81/29

EISCAT
TECHNICAL
NOTE

LECTURES FOR EISCAT PERSONNEL
VOLUME II

by
Kristen Folkestad

KIRUNA
Sweden

LECTURES FOR EISCAT PERSONNEL

VOLUME II

by

Kristen Folkestad

EISCAT
N-9027 Ramfjordbotn
Norway

EISCAT Technical Note 81/29
Printed in Sweden
EISCAT Scientific Association
Kiruna, April 1981
ISSN 0349-2710

In volume I of "Lectures for EISCAT personnel", EISCAT Techn Note 79/19, a forewarning was given that the report might be followed up by a second volume describing "basic working principles and design philosophy".

The present product represents this second volume. The author must admit, though, that the content is not in full accord with the original intentions as quoted above.

At an early stage it was felt that the lectures ought to deal with some basic aspects of signal theory and circuit analysis, as a prerequisite for a sound understanding of signal treatments as related to the EISCAT observations. Proceeding along this line the presentations grew, perhaps somewhat out of proportion, both in length and in complexity.

Three of the four chapters deal with topics like harmonic analysis, Fourier transform techniques, correlation and convolution, signal sampling, analog and digital filtering and signal descriptions. The fourth chapter gives features of general antenna theory, but also presents particulars of the EISCAT antenna systems.

Whether the lecture series will continue with a third volume, seems uncertain at present. There are certainly design topics and working principles still worthwhile expounding in written form. Only time will show whether ideas which the author might have in this respect, will be realized.

CONTENTS

CONTENTS (continued) Page

9 HARMONIC ANALYSIS

9.1 Periodic functions

9.1.1 Fourier series

For a periodic function f(t) of the independent variable t, expressed as a Fourier series, we have:

$$f(t) = \frac{a_o}{2} + \sum_{n=1}^{\infty} (a_n \cos n\omega_1 t + b_n \sin n\omega_1 t) \qquad (9.1)$$

The fundamental angular frequency ω_1 is related to the period T of the function by the following formula:

$$T = \frac{2\pi}{\omega_1} \qquad (9.2)$$

The validity of the Fourier series representation requires that the integral

$$\int_{-T/2}^{T/2} / f(t) / \, dt \qquad (9.3)$$

must be finite.

By multiplying equation (1) by $\cos n\omega_1 t$ and integrating over the period T we obtain

$$\int_{-T/2}^{T/2} a_n \cos^2 n\omega_1 t \, dt = a_n \frac{T}{2} = \int_{-T/2}^{T/2} f(t) \cos n\omega_1 t \, dt \qquad (9.4)$$

We note that all products of the form $\cos m\omega_1 t \cos n\omega_1 t$ or $\sin m\omega_1 t \cos n\omega_1 t$, where $m \neq n$, average out to zero in the integration. For the latter product the integral yields zero even for $m = n$.

If we had multiplied equation (1) by sin $n\omega_1 t$, and performed the integration as indicated above, we would have arrived at an expression for the coefficient b_n for the sine series:

$$a_n = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\cos n\omega_1 t\, dt \tag{9.5}$$

$$b_n = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\sin n\omega_1 t\, dt \tag{9.6}$$

By applying the formulas:

$$\cos n\omega_1 t = \frac{1}{2}(e^{jn\omega_1 t} + e^{-jn\omega_1 t}) \tag{9.7}$$

$$\sin n\omega_1 t = \frac{1}{2j}(e^{jn\omega_1 t} - e^{-jn\omega_1 t}) \tag{9.8}$$

we may write for (1):

$$f(t) = \frac{a_o}{2} + \frac{1}{2}\sum_{n=1}^{\infty}(a_n - jb_n)e^{jn\omega_1 t} + \frac{1}{2}\sum_{n=1}^{\infty}(a_n + jb_n)e^{-jn\omega_1 t} \tag{9.9}$$

Let us introduce negative values for n. We find:

$$a_{-n} = a_n \tag{9.10}$$

$$b_{-n} = -b_n \tag{9.11}$$

The last term of (9.9) becomes:

$$\frac{1}{2}\sum_{n=1}^{\infty}(a_n + jb_n)e^{-jn\omega_1 t} = \frac{1}{2}\sum_{n=-1}^{-\infty}(a_n - jb_n)e^{jn\omega_1 t} \tag{9.12}$$

so that (9.9) simplifies to:

$$f(t) = \sum_{n=-\infty}^{\infty} F(n)e^{jn\omega_1 t} \tag{9.13}$$

where $F(n) = \frac{1}{2}(a_n - jb_n)$ for $n = o, \pm 1, \pm 2, \ldots$ (9.14)

By inserting (9.10) and (9.11) in (9.14):

$$F(n) = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-jn\omega_1 t} dt \qquad (9.15)$$

Equation (9.15) represents the Fourier transform of the periodic function $f(t)$. $F(n)$ is the complex spectrum of $f(t)$. It contains full information of the amplitudes and phases of the sinusoids being summed. We note that since n only assumes discrete values, the spectrum is a line spectrum.

The absolute value of $F(n)$ is given as

$$|F(n)| = \frac{1}{2}\sqrt{a_n^2 + b_n^2} \qquad (9.16)$$

and the phase angle of $F(n)$:

$$\Theta_n = \text{Arc tg}(-\frac{b_n}{a_n}) \qquad (9.17)$$

$$f(t) = \sum_{n=-\infty}^{\infty} |F(n)| e^{j(n\omega_1 t + \Theta_n)} \qquad (9.18)$$

9.1.2 Correlation

Consider the integral

$$\frac{1}{T} \int_{-T/2}^{T/2} f_1(t) f_2(t+\tau) dt \qquad (9.19)$$

where $f_1(t)$ and $f_2(t+\tau)$ have the same fundamental angular frequency ω_1. Let us use the equation

$$\frac{1}{T}\int_{-T/2}^{T/2} f_1(t)\sum_{n=-\infty}^{\infty} F_2(n)e^{jn\omega_1(t+\tau)}dt \qquad (9.20)$$

By inverting the order of summation and integration:

$$\sum_{n=-\infty}^{\infty} F_2(n)e^{jn\omega_1\tau}\ \frac{1}{T}\int_{-T/2}^{T/2} f_1(t)e^{jn\omega_1 t}dt \qquad (9.21)$$

The integral term is seen to represent the complex conjugate of $F_1(n)$ and we have the Fourier relationship:

$$\frac{1}{T}\int_{-T/2}^{T/2} f_1(t)f_2(t+\tau)dt = \sum_{n=-\infty}^{\infty} \overline{F_1}(n)F_2(n)e^{jn\omega_1\tau} \qquad (9.22)$$

Equation (9.22) represents the <u>correlation theorem</u> for periodic functions. If $f_1(t) = f_2(t)$, equation (9.22) becomes:

$$\frac{1}{T}\int_{-T/2}^{T/2} f_1(t)f_1(t+\tau)dt = \sum_{n=-\infty}^{\infty} |F(n)|^2 e^{jn\omega_1\tau} \qquad (9.23)$$

For $\tau=0$ the equation is:

$$\frac{1}{T}\int_{-T/2}^{T/2} f_1^{\,2}(t)dt = \sum_{n=-\infty}^{\infty} |F_1(n)|^2 \qquad (9.24)$$

The mean square value of $f_1(t)$ is equal to the sum of the squares of the absolute value of the spectrum, over the entire range of harmonics. Equation (9.24) is also known as <u>Parsevals theorem</u> for periodic functions.

The left-hand side of (9.23) is the <u>autocorrelation function</u> of $f_1(t)$.
We represent it by the symbol $\varphi_{11}(\tau)$:

$$\varphi_{11}(\tau) = \frac{1}{T}\int_{-T/2}^{T/2} f_1(t) f_1(t+\tau)\,dt \tag{9.25}$$

For the power spectrum $|F_1(n)|^2$ we write $\Phi_{11}(n)$.
In terms of the new symbols:

$$\varphi_{11}(\tau) = \sum_{n=-\infty}^{\infty} \Phi_{11}(n) e^{jn\omega_1\tau} \tag{9.26}$$

and inversely:

$$\Phi_{11}(n) = \frac{1}{T}\int_{-T/2}^{T/2} \varphi_{11}(\tau) e^{-jn\omega_1\tau}\,dt \tag{9.27}$$

We may notice that $\varphi_{11}(\tau)$ is an even function of τ.

$$\varphi_{11}(-\tau) = \frac{1}{T}\int_{-T/2}^{T/2} f_1(t) f_1(t-\tau)\,dt \tag{9.28}$$

$$= \frac{1}{T}\int_{-T/2-\tau}^{T/2-\tau} f_1(s) f_1(s+\tau)\,dt \tag{9.29}$$

$$= \varphi_{11}(\tau)$$

9.2 Aperiodic functions

In the following we derive Fourier relations for aperiodic functions by assuming that a function represented by a Fourier series approaches an aperiodic function if the period of the series approaches infinity.

By equations (9.13) and (9.15):

$$f(t) = \sum_{n=-\infty}^{\infty} e^{jn\omega_1 t} \frac{1}{T} \int_{-T/2}^{T/2} f(s) e^{-jn\omega_1 s} ds \qquad (9.30)$$

Since $1/T = \omega_1/2\pi$

$$f(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{jn\omega_1 t} \omega_1 \int_{-T/2}^{T/2} f(s) e^{-jn\omega_1 s} ds \qquad (9.31)$$

If the period T grows without limit, f(t) tends to an aperiodic function and (9.31) approaches a limiting form. With T approaching infinity the fundamental angular frequency ω_1 becomes a differential $d\omega$ of angular frequency, and the n-th harmonic frequency $n\omega_1$ becomes the continuous angular frequency ω. The summation over all harmonics becomes an integration over the continuous frequency range $(-\infty,\infty)$:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} d\omega \int_{-\infty}^{\infty} f(s) e^{-j\omega s} ds \qquad (9.32)$$

The extreme right-hand integral is the complex continuous spectrum.

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \qquad (9.33)$$

$$f(t) = \frac{1}{2\pi} \int F(\omega) e^{j\omega t} d\omega \qquad (9.34)$$

9.2.1 Correlation

Let us asses the correlation function for aperiodic functions, as defined by the integral

$$\int_{-\infty}^{\infty} f_1(t) f_2(t+\tau)\,dt \tag{9.35}$$

By substituting for $f_2(t+\tau)$:

$$\int_{-\infty}^{\infty} f_1(t)\,dt \frac{1}{2\pi} \int_{-\infty}^{\infty} F_2(\omega)\, e^{j(\omega t+ \tau)}\, d\omega \tag{9.36}$$

Inverting the order of integration:

$$\int_{-\infty}^{\infty} F_2(\omega) e^{j\omega\tau} d\omega \frac{1}{2\pi} \int_{-\infty}^{\infty} f_1(t)\, e^{j\omega t}\, dt \tag{9.37}$$

The right-hand integral is the conjugate of the spectrum of $f_1(t)$.

$$\int_{-\infty}^{\infty} f_1(t) f_2(t+\tau)\,dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{F}_1(\omega) F_2(\omega)\, e^{j\omega\tau}\, d\omega \tag{9.38}$$

We find that

$$\int_{-\infty}^{\infty} f_1(t) f_2(t+\tau)\,dt \quad \text{and} \quad \bar{F}_1(\omega) F_2(\omega) \tag{9.39}$$

constitute a Fourier transform pair.
When $f_1(t) = f_2(t)$ we have the autocorrelation function for the

aperiodic function $f_1(t)$:

$$\varphi_{11}(\tau) = \int_{-\infty}^{\infty} f_1(t) f_1(t+\tau)\,dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F_1(\omega)|^2 e^{j\omega\tau} d\omega \qquad (9.40)$$

For $\tau=0$:

$$\varphi_{11}(0) = \int_{-\infty}^{\infty} f_1^{\,2}(t)\,dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F_1(\omega)|^2 d\omega \qquad (9.41)$$

If $f_1(t)$ represents a voltage or a current, and if a 1-ohm resistance is assumed, the total energy consumed by the resistance is clearly given by the left-hand integral. $|F_1(\omega)|^2$ is measured in watt-seconds per unit angular frequency. We note that the total energy of $f_1(t)$ is given by the zero lag autocorrelation function.

So far we have thought of $f_1(t)$ as representing a real signal. In measuring both amplitude and phase of the signals it may be convenient to treat $f_1(t)$ as a complex quantity, yielding an autocorrelation function . in general possessing both a real and an imaginary part. (The power spectrum, being the absolute value of the frequency spectrum squared, is, of course, always positive).

With this generalization in mind we may note the following:

(i) For $f_1(t)$ real:

$\varphi_{11}(\tau)$ is real
$|F(\omega)|^2$ is symmetric about DC-level ($\omega=0$)

(ii) For $f_1(t)$ complex:

$\varphi_{11}(\tau)$ is complex hermitian ($\varphi_{11}(\tau)=\varphi_{11}^*(-\tau)$)
$|F(\omega)|^2$ is not generally symmetric

In case $\varphi_{11}(\tau)$ is complex, its real and imaginary parts are easily derived from equation (9.40)

9.2.2 Convolution

The convolution integral is given as:

$$\int_{-\infty}^{\infty} f_1(t) f_2(\tau-t)\,dt \tag{9.42}$$

It resembles the crosscorrelation function in that it involves displacement, multiplication and integration. It differs from correlation because it includes folding, or, reflection of the displaced function. The function $f_2(\tau-t)$ is first displaced to the left by τ and then folded back with respect to the point $t = 0$.

Let us replace $f_2(\tau-t)$ in the integral by its Fourier transform.

$$\int_{-\infty}^{\infty} f_1(t) f_2(\tau-t)\,dt = \int_{-\infty}^{\infty} f_1(t)\,dt \frac{1}{2\pi} \int_{-\infty}^{\infty} F_2(\omega) e^{j\omega(\tau-t)}\,d\omega \tag{9.43}$$

$$\int_{-\infty}^{\infty} f_1(t) f_2(\tau-t)\,dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_2(\omega) e^{j\omega\tau}\,d\omega \int_{-\infty}^{\infty} f_1(t) e^{-j\omega t}\,dt$$

The integral on the extreme right is $F_1(\omega)$:

$$\int_{-\infty}^{\infty} f_1(t) f_2(\tau-t)\,dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_1(\omega) F_2(\omega) e^{j\omega\tau}\,d\omega \tag{9.44}$$

We find that

$$\int_{-\infty}^{\infty} f_1(t) f_2(\tau-t)\,dt \text{ and } F_1(\omega) F_2(\omega) \tag{9.45}$$

are Fourier transforms of each other.

The inverse of (9.44) becomes:

$$F_1(\omega)F_2(\omega) = \int_{-\infty}^{\infty} e^{-j\omega\tau}d\tau \int_{-\infty}^{\infty} f_1(t)f_2(\tau-t)dt \qquad (9.46)$$

Equations (9.44) and (9.46) constitute the convolution theorem for aperiodic functions.

The following relationship is easily derived:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} F_1(\omega)F_2(\omega_o-\omega)\,d\omega = \int_{-\infty}^{\infty} f_1(t)f_2(t)e^{-j\omega_o t}\,dt \qquad (9.47)$$

From (9.44) and (9.46) we conclude that convolution in the time domain corresponds to multiplication in the frequency domain and vice versa.

Examples

1. Rectangular pulse

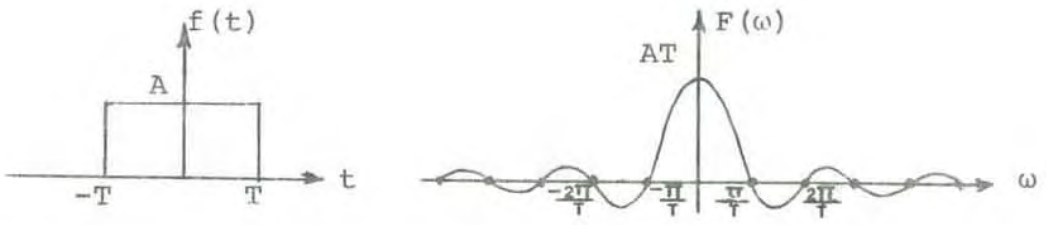


Figure 9.1 Rectangular pulse and its frequency spectrum

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt = A\int_{-T}^{T} e^{-j\omega t}dt \qquad (9.48)$$

$$F(\omega) = 2AT\,\frac{\sin \omega T}{\omega T} \qquad (9.49)$$

Let us assume that the area of the pulse, 2 AT, is equal to unity. If the pulse width is allowed to become infinitesimal, the height of the pulse approaches infinity, and we obtain an impulse at the origin, whose integral is unity and whose spectrum is uniform and equal to 1.

$$F(\omega) = \lim_{T \to 0} 2AT \frac{\sin \omega T}{\omega T} = 1 \qquad (9.50)$$

A function with the indicated properties is called a delta-function or a unit-impulse function.

