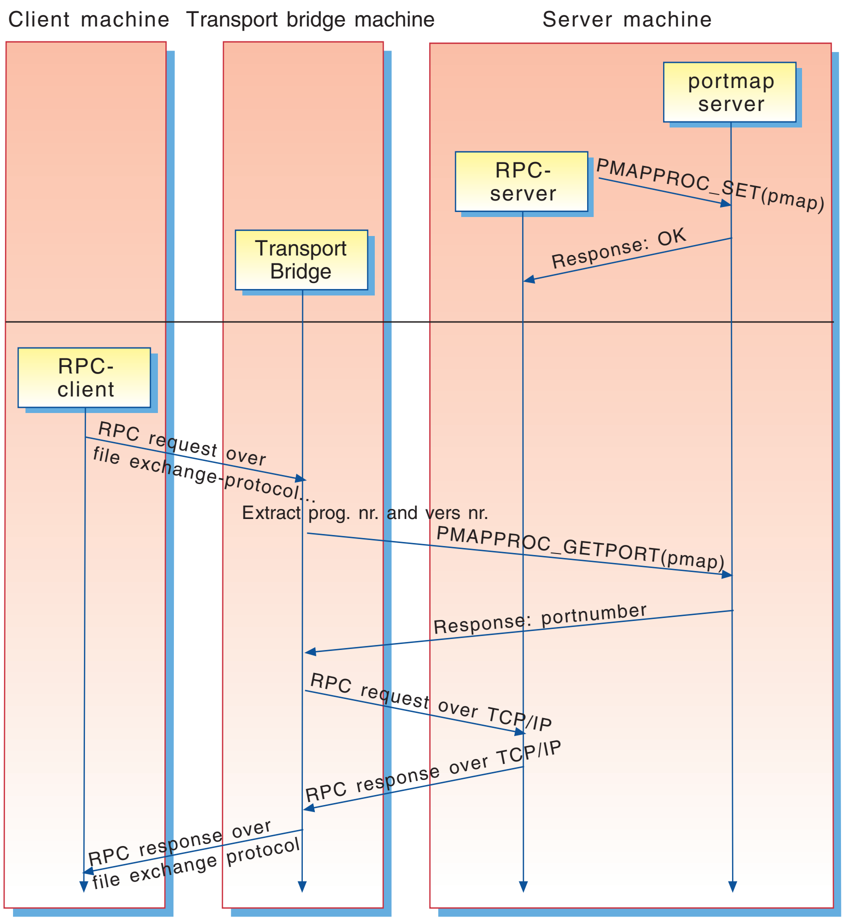


Figure 4 Client-server session using the Transport Bridge

done by some other PC-program or manually. When you are safely logged in to the host machine (the one with the “transport bridge”) the RPC client-stub takes over. The client itself does not have to be changed in any way other than specifying “modem” as the transport protocol (the last parameter to the standard *clnt_create()* procedure). Special *“dsz”* parameters like comport/speed settings go into the first *clnt_create()* parameter named “host” (normally carrying the IP host information). The portmap call is skipped (as noted before).

The changes in the client-stub limits to a new section in the file clnt_gen.c and a new file clnt_mod.c (builds on the file clnt_tcp.c).

7.2 The host-programs (transport bridge)

The host-programs were developed on a Unix machine. The program *“recxdr”* is started automatically by the sending zmodem program on the PC (the command *“rz”* is aliased to *“recxdr serverhost”*). *“recxdr”* administers the zmodem file-transfers via the public-domain *rz* and *sz* programs as well as the program *“rawxdr”* which is doing the TCP/UDP part. In order to avoid name-conflicts, I chose to tweak the *rz* program to ignore the received file-name and instead use the name supplied by a new -f option. The tweaked *rz* is called *“myrz”*. All file names are provided by *“recxdr”* via the *tmpnam()* routine. You get a good overview of the program flow of *recxdr* by just looking at the program code (Figure 5).

The *“rawxdr”* program includes a partial portmap-client and code to transmit/receive the client-stub data (XDR-encoded RPC-message) on a TCP socket. The RPC program- and version-numbers are easy to extract as they are always located at the same byte offset in the request-message.

7.3 Program interaction

To give an impression of how the programs interact, I have included a chart showing where the different programs are in operation (Figure 6).

```
#include <stdio.h>

main( int argc, char **argv ) {

char recfile[ 30 ];
char sendfile[ 30 ];
char cmdline[ 100 ];

        if (argc != 2) {
                fprintf( stderr, "Usage: %s serverhost\n", argv[0] );
                exit(1);
        }
        strcpy( recfile, tmpnam(NULL) );
        strcpy( sendfile, tmpnam(NULL) );

        sprintf( cmdline, "myrz -q -f %s", recfile );
        if (system( cmdline )) {
                unlink( recfile );
                exit(1);
        }

        sprintf( cmdline, "rawxdr %s < %s > %s", argv[1], recfile, sendfile );
        if (system( cmdline )) {
                unlink( recfile );
                unlink( sendfile );
                exit(1);
        }

        sprintf( cmdline, "sz %s", sendfile );
        if (system( cmdline )) {
                unlink( recfile );
                unlink( sendfile );
                exit(1);
        }

        unlink( recfile );
        unlink( sendfile );
        exit(0);
}
```

Figure 5 The “recxdr” program

8 References

1 Sun Microsystems, Inc. *XDR: External Data Representation Standard.* RFC 1014, June 1987.

2 Sun Microsystems, Inc., RPC. *Remote Procedure Call Protocol Specification Version 2.* RFC 1057, June 1988.

3 Postel, J (ed.). *Transmission control protocol – DARPA Internet program protocol specification.* RFC 793, September 1981.

4 Postel, J (ed.). *Internet protocol – DARPA Internet program protocol specification.* RFC 791, September 1981.

5 Postel, J. *User datagram protocol.* RFC 768, August 1980.

6 Postel, J. *Nonstandard for transmission of IP datagrams over serial lines: SLIP.* RFC 1055, June 1988.

7 Simpson, W. *The Point-to-point Protocol (PPP) for the transmission of Multi-protocol Datagrams over Point-to-Point Links.* RFC 1331, May 1992.

8 Tanenbaum, A S. *Computer Networks.* Prentice-Hall 1989. ISBN 0-13-166836-6.

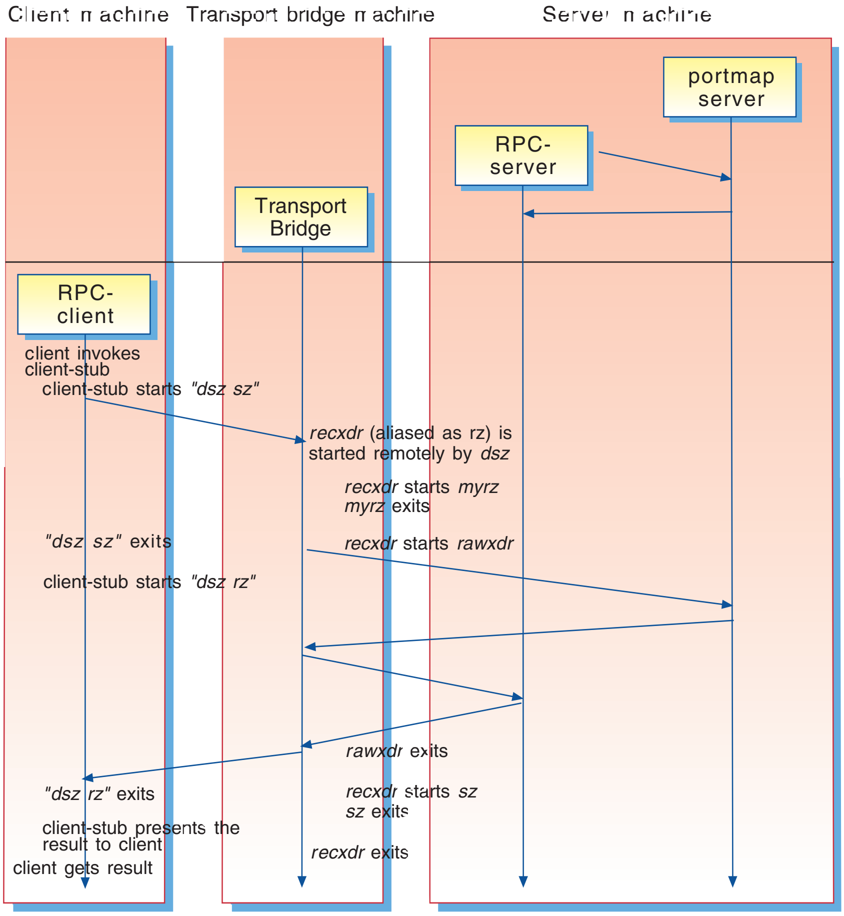


Figure 6 Program interaction during client-server session

Isolation levels in relational database management systems

BY OLE J ANFINDSEN AND MARK HORNICK

A shorter version of this article was published in *Platinum Edge Magazine,* vol 1, No. 1, 1994.

1 Introduction

Database management systems (DBMSs) evolved to allow multiple users safe access to shared data. Safe access implies that different user transactions do not interfere with one another. The activity necessary to do that is known as concurrency control (CC). Most commercially available DBMS we are aware of – including the DB2 Family, Ingres, Informix, Oracle, and Sybase – use a CC implementation technique called *locking*.

DBMS locking mechanisms are typically complicated yet fascinating. To make information systems run smoothly, database application programmers and database administrators working in high-transaction-volume environments need a solid understanding of these mechanisms. However, this article will not focus on locking per se, but on a higher level concept used by DBMSs to control the underlying protocols for acquiring and releasing locks.

2 Isolation levels

Applications have different requirements when it comes to CC: some need to execute as if they had the database all to themselves, others can tolerate some degree of interference from concurrently running applications. To meet the needs of different applications, DBMS vendors provide a way to specify different concurrency properties. In the case of relational DBMSs (RDBMSs), this is typically done by means of *isolation levels*. Although one of the nice things about isolation levels is that they are implementation independent, we will assume throughout this article that they are implemented by means of locking. Traditional isolation levels include (from least to most restrictive):

- **UR** (uncommitted read; also known as read uncommitted, read through, or dirty read) allows an application to read both committed and uncommitted data.

  UR applications do not acquire read locks; they simply read without locking. However, they do acquire write locks (see below).

  If you want to use UR, you need be sure you know what you are doing since the DBMS cannot guarantee the quality of the retrieved data with this isolation level. There are basically two situations where UR might be attractive: (1) when you don't need an exact answer, and (2) when you *know* that the data you want to read is not being updated by anyone else (although using *table* locks is probably a better solution in the latter case).

- **CR** (committed read; also known as read committed) allows an application to read only committed data.

  CR can be implemented by "zero duration" read locks. That is, if a CR application wants to read some database object, it suffices for the DBMS to check if a read lock *could* have been granted. If the answer is *yes,* the desired object is read (but no lock is acquired).

  Because of this, CR is much cheaper than CS (the next isolation level) in terms of CPU cycles. Unless your application needs the extra semantics offered by CS (many applications do not), CR is a far better alternative.

- **CS** (cursor stability) allows an application to read only committed data and guarantees that a row will not change as long as a cursor is positioned on it. This is useful, e.g., for an application that fetches a row from a cursor and then performs some database manipulation based on the current row's data values, before fetching the next row. In particular, we strongly recommend that you always use *at least* isolation level CS for cursors that you intend to use for UPDATE or DELETE operations (otherwise the row in question could change after being fetched but before execution of UPDATE WHERE CURRENT OF / DELETE WHERE CURRENT OF).

  CS makes sense only if you are locking pages or rows (as opposed to locking the entire table). The row/page read lock is kept until the cursor moves to the next row/page.

- **RR** (repeatable read) allows an application to read only committed data and guarantees that read data will not change until the transaction terminates (i.e., a read that is repeated will return the original row unchanged). RR will not prevent the so called *phantom row* phenomenon, i.e., when a cursor is reopened a row not present the previous time may appear.

  Read locks covering rows retrieved by an RR application must be kept until the end of the transaction, i.e., until the transaction either commits or aborts.

- **TC** (transaction consistency; also known as serializable) allows an application to read only committed data and guarantees that the transaction has a consistent view of the database (as if no other transactions were active).

  We prefer to refer to this isolation level as TC rather than SR (serializable) because a TC transaction may be part of a non-serializable schedule where other transactions use lower levels of isolation. In other words, since serializability is a property of a schedule of database transactions, attaching the label "serializable" to a single transaction is misleading. On the other hand, if all transactions in a schedule use TC, that schedule will be serializable. (A schedule is serializable if it is equivalent to a serial schedule, i.e., a schedule where each transaction is allowed to run to completion before the next can start.)

  All read locks acquired by a TC application must be kept until the end of the transaction, i.e., until the transaction either commits or aborts.

Each of these isolation levels, except CS, are defined in the SQL-92 standard, see e.g. (Date and Darwen 1993) or (Melton and Simon 1993). Given that CS is very useful for some applications, we fail to see any good reason for leaving that isolation level out of the standard. We hope users will demand support for all of the above-mentioned isolation levels from DBMS vendors. We also hope ANSI/ISO will include CS and QC (see below) in the SQL standard. Not only is every single isolation level mentioned in this article useful for certain applications, it is also very much to RDBMS users advantage if different systems have uniform CC options (this is, e.g., an interoperability issue).

Readers should be aware that some DBMS vendors (e.g., IBM) implement RR in such a way that phantom rows are disallowed, thus providing full TC support – "disguised" as RR.

As pointed out in (Anfindsen and Hornick 1994), an important isolation level is missing from this set. Consider a banking application that executes the query

```
SELECT   SUM(Balance)
FROM     Accounts
```

If isolation level CS or weaker is used and money is transferred between accounts while the query executes, SUM(Balance) may return an incorrect result. Specifically, money can be transferred between an account read by the query (but no longer locked) and an account not yet read by the query (and thus not yet locked).