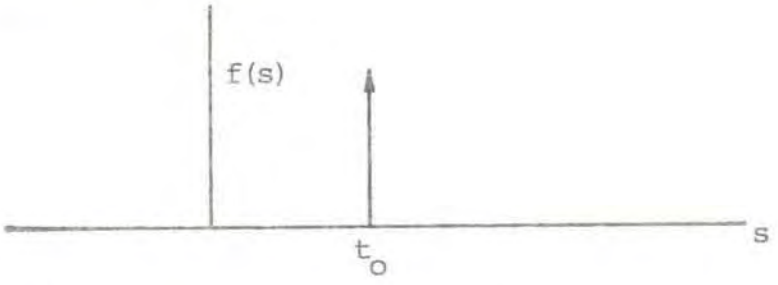


Figure 9.2 Delayed delta-function

Let us consider the integral

$$\int_{-\infty}^{\infty} f(s)\ \delta(t_0-s)\ ds \qquad (9.51)$$

where $\delta(t_0-s)$ is the delayed delta-function depicted in figure 9.2.

The integral differs from zero only at $s = t_0$ yielding:

$$\int_{-\infty}^{\infty} f(s)\ \delta(t_0-s)ds = f(t_0) \int_{-\infty}^{\infty} \delta(t-s)ds = f(t_0) \qquad (9.52)$$

since the integral $\int_{-\infty}^{\infty} \delta(t-s)ds$ equals 1 by definition.

The convolution of an aperiodic function with the unit-impulse function leaves the function unchanged. Due to the property of the unit-impulse function described by (9.52) this function is sometimes denoted the sampling function. Equation (9.52) is sometimes also denoted the sifting-

integral expression.

Mathematically the delta-function is not an "ordinary function". Some of its properties must be derived by complicated limiting arguments. It is comforting to know that several difficulties may be eliminated by the use of the theory of distributions (Papoulis, 1962).

2. Time shifted function

We assume the following Fourier transform pair:

$$f(t) \Leftrightarrow F(\omega)$$

and wish to find the spectrum of the time-shifted function $f(t-t_o)$:

$$F_s(\omega) = \int_{-\infty}^{\infty} f(t-t_o)e^{-j\omega t}dt = \int_{-\infty}^{\infty} f(s)e^{-j\omega s}e^{-j\omega t_o}dt \qquad (9.53)$$

where we have introduced $s = t-t_o$.

$$F_s(\omega) = F(\omega)e^{-j\omega t_o} \qquad (9.54)$$

The spectrum remains the same, but a linear term $-t_o\omega$ has been added to its phase angle.

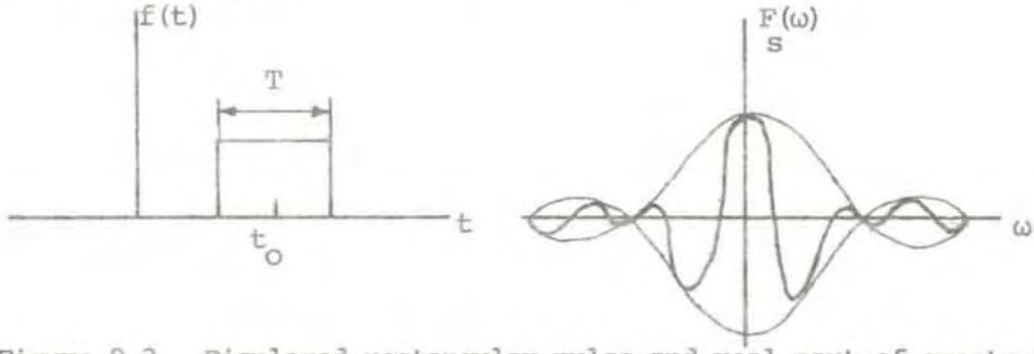


Figure 9.3 Displaced rectangular pulse and real part of spectrum

3. Sine/cosine signal

Assume a sine wave:

$$f(t) = \sin \omega_o t = \frac{1}{2j} \left(e^{j\omega_o t} - e^{-j\omega_o t} \right) \tag{9.55}$$

$$F(\omega) = \int_{-\infty}^{\infty} \frac{1}{2j} \left(e^{j\omega_o t} - e^{-j\omega_o t} \right) e^{-j\omega t} \, dt \tag{9.56}$$

$$F(\omega) = \frac{j}{2} \left[\delta(\omega + \omega_o) - \delta(\omega - \omega_o) \right] \tag{9.57}$$

With a cosine signal:

$$F(\omega) = \frac{1}{2} \left[\delta(\omega - \omega_o) + \delta(\omega + \omega_o) \right] \tag{9.58}$$

For a sinusoidal signal the spectrum is represented by two delta-functions of strength ½ appearing at $\pm \omega_o$.

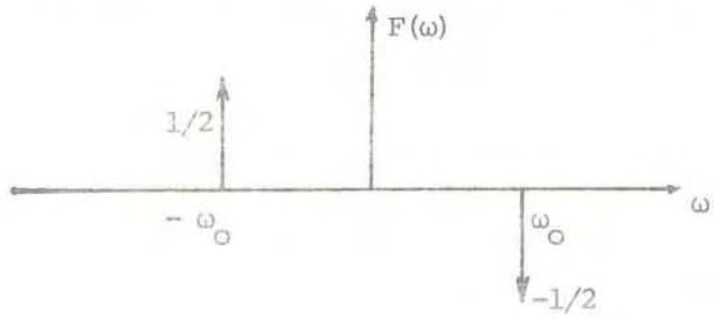


Figure 9.4 Spectral representation of sine wave

4. Modulated carrier

Consider the wave shown in figure 9.5

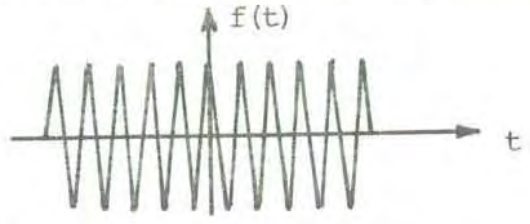


Figure 9.5 Modulated cosine

The modulated wave may be thought of as the product of a pure cosine wave $f_1(t)$ and the modulating function $f_2(t)$:

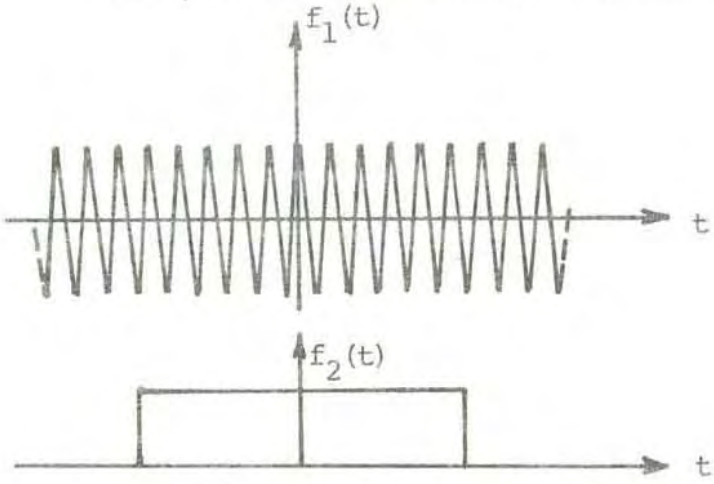


Figure 9.6 Carrier waveform and modulating function

For the Fourier transform we write the cosine wave in exponential form:

$$F(\omega) = \frac{1}{2}\int_{-\infty}^{\infty} (e^{j\omega_o t} + e^{-j\omega_o t})\, f_2(t)\, e^{-j\omega t}\, dt \qquad (9.59)$$

$$F(\omega) = \frac{1}{2} F_2(\omega - \omega_o) + \frac{1}{2} F_2(\omega + \omega_o) \qquad (9.60)$$

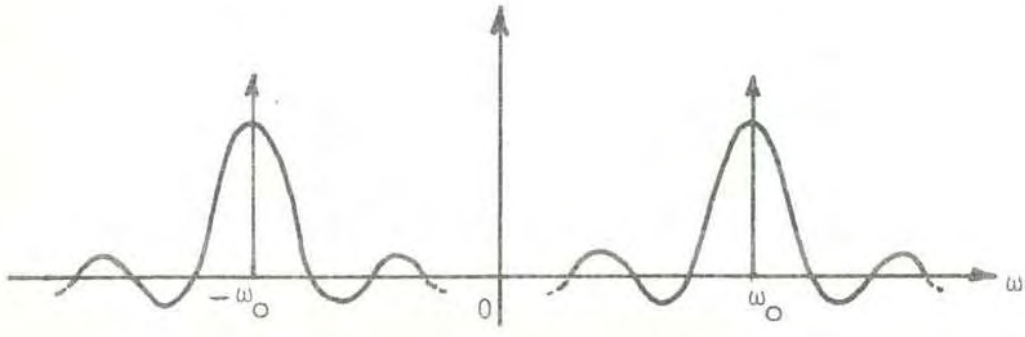


Figure 9.7 Spectrum of modulated cosine

5. Convolution of rectangular pulses

Consider the following rectangular waveforms:

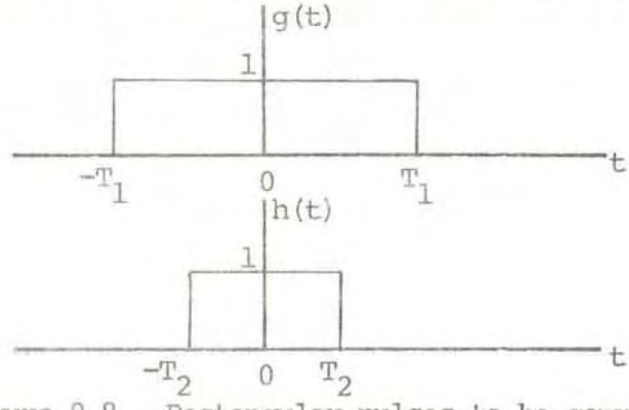


Figure 9.8 Rectangular pulses to be convolved

According to (9.42) the convolution of the two waveforms is given by:

$$f(t) = \int_{-\infty}^{\infty} g(s)\, h(t - s)ds \qquad (9.61)$$

Geometrically f(t) is represented in figure (9.9):

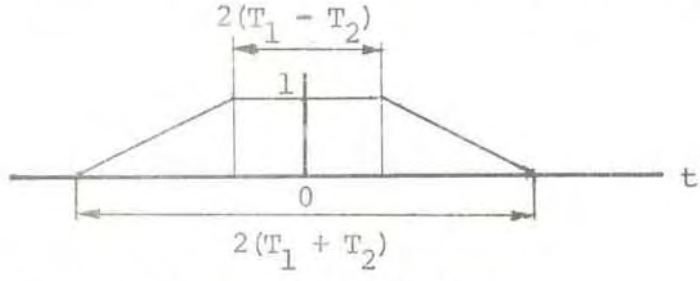


Figure 9.9 Convolved waveform

Using (9.45) we find for the spectrum of the convolved signal:

$$F(\omega) = F_1(\omega)\, F_2(\omega) \qquad (9.62)$$

where $F_1(\omega)$ and $F_2(\omega)$ have the form given by (9.49).

$$F_{1,2}(\omega) = 2T \frac{\sin \omega_{1,2} T_{1,2}}{\omega_{1,2} T_{1,2}} \qquad (9.63)$$

If $T_1 = T_2 = T$ we obtain a triangular waveform, as shown in figure (9.10).

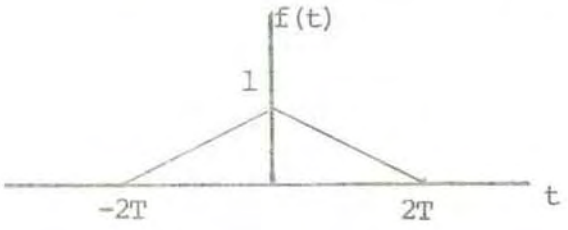


Figure 9.10 Result of folding rectangular pulse with itself

The spectrum is readily inferred from (9.62) and (9.63)

$$F(\omega) = 4T^2 \left(\frac{\sin \omega T}{\omega T} \right)^2 \qquad (9.64)$$

If the triangular form is convolved with the original rectangular pulse, a parabolic pulse form results, whose frequency spectrum is

$$F(\omega) = \left(\frac{\sin \omega T}{\pi f} \right)^3 \qquad (9.65)$$

In general, after n convolutions, the spectrum assumes the form

$$F(\omega) = \left(\frac{\sin \omega T}{\pi f} \right)^n \qquad (9.66)$$

It may be noticed that the subsidiary maxima in the spectrum become smaller and smaller as the number of convolutions increases.

The Gaussian waveform:

$$f(t) = e^{-\alpha t^2} \qquad (9.67)$$

has the spectrum:

$$F(\omega) = \sqrt{\frac{\pi}{\alpha}}\, e^{-\omega^2/4\alpha} \qquad (9.68)$$

which falls off monotonically from $\omega = 0$. It will be seen from the equations above that a Gaussian signal has no subsidiary maxima.

6. Differentiation

Equation (9.34) states that

$$f(t) = \frac{1}{2\pi} \int F(\omega) e^{-j\omega t} d\omega \qquad (9.69)$$

By performing repeated differentiations (n) of f(t), with the derivative operator under the integral, we obtain:

$$g(t) = \frac{d^n}{dt} f(t) = \int_{-\infty}^{\infty} (j\omega)^n F(\omega) e^{j\omega t} dt \qquad (9.70)$$

$$G(\omega) = (j\omega)^n F(\omega) \qquad (9.71)$$

Relation (9.71) may be used to derive Fourier transforms from previously found transforms. In the following we consider a few examples:

a) Differentiation of rectangular pulse

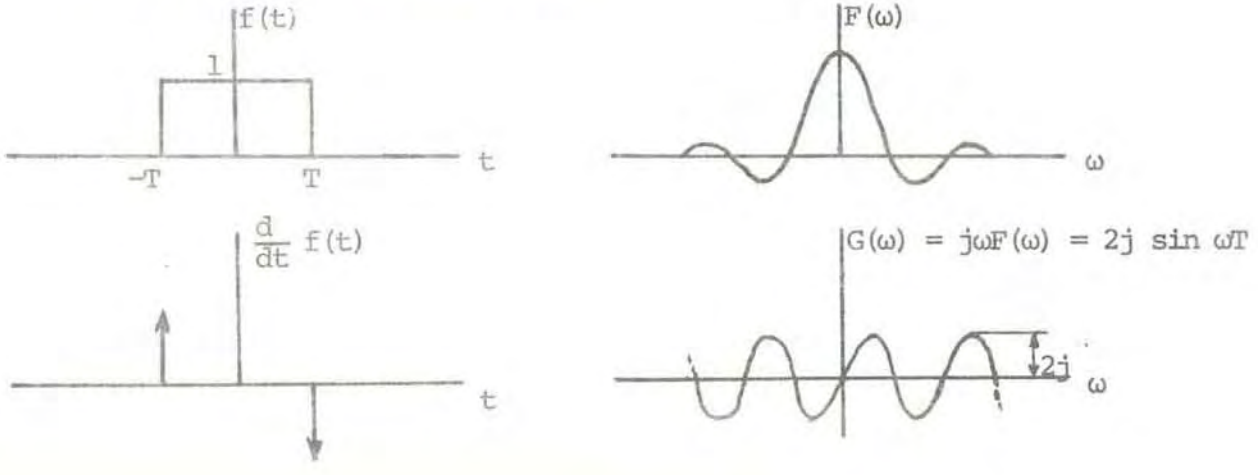


Figure 9.11 Differential pulse and Fourier transform

b) Sloping signal form

By integrating the rectangular pulse form we find:

$$h(t) = \int_{-\infty}^{t} f(t)dt \qquad (9.72)$$

with spectrum $H(\omega)$.

Using relations (9.70) and (9.71):

$$f(t) = \frac{d}{dt} h(t) \qquad (9.73)$$

and

$$F(\omega) = j\omega H(\omega) \qquad (9.74)$$

Since $F(\omega)$ is known from previous derivations, we readily find:

$$H(\omega) = 2T \frac{\sin \omega T}{j\omega^2 T} = -j2T^2 \frac{\sin \omega T}{(\omega T)^2} \qquad (9.75)$$

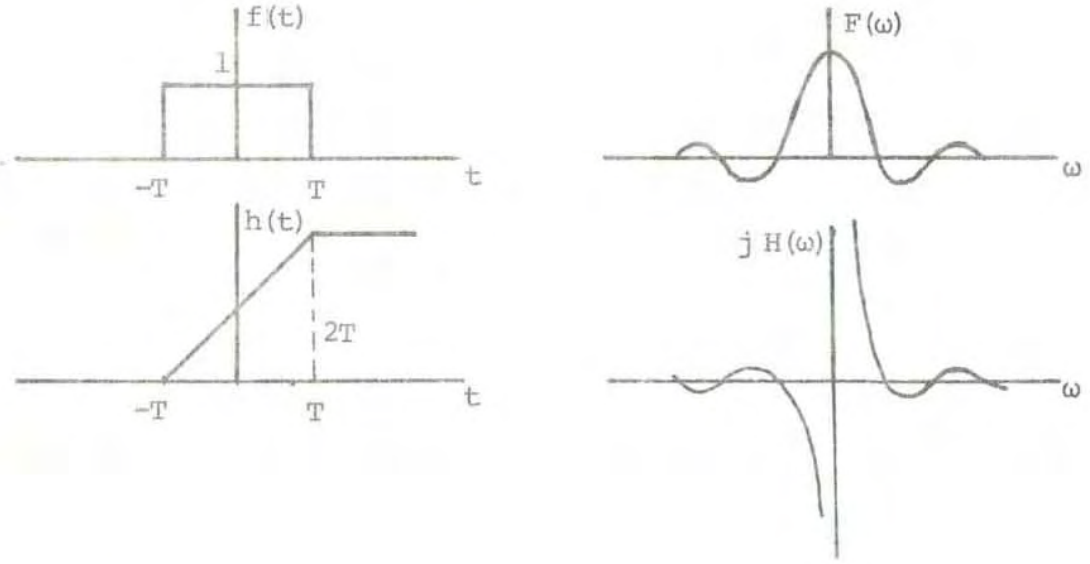


Figure 9.12 Sloping system and spectrum

c) Step function

Consider a delta pulse at t = 0.

$$g(t) = \int_{-\infty}^{t} \delta(t)dt \qquad (9.76)$$

$$G(\omega) = \frac{F(\omega)}{j\omega} = -\frac{j}{\omega} \qquad (9.77)$$

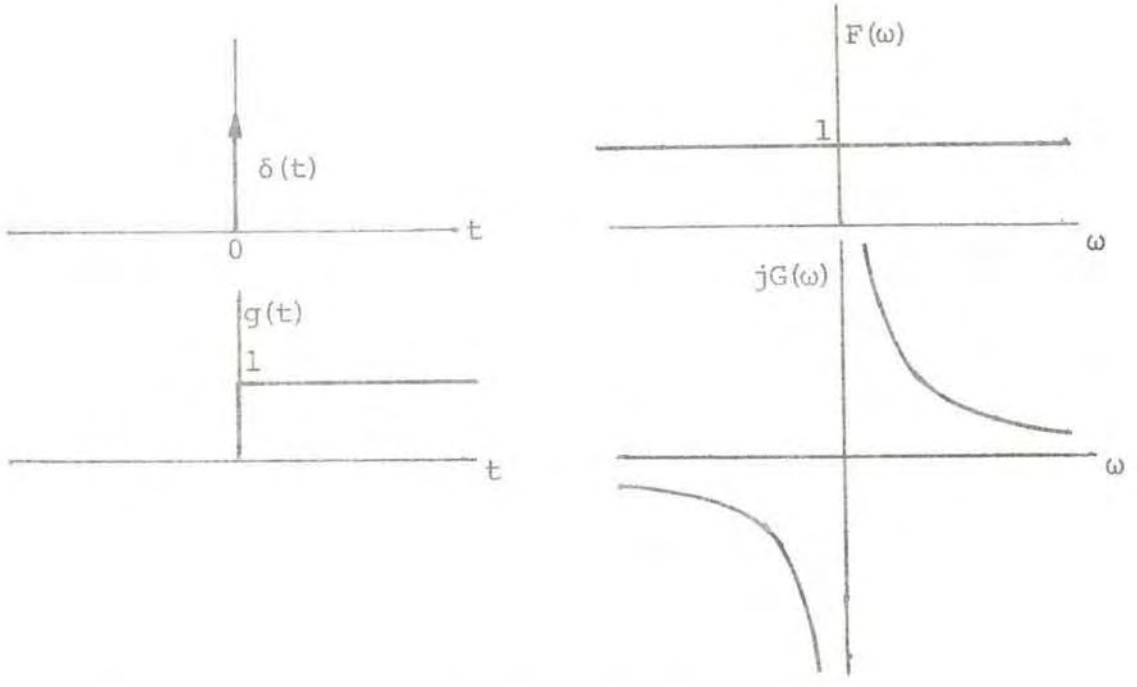


Figure 9.13 Step function with spectrum

9.3 Statistical ensembles

Frequently the quantities to be correlated are random variables, assuming either discrete or continuous values within certain ranges. Voltages, temperatures, velocities etc. in physical experiments are often of this nature. The same is true for signals scattered the thermal fluctuations in the ionosphere. An infinite aggregate of individual measurements of the random variable is called an ensemble.

Provided the ensemble is stationary, it may be permissible to replace the autocorrelation function

$$\varphi_{11}(\tau) = \int_{-\infty}^{\infty} f(t) f^{*}(t + \tau) dt \qquad (9.78)$$

by the ensemble average

$$\langle f(t) f^{*}(t + \tau) \rangle \qquad (9.79)$$

In any practical measurements one is, of course, bound to work with values averaged over a fininite number of ensemble members. The errors thereby introduced are of primary importance in assessing the accuracy of parameters derived from the observations.

The following example serves to illustrate the usefulness of assessing the correlation function as an ensemble average instead of deriving the function as a time averaged quantity.

The task is to provide an estimate of the mean power scattered by pulsed radar transmissions from a given height range in the remote part of the atmosphere. In addition the distribution of scattered averaged power in frequency is wanted. The treatment is confined to monostatic operation. From signal theory we know that the average power is given by the zero lag autocorrelation function of the received signal and the distribution ot power in frequency by the Fourier transform of the autocorrelation function.

As described in chapter 4 (volume I)the signals scattered from a given height range are observed in a specific time interval after the transmission of each pulse. A continuous monitoring of the return from the height range in question is not possible. Instead, as illustrated below, the autocorrelation function is derived by averaging over the ensemble aggregate made up by the returns from all the consecutive pulses. The procedure remains valid as long as the assumption of a stationary medium holds good.

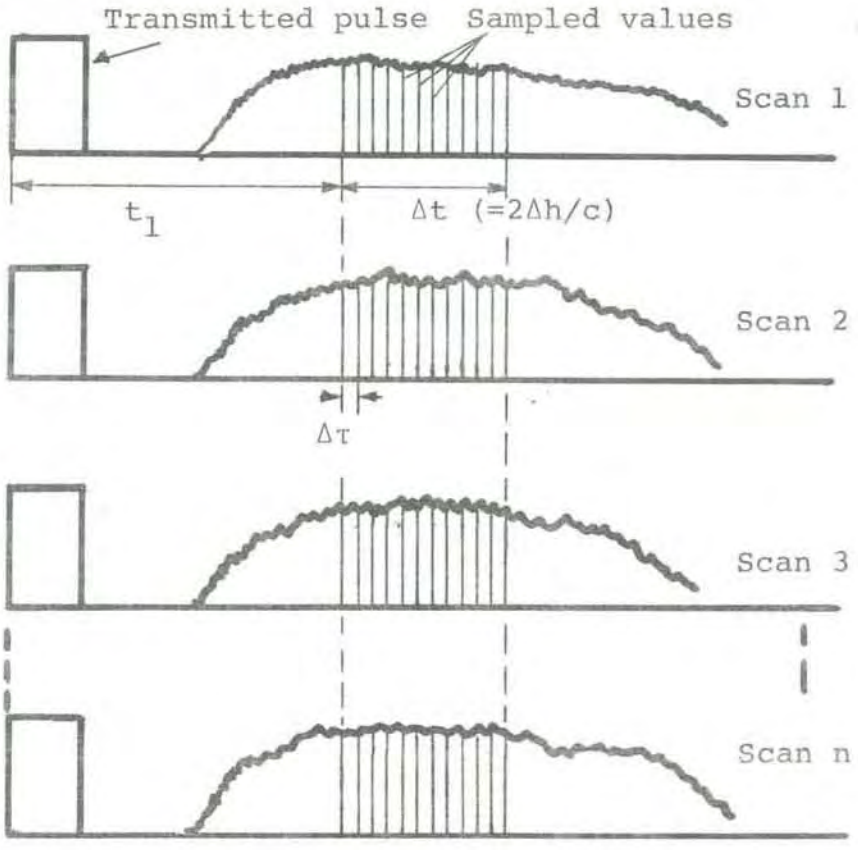


Figure 9.14 Autocorrelation function of scattered signal estimated as an ensemble average (The width of the range gate is drawn out of scale to illustrate pertinent points).

Strictly speaking the illustrated procedure is a mixture of time and ensemble averaging, in that $\rho(k \Delta \tau)$ $(k = 0,1,...m)$ is estimated for each member aggregate before being averaged over the ensemble. The number of element pairs involved in the averaging over the member aggregate depends upon the lag index, k. For instance, m - 1 (m is the number of samples within the range gate) element pairs are included in the member estimate of $\rho(\Delta\tau)$, whereas only one element pair is included in the estimate of $\rho(m \Delta \tau)$. It follows that the reliability of the final autocorrelation estimates decreases with increasing lag index.

References

Lee Y W - Statistical Theory of Communication, John Wiley & Sons, Inc., New York (1960)

Papoulis A - The Fourier Integral and its Applications, McGraw-Hill, Inc., New York (1962)

Ratcliffe J A - Lecture Notes, Cavendish Laboratory, Cambridge

10 SAMPLED SIGNALS AND FAST FOURIER TRANSFORM

The data processing and the extraction of physical parameters in the EISCAT system will be performed by handling digitized signals, derived from sequencies of samples taken at discrete intervals.

It is evident that the accuracy with which the sampled signals represent the original analog waveform must depend upon how frequently the sampling is made. In this chapter we will study under which conditions an accurate reconstruction of the waveform can be effected. We also intend to briefly outline the principles for the commonly applied technique, the Fast Fourier Transform (FFT), for transforming discrete signals.

Before embarking upon a more detailed treatment we might point out a general limitation which exists: In a strict sense signals cannot simultaneously be both band- and time-limited. We will assume that the waveforms dealt with in the following are essentially time- and frequency-limited, implying that the spectrum of a signal bounded in the time domain has a small magnitude outside some frequency bands.

10.1 Uniform, ideal sampling

In figure 10.1 is indicated a function sampled by a pulse train with pulses occurring at 0, T_1, T_2,- - - - -

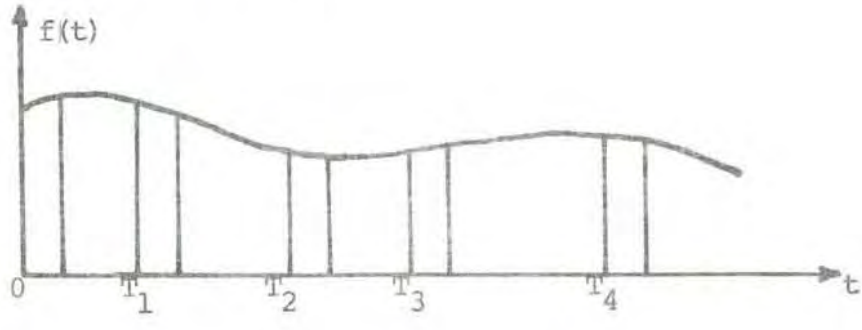


Figure 10.1 Sampling with rectangular pulses occurring at 0, T_1, T_2,- - - -

In the subsequent treatment we shall assume that the sampling is uniform, with the sampling pulses occurring at periodic intervals T. We shall further presuppose that the sampling is ideal, meaning that the sampling pulse train may be described as a sequence of impulse functions:

$$s(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT) \qquad (10.1)$$

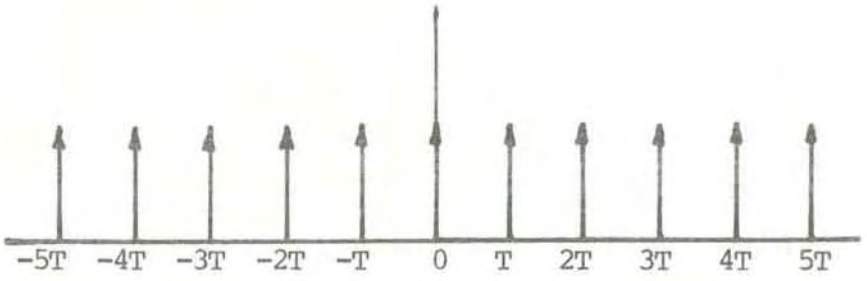


Figure 10.2 Sampling pulse train

It is clear that mathematically the sampled signal $f_s(t)$ may be represented by the product:

$$f_s(t) = f(t)\, s(t) = f(t) \sum_{n=-\infty}^{\infty} \delta(t - nT) \qquad (10.2)$$

From the preceding chapter we know that since s(t) is periodic, it can be represented by a Fourier series:

$$s(t) = \sum_{n=-\infty}^{\infty} C_n e^{jn\omega_s t} \qquad (10.3)$$

with

$$\omega_s = 2\pi/T \qquad (10.4)$$

and

$$C_n = \frac{1}{T} \int_{-T/2}^{T/2} \delta(t) e^{-j\omega t} dt = \frac{1}{T} \qquad (10.5)$$

Inserting (10.3) and (10.5) in (10.2) yields:

$$f_s(t) = \frac{1}{T} \sum_{n=-\infty}^{\infty} f(t) e^{jn\omega_s t} \qquad (10.6)$$

The Fourier transform of $f_s(t)$ becomes:

$$F_s(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} f(t) e^{-j(\omega - n\omega_s)t} dt \qquad (10.7)$$

$$F_s(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - n\omega_s) \qquad (10.8)$$

<u>The spectrum of the sampled signal is periodic with period ω_s, consisting of repetitive translations of the baseband transform $F_s(\omega)$.</u>

The spectrum of the sampled signal may alternatively be derived by transforming (10.2) directly:

$$F_s(\omega) = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(nT)\, \delta(t - nT) e^{-j\omega t} dt \qquad (10.9)$$

$$F_s(\omega) = \sum_{n=-\infty}^{\infty} f(nT) e^{-jn\omega T} \qquad (10.10)$$

We note that (10.10) represents the spectrum as a Fourier series with coefficients given by the sample values of $f(t)$.

If $F_s(\omega)$ is given, the sampled values $f(nT)$ may be expressed:

$$f(nT) = \frac{1}{\omega_s} \int_{-\omega_s/2}^{\omega_s/2} F_s(\omega) e^{jn\omega t} d\omega \qquad (10.11)$$

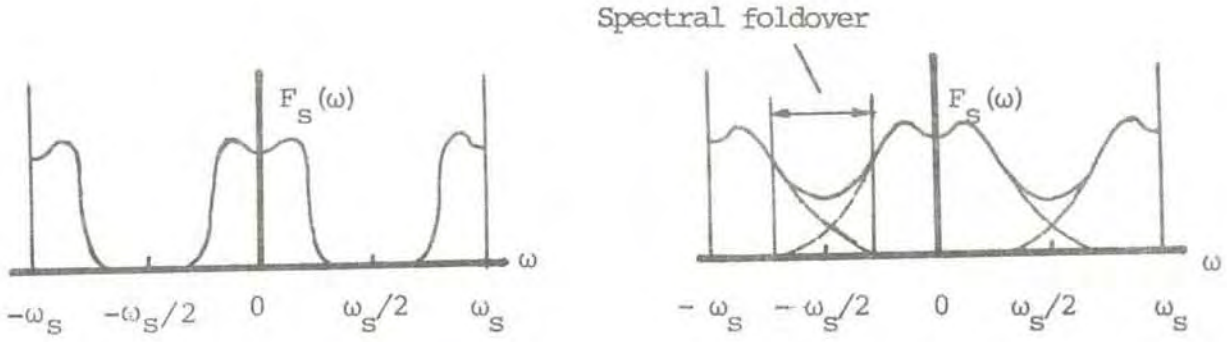


Figure 10.3 Spectra with: (a) $\omega_s > 2\omega_e$ and (b) $\omega_s < 2\omega_e$

Figure 10.3 shows spectra of sampled signals. In case (a) $\omega_s/2$, known as the Nyquist frequency, is higher than the highest significant baseband frequency, ω_e. The original signal may be recovered by passing $f_s(t)$ through an ideal low-pass filter with cut-off frequency ω_e. The second spectrum, (b), corresponds to $\omega_s < 2\omega_e$ and the repetitive baseband patterns are found to overlap. The phenomenon is denoted Aliasing or Spectral Folding.

When aliasing occurs, frequency components at the high end of the spectrum are reflected down to a lower frequency range, distorting the original spectrum. The high frequency components are said to "impersonate" lower frequencies.

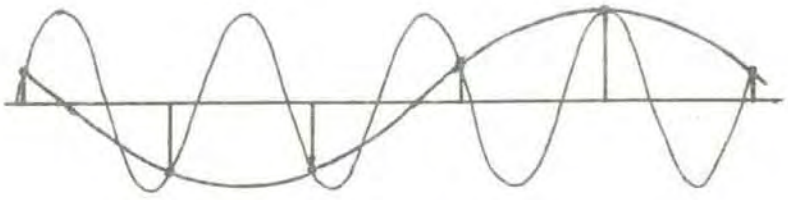


Figure 10.4 Sampling of sinusoid at $\omega_s < 2\omega_e$

For an illustration of the spectral folding effect study figure 10.4. The original signal is the high frequency sinusoid. With the sampling instants indicated, what will be represented is the low frequency waveform.