In general, isolation level CS can cause problems for any query that uses aggregation and for any application that processes a cursor where the answer set must be consistent. Using RR or TC solves this problem, but these isolation levels are unnecessarily restrictive for such applications, since getting the same answer set more than once is not the issue. To provide necessary and sufficient isolation, we introduce the isolation level *query consistency:*

- **QC** (query consistency) allows an application to read only committed data and guarantees that all data accessed in a single query is consistent

Since cursors are defined using queries, QC guarantees that all rows in a cursor's answer set are consistent. QC is also valuable when performing statistical analysis, or when comparing or otherwise using together data values from different cursor row occurrences. QC is weaker than RR but stronger than CS.

Implementing QC by means of locking is straightforward; all read locks must be kept until the query is completed (until the cursor is closed).

Oracle-specific remark: Oracle7 differs from most other commercially available RDBMSs in the CC area. It uses *multiversion CC* which causes read-write dependencies to be avoided, and it does not use the isolation level concept. It is interesting to note, however, that the two CC options offered by Oracle7 seem to correspond very closely with isolation levels QC and TC.

Figure 1 shows isolation level support in ten commercial systems.

3 Conclusions

Conventional RDBMSs provide applications an incomplete set of options for controlling concurrency. Partly because they support only a subset of the standard isolation levels, and partly because the SQL standard does not recognize the important isolation level *query consistency* (QC). From our discussion it is clear that each of these isolation levels satisfies common application requirements. The situation is unlikely to change unless DBMS users tell vendors and standard committees that fine-granularity control over CC options is of great value to them.

Until your RDBMS has QC support added to it, you must use isolation levels RR or TC to ensure consistent results for entire queries, e.g., when using aggregate functions. This reduces potential concurrency for your applications and overall DBMS throughput.

For readers interested in a deeper understanding of isolation levels and transaction management in general, we recommend (Gray and Reuter 1993).

4 References

Anfindsen, O J, Hornick, M. 1994. *Application-Oriented Transaction Management.* Kjeller, Norwegian Telecom Research (technical report TF R 24/94).

Date, C J, Darwen, H. 1993. *A guide to the SQL standard.* Reading, Mass., Addison-Wesley.

Gray, J, Reuter, A. 1993. *Transaction processing: concepts and techniques.* San Mateo, Calif., Morgan Kaufmann Publishers.

Melton, J, Simon, A R. 1993. *Understanding the new SQL: a complete guide.* San Mateo, Calif., Morgan Kaufmann Publishers.

Isolation level	DB2/MVS	DB2/2	DB2/6000	DB2/400	Oracle	Sybase	Ingres	Informix	SQL/DS	Rdb
Uncommitted read (UR)		●	●	●			●	●		
Committed read (CR)	●					●		●		●
Cursor stability (CS)	●	●	●	●				●	●	
Query consistency (QC) *W*					●					
Repeatable read (RR)				●		●				●
Transaction consistency (TC)	●	●	●		●		●	●	●	●

● Supported   *W* Proposed isolation level   ● Oracle read consistency corresponds to isolation level QC

Figure 1 Isolation level support in ten commercial DBMSs

Mathematical methods and algorithms in the network utilization planning tool RUGINETT

BY RALPH LORENTZEN

1 Introduction

Once a year Telenor makes plans for routing and grouping of circuits in the Norwegian long distance and regional networks for a planning period consisting of several years. Up to 1993 these plans were made partly manually and partly using the EDP-based planning tool RUGSAM.

In RUGSAM separate models were used for 64 kbit/s circuit planning and for planning of circuits with bit rate ≥ 2 Mbit/s. The model for circuits with bit rate ≥ 2 Mbit/s consisted of two integer programming submodels, namely a routing and a grouping model, which were run sequentially. These submodels were large and required several hours running times on a powerful work station. The partitioning of the problem into a routing and a grouping problem also lead to solutions which in many cases had to be modified by the planners. Furthermore, SDH was not supported in RUGSAM.

When the need arose for an EDP-based tool for planning the utilization of PDH/SDH networks the availability of new, powerful optimization software was taken into consideration. It was now possible to create a model which, in addition to treating a combined PDH and SDH network, could consider routing and grouping simultaneously. So, in lieu of extending the functionality of RUGSAM's models for planning of circuits with bandwidth ≥ 2 Mbit/s to encompass SDH, it was decided to develop a new model which got the name RUGINETT.

RUGINETT is thus an *integer programming optimization model* which attempts to route and group circuit demands in a given PDH/SDH transmission network for each year in the planning period such that the cost of new multiplexer equipment is minimized whilst at the same time securing that the routing and grouping plan for each year does not deviate unnecessarily from previous year's plan.

The current version of RUGINETT does not pretend to solve the cost minimization problem to optimality. However, the user interface allows the planner to make modifications to the solution found by RUGINETT and check feasibility and cost.

RUGINETT consists of several modules, each of which can be put into one of two categories:

- Modules which enable the user to enter, edit and modify input/output data via tables and which prepare the data into formats suitable for the optimization module,
- An optimization module which finds solutions to the routing and grouping problem by integer programming algorithms,
- An optimization module for the determination of stand-by groups.

In this paper the integer programming models and solution algorithms which form the basis for the optimization modules are described.

2 The routing and grouping problem

A PDH/SDH transmission network consists of a set of nodes which are called *transmission stations* and a set of edges which are called *transmission systems* or short, *systems*. The transmission systems have bit rate capacities measured in bits per second.

It is assumed that the reader is familiar with concepts like routing and grouping of circuits.

Groups can be routed on systems with the same bit rate as the group and on groups with higher bit rates than the group in accordance with a given group hierarchy.

An example of a group hierarchy is shown in Figure 1.

The arrows show which lower order groups that can be *directly routed* on which higher order groups, and the number beside the arrows indicate how many lower order groups can be directly routed on a higher order group.

In order to simplify the presentation we consider a system to be a special type of group which we call a *system group*. A system group is never routed on other groups. So with this notation a group may be routed on system groups with the same bit rate as the group.

Actually, RUGINETT works with a slightly more general concept of transmission network where the nodes are the transmission stations and the edges are groups which have the property that they are never routed on other groups. These groups are called *edge groups*. Of course, edge groups which are not system groups will in general be routed on other groups. This routing is, however, not modelled in RUGINETT.

The box on top of Figure 1 represents the bit rate 2 Mbit/s. The number 30 in the box indicates the number of 64 kbit/s single circuits which can be routed on a 2 Mbit/s group. The *T* in the box indicates that there exist transmission systems with bit rate capacity 2 Mbit/s.

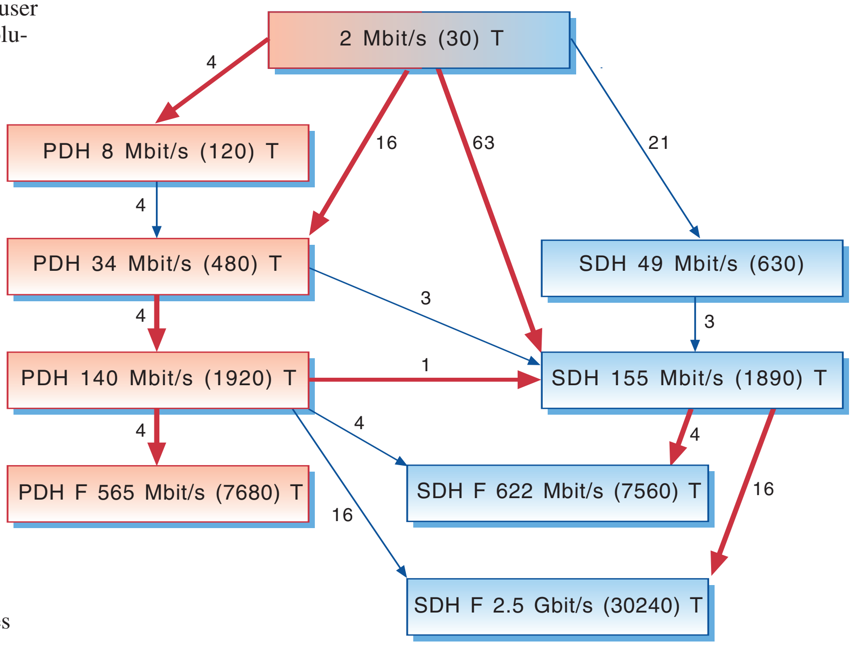


Figure 1 Example of a group hierarchy

The other boxes are given a label PDH or SDH depending on whether the bit rate belongs to the PDH or SDH subhierarchy. Again, a *T* in a box indicates that there exist transmission systems with the given bit rate capacity.

Some of the boxes have an *F* (for ‘fixed’) in them. This indicates that the groups with the corresponding bit rate can be directly routed on a single system only. It will be sufficient for us in this case to model the group and not the system it is routed on.

The fat arrows show the direct routing possibilities RUGINETT currently will choose from when new routings are proposed.

The edge groups may have one or more numbers called *reliability codes* associated with them. The idea is that edge groups sharing a common reliability code can fail due to the same ‘cause’. Groups inherit reliability codes from the groups they are routed on.

We shall now define *group sets*. Edge groups which connect the same two transmission stations, have the same capacity and have the same reliability codes, make up a group set. Furthermore, groups which have the same capacity and which are routed on groups belonging to the same group sets, make up a group set.

Thus, we can consider reliability codes as being attributes of group sets rather than groups.

A class of group sets having at least one reliability code in common is called a *reliability class*.

In addition to talking about groups routed on groups we can now talk about group sets routed on group sets.

The routing and grouping problem is to establish a routing and grouping scheme for a given set of *circuit demands* with various bit rates ≥ 2 Mbit/s between pairs of transmission nodes.

Circuit demands are *satisfied* or *covered by demand group sets* or, shorter, *demand sets* with corresponding bit rate. A 2 Mbit/s circuit demand must thus be covered by 2 Mbit/s demand sets while, for example, a 155 Mbit/s demand must be covered by 155 Mbit/s demand sets. No group set can be routed on demand sets.

Circuit demands may be *directed* or *undirected.* The corresponding demand group sets are also denoted as directed or undirected. A directed demand set utilizes capacity in one direction only in the group sets it is directly routed on.

An undirected demand may be subject to a *diversity requirement*. This means that the fraction of the corresponding demand sets in the same reliability class has an upper bound which is less than one. This upper bound is called the *diversity factor* of the demand. A typical diversity factor value is .5.

Similarly, a family of undirected demands may be subject to a *load sharing requirement*. This means that demand sets for at most one of the circuit demands in the family can belong to any particular reliability class. Such a family is called a *load sharing family*.

A demand cannot both be subject to a diversity requirement and be a member of a load sharing family.

Amongst all routing and grouping schemes which satisfy given demands with their diversity and load sharing requirements the problem is to find one which minimizes the cost of new multiplexer equipment and which at the same time does not deviate unnecessarily from previous year’s scheme. This routing and grouping problem will be solved using integer programming.

The planner has a problem of a dynamical nature. In establishing the best way of using the network he must take into consideration the development over time of network structure and circuit demands. So, ideally the problem should be solved dynamically where network topology and capacity and circuit demands are taken into consideration simultaneously for all the years in the planning period. This has been beyond our capability, so we have elected to apply a static model which is run separately for each year in the planning period. A simple feature which supports ‘dynamical use’ is, however, built into RUGINETT. The planner can label certain group sets as *target group sets* which are then given a bonus. It is recommended that the planner first make a ‘future run’ where the input represents the situation in a ‘target year’ some time into the future. The group sets selected by RUGINETT in this run can then be copied into the regular RUGINETT runs and labelled as target group sets.

It is important for the planner to be able to exercise a certain control over the characteristics of the solution. There may exist group sets which he would label as *fixed* and which he would require the solution to contain, and there may be other group sets which he would label as *recommended* and which he would give a special bonus.

Topologically, group sets connect pairs of transmission stations and follow a path in the transmission network. A group set is said to have *complete routing* if the group sets it is routed on completely covers the path of the group set. Otherwise, it is said to have *incomplete routing*. A group set is said to have *no routing* if it is not routed on any group set. Edge group sets have no routing. There may exist other fixed group sets and recommended group sets which have incomplete or no routing. The new group sets which RUGINETT suggests will, however, always have complete routing.

3 The integer programming model

3.1 Definitions and notation for the formulation of the constraints

First we give some definitions.

A *slot* in a group is the largest common divisor of the maximum number of groups of each type which can be directly routed on the group. So every group set which is directly routed on another group set occupies an integer number of slots in the latter group set.

We write $g \in u$ ($g \in d$) to denote that group set g is an undirected (directed) demand set which contributes to the satisfaction of the undirected demand u (directed demand d).

Furthermore, we write $g \in g'$ to denote that a (directed or undirected) group set g is directly routed on group set g'.

We also write $g \in g'^{+}$ to denote that g is a directed demand set directly routed on g' in *positive direction,* and $g \in g'^{-}$ to denote that g is a directed demand set directly routed on g' in *negative direction*.

Now we introduce the notation we need for the formulation of bounds and constraints in the integer program. We will later introduce some more notation which is required for the formulation of the cost function to be minimized.

Subscripts:

u	undirected circuit demand
d	directed circuit demand
g	group set
r	recommended
t	target
c	reliability class
f	load sharing family

Constants:

F_g	number of fixed groups in group set g ($F_g > 0$ for edge group sets)
R_g	number of recommended groups in group set g
T_g	number of target groups in group set g
D_u	number of demand groups for undirected circuit demand u
D_d	number of demand groups for directed circuit demand d
$k_{gg'}$	number of slots a group in set g occupies in a group in set g'
H_g	number of slots in a group in group set g
P_u	diversity factor for undirected demand u

3.2 Problem variables

Here we define the variables used in the integer programming model.

x_{rg}	number of recommended groups which are used in group set g (generated only if $R_g > 0$)
x_{tg}	number of target groups which are used in group set g (generated only if $T_g > 0$)
x_g	number of other groups which are used in group set g
w_{uc}	violation of diversity requirement of demand u for reliability class c
w_{fc}	violation of load sharing requirement of family f for reliability class c
w_u	unsatisfied part of undirected demand u
w_d	unsatisfied part of directed demand d

All x-variables are required to be integer.

3.3 Bounds and constraints

Here we list the bounds and constraints used in the model.

Bounds:

$x_g \geq F_g$ if group set g can be extended (1)

$x_g = F_g$ if group set g cannot be extended, or if g is a demand set and there is a demand set $\bar{g}$ with $F_{\bar{g}} > 0$ in the same reliability class and load sharing family as g (2)

$$x_{rg} \leq R_g \quad (3)$$

$$x_{tg} \leq T_g \quad (4)$$

(1) together with (2) express that the number of groups in group set g cannot be lower than the number of fixed groups in the group set.

(2) sharpens the inequality to equality for group sets which cannot be extended or where an increase in x_g would lead to an increase in load sharing violation.

(3) limits the number of recommended groups in group set g to R_g.

(4) limits the number of target groups in group set g to T_g.

Constraints:

$\Sigma_{g\in u}\,(x_{rg} + x_{tg} + x_g) + w_u \geq D_u$ for all undirected demands u (5)

$\Sigma_{g\in d}\,(x_{rg} + x_{tg} + x_g) + w_d \geq D_d$ for all directed demands d (6)

$H_{g'}(x_{rg'} + x_{tg'} + x_{g'}) - \Sigma_{g\in g'\setminus(g'^{+}\cup g'^{-})}\, k_{gg'}(x_{rg} + x_{tg} + x_g) - \Sigma_{g\in g'^{+}}\, k_{gg'}(x_{rg} + x_{tg} + x_g) \geq 0$ for all group sets g' (7)

$H_{g'}(x_{rg'} + x_{tg'} + x_{g'}) - \Sigma_{g\in g'\setminus(g'^{+}\cup g'^{-})}\, k_{gg'}(x_{rg} + x_{tg} + x_g) - \Sigma_{g\in g'^{-}}\, k_{gg'}(x_{rg} + x_{tg} + x_g) \geq 0$ for all group sets g' (8)

$-\Sigma_{g\in u\cap c}\,(x_{rg} + x_{tg} + x_g)/(P_u D_u) + w_{uc} \geq -1$ for undirected demands u subject to diversity requirements and reliability classes c (9)

$-\Sigma_{u\in f}\,\Sigma_{g\in u\cap c}\,(x_{rg} + x_{tg} + x_g)/D_u + w_{fc} \geq -1$ for load sharing families f of undirected demands and reliability classes c where $\Sigma_{g\in f\cap c}\, F_g = 0$ (10)

(5) and (6) express that the number of groups in demand sets covering demands u and d cannot be less than D_u - w_u and D_d - w_d respectively.

(7) and (8) express that the group sets directly routed on group set g' do not exceed the capacity of group g'. For those group sets g' on which no directed demand sets will be considered to be directly routed, (7) and (8) will coincide and will of course not be duplicated in the integer program.

(9) expresses that the fraction of undirected demand u covered by group sets in reliability class c should not be greater than D_u.

(10) expresses (as well as possible within the linear relaxation) that for at most one demand in load sharing family f corresponding demand sets may belong to reliability class c.

It is important to realize that the bounds and constraints given above are not sufficient to describe the problem completely. The constraints (7) and (8) put limits on the number of group sets which can be directly routed on a group set, but they do not model the fact that a group set which is not an edge group set must be directly routed on a sequence of consecutive higher level group sets. This has one advantage, namely that the integer programming problem permits the existence of user defined group sets which have incomplete routing. This is useful when a group set is partially routed in a network where the planner has no responsibility/authority. Furthermore, as will be seen later, the non-edge group set variables x_g will be generated dynamically in the course of the solution process in a way that guarantees that they are directly routed on a sequence of consecutive higher level group sets.

3.4 The cost function

In order to formulate the cost function to be minimized in RUGINETT we need to introduce some additional notation:

C_g^{term}	termination cost of group g
$C_{gg'}$	cost of routing group set g on group set g'
F_r	bonus factor for recommended groups
F_t	bonus factor for target groups
M^1	penalty factor for not satisfying diversity requirements
M^2	penalty factor for not satisfying load sharing requirements
M^3	penalty factor for not satisfying undirected demands
M^4	penalty factor for not satisfying directed demands

The cost function to be minimized is as follows:

$$\Sigma_g C_g x_g + \Sigma_g F_r C_g x_{rg} + \Sigma_g F_t C_g x_{mg}$$
$$+ M^1 \Sigma_{dc} w_{dc} + M^2 \Sigma_{fc} w_{fc} + M^3 \Sigma_u w_u + M^4 \Sigma_d w_d$$

where $C_g = C_g^{term} + \Sigma_{\{g';g \in g'\}} C_{gg'}$

We will not treat in detail how the individual cost elements are derived from the input data to RUGINETT here, but limit ourselves to the following remarks:

C_g^{term} reflects the cost of multiplexing equipment used in both ends of group g, and depends only on the bit rate of g.

$C_{gg'}$ reflects the cost of multiplexing group g into group g' and depends only on the bit rates of g and g'.

F_r and F_t lie between 0 and 1. Small values of F_r and F_t imply high desirability of sticking to recommended groups and target groups respectively.

M^1, M^2, M^3 and M^4, are large positive numbers. M^3 and M^4 should be much larger than M^1 and M^2 because it is more important to cover the demands than to satisfy diversity and load sharing requirements.

3 Solution method

3.1 General

The integer program is too large and complex to be solved to a theoretical optimum, so some heuristics are used. To compensate for this the user interface contains a wide variety of options for inspection and manipulation of the solution.

The solution method consists of a combination of

- linear programming with dynamic generation of constraints and variables
- variable generation and variable bounding heuristics.

We start by solving the linear programming relaxation of the integer program.

3.2 Linear programming with dynamic generation of rows and columns

The number of group variables x_g and constraints of type (7) and (8) is so enormous that it is unrealistic to include them all explicitly in the model. Furthermore, the number of constraints of types (9) and (10) is also very large, and only a few of them will be binding in the optimum solution. We therefore solve the linear programming relaxation by using dynamic row and column generation.

Initially, a restricted version of the linear program is set up and solved which has

- all variables of types x_{rg} and x_{tg}
- variables of type x_g with $F_g > 0$
- all constraints of types (5) and (6)
- constraints of type (7) and (8) corresponding to sets g' with $F_{g'} > 0$ or $R_{g'} > 0$ or $T_{g'} > 0$ (no duplication if (7) and (8) coincide for a given g')
- constraints of types (9) or (10) containing at least one of the above-mentioned variables
- all relevant bounds.

Based on the shadow prices associated with the optimal solution of this restricted linear program new groups with associated variables of type x_g and new constraints of types (7), (8), (9) and (10) are generated, and a new (restricted) linear program is solved. This process is continued until it can be ascertained that we have an optimal solution to the complete linear program.

In order to describe the details of this iterative process we assume that we are at a stage where we have just found an optimal basis for a restricted linear program which, in addition to the initial constraints and variables, may include

- some additional variables of type x_g
- the constraints of type (7) and (8) corresponding to the additional $x_{g'}$ variables (no duplication if the constraints coincide)
- constraints of types (9) or (10) which contain at least one of the additional x_g variables.

First, we check our basic solution to see if it violates any potential constraints of type (9) and (10). If it does, such violated constraints are added to the linear program which is reoptimized. This process is continued until all potential constraints of type (9) and (10) are satisfied.

The following notation is used for the shadow prices associated with a basis for the restricted and the complete linear program:

Constraint type:	Shadow price:
(5)	π_u
(6)	π_d
(7)	$\pi_{g'+}$
(8)	$\pi_{g'-}$
(9)	π_{cu}
(10)	π_{cf}

If (7) and (8) coincide, their common shadow price is denoted by $\pi_{g'}$.

The optimal solution to the restricted linear program is obviously a feasible solution to the complete linear program. We shall now establish whether it is optimal or whether we can find one or more not yet generated x_g variables with negative reduced cost which we can introduce into the linear program together with the necessary constraints and continue the solution process.

First, we shall extend our optimal basis for the restricted linear program to an imagined feasible basis for the complete linear program. We imagine for a moment that we have added all the not yet generated variables and constraints to the restricted program. We may consider the slack variables of the not yet generated constraints of type (9) and (10) to be basic. If, for a group set g', one of (7) and (8) has been generated but the other not, we consider the slack variable of the not generated constraint to be basic. If (7) and (8) coincide for a not generated group set g', we consider the corresponding $x_{g'}$ variable to be basic. This implies that $\pi_{g'}$ is non-negative. If they do not coincide, we still consider the corresponding $x_{g'}$ variable, and, in addition, the slack variable of one of the constraints of type (7) or (8) of our choice to be basic. We shall return to this choice later. This implies that one of the shadow prices $\pi_{g'+}$ and $\pi_{g'-}$ is zero and that the other is non-negative, which in turn implies that the reduced costs of the slack variables in (7) and (8) are non-negative.

Since all not generated $x_{g'}$ variables are basic, their reduced cost is zero. Furthermore, since the slack variables for the not yet generated constraints of types (9) and (10) are basic, the shadow prices for these constraints are 0 since the slack variables have reduced cost zero. Thus, the diversity and load sharing violation variables have positive reduced cost. Since the shadow prices for the constraints (7) and (8) are non-negative, the slack variables of these constraints have non-negative reduced cost.

Therefore, when we search for new candidates in the column generation process it suffices to consider x_g variables corresponding to demand sets.

The demand u may be subject to a diversity requirement or belong to a load sharing family f (but not both).

We shall now establish the reduced cost for a new demand set variable for a given demand relative to this basis for the complete linear program.

3.3 Finding reduced costs for demand sets

Let

G^+ denote the group sets g' for which (7) has already been generated

G^- denote the group sets g' for which (8) has already been generated

G denote the group sets g' for which (7) and (8) coincide and has been generated

$\overline{G}$ denote the group sets g' for which neither (7) nor (8) has been generated.

In section 3.2 we stated that we could choose whether we would consider the slack variable in (7) or (8) corresponding to a g' in $\overline{G}$ to be basic. When we establish the reduced cost for an undirected demand set variable directly routed on a g' in $\overline{G}$, (7) and (8) corresponding to g' coincide. When we establish the reduced cost for a directed demand set variable x_d routed on $g'+$, we elect to consider the slack variable associated with (8) to be basic and vice versa. This choice will make the reduced cost as large as possible. We permit ourselves to denote the shadow price $\pi_{g'+}$ or $\pi_{g'-}$ corresponding to the not generated constraint (7) or (8) whose slack variable is non-basic, and thus may be non-zero, by $\pi_{g'}$.

The reduced cost for of the demand group variable x_g contributing to the directed demand d can be written:

$$
\begin{aligned}
-\pi_d &+ C_g^{term}\\
&+ \Sigma_{\{g'\in G^+;g\in g'+\}}(C_{gg'} + k_{gg'}\pi_{g'+})\\
&+ \Sigma_{\{g'\in G^-;g\in g'-\}}(C_{gg'} + k_{gg'}\pi_{g'-})\\
&+ \Sigma_{\{g'\in G;g\in g'\}}(C_{gg'} + k_{gg'}\pi_{g'})\\
&+ \Sigma_{\{g'\in \overline{G};g\in g'\}}(C_{gg'} + k_{gg'}\pi_{g'})
\end{aligned}
\tag{11}
$$

whilst the reduced cost for the demand group variable x_g contributing to the undirected demand u can be written:

$$
\begin{aligned}
-\pi_u &+ C_g^{term}\\
&+ \Sigma_{\{g'\in G^+;g\in g'\}}(C_{gg'} + k_{gg'}\pi_{g'+})\\
&+ \Sigma_{\{g'\in G^-;g\in g'\}}(C_{gg'} + k_{gg'}\pi_{g'-})\\
&+ \Sigma_{\{g'\in G;g\in g'\}}(C_{gg'} + k_{gg'}\pi_{g'})\\
&+ \Sigma_{\{g'\in \overline{G};g\in g'\}}(C_{gg'} + k_{gg'}\pi_{g'})\\
&+ \Sigma_{\{c;g\in c\}}\pi_{cu}/(P_uD_u) \quad \text{(if } u \text{ is subject to a diversity requirement)}\\
&+ \Sigma_{\{c;g\in c\}}\pi_{cf}/D_u \quad \text{(if } u \text{ belongs to a load sharing family } f\text{)}
\end{aligned}
\tag{12}
$$

The values of the shadow prices $\pi_{g'+}$, $\pi_{g'-}$ and $\pi_{g'}$ are known for all $g' \in G^+, G^-$, and G respectively. Furthermore, since the slack variables for the not yet generated constraints of types (9) and (10) are basic, the shadow prices π_{cu} and π_{cf} for these are 0. It remains to establish the values of the relevant shadow prices $\pi_{g'}$ for $g' \in \overline{G}$ (if such g'-s exist). Since the corresponding $x_{g'}$ variables are basic, they price out zero. Furthermore, since the g'-s cannot be demand groups, the $x_{g'}$ variables do not appear in any constraints of types (9) or (10). This means that:

$$
\begin{aligned}
&H_{g'}\pi_{g'} - \Sigma_{\{g''\in G;g'\in g''\}}k_{g'g''}\pi_{g''}\\
&- \Sigma_{\{g''\notin G;g'\in g''\}}k_{g'g''}\pi_{g''} = C_{g'}
\end{aligned}
\tag{13}
$$

which implies that we can express the unknown shadow price $\pi_{g'}$ in terms of the shadow prices $\pi_{g''}$ for the groups g'' which the g'-s are routed on:

$$\pi_{g'} = (1/H_{g'}) [C_{g'}^{term} + \Sigma_{\{g'' \in G; g' \in g''\}} (C_{g'g''} + k_{g'g''}\pi_{g''}) \\ + \Sigma_{\{g'' \notin G; g' \in g''\}} (C_{g'g''} + k_{g'g''}\pi_{g''})] \tag{14}$$

Again, the $x_{g''}$ variables for $g'' \notin G$ price out zero, which means that

$$H_{g''}\pi_{g''} - \Sigma_{\{g''' \in G; g'' \in g'''\}} k_{g''g'''}\pi_{g'''} \\ - \Sigma_{\{g'' \notin G; g'' \in g'''\}} k_{g''g'''}\pi_{g'''} = C_{g''} \tag{15}$$

which in turn implies that we can express the unknown shadow prices $\pi_{g''}$ in terms of the shadow prices $\pi_{g'''}$ for the groups g''' which the g''-s are routed on:

$$\pi_{g''} = (1/H_{g''})[C_{g''}^{term} \\ + \Sigma_{\{g''' \in G; g'' \in g'''\}} (C_{g''g'''} + k_{g''g'''} \pi_{g'''}) \\ + \Sigma_{\{g''' \notin G; g'' \in g'''\}} (C_{g''g'''} + k_{g''g'''}\pi_{g'''})] \tag{16}$$

In this way we can continue until we get to a level where all group sets and their corresponding inequalities are generated and we know the value of all the shadow prices.