The cure to the problem of spectral foldover, is clearly to sample at a rate at least twice as high as the highest frequency component of the original signal. This is the essential content of the Sampling Theorem. In practice low-pass filters with cut-off frequencies set to 35-40 % of the sampling rate are often used to prevent aliasing. Some degrading of the signal may result from the filtering, but this will usually have a less detrimental effect than the foldover which may otherwise occurs.

A mathematical expression of the sampling theorem may be obtained by noting from (10.8) that for $|\omega| < \omega_e$ and $\omega_s = 2\omega_e$ $F_s(\omega) = F(\omega)/T$. Let us pass $f_s(t)$ through a low-pass filter with the following characteristics:

$$G(\omega) = \begin{cases} T & \text{for} \quad |\omega| < \omega_e \\ 0 & \text{for} \quad |\omega| \geq \omega_e \end{cases}$$

We have for the filter output:

$$F(\omega) = F_s(\omega)\ G(\omega) \qquad (10.12)$$

Using (10.10):

$$F(\omega) = \sum_{n=-\infty}^{\infty} G(\omega)\ f(nT) e^{-jn\omega T} \qquad (10.13)$$

By Fourier transformation:

$$f(t) = \sum_{n=-\infty}^{\infty} f(nT)\ g(t - nT) \qquad (10.14)$$

with

$$g(t - nT) = \frac{\sin \omega_e (t - nT)}{\omega_e (t - nT)} \qquad (10.15)$$

At the sampling instants:

$$g(t - nT) = g(kT - nT) = \begin{cases} 1 & \text{for} \quad k = n \\ 0 & \text{for} \quad k \neq n \end{cases} \qquad (10.16)$$

Only one term in the summation is nonzero at the sampling instants nT, no interference from the samples before or after f(nT) being noted at time nT. This effect is demonstrated in figure 10.5.

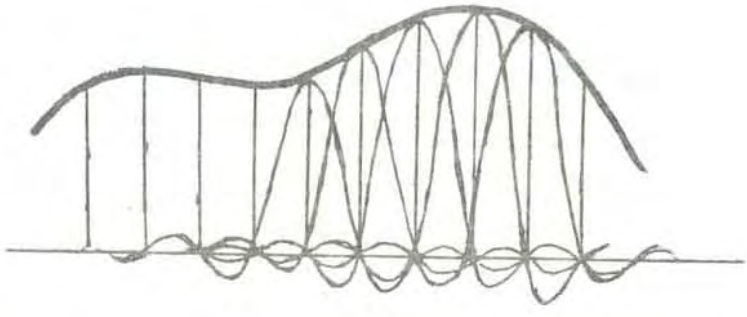


Figure 10.5 Property of the interpolating function

The response characteristics displayed in figure 10.5 are said to cause no "intersymbol interference".

Sampling at $\omega_s < 2\omega_e$ is called Undersampling. Some estimate of the effect of undersampling may be obtained by computing the spectral distribution of the energy within the band Ω centred on $\omega = 0$ (or any $n\omega_s$).

$$E(\Omega) = \int_{-\Omega/2}^{\Omega/2} |F(\omega)|^2 d\omega \qquad (10.17)$$

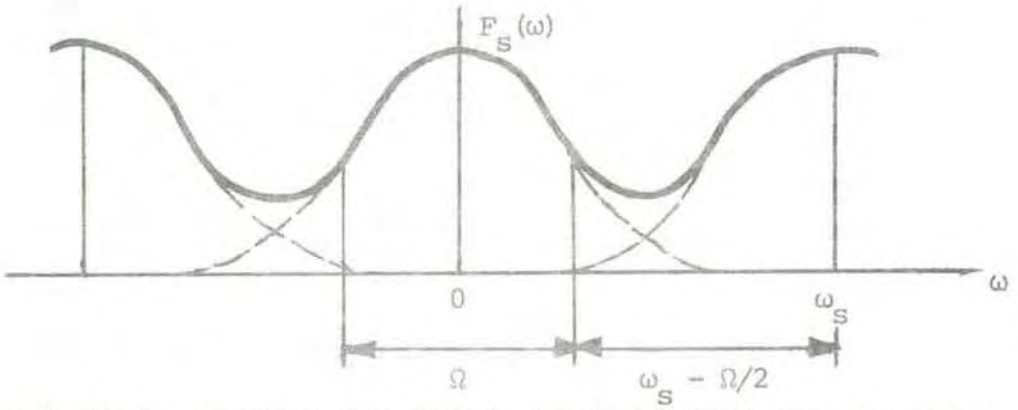


Figure 10.6 Spectrum of sampled signal and band used for computation of spectral energy distribution

The relative spectral energy outside of the band of width Ω:

$$D(\Omega) = 1 - \frac{E(\Omega)}{E(\infty)} \qquad (10.18)$$

In figure 10.6 we observe that each edge of the band of interest is at a distance $\omega_s - \Omega/2$ from the mean of the nearest adjacent repeated spectrum. Since all the repeated spectra are identical, the relative energy which folds into the band of interest, provided only the nearest spectra contribute, is given as:

$$D(2\omega_s - \Omega) \qquad (10.19)$$

This expression may serve as a guideline in selecting an appropriate sampling rate.

In the discussion of sampling so far we have tacitly assumed that the frequency spectra in question have been of the baseband type, including $\omega = 0$. A signal is said to be of the Bandpass type when it is bandlimited, but does not include the zero frequency. An example is shown in figure 10.7.

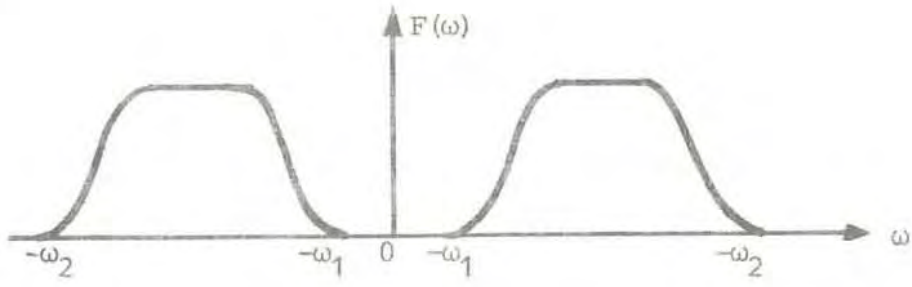


Figure 10.7 Spectrum of bandpass-signal

The spectrum for negative frequencies is entirely determined by the spectrum for positive frequencies, and there is no loss of information if the negative frequency portion is removed.

It is customary in mathematical treatments of band-pass signals to identify the physical signal f(t) with the real part of the complex function

$$\xi(t) = f(t) + j\,\hat{f}(t) \qquad (10.20)$$

where $\hat{f}(t)$ is the Hilbert transform of f(t). Physically the Hilbert transform is obtained by passing f(t) through an ideal 90° phase shifter. $\xi(t)$ is the complex analytic waveform. Its spectrum exists only for positive frequencies. A band-pass signal is also uniquely determined by its in-phase and quadrature components, or from samples of its envelope and instantaneous phase. The necessary sampling rate is ω_e, half the rate required for sampling a base-band signal. Since both signal components are sampled, however, in the band-pass case, the required total number of sampling points is the same for both signal types.

10.2 Fast Fourier Transform (FFT)

Let us consider a signal f(t) in the time domain. We sample this signal at discrete times nT and form the spectrum $F_s(\omega)$ according to (10.10), but only at discrete frequencies:

$$F_s(k\Delta\omega) = \sum_n f(nT)e^{-jnk\Delta\omega T} \qquad (10.21)$$

For the moment we leave n unspecified. As we know from (10.8) $F_s(\omega)$ is repetitive over the interval ω_s. We divide this interval into N equal increments:

$$\Delta\omega = \frac{\omega_s}{N} \qquad (10.22)$$

Since $\omega_s = 2\pi/T$, we have:

$$T\Delta\omega = \frac{2\pi}{N} \qquad (10.23)$$

We will insert the expression (10.23) in (10.21):

$$F(k) = \sum_n f(n)e^{-jnk2\pi/N} \qquad (10.24)$$

To conform with customary notations in the current literature we omit the subscript s and make the substitutions:

$k\Delta\omega \rightarrow k$

$nT \rightarrow T$

It is understood, though, that $F_s(k)$ and $f(n)$ are the k-th and n-th samples of $F_s(\omega)$ and $f(t)$ and to identify their positions on the frequency- and time-axes, we have to multiply k and n with the frequency- and time increments, $\Delta\omega$ and T.

We have already found that a function sampled in time is repetitive in the frequency domain. Likewise, the sampled version of $F_s(\omega)$ must correspond to a function which repeats itself in the time domain. The repetition period is:

$$t_r = 1/\Delta\omega = \frac{NT}{2\pi} \qquad (10.25)$$

If the Fourier transform of $F_s(k)$ is to correspond to the samples of $f(t)$, then $f(t)$ must be limited to a period not longer than t_r. This implies that there must not be more than N samples of the original signal:

$$F(k) = \sum_{n=0}^{N-1} f(n)e^{-jnk2\pi/N} \qquad (10.26)$$

This is the Discrete Fourier Transform (DFT) of the sampled sequence $f(0), f(1), - - - - f(N-1)$. The inverse transform may be written:

$$f(n) = \frac{1}{N} \sum_{k=0}^{N-1} F(k)e^{jnk2\pi/N} \qquad (10.27)$$

The Fast Fourier Transform (FFT) is a collective, and somewhat loose, reference for a set of algorithms effecting the summation (10.27) in a very efficient way.

Equation (10.26) may be written:

$$F(k) = \sum_{n=0}^{N-1} f(n)\, W_N^{jnk} \qquad (10.28)$$

where we have introduced the customary convenient notation

$$W_N = e^{-j2\pi/N} \qquad (10.29)$$

The basis for the time saving offered by the various FFT-algorithms are the following properties of the quantity W_N:

(i) $W_N^{k(N-n)} = (W_N^{kn})^*$

(ii) $W_N^{kn} = W_N^{k(n+N)}$

In particular the cyclic character expressed by (ii) is important. A representation of W_N^n in the complex plane for N = 8 is given in figure 10.8.

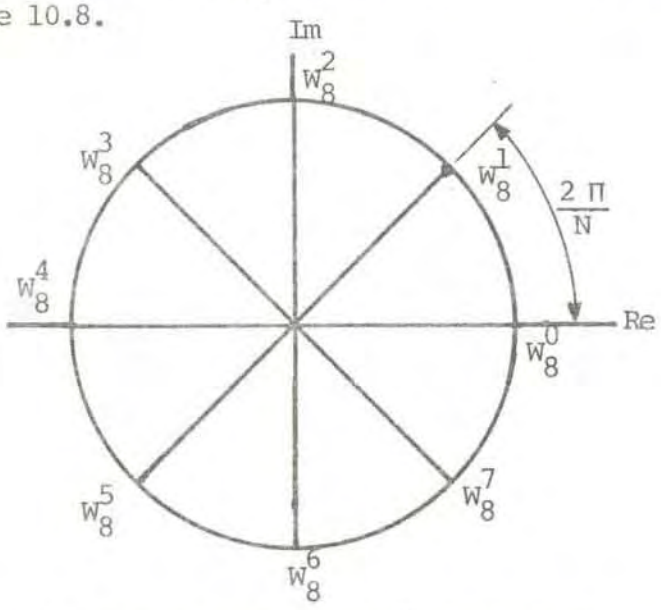


Figure 10.8 Presentation of W_8^n in complex plane

In figure 10.8 we note (neglecting the subscript 8) that $W^8 = W^0 = 1$, $W^9 = W^1$, $W^{10} = W^2$ etc. Furthermore, we find $W^1 = -W^5$, $W^{10} = -W^6$ etc.

In the following we will deal only with the evaluation of equation (10.26). The procedures may be applied also to (10.27) with straightforward modifications.

If we are to evaluate $F_s(k)$, for N values of k, by direct computations about N^2 complex multiplications (f(n) assumed to be complex) and additions are required. The time-saving feature of the various FFT-algorithms arises by decomposing the transformation of a long sequence into a combination of Fourier transforms of several shorter sequences.

The methods rely on N being a product of factors

$$N = p_1 p_2 - - - - p_s \qquad (10.30)$$

In most cases N is taken to be a power of 2

$$N = 2^{\nu} \qquad (10.31)$$

With N obeying the power law (10.31) two routines are in current use, designated: (i) transformation by decimation in frequency and, (ii) transformation by decimation in time. In the following we will briefly describe the main principles of both methods.

10.2.1 Decimation in frequency

In this procedure we can consider dividing the output sequence F(k) into smaller and smaller subsequences. As a first step the input sequence { f(n) } is divided into the first half and the last half of the points so that

$$F(k) = \sum_{n=0}^{N/2-1} f(n)\, W_N^{nk} + \sum_{n=0}^{N/2-1} f(n + N/2)\, W_N^{(n+N/2)k} \qquad (10.32)$$

It proves convenient to consider k even and k odd separately. Exploiting the properties of the quantity W_N^{nk} the two summation terms may be written:

$$F(2r) = \sum_{n=0}^{N/2-1} [f(n) + f(n + N/2)]\, W_N^{2rn} \qquad (10.33a)$$

$$F(2r + 1) = \sum_{n=0}^{N/2-1} [f(n) - f(n + N/2)]\, W_N^{n}\, W_N^{2rn} \qquad (10.33b)$$

$$r = 0, 1, - - - , N/2-1$$

Each equation is recognized as a DFT with N/2 points.
Since N is a power of 2, N/2 is an even number and each of the N/2-points DFTs can be computed in the same manner as indicated above by subdividing the N/2 sequence into the first and latter half of the points. Proceeding in this manner the computation is finally reduced to the computation of two-points DFTs.

Counting the arithmetic operations required we find that the computations have been reduced to N/2 $\log_2 N$ complex multiplications and N $\log_2 N$ complex additions.

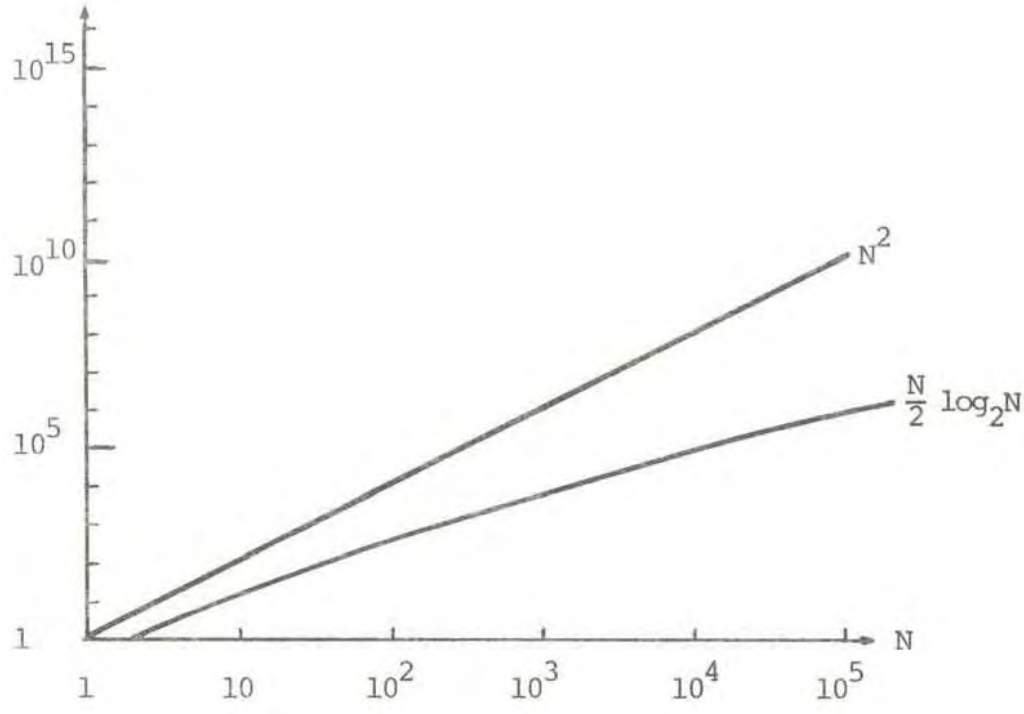


Figure 10.9 Comparison of no. of operations (no. of complex multiplications) required in straightforward computation of DFT and FFT

10.2.2 Decimation in time

Here the subdivision of the input sequence is into even and odd values of the index n, substituted in the following by 2r and 2r + 1 respectively. Again considering the cyclic properties of W_N^n equation (10.28) may be arranged as:

$$F(k) = \sum_{r=0}^{N/2-1} f(2r)\, W_{N/2}^{rk} + W_N^k \sum_{r=0}^{N/2-1} f(2r+1)\, W_{N/2}^{rk} \qquad (10.34)$$

$$F(k) = A(k) + W_N^k B(k) \qquad (10.35)$$

Each sum represents a DFT with N/2 points, the first including the even-numbered points of the original sequence, the second sum the odd-numbered points. Both A(k) and B(k) are periodic with period N/2. The same subdivision as made above for the complete sequence may be repeated for the subsets. It turns out that the required number of computational operations is the same irrespective of which method, decimation in frequency or time, is used. Programming details will determine which procedure is optical in terms of time saving. Some kind of an ordering routine must ordinarily be included to present the final data in a desired form.

10.3 Examples and discussion

In the last section of this chapter we intend to show and discuss two examples of signal transformation, partly because they serve to underline the effect of truncation, inherent in any observational procedure and partly because they expose pitfalls which must be considered in applying the DFT-technique.

For the first example we use a set of illustrations given by Bergland (1969). The second example is taken from Ramirez (1975).

Example 1

DFT used to approximate continuous Fourier transform (CFT). Cosine-wave input.

From (9.58) we know that the spectrum of a cosine waveform consists of two impulse functions, as indicated in the righthand portion of line (a) of figure 10.10.

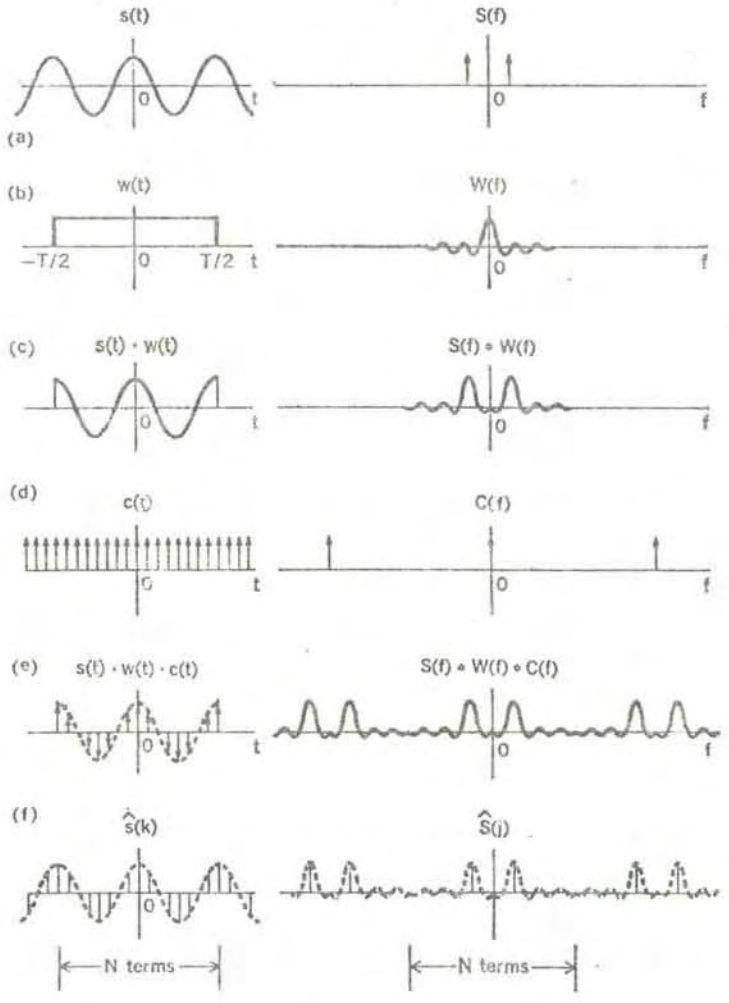


Figure 10.10 Discrete Fourier transform (after Bergland, 1969)

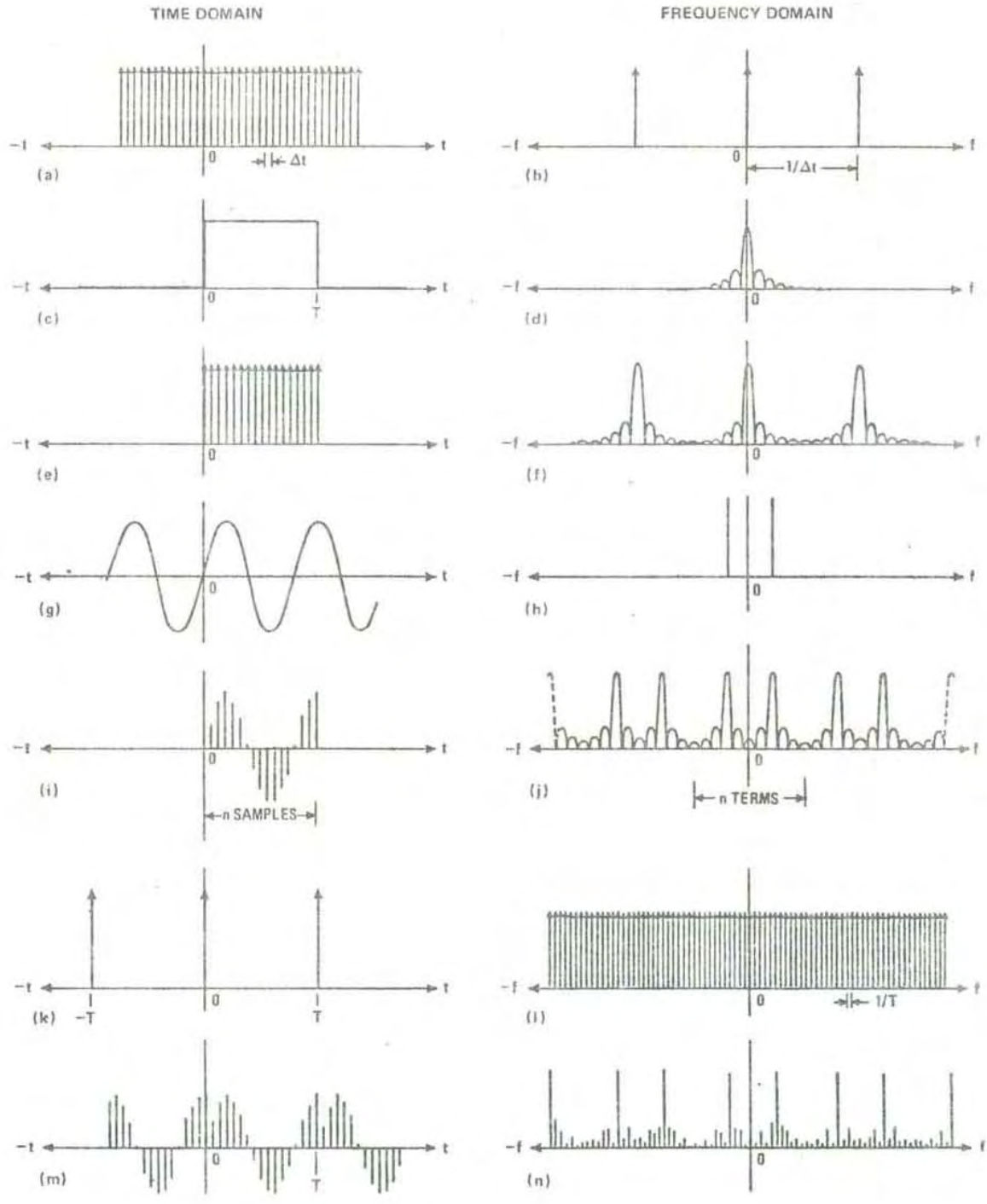


Figure 10.11 Discrete Fourier transform relations (Ramirez, 1975)

However, a pure cosine wave extends over the entire time axis and is nothing but a theoretical concept. Every observed waveform is necessarily a truncated version, and this serves to modify the spectral representation.

In the selected example the wave portion observed extends from +T to -T, equivalent to viewing the wave through a unity amplitude data window whose spectrum is drawn in line (b). Mathematically the truncated wave is described as the product of the two functions s(t) w(t).

From the preceding chapter we know that multiplication in the time domain corresponds to convolution in frequency. The result of convolving the spectra of s(t) and w(t) is described by the double-humped spectrum in line (c). The necessity of observing the waveform over a finite time interval has turned the spectral components at $\pm f_o$ into other frequencies as well. This transfer of signal energy brought about by the truncation in time is referred to as leakage. The truncation corruption evidently depends upon the width of the observing window, and will vanish when $T \to \infty$. The corruption also varies with the form of the data window. Frequently used window functions are the Hamming and Hanning windows tapering off in a specified fashion at the ends of the observing interval.

Sampling is effected by multiplying the truncated wave function with an impulse train function and the spectrum evolves by a folding of the truncated wave spectrum with the impulse train spectrum. The result of this operation is the continuous spectrum depicted in line (e).

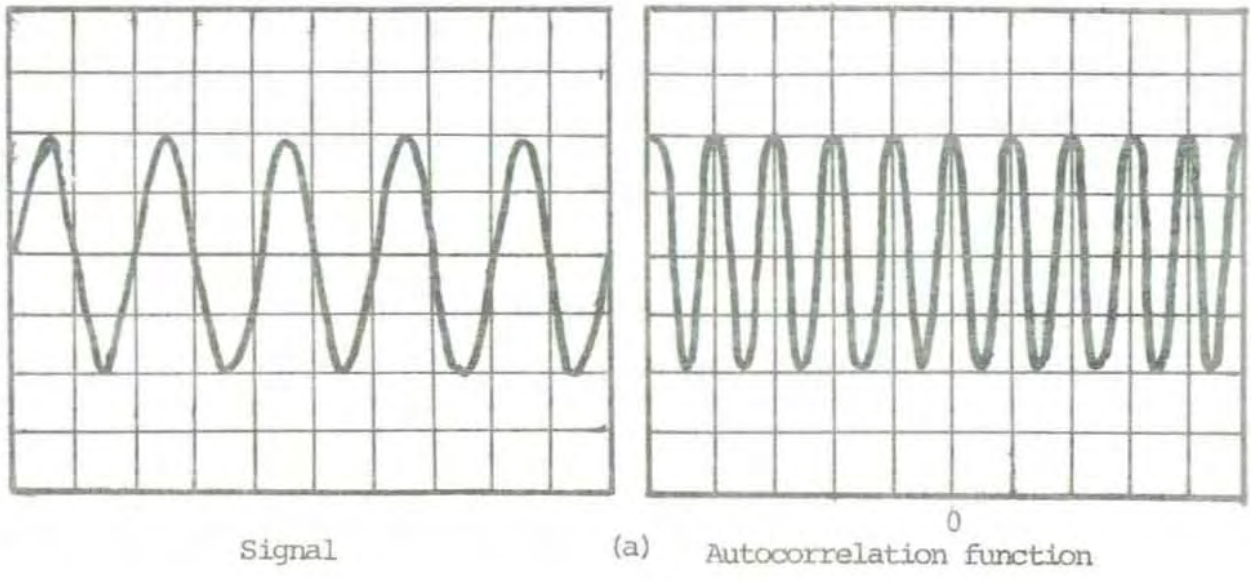
To obtain a discrete frequency spectrum we will have to sample the continuous spectrum, but this forces the time-function to become periodic.

In figure 10.10 the window is an integral multiple of the wave period and the retransformed discrete points in line (f) are seen to represent the input wave fairly well. In practice the sample window will not be an integral multiple of the wave period (in experimental work the input signal will frequently be composed of a large number of sinusoidal waves) and the peridicity assumed by the FFT may lead to curious results. Figure 10.11, taken from Ramirez (1975) illustrates this effect. Further comments are given in the discussion of the next example.

Example 2

Correlation using FFT.

As mentioned above the example is treated by Ramirez (1975). The signal selected is the sine wave portion shown in the lefthand frame of figure 10.12a.



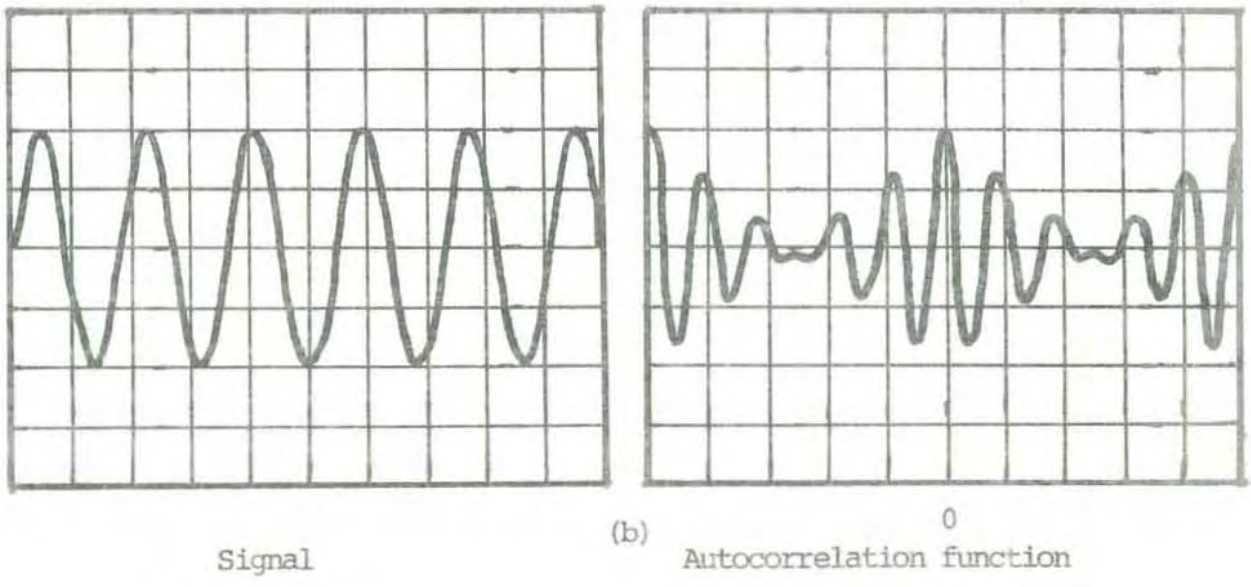


Figure 10.12 Correlation of sine wave with itself:
a) Data window an integral number of wave lengths,
b) Data window an non-integer number of wave lengths

The autocorrelation function is taken to be represented by

$$\varphi(k) = \frac{1}{N} \sum_{k=0}^{N-1} f(t)\ f(t+k) \qquad (10.36)$$

As we know from chapter 9 correlation in the time domain corresponds to complex conjugate multiplication in the frequency domain. This appears to open for a time-saving calculation: The function f(t) is sampled and transferred to the complex frequency domain by an FFT algorithm. Then the signal is conjugated and multiplied by the original transform. Subsequently the signal is inverse-transformed to the time domain.

The retrieved autocorrelation shown in figure 10.12(a) behaves as expected. Maximum value is obtained for zero lag, corresponding to the centre of the screen. The area of the product decreases with increasing value of the lag, is zero for T/4 (90^{O} phase shift) and reaches a negative peak for T/2. As the sine ways exactly overlay at a shift of a full period, T, a maximum is again attained.

In case 10.12(b) the frequency is increased, whereas the window (represented by the screen) remains the same. The autocorrelation function in this case does not at all resemble the sinusoidal pattern obtained in the first case. The explanation is given in terms of the non-integral periods in the window and the periodicity assumed by the FFT algorithm used. Noting the previous discussion the illustration offered by figure 10.13 is selfexplanatory.

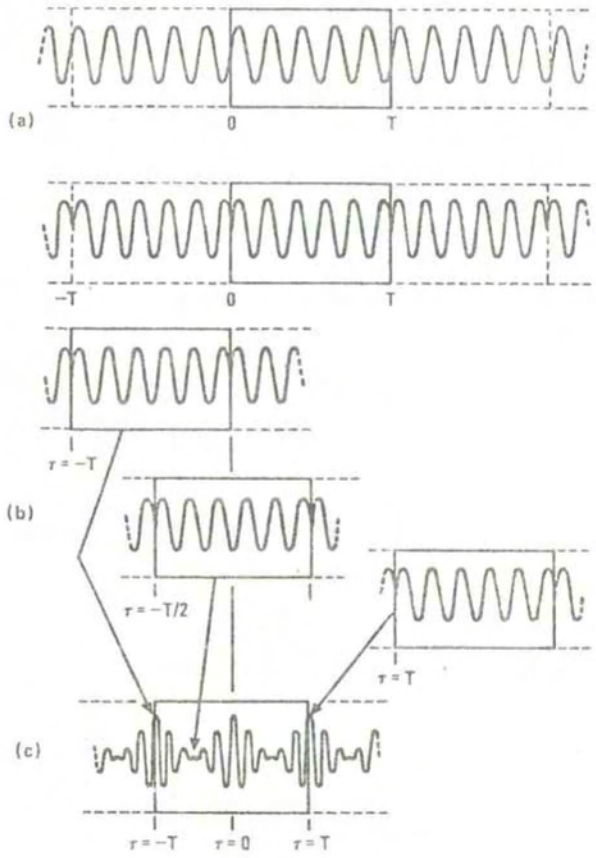


Figure 10.13 Corruption introduced by a non-integer number of wave periods in the window

In closing the chapter we might mention that in observational works a problem arises in that short data segments are desirable for good temperal resolution but this tends as we have noted in this chapter to adversely affect the reliability of the spectral information derived. Assumptions inherent in current methods of analysis, that: (i) the observed waveform repeats itself outside the data window or (ii) the signal is zero outside the window, are both likely to introduce corrupting effects in the derived spectra.