3.4 Finding demand set variables with minimum reduced cost

We shall here show how for each demand we find a demand set variable with minimum reduced cost by solving shortest path problems.

We focus first on a directed demand d. From (11) we see that a corresponding demand set variable x_g with minimum reduced cost can be considered as a shortest path between the two nodes associated with d in a directed network where the arcs correspond to group sets g' connecting pairs of nodes in the transmission network. The length of an arc corresponding to group set g' is given by

$C_{gg'} + k_{gg'}\pi_{g'+}$ in positive direction if g' belongs to G^+,

$C_{gg'} + k_{gg'}\pi_{g'-}$ in negative direction if g' belongs to G^-,

$C_{gg'} + k_{gg'}\pi_{g'}$ in both directions if g' belongs to either G or $\overline{G}$.

For the generated group sets g' these lengths are known. For the group sets g' not yet generated they must be found. However, for every pair of nodes it is sufficient to find the group set g' at a given level in the group hierarchy with minimal value of $\pi_{g'}$. (14) shows us that this minimum can be found by solving a shortest path problem between the relevant pair of nodes in an undirected network where the edges correspond to group sets g'' connecting pairs of nodes in the transmission network where the length of an edge corresponding to a group set g'' is given by

$$(1/H_{g'})(C_{g'g''} + k_{g'g''}\pi_{g''}).$$

For the generated group sets g'' these lengths are known. For the group sets g'' not yet generated they must be found. However, for every pair of nodes it is sufficient to find the group set g'' at a given level in the group hierarchy with minimal value of $\pi_{g''}$. (16) shows us that this minimum can be found by solving a shortest path problem between the relevant pair of nodes in the network of group sets g''' connecting pairs of nodes in the transmission network where the length of a group set g''' is given by

$$(1/H_{g''})(C_{g''g'''} + k_{g''g'''}\pi_{g'''}).$$

This reasoning can be continued until we encounter shortest path problems in networks of edge group sets where the lengths are known.

So, the procedure will be to start using Floyd's algorithm to establish shortest paths between all pairs of nodes in networks where the edges are edge group sets at a given level. The results are then used to set up a new network where the edges with their lengths correspond to the shortest paths with their lengths. Then we again use Floyd's algorithm to establish shortest paths for group sets at different levels which can be routed in this network. This is continued recursively until we get to the establishment of the shortest paths for our demand sets. This is done using Dijkstra's algorithm.

We observe that the lengths of the arcs and edges in the different networks depend on group levels only so that it is sufficient to execute Floyd's algorithm once for each level in the group hierarchy.

For an undirected demand u the situation is complicated somewhat by the existence of diversity and load sharing constraints. From (12) we see that a corresponding demand set variable x_g with minimum reduced cost can be considered as a shortest path between the two nodes associated with d in the undirected network of group sets g' connecting pairs of nodes in the transmission network. The 'length' of a group set g' is given by a sum of the following terms:

$C_{gg'} + k_{gg'}\pi_{g'+}$ if g' belongs to G^+,

$C_{gg'} + k_{gg'}\pi_{g'-}$ if g' belongs to G^-,

$C_{gg'} + k_{gg'}\pi_{g'}$ if g' belongs to either G or $\overline{G}$,

$\Sigma_{\{c; g' \in c, (9) \text{ already generated for } (c,u)\}} \pi_{cu}/(P_u D_u)$,

$\Sigma_{\{c; g' \in c, (10) \text{ already generated for } (c,f)\}} \pi_{cf}/D_u$

For the generated group sets g' these lengths are known. For the group sets g' not yet generated they must be found. However, for every pair of nodes it is sufficient to find the group set g' at a given level in the group hierarchy which minimizes the sum of the last three terms in the sum given above.

(14) shows us that this minimum can be found by solving a shortest path problem between the relevant pair of nodes in the network of group sets g'' connecting pairs of nodes in the transmission network where the length of a group set g'' is given by:

$$(1/H_{g'})(C_{g'g''} + k_{g'g''}\pi_{g''}) \\ + \Sigma_{\{c; g'' \in c, (9) \text{ already generated for } (c,u)\}} \pi_{cu}/(P_u D_u) \\ + \Sigma_{\{c; g'' \in c, (10) \text{ already generated for } (c,f)\}} \pi_{cf}/D_u \tag{17}$$

For the generated group sets g'' these lengths are known. For the group sets g'' not yet generated they must be found. However, for every pair of nodes it is sufficient to find the group set g'' at a given level in the group hierarchy which minimizes (17).

(16) shows us that this minimum can be found by solving a shortest path problem between the relevant pair of nodes in the

network of group sets g''' connecting pairs of nodes in the transmission network where the length of a group set g''' is given by:

$$(1/H_{g'})(C_{g'g''} + k_{g'g''}(1/H_{g''})(C_{g''g'''} + k_{g''g'''}\,\pi_{g'''}))$$
$$+ \Sigma_{\{c;g'''\in c,(9)\ \text{already generated for}\ (c,u)\}}\ \pi_{cu}/(P_u D_u)$$
$$+ \Sigma_{\{c;g'''\in c,(10)\ \text{already generated for}\ (c,f)\}}\ \pi_{cf}/D_u$$

This reasoning can be continued until we encounter shortest path problems in networks of edge group sets where the lengths are known.

So, the procedure will be to start using Floyd's algorithm to establish shortest paths between all pairs of nodes in networks where the edges are edge group sets at a given level. The results are then used to set up a new network where the edges with their lengths correspond to the shortest paths with their lengths. Then we again use Floyd's algorithm to establish shortest paths for group sets at different levels which can be routed in this network. This is continued recursively until we get to the establishment of the shortest paths for our demand sets. This is done using Dijkstra's algorithm.

For demands with diversity requirements the lengths of the edges in the networks may vary from demand to demand, so that Floyd's and Dijkstra's algorithms must be run separately for each demand. Likewise, for demands with load sharing requirements the edge lengths may vary from load sharing family to load sharing family, so that Floyd's and Dijkstra's algorithms must be run separately for each family.

We see that we pay for not having established the complete linear programming matrix by having to solve a very large number of shortest path problems. Therefore, care has been taken to code the establishment of the networks and the shortest path algorithms efficiently.

3.5 Introducing new variables and constraints into the linear program

By the procedure described above we establish in general some demand set variables with negative reduced costs. These demand set variables, together with the variables corresponding to all sets the demands are routed on, are introduced into the linear program. Furthermore, all corresponding constraints (7) and (8) relating to new variables $x_{g'}$ are introduced.

When no more demand set variables with negative reduced costs are found, it is checked whether all potential constraints of types (9) and (10) are satisfied. If not, the unsatisfied constraints are established and introduced into the linear program which is resolved again with dynamic generation of demand set variables and corresponding constraints. This process terminates when all minimal reduced costs of demand set variables are non-negative and all potential constraints of type (9) and (10) are satisfied.

3.6 Optimality criterion for the linear program

The algorithm will terminate with the minimum reduced costs corresponding to all demands being non-negative. Since the basis relative to which a directed demand set variable is priced out depends on the variable, it is not immediately obvious that we have obtained an optimal solution to the complete linear program. We make the assumption that all directed demand sets corresponding to a directed demand are directly routed in he same direction on any given group set.

To see that we really have obtained the optimal solution we consider a not generated group set g' on which directed demand sets contributing to a given directed demand d may be directly routed. We can have two situations:

In an optimal solution no directed demand set contributing to d is directly routed on g'. We can then ignore the reduced costs for demand set variables contributing to d which cross the not generated constraints (7) and (8) corresponding to g'.

In an optimal solution directed demand sets contributing to d are directly routed on g' in positive (negative) direction. We can then ignore the reduced costs for demand set variables contributing to d which cross the not generated constraints (8) ((7)) corresponding to g'.

In both situations we see that we may assume that the non-basic directed demand set variables have non-negative reduced costs relative to a unique basis.

3.7 Integer programming heuristics

When the linear programming relaxation of the integer program has been found, a lot of the group set variables will normally be fractional. In order to reduce cost the solution 'cheats' us by using fractional group sets.

The basic principle used in the heuristics is to solve the linear program and inspect the solution to see if it contains fractional group set variables or load sharing violations. If so, new bounds are added to the linear program. The linear program is solved again and checked for fractional group set variables or load sharing violations which give rise to additional bounds, etc. This process is continued until all the group set variables are integer and all load sharing requirements are satisfied.

A typical situation is illustrated in Figure 2. Three system group sets of value 1 with the same bit rate meet at a transmission node *A*. Group sets 1 and 2, both with the same bit rate as the system groups and with values .75 and .25, are routed on the system sets as indicated in Figure 2. This is illegal since there is of course room for at most one of the group sets 1 and 2 on the system set to the left. The corresponding constraint of type (6) for the system set to the left is, however, satisfied.

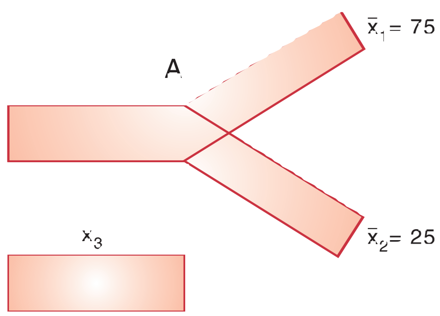


Figure 2 Example of an illegal fractional solution

Let in general the variable names with bars above them indicate the values they take in the current solution. A simple, but of course rough, heuristic would be, since $\max(\bar{x}_1, \bar{x}_2) = \bar{x}_1$, to force x_1 to 1 and then solve the problem again. We have, however opted for a more refined heuristic where group set 2 is excluded, and RUGINETT is given the choice between group set 1 and a group set 3 which has the same bit rate as the others and which is routed on the system group to the left only (see Figure 2).

We shall now give a general description of the heuristic. Assume that we are at a stage where we have solved the restricted linear program and obtained a solution $(\bar{x})$ where all potential constraints of types (9) and (10) are satisfied.

1 First we check if the load sharing requirements are satisfied. If not, then we pick one violated requirement (if any exists) from each load sharing family, put an upper bound of zero on all demand set variables except the one with the largest value amongst those routed on groups belonging to the relevant reliability class and reoptimize.

2 Next we set a lower bound of $\lfloor \bar{x} \rfloor$ on all variables x.

3 We then identify the highest level with fractional group set variables. We shall now select a fractional group set variable at this level. If there are fractional demand set variables, the selection will be amongst those. Amongst the sets which are directly routed on the lowest number of sets we select a group set s where $\bar{x}_s$ has maximum fractional part.

   The group set s is directly routed on a class G' of group sets. Define

$$S = maks_{g' \in G'} \frac{\sum_{g \in g'} k_{gg'} \lceil \bar{x}_g \rceil}{H_{g'} \bar{x}_{g'}} \tag{12}$$

   If $S \leq 1$, the bound

$$x_s \geq \lceil \bar{x}_s \rceil \tag{13}$$

   is added to the linear program which is then solved again, after which we return to 1.

4 If $S > 1$, we temporarily add the bound (13) and run the linear program again without column and row generation.

5 Assume first that this linear program is infeasible. We track down this infeasibility by removing the bound (13) and then, for one set g at a time, run the linear program (without column and row generation) with the additional bound $x_g \geq 1$ where g is a group set routed on exactly one of the sets $g'(g)$ in G' which s is routed on. For each group set g which results in an infeasibility (there must be at least one) we block the possibility of routing on $g'(g)$ group sets with level greater or equal to the level of g. We also fix the value of x_s to $\lfloor \bar{x}_s \rfloor$, run the linear program again and return to 1.

6 Assume that the linear program is feasible. If s is a demand set or a set routed on one group set only, the bound (13) is added to the linear program which is then solved again, after which we return to 1. Otherwise, we keep (13) temporarily and, for one set g' at a time on which s is routed together with other fractional sets on s's level, we temporarily introduce the bound $x_{g(g')} \geq 1$ where $g(g')$ is the set on s's level which is routed on g' only. We then run the linear program without column or row generation for each such bound. If all these linear programs are feasible, we remove the bounds $x_{g(g')} \geq 1$, introduce permanently the bound (13), solve the linear program again and return to 1. Otherwise, the bound (13) is removed, the bounds $x_{g(g')} \geq 1$ are introduced corresponding to each infeasible linear program, solve the linear program again and return to 1.

7 If the highest level fractional group set variable represents a demand group or if there are no fractional group set variables at the level in question, we check the load sharing requirements for demands at the level in question. For each load sharing family we find (if possible) a reliability class where the load sharing requirement is violated and block the class for all demands except one. The demand which is not blocked is either one whose routing on a set in the reliability class has received a positive lower bound earlier in the heuristic, or, if no such bound has been set, a demand with maximum number of groups routed on group sets in the reliability class. Then the linear program is solved again.

8 If all load sharing requirements are satisfied, we pick a demand set g where $\bar{x}_g$ has largest fraction, include the bound $x_g \geq \lceil \bar{x}_g \rceil$ and solve the linear program again.

This process is continued until we obtain a solution of the linear program with all group set variables integer.

Finally, a post processing is done where adjacent group sets on which the same group sets are routed are concatenated to form longer group sets.

3.8 Typical problem size and solution time

A typical real problem will have many groups fixed from previous years and have around 1500 initial constraints and 4000 initial variables. The path variable generation process will typically generate around 500 additional constraints and around 1000 additional variables. The solution time for such a problem is around 50 minutes on a 486 33 MHz PC.

4 Stand-by capacity planning

4.1 Introduction

According to Telenor standards every group g of level 140 or 155 Mbit/s shall have a *stand-by group* which can be set up if the group g fails. This means that g's stand-by group cannot belong to any of the reliability classes of g.

Ideally, the establishment of stand-by group sets should be done simultaneously with the routing and grouping of the demand sets. This has not yet been implemented. Furthermore, the planners have expressed the wish for a tool which plans the stand-by groups based on the capacity in the network which remains after the routing and grouping of the demand sets have been determined.

The stand-by group sets will be directly routed on so-called *reserve group sets* which has level 140 or 155 Mbit/s. 140 Mbit/s stand-by group sets must be directly routed on 140 Mbit/s reserve group sets, and 155 Mbit/s stand-by group sets must be directly routed on 155 Mbit/s reserve group sets.

However, 140 Mbit/s reserve group sets may be directly routed on 155 Mbit/s group sets. The reserve group sets are permanently set up and can be used by different stand-by group sets corresponding to different failure situations.

4.2 The integer programming model

An integer programming model has been implemented for stand-by group set planning. In order to describe this model the following notation is used:

Subscripts:

g	group set for which a stand-by group set is to be found
s	stand-by group set
r	reserve group set
c	reliability class

Constants:

$\lvert g \rvert$	number of groups in the group set g
$g(s)$	the group set for which the group set s is stand-by
C_r	cost per group in reserve group set r
C_s	cost per group in stand-by group set s
$F_{g'}$	number of free 140/155 Mbit/s channels in group set g'
N^C	number of reliability classes

Problem variables:

x_s	number of groups in stand-by group set s
x_r	number of groups in reserve group set r

Some of the x_s and x_r variables may have lower bounds L_s and L_r respectively, which may have been carried over from earlier time periods or otherwise specified by the user:

$$x_s \geq L_s$$

$$x_r \geq L_r$$

We write:

$s \in g$	if the group set s is stand-by for groups in group set g
$s \in r$	if the group set s is directly routed on group set r
$s \in c$	if the group set s is stand-by for a group set in reliability class c

The constraints of the integer program are:

$$\Sigma_{s \in g}\, x_s \geq \lvert g \rvert \quad \text{for all group sets } g \qquad (14)$$

$$x_r - \Sigma_{s \in r, s \in c}\, x_s \geq 0 \quad \text{for all reserve group sets } r \text{ and reliability classes } c \qquad (15)$$

$$-\Sigma_{r \in g'}\, x_r \geq -F_{g'} \quad \text{for all group sets } g' \text{ on which reserves may be routed} \qquad (16)$$

All variables are required to be integer.

The objective function to be minimized is

$$\Sigma_r\, C_r x_r + \Sigma_s\, C_s x_s \qquad (17)$$

We shall assume that the cost coefficients C_r and C_s have the following structure:

$$C_r = C^{R0} + C^{R1} \times \#(\text{group sets which } r \text{ is directly routed on})$$

$$C_s = C^{S0} + C^{S1} \times \#(\text{reserve group sets which } s \text{ is directly routed on})$$

This integer program will be solved by first solving the linear programming relaxation of the problem. This is followed by an integer programming heuristic.

4.3 Solving the linear programming relaxation

The number of x_r and x_s variables and type (15) constraints in the linear program is too large to be included in the model from the outset. They will be generated dynamically during the solution procedure.

Initially, a restricted version of the linear program is set up and solved which has

- all variables of type x_s with $L_s > 0$
- all variables of type x_r on which x_s variables with $L_s > 0$ are directly routed
- all variables of type x_r with $L_r > 0$
- for each reserve group set level, at least one variable of type x_r between every pair of nodes where it is possible to establish a reserve group set at the relevant level
- all constraints of type (14) and (16)
- all constraints of type (15) involving x_s variables with $L_s > 0$
- all relevant bounds.

Initially, there may, for each reserve group set level, exist node pairs with no x_r on which x_s variables with $L_s > 0$ are directly routed and no x_r with $L_r > 0$. An x_r variable for such a node pair can if possible be established by finding a shortest path in the network of group sets with edge lengths C^{R1}.

Based on the shadow prices associated with the optimal solution of this restricted linear program new groups with associated variables of types x_r and x_s and new constraints of type (15) are generated, and a new (restricted) linear program is solved. This process is continued until it can be ascertained that we have an optimal solution to the complete linear program.

In order to describe the details of this iterative process we assume that we are at a stage where we just have found an optimal basis for a restricted linear program which, in addition to the initial constraints and variables, may include

- some additional variables of type x_r and x_s
- the constraints of type (15) involving the additional x_s variables.

The following notation is used for the shadow prices associated with a basis for the restricted and the complete linear program:

Constraint type:	Shadow price:
(14)	π_u
(15)	π_{cr}
(16)	$\pi_{g'}$

The optimal solution to the restricted linear program is obviously a feasible solution to the complete linear program. We shall now establish whether it is optimal or whether we can find one or more not yet generated x_s variables with negative reduced cost which we can introduce into the linear program together with the necessary constraints and continue the solution process.

First, we shall give a procedure for extending our optimal basis for the restricted linear program to an imagined feasible basis for the complete linear program.

We imagine for a moment that we have added all the not yet generated variables and constraints of type (15) to the restricted program. We introduce a basic variable corresponding to each of the not generated constraints. For each not generated type (15) constraint corresponding to a generated reserve group r we assume that the corresponding slack variable is basic. This implies that the corresponding $\pi_{cr} = 0$. Amongst the not generated type (15) constraints corresponding to a not generated reserve group r we arbitrarily pick one constraint corresponding to which we let the variable x_r be basic. For the other constraints we let the slack variables be basic.

We see that using the procedure described above we can arrive at a variety of bases depending on which constraints we choose to correspond to the basic x_r variables. Each of these bases has its own set of shadow prices associated with it. When we want to price out a not yet generated x_s variable, we will not use any of these sets of shadow prices, *but a convex combination of them*. This convex combination (for which we shall use the same π notation) will be specified later. From linear programming theory we know that if all not generated variables price out non-negatively using such a convex combination of shadow prices, the solution is optimal.

Any convex combination of the shadow prices for a not generated constraint is non-negative. Therefore, the slack variables of the not generated constraints price out non-negatively.

All the not generated x_r variables are basic and consequently price out zero.

Therefore, when we search for new candidates in the column generation process it suffices to consider x_s variables.

If s is a stand-by group set, the reduced cost for x_s using the convex combination is

$$C_s - \pi_{g(s)} + \Sigma_{c \in g, r \ni s} \pi_{cr}$$

Here $\pi_{g(s)}$ and π_{cr} for generated x_r variables are known.

Since not generated x_r variables are basic, they have reduced cost equal zero:

$$C_r - \Sigma_c \pi_{cr} + \Sigma_{g' \ni r} \pi_{g'}$$

which gives

$$\Sigma_c \pi_{cr} = C_r + \Sigma_{g' \ni r} \pi_{g'} = C^{R0} + \Sigma_{g' \ni r} (C^{R1} + \pi_{g'})$$

For each of the basis extensions described above, only one π_{cr} in the sum on the left hand side is different from zero.

We will now specify the convex combination mentioned earlier. We pick a reserve group set between the end nodes of r, $\bar{r}$ say, with maximum value, and set

$$\pi_{cr} = (\pi_{c\bar{r}} / \Sigma_c \pi_{c\bar{r}}) \Sigma_c \pi_{cr} = \pi_{c\bar{r}} [C^{R0} + \Sigma_{g' \ni r} (C^{R1} + \pi_{g'})] / [C^{R0} + \Sigma_{g' \ni r} (C^{R1} + \pi_{g'})]$$

This convex combination does not have any scientific justification, actually any combination would do in principle. It has been chosen in order to try to keep the number of generated rows and columns down.

For each group set g for which a stand-by group set is to be set up, the process of finding the stand-by group set with the least reduced cost is carried out in two steps.

First, the least cost reserve group set between every node pair is found using Floyd's shortest path algorithm on the network of admissible group sets (i.e. group sets which do not share reliability codes with g) where the costs are given by π_{cr} for already generated reserve group sets r and $C^{RI} + \pi_{g'}$ for group sets g' on which new reserve group sets can be directly routed.

Then Dijkstra's algorithm is used to find the least cost stand-by group directly routed on the reserve groups where the costs of the reserve groups are given by $C^{SI} + \Sigma_{c \ni g} \pi_{cr}$.

4.4 Integer programming heuristics

When the linear programming relaxation of the integer program has been found a lot of the group set variables will normally be fractional.

The basic principle used in the heuristics is to solve the linear program and inspect the solution to see if it contains fractional group set variables. If so, new bounds are added to the linear program. The linear program is solved again and checked for fractional group set variables which give rise to additional bounds, etc. This process is continued until all the group set variables are integer.

We shall now give a general description of the heuristic.

Assume that we are at a stage where we have solved the linear program where we denote the values of the variables with bars over the variable name. We check if there are any group set variables with fractional value. If so, we select one of these, r say, whose fractional part is maximum where priority is given to the reserve group set variables. We then solve the linear program again with the additional bound $x_r \geq \lceil \bar{x}_r \rceil$ (or $x_s \geq \lceil \bar{x}_s \rceil$).

This procedure is repeated until all group set variables have integer values.

5 Related work

RUGINETT contains also a module which models SDH ring structures. This module will be described elsewhere.

RUGINETT can easily be extended to include decision variables for system sets. If system set variables are included in the model, RUGINETT would, in addition to suggesting a routing and grouping plan, come up with proposals for the establishment of new transmission systems. This will be pursued later.

6 Acknowledgement

The author wishes to thank Ø. Eriksen for valuable assistance in setting up and implementing RUGINETT.

Status

International research and
standardization activities
in telecommunication

Editor: Tom Handegård

Introduction

BY TOM HANDEGÅRD

This is the fifth issue of Telektronikk containing the Status section on international research and standardisation activities. The contributions made so far within the different telecommunication areas are listed in Table 1. In addition, overview papers on ITU, EURESCOM, TINA-C and ACTS have been provided. The papers all conveys information about important organisations, projects and ongoing processes often difficult to acquire for those not directly involved.

In this issue, status reports on the following telecommunications areas are included:

- Broadband
- Terminal equipment and user aspects
- Telecommunications management network (TMN).

Table 1 List of contributions to the Status section

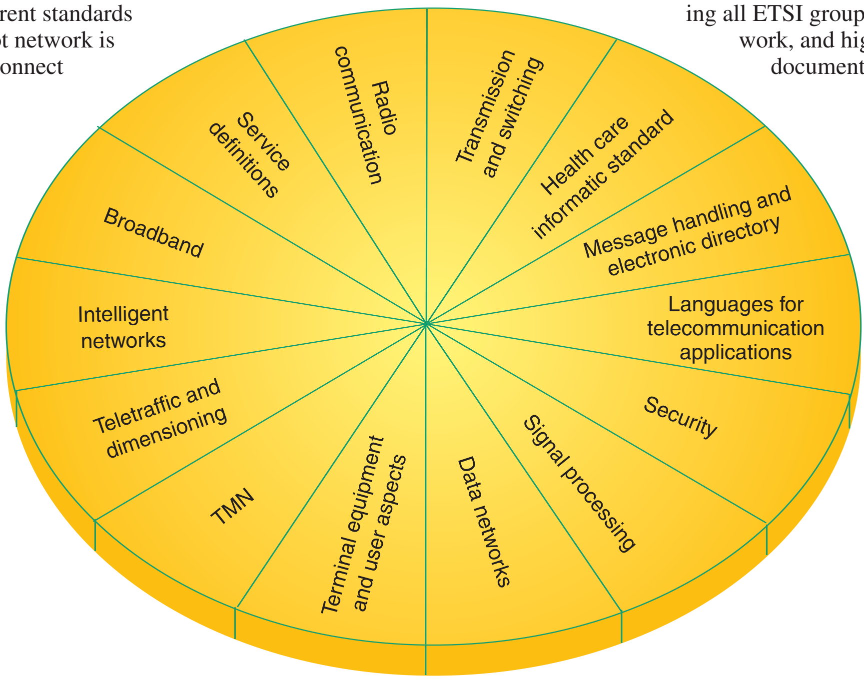
Issue No.	Study area	Editor
4.93, 3.94	Service definitions	Elisabeth Tangvik
4.93, 3.94	Radio communications	Ole D Svebak
4.93	Transmission and switching	Bernt Haram
4.93, 3.94	Intelligent Networks	Endre Skolt
4.93, 4.94	Broadband	Inge Svinnset
1.94, 4.94	Terminal equipment and user aspects	Trond Ulseth
1.94	Signal processing	Gisle Bjøntegaard
1.94, 4.94	Telecommunications Management Network (TMN)	Ståle Wolland
1.94	Teletraffic and dimensioning	Harald Pettersen
1.94	Data networks	Berit Svendsen
2.94	Languages for telecommunication applications	Arve Meisingset
2.94	Message Handling and Electronic Directory	Geir Thorud
2.94	Security	Sverre Walseth

The paper on broadband is provided by Mr. *Inge Svinnset*, and presents the European ATM pilot network that was put into operation in November 1994, and the effort undertaken by EURESCOM to provide specifications for a European ATM network. EURESCOM has provided specifications for a VP crossconnect network and is currently working on specifications for a VC switched network. Both sets of specifications extends current standards from ITU-T. The pilot network is currently a VP crossconnect network with no signalling. Connection establishment thus involves manual procedures and management systems. The pilot may be followed by a switched network, but not decision has been made on this yet. The pilot currently involves 18 public network operators, each responsible for one or more crossconnect nodes.