In recent years there has been considerable interest for the so-called maximum-entropy method in spectral analyses. (Ulrych and Bishop, 1975; Moorcroft, 1978). A given time-limited observational sequence may clearly

be combined with several possible records outside the data window. In applying the method quoted above one searches to find that autocorrelation function which is consistent with the greatest number of possible records. It has been verified that under certain conditions the maximum entropy procedure leads to more reliable results than conventional analysing techniques.

Literature:

Stark, H and F B Tuteur - Modern Electrical Communications, Theory and Systems, Prentice-Hall, Inc., New Jersey (1979)

Oppenheim A V and R W Schafer - Digital Signal Processing, Prentice Hall, Inc., New Jersey (1975)

Tretter S A - Introduction to Discrete-Time Signal Processing, John Wiley & Sons, Inc., (1976)

Bergland C D - IEEE Spectrum, p. 41, July (1969)

Ramirez R - Electronics, p. 98, June (1975)

Ulrych T J and T N Bishop - Reviews of Geophysics and Space Physics 13, 183 (1975)

Moorcroft D R - Radio Science 13, 649 (1978)

Mitchell R L - Radar Simulation Course Notes, Technology Service Corporation (1973)

11 LINEAR CIRCUITS AND FILTERS

From time to time in the two preceding chapters we have referred to filters and filtering effects. In this chapter we wish to summarize, in a more systematic fashion, some basic elements of filtering circuit theory. We start with a brief discussion of linearity, time invariance and causality. Thereafter we present some features of analog and digital filtering theory. The z-transform, of central importance in studies of digital filters, is briefly treated.

Needless to say, any presentation within the scope of these lectures must necessarily be a rather compressed, elementary and descriptive account of a subject matter which covers a wide area both in theory and in applications.

11.1 Linearity, time invariance and causality

Consider the following transducer circuit where the output terminal effect, $g(t)$, is determined uniquely by the input $f(t)$.

Figure 11.1 Transducer circuit

The output can be represented by

$$g(t) = \mathcal{K}\left[f(t)\right] \tag{11.1}$$

$\mathcal{K}$ is an operator, signifying the law of determining $g(t)$ from the input $f(t)$. The operator is linear if the circuit obeys the superposition principle:

$$g(t) = L\left[a_1 f_1(t) + a_2 f_2(t)\right] = a_1 L\left[f_1(t)\right] + a_2 L\left[f_2(t)\right] \tag{11.2}$$

The system is time invariant if the parameters of the system are independent of time. Then we may write:

$$g(t-t_1) = L\left[f(t-t_1)\right] \qquad (11.3)$$

where t_1 is an arbitrary constant.

For the following discussion consider figure 11.2

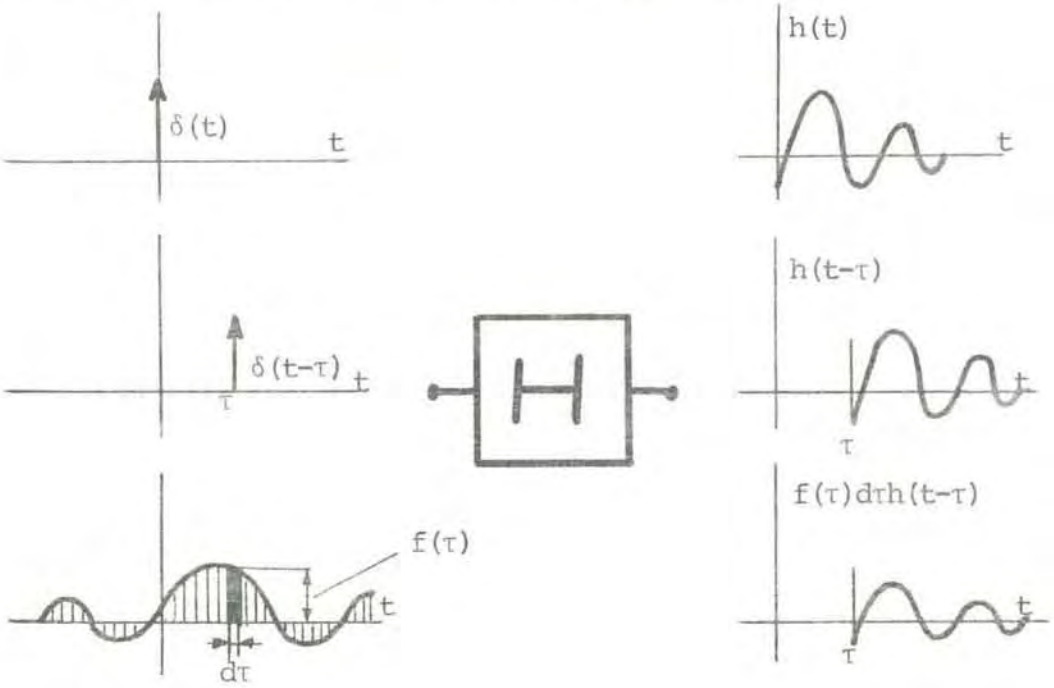


Figure 11.2 Derivation of superposition theorem

h(t) designates the response of the network to an impulse function δ(t) applied at t=0. h(t-τ) is the corresponding response to an impulse occurring at τ. An arbitrary function f(t) may be regarded as composed of an infinite number of impulse functions of strength f(τ)dτ When the superposition principle remains valid, the output is obtained by summing up all the infinitesimal outputs from the beginning to the time t:

$$g(t) = \int_a^t f(\tau)\, h(t-\tau)\, d\tau \qquad (11.4)$$

By this procedure the output appears as the summation of the weighted past. If f(t) is zero outside the integration interval (a,t), nothing is changed by replacing the integration limits by $(-\infty,\infty)$. With this modification we obtain:

$$g(t) = \int_{-\infty}^{\infty} f(\tau)\, h(t-\tau)\,d\tau = \int_{-\infty}^{\infty} f(t-\tau)\, h(\tau)\,d\tau \qquad (11.5)$$

Expression (11.5) can also be written:

$$g(t) = f(t) * h(t) = h(t) * f(t) \qquad (11.6)$$

The output is given as the folding (or convolution) of the input signal with the network's impulse response function:

By Fourier transforming both sides of (11.5) we obtain for a linear, time-invariant system:

$$G(\omega) = \int_{-\infty}^{\infty} g(t) e^{-j\omega t} dt = F(\omega)\, H(\omega) \qquad (11.7)$$

In general a time-invariant system is described by a differential equation with constant coefficients, whose solution consists of two parts: (i) the complementary or transient solution and, (ii) the driven or steady-state solution. The Fourier transform of the impulse response, $H(\omega)$, is sometimes termed the system function, sometimes also frequency response function or the transfer function of the system, and corresponds to the steady-state solution of the differential equation. More generally the transfer function is defined in terms of the Laplace transform and the complex variable

$$s = \sigma + j\omega.$$

A system which cannot respond to a signal prior to its application is termed a causal system. For such a system

$$h(t - \tau) = 0 \quad \text{for} \quad t < \tau \qquad (11.8)$$

11.2 Linear, analog filters

In the following we assume that our filters are ideal, in the sense that they pass a part of the input spectra unattenuated and reject the unwanted parts completely. Any practical circuit can only approximate this behaviour.

In the ideal case the output is a time delayed, scaled replica of the input:

$$g(t) = K\left[f(t-\tau)\right] \tag{11.9}$$

By Fourier transformation:

$$G(\omega) = K\, e^{-j\omega\tau} F(\omega) = H(\omega)\, F(\omega) \tag{11.10}$$

The system function is:

$$H(\omega) = K\, e^{-j\omega\tau} = K\, e^{-j\Theta(\omega)} \tag{11.11}$$

The filter is amplitude - distorted if K varies over the pass-band. Phase-distortion results when $\Theta(\omega)$ varies in a nonlinear fashion.

In the following illustrations we show the frequency responses for a set of conventional, ideal filters.

11.2.1 Lowpass filter

The transfer function is written

$$H(f) = K\,\mathrm{rect}\left(\frac{f}{2f_o}\right)e^{-j2\pi f\tau} \tag{11.12}$$

$$\text{where } \mathrm{rect}\left(\frac{f}{2f_o}\right) = \begin{cases} 1 & \text{for } |f| \leq f_o \\ 0 & \text{for } |f| > f_o \end{cases} \tag{11.13}$$

11.2.2 Highpass/bandstop

For the transfer functions of the two filter types we have:

$$\text{Highpass:} \quad H(f) = K \left[1 - \text{rect} \left(\frac{f}{2f_o} \right) \right] e^{-j2\pi f\tau} \qquad (11.14)$$

$$\text{Bandstop:} \quad H(f) = K \left[1 - \text{rect} \left(\frac{f-f_1}{f_o} \right) - \text{rect} \left(\frac{f+f_1}{f_o} \right) \right] e^{-j2\pi f\tau} \qquad (11.15)$$

The characteristics are graphically displayed in figure 11.4.

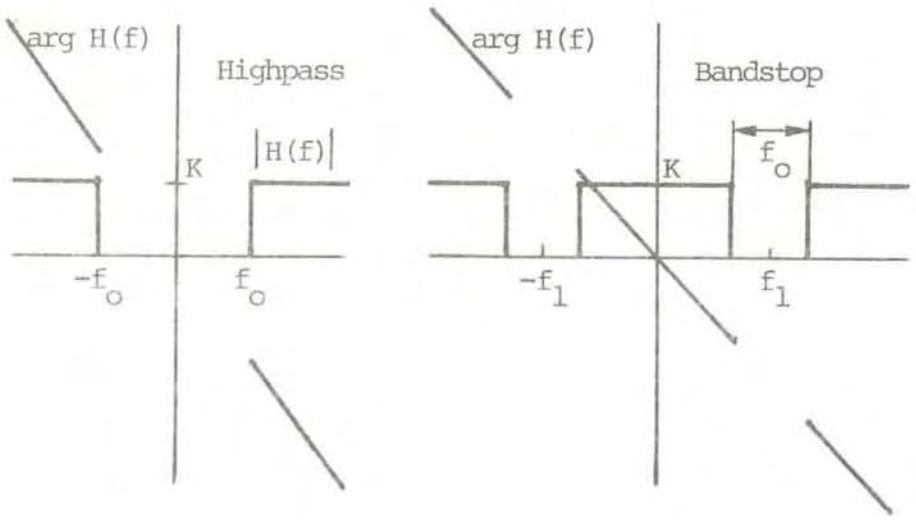


Figure 11.4 Highpass and bandstop characteristics

11.2.3 Bandpass filter

The ideal bandpass characteristic is given by:

$$H(f) = K \left[\text{rect} \left(\frac{f-f_1}{f_o} \right) + \text{rect} \left(\frac{f+f_1}{f_o} \right) \right] e^{-j2\pi f\tau} \qquad (11.16)$$

Response characteristics are shown in figure 11.5.

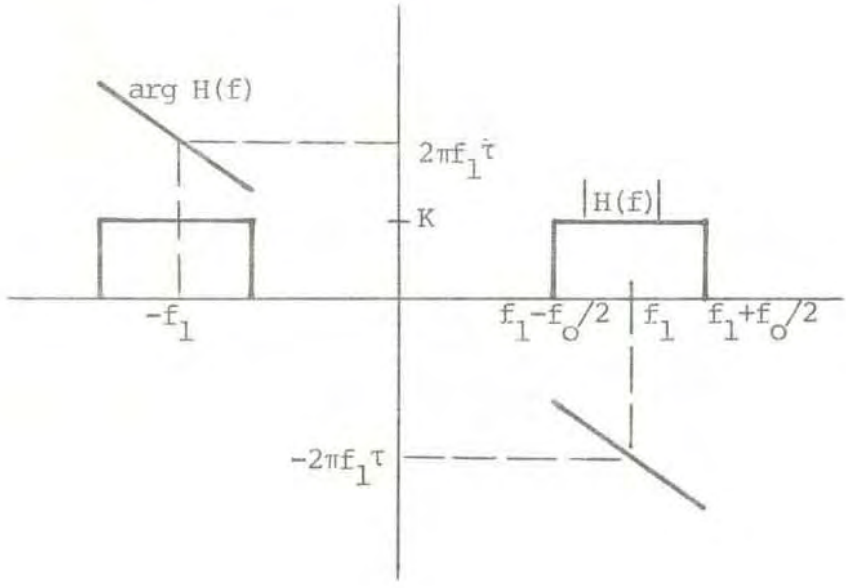


Figure 11.5 Bandpass filter characteristics

The envelope of the bandpass filter, superposed on the high-frequency signal $\cos 2\pi f_1(t-\tau)$, is similar to the impulse response of the lowpass filter.

11.2.4 Causality

None of the filters referred to in 11.2.1 to 11.2.3 are causal filters. Their impulse responses are found to extend beyond $t < 0$. The filters are, therefore, not physically realizable. However, for $f_o\tau >>$, the tail of the impulse response extending to negative values of t may be negligible and without any practical significance. In general, if $h(t) = 0$ for $t < 0$, the frequency response extends beyond the band limits, with ripples occurring in the passband.

11.3 The z-transform

A most important tool for analysing linear, discrete, time-invariant systems is the z-transform, which for sampled, continuous-time waveforms is defined as:

$$F(z) = \sum_{n=-\infty}^{\infty} f(nT)\, z^{-n} \qquad (11.17)$$

where T is the time between consecutive samples and z is a complex quantity. f(nT) denotes the sequence of sampled values. With unit time interval between samples the sequence becomes f(n). In this case the sampling time is numerically equal to the time index n. With a few exceptions we will use this notation in the following. A sequence does not necessarily arise from sampling continuous waveforms. On some occasions f(n) will only be defined for integer values of n. It is customary, though, even in such cases, to refer to f(n) as the n-th sample value.

In mathematical parlance F(z) is the Laurent series expansion of F(z) around z = 0, the coefficients of the series being given by the sampling sequence f(n). For a causal signal, with f(n) = 0 for n < 0, the expansion becomes:

$$F(z) = \sum_{n=0}^{\infty} f(n) z^{-n} \qquad (11.18)$$

In figure 11.6 we illustrate the creation of the z-transform of a causal sequence.

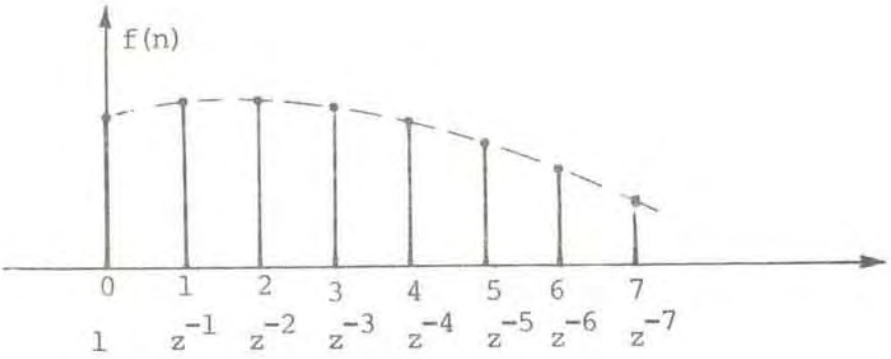


Figure 11.6 Procedure for establishing the z-transform for a causal signal. Each sample is multiplied with an exponential of (1/z), the exponent given by the time index

Equations (11.17) and (11.18) are commonly termed the two-sided (or bilateral) and one-sided z-transforms respectively.

Readers who have worked with the Laplace-transformation, know that stable networks have their poles in the left half of the s-plane (where $s = \sigma + j\omega$).

Taking

$$z = e^{sT} \tag{11.19}$$

the left half of the s-plane maps into an area bounded by a circle in the z-plane, the circle itself corresponding to $s = j\omega$.
For a stable digital filter all poles are within the unity circle in the z-plane.