In his paper on terminal equipment and user aspects, Mr *Trond Ulseth* covers four ETSI/UTU-T issues:

- Channel aggregation, which is the technique used to implement applications in the ISDN requiring a bandwidth larger than 64 kbit/s,
- File transfer in ISDN, where ETSI has standardised two protocol profiles,
- Pictograms for videotelephony, where ETSI now has a standard ready, and
- PSTN telephones, where ETSI is working on a standard on test principles, enabling, if successful, providers to use a single set of test procedures to test all PSTN telephones in Europe.

The last paper, provided by Mr *Ståle Wolland*, gives an up to date overview of the TMN-work in ETSI, listing all ETSI groups involved in TMN-work, and highlighting important documents.



Eurescom work on ATM Network Studies and the European ATM Pilot Network

BY INGE SVINNSET

Background

Availability of high capacity fibre optics and advanced switching technology is now making broadband service provision feasible at data rates up to 155 Mbit/s (later 565 Mbit/s). This has resulted in intense international activity for standardisation and realisation of broadband services and networks in the context of B-ISDN (Broadband Integrated Services Digital Network) and ATM (Asynchronous Transfer Mode). A B-ISDN is a network that can support services requiring a capacity greater than the primary rate of ISDN. ATM has been chosen by ITU-T as the transfer mode for B-ISDN. In ATM the user information is segmented into cells of 53 octets. Cells belonging to one connection do not occupy pre-allocated time slots but may arrive in random order. The information identifying which connection a cell belongs to (the routing field) can be found in the cell header which is the first 5 octets of the cell. Of the other 48 octets some are used for adaptation layer functions, the exact number depending on the service.

The routing field is divided in two parts, the Virtual Path Identifier (VPI) field and the Virtual Channel Identifier (VCI) field. The VPI field is 8 bits at the User Network Interface (UNI) and 12 bits at the Network Node Interface (NNI). The VPI identifies the Virtual Path (VP) that the cell belongs to and has only local significance. The VCI field is 16 bits at both the UNI and the NNI. It identifies the Virtual Channel (VC) within the VP that the cell belongs to and also the VCI has only local significance.

The bundling of VCs into VPs enables us to organise a two level network where VC connections can be switched separately in a VC switch or the whole bundle of VCs in a VP may be switched altogether in a VP switch (‘bundle switching’). VP connections may be established by the Network Operator as a tool for managing the VC connections in the network or may be established end-to-end between users. The last case opens up for possible Virtual Private Networks inside a public ATM network.

The European ATM Pilot Network

BT, DBP Telekom, France Telecom, STET (Italy), IRITEL (Italy) and Telefonica (Spain) signed in November 1992 the ‘Memorandum Of Understanding among the participating Operators for the installation of at ATM Pilot’. The intention behind the launching of this pilot network was to support an evaluation of ATM technology, to check proposed ATM standards and the specification work of the Eurescom project P105 (see below) and, not the least, start the process towards the first trans-European broadband network. Norwegian Telecom joined the pilot in March 1993 together with Belgacom, PTT Telecom Nederland, Telecom Finland, Telefones de Lisboa e Porto, Telia AB (Sweden) and Swiss PTT Telecom. Later, Tele Danmark, Telecom Portugal, Telecom Eireann, Austrian PTT and ATC Finland have also joined the Pilot, making it a total of 18 partners. Telefones de Lisboa e Porto and Telecom Portugal have later merged to form Portugal Telecom SA, and STET and IRITEL have merged into Telecom Italia.

Each participating country is in the Pilot responsible for one international VP Crossconnect node, but in most countries extra Crossconnect equipment from Different manufacturers have been provided constituting a national branch of the network.

The VP crossconnect is like a switch but lacks signalling. The establishment of connections will thus take more time since manual procedures and a management system will be involved. To allow for use of already installed transmission systems the minimum requirement has been set to bi-directional 34 Mbit/s PDH links for interconnection of the international nodes. Additional links may use 140 Mbit/s PDH or 155 Mbit/s SDH. The international node in Norway is connected both to Gothenburg and to Copenhagen with a 34 Mbit/s PDH link.

The official opening of the Pilot will be 24 November this year. The Pilot users will be selected by the National Operators. Among others, various RACE[1] testbeds are planned to be interconnected. Since the Pilot will not offer a commercial service, the uses are picked out to gain experience with various services over an ATM network, and the users should themselves be interested in gaining this experience from an application or research point of view. Feedback from the users will be very valuable for the network operators when later launching an ATM commercial service.

In the Norwegian branch of the network, Norwegian Telecom Research is interconnected to the Norwegian international ATM crossconnect with a 155 Mbit/s SDH link. In addition to this Alcatel Norway, Backup Centralen, the University of Oslo and the Norwegian Computing Center are for the time being potential Pilot users.

The international ATM crossconnect node in Oslo is an Alcatel A1000 crossconnect. The other participating operators have chosen crossconnects from Alcatel, AT&T, Netcom (GDC) and Siemens. For the time being the Pilot is scheduled to last until June 30, 1995, with a possible extension to allow ore users to take part. After that the Pilot may be followed by a switched ATM network, but no decision has been taken on that matter yet. The benchmark services for the network are SMDS/DBDS, Frame Relay, ATM circuit emulation and a direct ATM VP service which provides ATM to the customer premises.

During the lifetime of the ATM Pilot, the operators will use the network themselves for service trials and for performance measurement experiments. The performance measurements have two aspects. Firstly, we are interested in measuring live traffic generated by the users. This means to register, at certain measurement points, the time between arrival of ATM cells belonging to a certain connection. This will be a basis for characterising the user generated cell stream, and will increase our knowledge about how much traffic (i.e. how many users) we can allow to share the capacity of for instance and SDH link. This will again be a basis for dimensioning an ATM network knowing how much traffic demand we can expect. It will also be a basis for forming the contract between a user of an application and the Network Operator. This contract should be formed to regulate the intensity of the information flowing from the user into the network in an optimal way. This is not only important for the accounting, but also to avoid that a too heavy information flow disturbs other users’ service because of buffer overflow in the network.

[1] *Research and development in Advanced Communications technologies in Europe – the EU research project on telecommunications.*

The other aspect of the measurement experiments is to observe the amount of errored and lost ATM cells and how this affects the quality of service. This will give input to the quality of service requirements for the various applications and will also have impact on dimensioning of the necessary equipment.

The VP crossconnect network specifications

EURESCOM (European Institute for Research and Strategic studies in Telecommunications) was established in 1991. It is a research and development institute jointly owned by 26 public network operators. A brief description of EURESCOM and some of its activities can be found in (1) and (2). The work within EURESCOM on ATM started in October 1991 with the project P105 on European ATM Network Studies. The main objective was to provide specifications beyond existing standards for the development of an early European ATM pilot network. The project had 13 partners and all the major public network operators took part. The project leader was Pierre Adam, France Telecom. This project was limited to the specifications for a VP crossconnect network. A major effort in this project was therefore to provide specifications for handling of establishment, modifications and release of VP connections by use of a management system.

The flexibility of ATM allows the user to ask for a wide range of bandwidth or capacities, only limited by the maximum system bandwidth and the available resources. The bandwidth given to the user for a specified connection is part of a contract between the user and the network. In this project and in the ATM Pilot the bandwidth specification is limited to the maximum rate (peak rate) of the connection and the cell delay variation (CDV) tolerance. The CDV tolerance allows for network induced variations in the rate due to buffers in network components (including the effect of a local network). The user may use the bandwidth as he wants within the limits defined by the contract with the Network Operator. Traffic contracts between the Network Operator and users, procedures for deciding if a bandwidth demand can be met (i.e. Connection Acceptance Control – CAC) and procedures for controlling that the actual used bandwidth is within the contract (Usage Parameter Control – UPC) have thus also been a necessary and a very important part of the specifications.

Dimensioning methods, routing and addressing were also studied within this project and key parameters for charging and accounting were identified. Since the user naturally is expecting a certain Quality of Service, a study was performed within the project defining ATM network performance parameters, i.e. requirements for cell error ration, cell loss ration, cell transfer delay, etc. Measurement procedures for the identified performance parameters were studied in a separate task.

Another aspects that has been studied is the interworking with other networks such as ISDN, the potential use of satellites in the network, physical interfaces and protocols over the interfaces. Norwegian Telecom participated in various subtasks in the project.

Specifications for VC switched network

In September 1993 the EURESCOM P302 project ‘European switched VC-based ATM network studies’ was started up as a follow-up of the P105 project. The project has 16 partners and all the major public network operators are participating. The project leader is Pedro Chas Alonso, Telefonica.

The P105 project treated only Virtual Paths set up by management procedures. Since then the first recommendations on signalling have been released by ITU. The recommendations Q.2931 for the B-ISDN access signalling and Q.2764 for the network signalling covers capability set 1 of B-ISDN (see (3)). These recommendations build on the SS7 for ISDN and are limited to the establishment and release of point-to-point VC connections within an already established VP connection. P302 will build on these recommendations and the work in ITU-T SG 11 to extend them to multi-connection calls and to allow for renegotiation of connection parameters (i.e. B-ISDN capability set 2 step 1).

Of the areas of study in the P302 project we can mention signalling interfaces, network elements, service identification, network aspects of a direct connectionless service, management aspects, resource allocation and network performance.

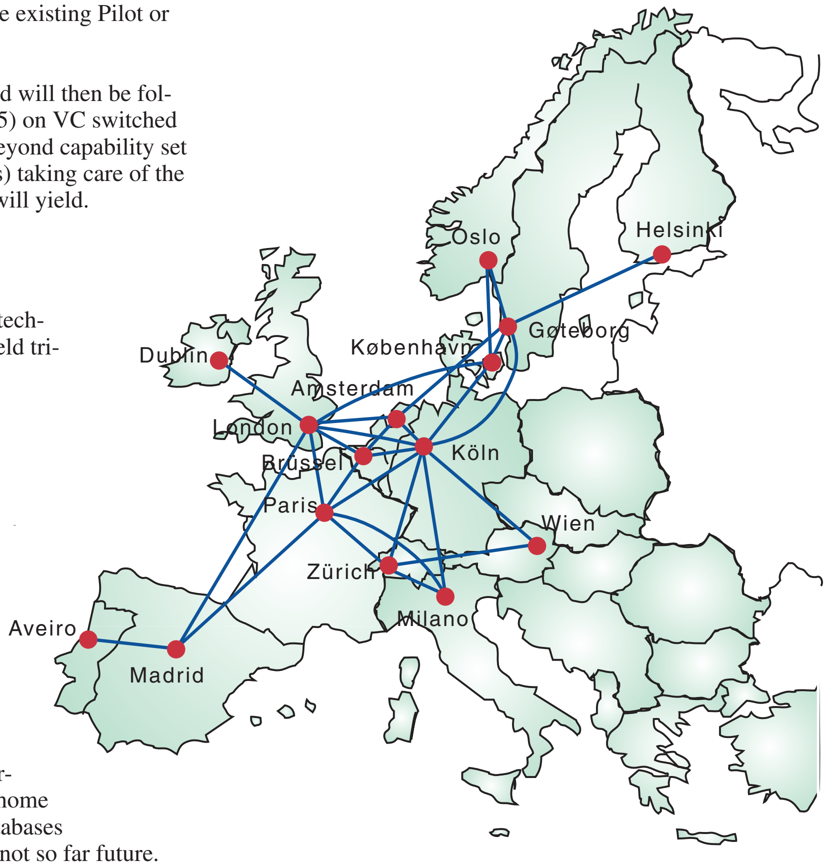
While the set-up procedure in the ATM Pilot involves a telefax procedure, manual operation and the management system, set-up of a switched connection should be possible end-to-end in only a few seconds by means of signalling. The new capabilities also include the possibility to change bandwidth during the lifetime of a connection by means of signalling and to set up multi-connection calls between two end users (i.e. multi-media applications).

Another important aspect is to take care of services with different requirements for quality of service. Data applications are loss sensitive but the requirements on transfer delay and CDV is not so strict. This may be taken care of by using long buffers in the switching and crossconnect nodes. Voice and video services are delay sensitive (apart from broadcasting), CDV sensitive but not so loss sensitive. These services should thus see a short buffer. In addition to this a ‘best effort’ service has been a study item in international forums (e.g. ATM Forum). This service is a very important candidate user of the network and should be a low cost service with low quality of service guarantee and using momentarily free capacity in the network.

The implementation of different services with so different quality of service demands and services for which the information flow rate of the connections varies in time (variable bit rate sources) is a great challenge for the future. To fully exploit this we will have to introduce more advanced contracts between user and network than the one used in the Pilot. We will have to see buffer management schemes in the switching and crossconnect nodes to allow for different treatment of cells belonging to connections with different service needs. This means that more advanced CAC procedures and parameter control procedures will have to be defined in this project to allow for a good utilisation of the network resources.

The purpose of the P302 project is to provide specifications for a first European switched ATM network. This network can either be an enhancement of the existing Pilot or a new Pilot network. The time schedule for enhancing the existing Pilot or launching of a new network is mid 1996.

The P302 project will end by March 1995 and will then be followed by another EURESCOM project (P515) on VC switched ATM networks with enhanced capabilities beyond capability set 2 step 1 (e.g. point-to-multipoint connections) taking care of the capabilities that the next signalling releases will yield.



Concluding remarks

The first international field trials with ATM technology are now taking place. One of these field trials is the European ATM Pilot network involving 18 public network operators. The first initiative for this Pilot was the research and development institute EURESCOM, jointly owned by 26 public network operators, with the project on European ATM network studies. The Pilot is intended to be followed up as the ATM technology emerges, and specifications are already being made within EURESCOM for a switched ATM pilot network.

The need for interconnection of LANs is only one of the driving forces for a broadband network. Video telephony, video conferences, multi-media calls, video-on-demand, home shopping, network access to libraries and databases are some applications that we will see in the not so far future. The different services will have different requirements for quality of service. This will be a great challenge for the future if the services are to be carried by the same network. Today, services like telephony, data and video (broadcasting) are carried in different networks. Even though these services have very different bandwidth, real time and quality demands, the flexibility of ATM will make it possible to integrate such services in the same network. The economy of this integration together with the need for providing new services, will be a driving force towards a mature broadband network needing all the functionality that ATM can give.

References

1 Fogerty, K D, Katzeff, K. EURESCOM: a new opportunity for collaborative R&D in European telecommunications. *British Telecommunications Engineering,* 12, 294–301, 1994.

2 Skolt, E. A presentation of EURESCOM. *Telektronikk,* 1 (90), 196–198, 1994.

3 ITU-T. *Study Group 11.* Report on the meeting held in Geneva, September 1994.

Terminal equipment and user aspects

BY TROND ULSETH

Most of the standardisation work on terminal equipment is addressing terminals and applications for the ISDN. However, there are applications that will give better performance at higher bitrates than 64 kbit/s, e.g. audio-visual applications. To offer to the customers higher bandwidth in the narrowband ISDN, channel aggregators are required. This presentation presents the status of ITU-T and ETSI work on recommendations/standards on channel aggregation. Two new ETSI standards on file transfer in the ISDN, and an ETSI standard on pictograms for point-to-point videotelephony are also presented. Finally, status for the work on ETSI standards for PSTN telephones are presented.

Channel aggregation

The bandwidth of a B-channel in the ISDN is 64 kbit/s. There are a number of applications that require larger bandwidth, e.g. transmission of moving images. In the future B-ISDN will offer larger bandwidth, and there is also standardisation work on a N x 64 kbit/s bearer service in the ISDN. However, to implement applications requiring a bandwidth larger than 64 kbit/s today, terminal based channel aggregation is required.

Even though a videotelephony connection can be established using one B-channel, there are three videotelephony modes that require use of two B-channels. This is the simplest form of channel aggregation, aggregating two B-channels. To aggregate these two B-channels the principles specified in ETS 300 144, the ETSI equivalent to ITU-T Recommendation H.221, are used. The present version of ETS 300 144 describes codepoints and a numbering scheme for aggregating up to 6 B-channels. For each B-channel 1.6 kbit/s is allocated to inband signalling and synchronisation of the B-channels.

In North America an industry consortium developed a system which is known as “Bonding”. The system can aggregate up to 63 B-channels, and split one B-channel into sub-channels. The system is standardised by Telecommunications Industry Association (TIA). Five different modes are defined:

- Mode 1. After parameter negotiation no monitoring of the status of each B-channel is provided. The delay difference between each B-channel is compensated.
- Transparent mode. No equalisation or parameter negotiation takes place.

These modes are mandatory for a channel aggregation unit implementing the TIA standard.

The following modes are optional:

- Mode 0. After parameter negotiation no delay equalisation is performed.
- Mode 2. This mode supports data rates that are multiples of 63/64 of the bearer rate. A monitoring function of each B-channel provides a continuous check for equalisation of the delay and an end-to-end bit error rate test for each B-channel.
- Mode 3. This mode supports data rates that are integral multiples of 8 kbit/s. An inband monitoring function provides a continuous check for delay equalisation and end-to-end bit error rate test. The overhead is provided by adding bandwidth, most likely an additional B-channel.

The parameter negotiation of the B-channels are not compatible with the ITU-T/ETSI standards for audio-visual services. Some of the ‘Bonding’ modes can be used for audio-visual communication, but the main applications are data communications.

ISO/IEC JTC 1/SC 6/WG 6 has prepared a Committee Draft (CD 13871). The draft is based on the principles of the TIA standard.

ITU-T SG 15 is working on a recommendation on channel aggregation based on the principles used for audio-visual communication. The number of B-channels aggregated can be extended to 24 in a flexible system.

ETSI STC TE4 is working on a similar ETSI standard. The European Commission has given a standardisation mandate to ETSI on channel aggregation (BC-T-310). The mandate recommends to align the ETSI standard with standards and recommendations of ISO/IEC and ITU-T, a recommendation which is fully supported by ETSI.

The ITU-T Recommendation/ETSI Standard defines four different modes of operation:

- Mode B1. This mode is corresponding to mode 1 as defined in ISO/IEC CD 13871.
- Mode H2. This is a mode using the framing specified in ETS 300 144. For all terminals intended for communication with terminals conforming to ETS 300 144, this mode is mandatory. The in-band signalling occupies 1.6 kbit/s on each B-channel.
- Mode B2. This mode is corresponding to mode 2 as defined in ISO/IEC CD 13871.
- Mode B3. This mode is corresponding to mode 3 as defined in ISO/IEC CD 13871.

The channel aggregation function may be implemented in the terminal equipment, a *Multiple Channel Equipment* (MCE), or in a separate unit, a *Channel Aggregation Unit* (CAU). The CAU will provide the necessary channel aggregation functions for *Single Channel Equipment* (SCE). An example of interconnection between a Single Channel Equipment and a Multiple Channel Equipment is described in Figure 1, and the interconnection between two Single Channel Equipment is described in Figure 2.

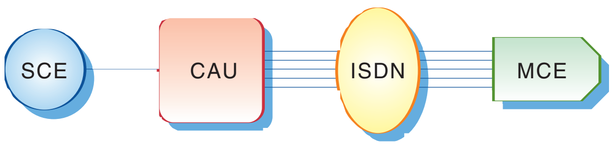


Figure 1 Interconnection between a Single Channel Equipment and a Multiple Channel Equipment

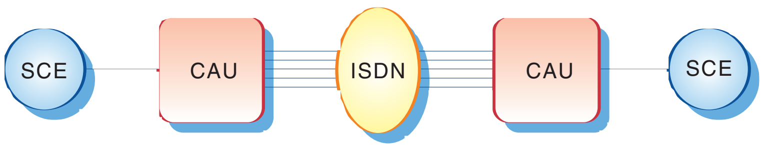


Figure 2 Interconnection between two sets of Single Channel Equipment

	Mode B1	Mode H2	Mode B2	Mode B3
Aggregation overhead	None	2.5 % in the additional channel	1.5 %	64/56 kbit/s
Dynamic rate changes	No	Yes	Yes	Yes
Exact multiples of 64/56 kbit/s	Yes	Yes (only framed according to ETS 300 144/ITU-T Recommendation H.221)	No	Yes
Interworking with MCE (audio-visual)	No	Yes	not applicable	No

Figure 3 Mode properties

The mode properties are summarised in Figure 3.

The applicability of the different modes defined in the ETSI standard is described in Figure 4.

The ITU-T Recommendation (H.244) is scheduled for resolution 1 approval in February 1995.

The ETSI standard was approved at STC level (TE4) in September 1994 for presentation to TC TE for approval to be submitted on PE.

File transfer over the ISDN

ETSI has standardised two file transfer profiles for use over ISDN. One profile is based on a full OSI protocol stack using the file transfer protocol FTAM. The FTAM profile is described in ETS 300 388. The second file transfer profile is based on the Eurofile file transfer protocol. The Eurofile profile is described in ETS 300 383. In the standardisation work attempts were made to agree on a single protocol to be standardised by ETSI. It became obvious at an early stage that there was equal support among the members of ETSI for the two alternatives. It was therefore decided to standardised both profiles. These profiles (layers 1 to 7) are described in Figures 5 and 6.

The lower layer (layer 1 to layer 3) protocols are as specified in ETS 300 080 for the DTE/DTE connection using ISDN circuit mode.

Standardisation of Pictograms for point-to-point videotelephony

Modern telecommunication terminals may offer a number of functions to the user. The use of these terminals may be complicated. It is important to create user friendly man-machine interfaces.

The work of ETSI TC HF is addressing these issues. Some Human Factors issues are not suitable for standardisation, but recommendations and guidelines are presented in ETSI Technical reports.

ETS 300 375 on pictograms for point-to-point videotelephony is the result of a number of experiments carried out by ETSI STC

Mode	AV/non-AV	Mode applicable when ...
B1	Audio-visual	the remote end is not an MCE, and the remote CAU does not support Mode H2, and Mode B3 is considered too inefficient or is not available; bitstream treated as Unrestricted Digital Information only; No inband management communication between SCE and CAU, so external means must be used to set bitrate
	Non-audio-visual	exact multiple of 64/56 kbit/s required, absence of dynamic rate change tolerable, B3 considered too inefficient
B2	Audio-visual	not applicable – B2 does not provide suitable bitrates
	Non-audio-visual	exact multiple of 64/56 kbit/s not essential, dynamic rate change desired
B3	Audio-visual	the remote end is not an MCE, and the remote CAU does not support Mode H2; dynamic rate change more important than efficiency; bitstream treated as Unrestricted Digital Information only; No inband management communication between SCE and CAU, so external means must be used to set bitrate
	Non-audio-visual	exact multiple of 64/56 kbit/s required, dynamic rate change more important than efficiency
H2	Audio-visual	remote end is an MCE, or a CAU supporting Mode H2; Inband signalling according to ETS 300 144/ITU-T Recommendation H.221, so no external control is needed
	Non-audio-visual	not applicable (unless terminal equipment is conformant to ETS 300 144/ITU-T Recommendation H.221)

Figure 4 Applicability of channels aggregation modes

HF1. In total 14 organisations have contributed to the work either by providing pictogram proposals or by participating in the evolution process of the proposals. The standard defines pictograms for eight point-to-point videotelephony functions,

1 Videotelephone/telephone modes (switching between these modes)

2 Videophone camera ON/OFF

3 Videophone microphone ON/OFF

4 Videophone selfview ON/OFF

5 Videophone still picture ON/OFF

6 Videophone document camera ON/OFF

7 Videophone handsfree ON/OFF

8 Videophone loudspeaking ON/OFF.

The results of the evaluation that form the basis for this standard are reported in ETSI ETR 113.

The pictograms are shown in Figure 7.

PSTN telephones

Telephones for attachment to the Public Switched Telephone Network (PSTN) are an important part of the telecommunications terminal market in Europe. This fact is recognised by the European Commission, and large resources are spent on trying to harmonise the PSTN attachment standards in Europe.

The specifications for PSTN telephones in Europe are different both in terms of characteristics and in terms of test methods. ETSI STC TE4 is working on standards for PSTN telephones. A standard on test principles is now ready for Public Enquiry. If this standard is approved, a harmonised set of test principles for testing of PSTN telephones has been developed. This is an important improvement for the suppliers, one single set of tests can be used for testing the PSTN telephones in Europe. The next step is to prepare a standard where all requirements are based on the test principles of the test standard. The requirements standard will hardly be a harmonised standard, but efforts will be made to try to limit the number of national variations.

Layer	D-channel	B-channel
7	FTAM ISO 8571 profiled according to ETS 300 388 ACSE ISO 10607	
6	ISO 8823 (ITU-T X.226)	
5	ISO 8327 (ITU-T X.225)	
4	ISO 8073 (ITU-T X.224)	
3	ETS 300 102-1 (Q.931)	ISO 8208
2	ETS 300 125 (Q.921)	X.75 (NOTE 1) or ISO 7776
1	ETS 300 011 (I.431) or ETS 300 012 (I.430)	

Figure 5 Protocol Pillar for FTAM

Layer	D-channel	B-channel
7	ETS 300 075 profiled according to ETS 300 383 ETS 300 079	
6	null (NOTE 2)	
5	null	
4	null	
3	ETS 300 102-1 (Q.931)	ISO 8208
2	ETS 300 125 (Q.921)	X.75 (NOTE 1) or ISO 7776
1	ETS 300 011 (I.431) or ETS 300 012 (I.430)	

Figure 6 Protocol Pillar for Eurofile

NOTE 1: ITU-T Recommendation X.75 as modified in ETS 300 080

NOTE 2: The main purpose of layer 6, the conversion from the "abstract syntax" to the "transfer syntax" is not necessary, because in this case the abstract data syntax in layer 7 is identical to the transfer data syntax. Also all other features of the layer 6 are not used and therefore "null" is inserted for layer 6. The abstract syntax in layer 7 and the coding in layer 6 correspond to the data syntaxes DS I, DS II and DS III in ITU-T Recommendation T.101.