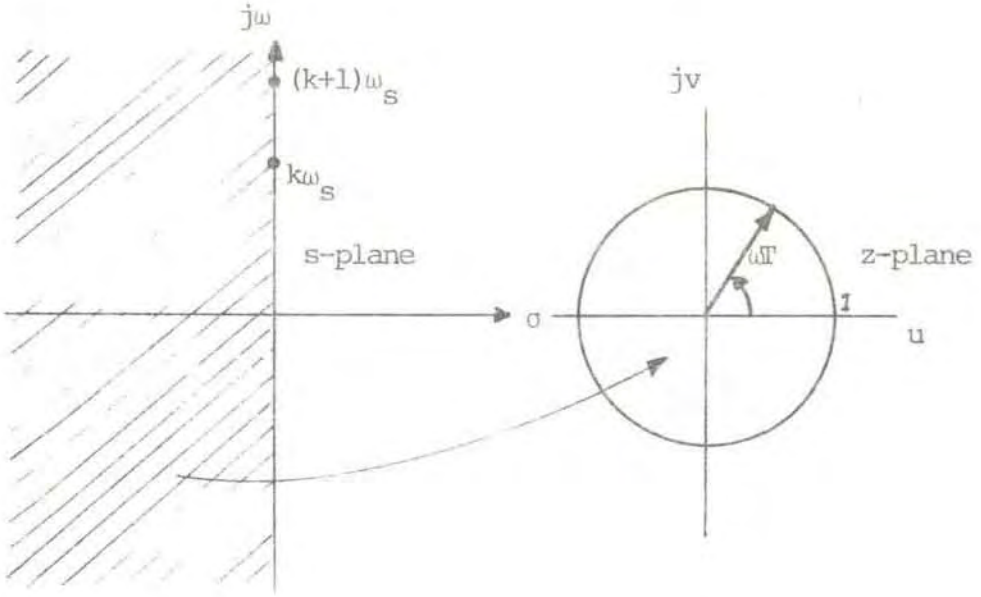


Figure 11.7 Transition from s-plane to z-plane

Note that an increase of the angular frequency from $k\omega_s$ to $(k+1)\omega_s$, where $\omega_s = 2\pi f_s = 2\pi / T$, corresponds to one complete rotation along the unity circle in the z-plane.

Consider the convolution of two sequences, f(n) and s(n):

$$g(n) = \sum_{k=-\infty}^{\infty} f(k)s(n - k) \tag{11.20}$$

The z-transform becomes:

$$G(z) = \sum_{n=-\infty}^{\infty} g(n)z^{-n} = \sum_{k=-\infty}^{\infty} f(k)z^{-k} \sum_{p=-\infty}^{\infty} s(p)z^{-p} \qquad (11.21)$$

$$G(z) = F(z)S(z) \qquad (11.22)$$

The transform of the convolution of two sequences is the product of the transforms of each sequence.

Equation (11.22), termed the discrete-time convolution theorem, is the basis for a treatment of digital filtering.

A few useful transform relationships are compiled in table 11.1.

Sequence	z-transform	Theorem/Remarks
f(n)	F(z)	
s(n)	S(z)	
a f(n) + b s(n)	a F(z) + b S(z)	Linearity
$f(n + n_o)$	$z^{n}o\ F(z)$	Shift of sequence
$a^n f(n)$	$F(a^{-1}z)$	Multiplication by exponential sequence
f(n) * s(n)	F(z) S(z)	Convolution of sequences
n f(n)	-z d F(z)/dz	Differentiation of F(z)

Table 11.1 z-transforms

Consider a discrete-time system with input f(n) and output g(n).

The output can be represented by the difference equation:

$$g(n) = \sum_{k=0}^{N} a_k f(n-k) - \sum_{k=1}^{M} b_k g(n-k) \qquad (11.23)$$

or equivalently:

$$\sum_{k=0}^{M} b_k\, g(n-k) = \sum_{k=0}^{N} a_k\, f(n-k) \qquad (11.24)$$

with $b_o = 1$.

Taking the z-transform of both sides of equation (11.23) yields:

$$G(z) = F(z) \sum_{k=0}^{N} a_k\, z^{-k} - G(z) \sum_{k=1}^{M} b_k\, z^{-k} \qquad (11.25)$$

Introducing the z-transform, H(z), of the impulse-response:

$$H(z) = \sum_{n=0}^{\infty} h(n)\, z^{-n} = \frac{G(z)}{F(z)} = \frac{\sum_{k=0}^{N} a_k\, z^{-k}}{1 + \sum_{k=1}^{M} b_k\, z^{-k}} = \frac{\sum_{k=0}^{N} a_k\, z^{-k}}{\sum_{k=0}^{M} b_k\, z^{-k}} \qquad (11.26)$$

Here the impulse-response has been derived by transforming the difference equation. Conversely, a discrete-time system which has an impulse-response given as a ratio of polynomials, as in equation (11.26), is described by a constant coefficient difference equation.

Given a z-transform it may be desirable to transform back to the original sequence. Formally the inverse z-transform is given by the contour integral:

$$f(n) = \frac{1}{2\pi j} \oint_c F(z) z^{n-1} dz \qquad (11.27)$$

c being a closed contour in the region of convergence of F(z), encircling z = 0. From complex function theory we know that the integral may be evaluated as the sum of the residues of $F(z)z^{n-1}$

at the poles within c. If the z-transform is given as a power series

$$F(z) = \sum_{n=-\infty}^{\infty} f(n) z^{-n} \qquad (11.28)$$

we note that f(n) is simply given as the coefficient of the term involving z^{-n}. Other methods for evaluating the inverse z-transform are also available. The methods are listed below:

- By using the residue theorem
- By power series expansion
- By partial fractional expansion of F(z)
- By long division.

11.4 Digital filters

A digital filter is a device which implements an algorithm for converting a sequence f(n) into another sequence g(n). A digital filter consists of multipliers, delay circuitries and adders, rather than resistors, capacitors and operational amplifiers, from which conventional analog filters are constructed.

Frequently several options are available for realizing the algorithms, including a variety of hardware and software solutions. Which solution shall be selected, often depends upon economy and the required precision. Any practical system will have finite length registers and the following reduction in precision is a primary problem for the filter designer.

Restricting our study to linear, time-invariant filters or processors we let the relationship expressed by equation (11.23) represent our system. We note immediately that if $b_k \neq 0$ for some value of k, the output depends upon the present value of the input, as well as upon the preceding values of both input and output. In this case the system is said to be recursive. If $b_k = 0$ for k = 1, 2, ..., N, the system is called nonrecursive. A nonrecursive system is sometimes denoted a transversal filter or, if N is finite, a finite-duration impulse response filter (FIR).

If G(z) (equation (11.25)) cannot be reduced to a polynomial of finite degree in z^{-1}, then g(n) has an infinite duration and the filter is frequently designated an infinite-duration impulse response filter (IIR).

In digital filtering the processes take place in the time domain. Digital filters are, therefore, sometimes called time-domain filters.

It follows from the classification criteria noted above that nonrecursive filters are without feedback. Such filters are inherently stable. Recursive filters operate with some kind of feedback paths and, as with similar analog structures, the stability of the filters depends upon the gain of the feedback loop.

A schematic model of a filtering arrangement is shown in figure 11.8. In this case the filter acts on a sequence created by sampling a continuous waveform. It may be noted that digital filtering sometimes is applied to processing data with no time dependence.

Figure 11.8 Model of filter system

As stated before, there are numerous ways of approximating a desired filter characteristic. It will, however, be beyond the scope of these lectures to go into a detailed presentation of design principles or practicable realization modes.

For $b_k = 0$ ($k = 1, 2, \ldots$) equation (11.23) describes a nonrecursive

structure. Since the output sequence is given by the convolution of the input- and the pulse response sequences, we have

$$g(n) = \sum_{m=-\infty}^{\infty} h(m)\, f(n-m) \qquad (11.29)$$

Assuming that: $h(m) = 0$ for $m < 0$ and for $m > N-1$ the equation becomes:

$$g(n) = \sum_{m=0}^{N-1} a_m\, f(n-m) \qquad (11.30)$$

The scheme in figure 11.9 implements this relation by a tapped delay line.

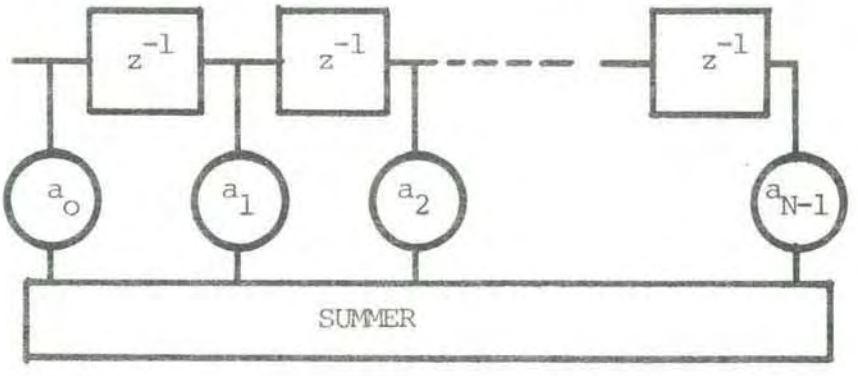


Figure 11.9 Nonrecursive filter (transversal filter)

The response of the filter is determined by the setting of the gains a_0, a_1, ... Note that boxes containing the symbol z^{-1} represent one-unit time delays, in practice they are implemented by shift registers.

Transversal filters can be designed to have precise phase characteristics. Their disadvantage, becoming less series with the rapid advance of large-scale integration, is the large number of multipliers and delay elements needed.

A scheme for realizing the general difference equation (11.23) is given in figure 11.10. The configuration shown is an example of the so-called "direct form 1" of digital filter design.

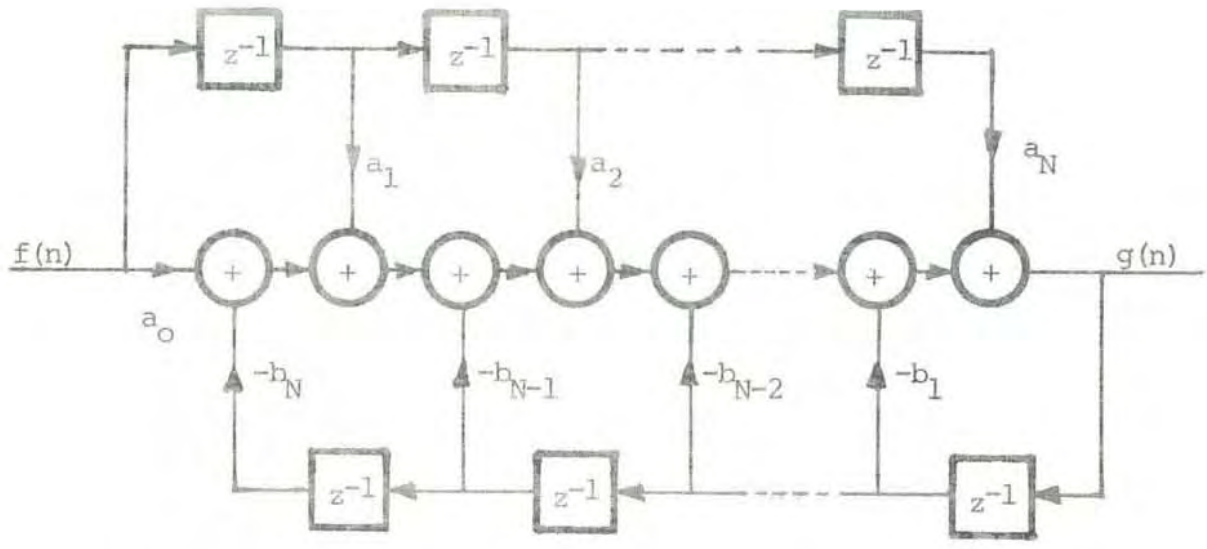


Figure 11.10 Direct form 1

The z-transform of the filter is described by equation (11.26). It is possible to rewrite this equation in different forms, allowing for alternative filter realizations, many of which are described in literature referenced at the end of this chapter.

11.4.1 Some general design considerations

The important role of the Laplace transformation in analog circuit design is well established. In terms of this transformation the system's transfer function is written:

$$H(s) = \int_0^\infty f(t)\, e^{-st} dt \qquad (11.30)$$

where $s = \delta + j\omega$

The behaviour of the system is largely determined by the locations of the poles in zeros in the complex s-plane, stable circuits having their poles in the left half-plane. For digital filters the locations of poles and zeros in the z-plane are of equal importance.

As stated in section 11.3, the lefthand half of the s-plane maps into the interior of the unit circle in the z-plane. In the s-plane

frequency is represented by a point on the imaginary axis; in the z-plane frequency is given by a point on the unit circle. Zero frequency corresponds to the point +1, a quarter of the sampling frequency to +j etc. Once the positions of the poles and zeros inside the unit circle are defined, the frequency response of the network may be determined by the pole and zero vectors as indicated in figure 11.11.

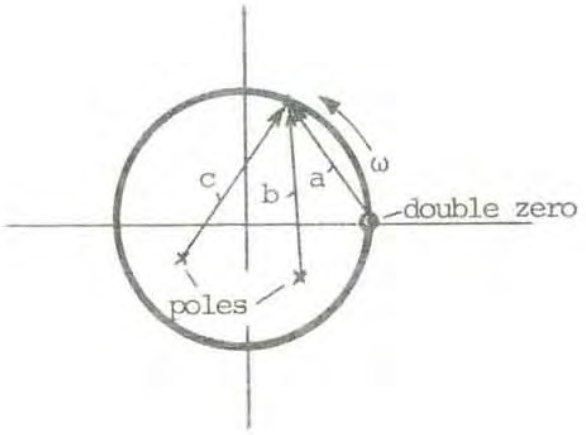


Figure 11.11 Graphical determination of frequency response. Amplitude response a^2/bc. Note that response peaks in the vicinity of the poles.

Although digital filters may be designed without reference to analog equivalents, it is a common procedure to start with an analog filter and subsequently derive the coefficients of the digital counterpart by convenient analog-digital transformations. Textbooks usually list four specific transforming techniques:

- Bilateral transformation
- Impulse invariant transformation
- Mapping of differentials
- Matched z-transform technique

Here we shall have no occasion to discuss the merits of the various techniques. Suffice it to briefly indicate the procedure involved in the use of the impulse invariance method.

The transfer function of an m th-order analog filter may be expanded in terms of partial fractions:

$$H_a(s) = \sum_{i=1}^{M} \frac{c_i}{s+s_i} \qquad (11.31)$$

By use of the inverse Laplace transform:

$$h_a(t) = \sum_{i=1}^{M} c_i e^{-s_i t} \qquad (11.32)$$

The impulse response sequence of the digital filter is derived from the response of the analog filter:

$$h(n) = h_a(nT) = \sum_{i=1}^{M} c_i e^{-s_i n T} \qquad (11.33)$$

We obtain for H(z):

$$H(z) = \sum_{n=0}^{\infty} \left(\sum_{i=1}^{M} c_i e^{-s_i nT} \right) z^{-n} \qquad (11.34)$$

$$H(z) = \sum_{i=1}^{M} \frac{c_i}{1-e^{-s_i T} z^{-1}} \qquad (11.35)$$

By comparing (11.35) and (11.31) we find the mapping relation:

$$\frac{1}{s + a + ib} \quad \rightarrow \quad \frac{1}{1 - e^{-aT} e^{-ibT} z^{-1}} \qquad (11.36)$$

A straightforward approach in designing FIR filters is to truncate an ideal impulse response (of infinite duration) by use of a finite

duration window w(n):

$$h(n) = h_d(n)\ w(n) \qquad (11.37)$$

where $h_d(n)$ represents the ideal response sequence. "Direct" truncation (applied in eq. 11.30) corresponds to a rectangular window with no modification of the impulse sequence embraced by the window. Mathematically the treatment of windows poses problems identical with those met in studies of convergence in the theory of Fourier series. The shape and duration of the window is often determined by a trade-off between a desire to have w(n) as short as possible and the requirement to keep the window's frequency spectrum narrow to accurately reproduce the wanted frequency response. Interested readers will find various window forms described in the literature.

11.5 Some realizable filter types

Here we briefly present the characteristics of some well-proven filter designs, used in analog circuits and also applied as base structures for analog-digital transformations in the construction of digital filters.

(i) Butterworth filters

The filter characteristics can be approximately described by an n th-order (Butterworth) polynomial, $B_n(\omega)$, in ω, being maximally flat at the origin, meaning that as many derivatives of the polynomial as possible go to zero as $\omega \to 0$. For the Butterworth low-pass filter the frequency-response is:

$$|H(\omega)| = \frac{1}{\sqrt{1 + (f/f_o)^{2n}}} \qquad (11.38)$$

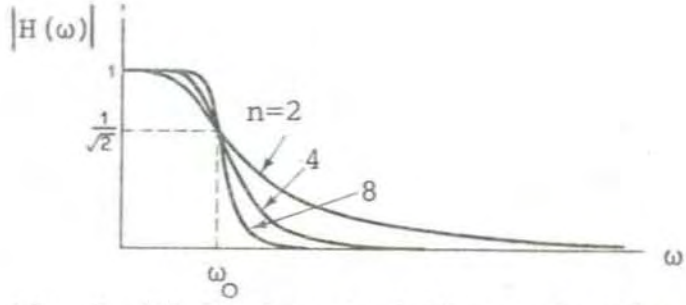


Figure 11.12 Amplitude characteristics as function of n

The Butterworth approximation is seen to be mathematically simple, but has a clear disadvantage in the slow transition between the pass- and the stop-bands.