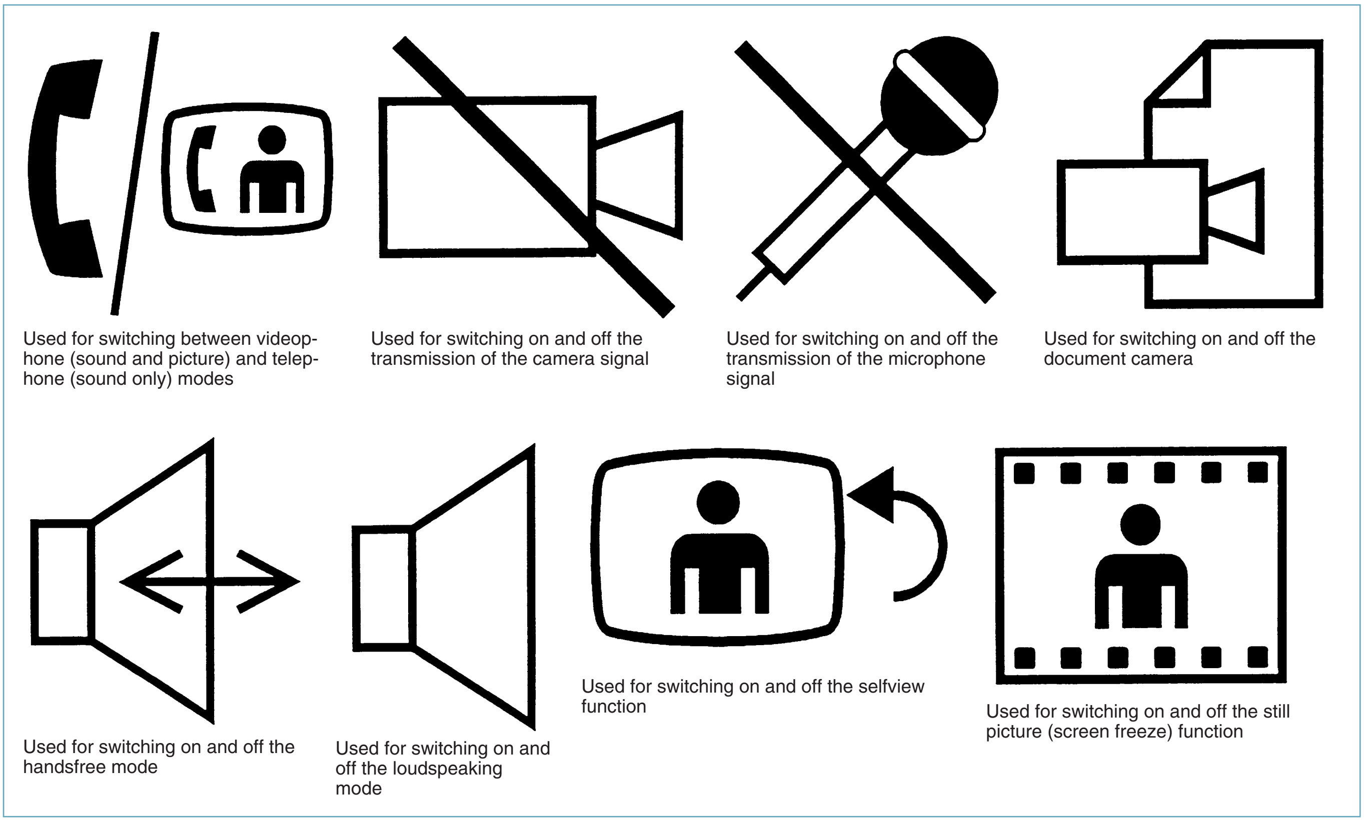


Figure 7 Pictograms for point-to-point videotelephony functions

Telecommunications Management Network

BY STÅLE WOLLAND

Introduction

Telecommunications Management Network (TMN) is an architectural concept for communication and processing of management information for tele- and data communications networks. The objective is to be able to implement management capability over a distributed management network (logically and/or physically) in a standardised way so as to be system and vendor independent. Interoperability and reuse of management resources are also important objectives.

In Telektronikk 1/94 the ISO/IEC JTC1 STC21 work within the management area was described. The following section will give an overview over some of the ETSI work on management. A number of Sub-Technical Committees (STCs) are actively doing work within Telecommunications Management Network (TMN).

There is related management work going on in Eurescom, ITU-T, Network Management Forum, RACE, ANSI, IEEE, TINA-C and other bodies.

ETSI

Work Items on TMN

ETSI has presently no work item pertaining to the X (inter domain TMN) and F (Work Station) interfaces. Only the Q (intra domain TMN) interface of TMN is addressed. Work on the Network Element view has received most attention so far. The first proposals for a Network Level view were proposed during the last year. At the last meeting of NA4 in September 1994 a first proposal for Managed Objects for a Service Management view was input based on RACE results (this work has also been submitted to ITU-T SG IV).

Presently, there is very limited security work additional to the STAG work within NA. Efforts to establish a Rapporteur's Group on security in NA4 was made at the last meeting in September. There are also initiatives to establish a security group at the STC NA level. The outcome of these initiatives remains to be seen.

Sub-Technical Committees working on TMN

The following is a brief introduction to some of the groups within ETSI that among other tasks also address management. Only the management part of their objectives has been described.

NA (Network Aspects)

NA4 (Network Architecture, Operation & Maintenance, Principles and Performance)

STC NA4 has overall responsibility within ETSI for co-ordination and general principles for operation, maintenance and network management in telecommunications networks and systems.

A number of so-called Rapporteur`s Groups are doing the detailed technical work in various areas:

- MNT (ISDN Customer access maintenance), VTM (Vocabulary for TMN)
- MAS (Definition of management architecture and services)
- GOM (Generic Object Model)
- SOC (Support Object Classes for TMN interfaces)
- OCA (Object Model for Customer Administration on TMN interfaces)
- ORM (Object Model for Routing Management on TMN interfaces)
- OTM (Object Model for Traffic Management on TMN interfaces)
- INM (TMN principles for IN).

Documents relevant to management under processing or issued by NA4 include:

- Final Draft prI-ETS 300 291 : 1994-05 – Functional Specification of Customer Administration (CA) on the Operations System/Network Element (OS/NE) Interface (presented to the NA plenary in October 1994)
- prI-ETS 300 292 : 1994-02 – Functional specification of call routing information management on the Operation System / Network Element (OS/NE) interface (not stable yet, further work to be done)
- Draft prI-ETS 300 293 : 1993-05 – Generic Managed Objects
- Work Item DI/NA43314 (proposed I - ETS) – Functional specification of Traffic Management on the NE/OS interface
- Work Item DI/43316 (proposed I - ETS) – TMN – Network Level Generic Object Classes
- Work item DTR/NA43315 (proposed TC-TR) – TMN Guidelines (second issue of DTR/NA43303)
- Work item DI/NA43319 (proposed I-ETS) – Technology independent path set-up
- Work Item DI/NA43320 (proposed I-ETS) DE/TM-2209 – Resource Management
- Work Item DI/NA43321 (proposed I-ETS) – Specification of the usage information on the OS/NE interface
- ETR 008 : 1990-12 – The method for the characterisation of the man-machine interface utilised by a Telecommunications Management Network (TMN)
- ETR 046 : 1992-07 – Telecommunications management networks information modelling guidelines
- ETR 078 : 1993-12 – TMN interface specification methodology [CCITT Recommendation M.3020 (1992)]
- ETR 037 : 1992-02 – Telecommunications Management Network (TMN) Objectives, principles, concepts sand reference configurations
- Work Item DTR/NA43206 (proposed ETR) – Migration to TMN
- Work Item DTR/NA43207 (proposed ETR?) – TMN Standardisation overview
- ETR 047 : 1992-09 – Telecommunications Management Network (TMN) Management Services
- ETR 048 1992-09 – Telecommunications Management Network (TMN) Management Services: Prose Descriptions

- ETR 088 : 1993-07 – Time / type of day dependent scheduling function support object classes
- TCRTR 008 : 1993-08 – Telecommunications Management Network (TMN) Vocabulary of terms.

NA5 (Broadband Networks)

STC NA5 has responsibility for co-ordination of studies of B-ISDN within ETSI including CPN and services.

Documents relevant to management under processing or issued include:

- Draft prETS 300 404 : 1994-03 – B-ISDN Operation and Maintenance (OAM) principles and functions
- ETR 089 : 1993-08 – Principles and requirements for signalling and management information transfer
- DE/NA-52210 B-ISDN – Management Architecture and Management Information Model for the ATM cross connect
- ETS 300 273 : 1994-03 – Metropolitan Area Network (MAN) Medium Access Control (MAC) layer management.

NA6 (Intelligent Networks)

STC NA6 has responsibility for defining the TMN principles for IN.

Documents relevant to management under processing or issued include:

- ETR 062 : 1993-05 – Baseline document on the integration of Intelligent Network (IN) and Telecommunications Management Network (TMN)
- TC-TR/NA60801 – IN Intra Domain Management Requirements for CS-2, DTR/NA6101 – Service and service feature interaction: service creation aspects, service management aspects and service execution aspects
- TC-TR/NA6- 61201 – Security Requirements for Global IN Systems (approved by STC NA6 in September 1994 – now for approval by TC NA).

NA7 (Universal Personal Telecommunications)

STC NA7 has among other objectives responsibility for identifying the management requirements for UPT.

Documents relevant to management under processing or issued include:

- Draft ETR NA-70306 – UPT Management Aspects, Draft ETR NA-05001 – UPT Management Services and Management Information Model.

SPS (Signalling, Protocols and Switching)

SPS3 (Digital Switching)

STC SPS3 has among other objectives responsibility for defining V5 interface management.

Documents relevant to management under processing or issued include:

- Work Item DE/SPS-33019 – Management of ATM switches
- Draft prETS 300 376-1 – Q3 interface at the Access Network (AN) for configuration management of V5 interfaces and associated user ports Part 1: Q3 interface specification
- Draft prETS 376-2 – Q3 interface at the Access Network (AN) for configuration management of V5 interfaces and associated user ports Part 2: Managed Objects Conformance Statement (MOCS) proforma
- Draft prETS 300 377-1 – Q3 interface at the local exchange (LE) for configuration management of V5 interfaces and associated customer profiles
- Draft prETS 300 377-2 – Q3 interface at the Local Exchange (LE) for configuration management of V5 interfaces and associated customer profiles Part 2: Managed Object Conformance Statement (MOCS) proforma
- Draft prETS 300 378-1 – Q3 interface at the Access Network (AN) for fault and performance management of V5 interfaces and associated user ports Part 1: Q3 interface specification
- Draft prETS 300 379-1 – Q3 interface at the Local Exchange (LE) for fault and performance management of V5 interfaces and associated customer profiles Part 1: Q3 interface specification.

TM (Transmission and Multiplexing)

TM2 (Transmission Networks Management, Performance and Protection)

STC TM2 has among other objectives responsibility for defining general principles for operation and maintenance of transmission networks.

Documents relevant to management under processing or issued include:

- DTR/TM-2207 SDH Network Level Information Model – Configuration Management Ensemble
- prETS DE/TM-2209 : 1994 Operations and Maintenance of Optical Access Networks (for STC approval in September 1994)
- ETS 300 150 : 1992-11 – Protocol suites for Q interfaces for management of transmission systems
- RE/TM-2202 – Revision of ETS 300 150 : 1992-11
- ETS 300 304 : 1993-04 – Synchronous Digital Hierarchy (SDH) Information model for the Network Element (NE) view
- RE/TM-2213 – Maintenance and enhancement of ETS 300 304 : 1993-04
- ETS 300 411 : 1994-03 – Performance monitoring information model for the network element view
- ETS 300 412 : 1994-03 – Payload configuration information model for the network element view
- ETS 300 413 : 1994-05 – Multiplex section protection information model for the network element view

- ETS 300 371 : 1993-08 – Plesiochronous Digital Hierarchy (PDH) information model for the Network Element (NE) view
- RE/TM-2223 – Maintenance and enhancement of ETS 300 371 : 1993-08
- DTR/TM-2223 – Software downloading process for use on Q interfaces
- DTR/TM-2221 – The application of ODP to the management of a transmission network level model
- DE/TM-2210 – Management of SDH transmission equipment
- DE/TM-2220 – SDH information model connection supervision function (HCS/LCS) for the Network Element view
- DE/TM-2218 – SDH Radio relay equipment information model for use on Q interfaces
- DE/TM-2216 – Object oriented information model for the protection of SDH NEs
- DE/TM-2207 – SDH network information model – configuration management ensemble
- DTR/TM-2222 – Management of Access Networks
- DE/T M-2224 – Management of fixed radio Access Networks
- DE/TM-2??? – Management of Access Networks providing interactive and / or distributive services
- DE/TM-2209 – Operation and maintenance of Optical Access Networks (OANs).

BTC (Business Telecommunications)

BTC1 (Private Networking Aspects of 64 kbit/s ISDN Business Communications)

STC1 has among other objectives responsibility for defining general principles and requirements on Corporate Network Management functions including control by the private operator of such networks.

BTC4 (Private Network Aspects of High Speed Business Communications)

STC4 has among other objectives responsibility for integration and co-ordination of high-speed services (LAN-interconnection etc.) by identifying the requirements for signalling, security and management services.

SMG (Special Mobile Group)

SMG6 (Operation & Maintenance)

STC SMG6 has among other objectives responsibility for specifying the network management functions of the GSM, DCS 1800 and UMTS systems.

Documents relevant to management under processing or issued include:

- D-ETR/SMG50501 Objectives and Framework for the TMN of the UMTS (to be finalised by the end of 1994).

SES (Satellite Earth Stations and Systems)

SES is responsible for all aspects relative to satellite communications including management.

Documents relevant to management under processing or issued include:

- ETS 300 160 : 1992-11 – Control and monitoring functions at a Very Small Aperture Terminal (VSAT)
- ETS 300 161 : 1992-11 – Centralised control and monitoring functions for VSAT networks.

ECMA TC32 (Communications, Networks and System Interconnection)

ECMA TC32 TG12 (Private Telecommunications Networks – Management)

TG12 is responsible for developing technical reports and standards for the management of Private Telecommunications Networks (PTNs) by extending and adapting general TMN to PTNs.