(ii) Tchebycheff filters

The squared magnitude of the transfer function for this filter type is written

$$|H(\omega)|^2 = \frac{1}{1 + \varepsilon^2 T_n^2 (f/f_o)} \qquad (11.39)$$

where ε is a constant and $T_n(s)$ is the so-called Tchebycheff polynomial of degree n. For $|s| < 1$:

$$T_n(s) = \cos (n \text{ Arc cos } s) \qquad (11.40)$$

The class of Tchebycheff filters has the property that the error of the approximation is uniformly distributed over the passband or the stopband or both. Response graphs for various values of n are shown in figure 11.13.

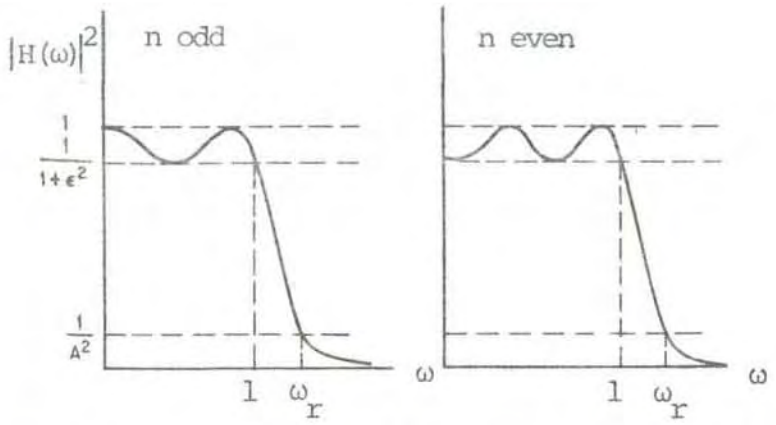


Figure 11.13 Response curves for Tchebycheff low-pass filters

(iii) Elliptic filters

The response characteristics for this class of filters are given by formula (11.39) with the polynomial $T_n(f/f_o)$ replaced by a ratio of polynomials. The elliptic filters are characterized by equiripple magnitude responses both in the passband and in the stopband and a rapid transition between the two bands (see figure 11.14).

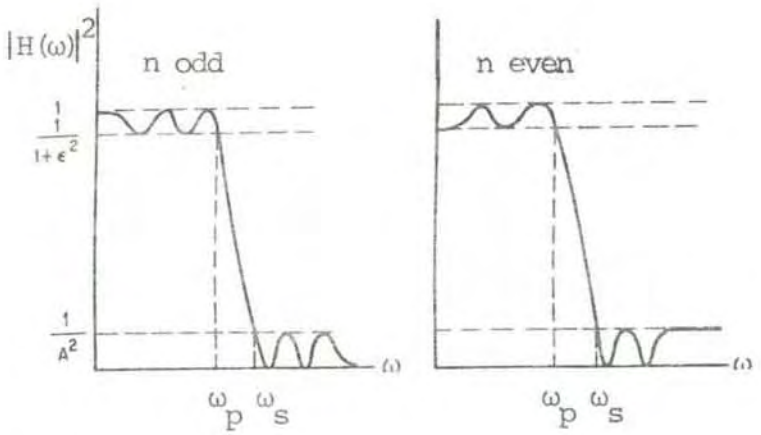


Figure 11.14 Magnitude-squared response of elliptical filters

11.6 Optimum filtering under noisy conditions

In this section we summarily describe two filter types which are both designed for optimum performance, according to criteria to be cited, in a noisy environment. The first is the matched filter and the second the so-called optimum linear filter.

11.6.1 Matched filters

Matched filter theory was treated in fair depth by Leppänen at the Abisko-meeting in March 1979 (Volume I of the meeting report). The presentation here is included for completeness and does not pretend more than underlining a few basic properties of this filter category.

The matched filter has the property of maximizing the signal-to-noise ratio at the output at a given time. The input is a known (deterministic) signal, $s(t)$, corrupted by noise, $n(t)$. The optimizing procedure does not generally allow for a reproduction of the signal form. The primary aim is to recognize whether the signal is present in the noisy environment.

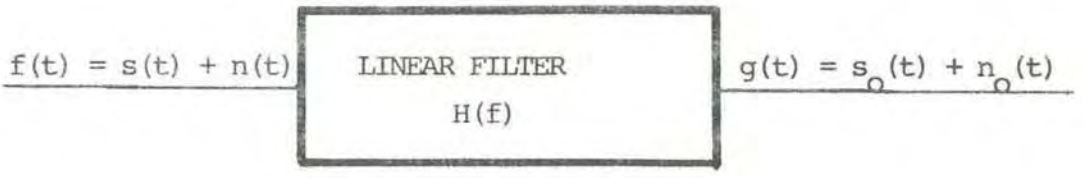


Figure 11.15 Schematic model

The noise is assumed to be stationary and white, the latter term meaning that the spectral density is uniform.

We denote the output signal spectrum

$$G(f) = F(f)\ H(f) \qquad (11.41)$$

where H(f) is to be optimized. By taking the Fourier transform:

$$g(t) = \int_{-\infty}^{\infty} G(f)e^{j2\pi ft}df = \int_{-\infty}^{\infty} F(f)\ H(f)e^{j2\pi ft}df \qquad (11.42)$$

We wish to maximize the envelope of the response, assumed to have its maximum at t_o.

$$|g(t_o)|^2 = \left|\int_{-\infty}^{\infty} F(f)\ H(f)\ e^{j2\pi ft_o}\,df\right|^2 \qquad (11.43)$$

With N_{in} being the input noise power per unit bandwidth, we may write for the mean noise at the output (N_{in} is here the two-sided spectrum, defined for positive and negative frequencies):

$$N = \frac{N_{in}}{2}\int_{-\infty}^{\infty} |H(f)|^2 df \qquad (11.44)$$

By the Schwartz' inequality in complex function theory we have:

$$|g(t_o)|^2 = \left|\int_{-\infty}^{\infty} F(f)\ H(f)\ e^{j2\pi ft_o}df\right|^2 \leq \int_{-\infty}^{\infty} |F(f)|^2 df \int_{-\infty}^{\infty} |H(f)|^2 df \qquad (11.45)$$

The first integral on the right represents the energy, S, of the incoming signal, assumed to be finite; the second integral is $2N/N_{in}$. We obtain:

$$\frac{|g(t_o)|^2}{N} = \frac{S}{N} \leq \frac{2S}{N_{in}} \qquad (11.46)$$

The equality sign in (11.45) applies for

$$H(f) = K\,F(f)^* e^{-j2\pi ft_o} \qquad (11.47)$$

where K is a constant scale factor. t_o is the delay inserted to make the filter realizable.

From Expression (11.46) we find that the optimum signal-to-noise ratio (SNR) at the output is twice the signal energy divided by the input spectral noise density per unit bandwidth. The SNR does not depend upon the particular signal waveform chosen.

Without imposing any loss in the generality of the treatment we set K in (11.47) equal to 1. By inserting (11.47) in (11.42) we have:

$$g(t) = \int_{-\infty}^{\infty} |F(f)|^2 e^{j2\pi f(t-t_o)} df \qquad (11.48)$$

We note that the maximum signal output is obtained for $t = t_o$:

$$g(t_o) = \int_{-\infty}^{\infty} |F(f)|^2 df \qquad (11.49)$$

The peak value of g(t) is thus found to be independent of t_o. The occurence of the peak output may consequently be varied by adjusting the phase term of the matched filter characteristic.

By Fourier transforming (11.47) we find the impulse response of the matched filter:

$$h(t) = f^*(t_o - t) \qquad (11.50)$$

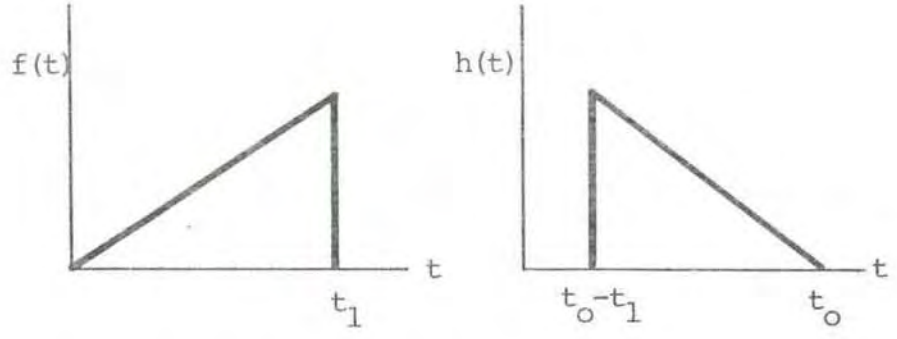


Figure 11.16 Signal and its matched filter response

We note that the impulse response of the matched filter is a time inversed replica of the input signal, meaning that the structure of the filter is determined by the properties of the transmitted waveform. To make the filter physically realizable a time delay t_o ($> t_1$) must be introduced.

The output signal can be written as the convolution of the input, f(t), with h(-t):

$$g(t) = \int_{-\infty}^{\infty} f(\tau)\, h(-t+\tau)\,d\tau = \int_{-\infty}^{\infty} f(\tau)\, f(t - t_o + \tau)\,d\tau = \varphi_{11}(t - t_o) \qquad (11.51)$$

$\varphi_{11}(t - t_o)$ is the delayed autocorrelation function of the input signal, having its maximum value for $t = t_o$.

In radar operational systems matched filters are usually applied in conjunction with pulse compression waveforms. A simple compression scheme is depicted in figure 11.17.

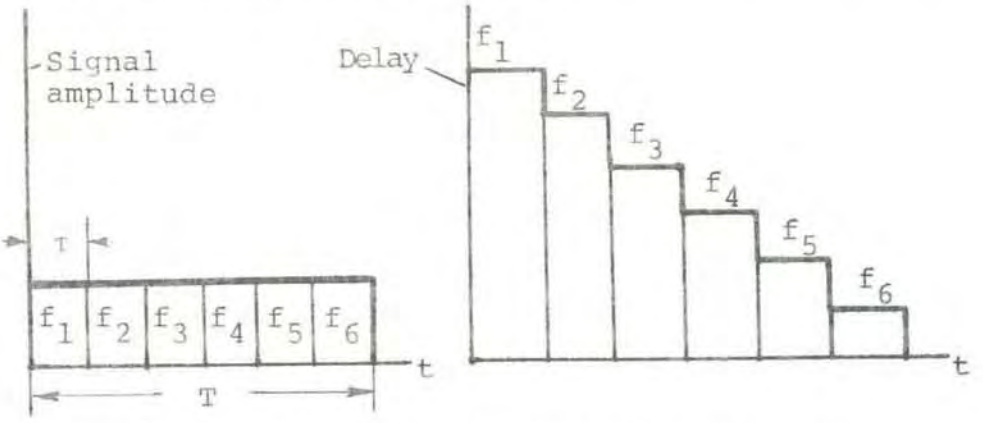


Figure 11.17 Chirp radar signal and delay equalizer

The transmitted waveform consists of adjacent short pulses each of a unique carrier frequency, f_n. Upon reception the pulses are delayed as shown in the righthand sketch, the first frequency having the longest delay. Since the network is assumed to preserve the signal energy, the compressed peak power builds up to the level:

$$P_o = P_i \frac{T}{\tau} = P_i\, D \qquad (11.52)$$

where P_i is the peak power of the input pulse and T/τ the compression ratio.

Practical chirp systems utilizes linearly swept frequencies instead of discrete frequency patterns. Alternative pulse compression schemes are designed for phase-coded waveforms, an example of which being shown in figure 11.18.

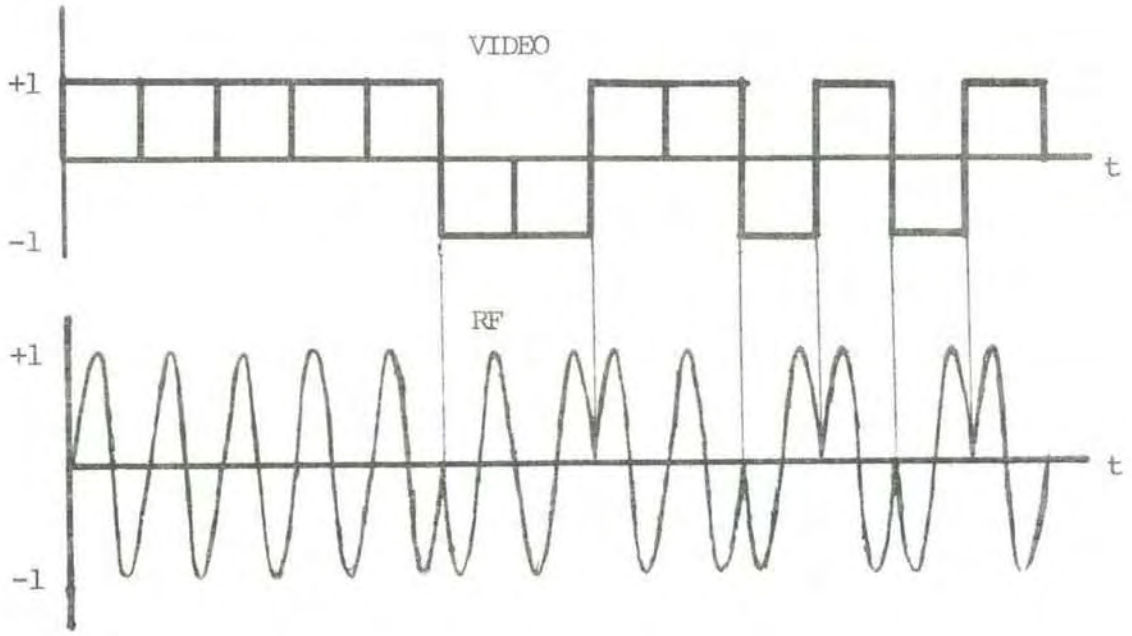


Figure 11.18 Phase-coded signal. Note that pulse amplitudes of + 1 or - 1 correspond to carrier signal phases of 0^o or 180^o respectively

The illustrated waveform is a so-called 13 bit Barker sequence. Barker codes are the most frequently applied binary codes. They are characterized by sidelobes falling between +1 and -1 and a main lobe amplitude N, where N denotes number of bits in the sequence. The autocorrelation function for N=13 is given in figure 11.19.

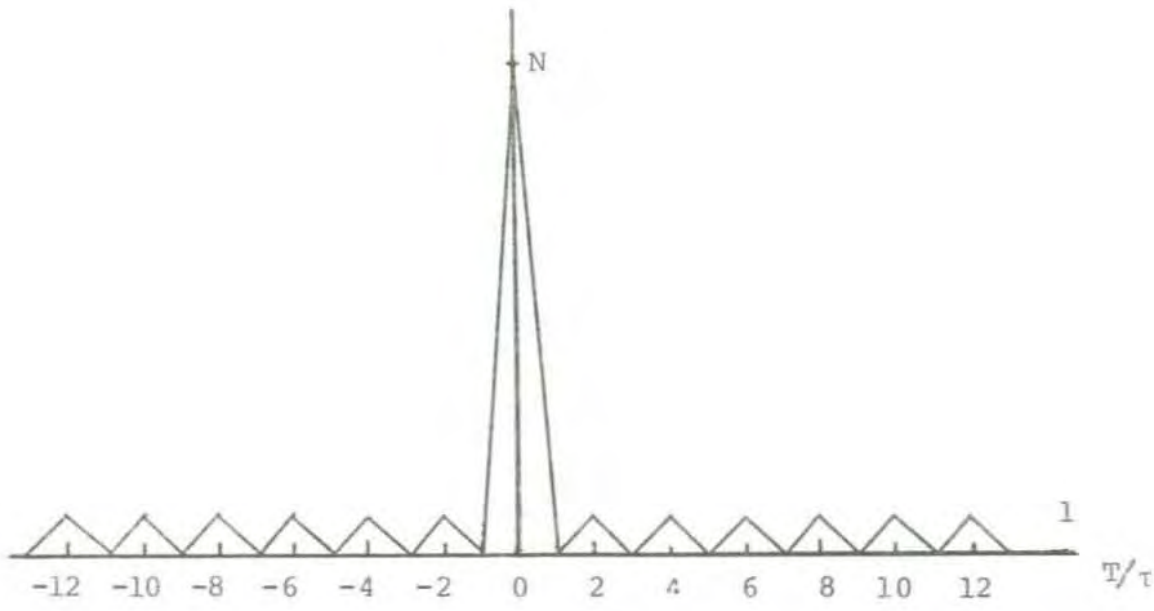


Figure 11.19 Autocorrelation function for 13-bit Barker code

The range resolution offered by the pulse compression technique corresponds to the width of the element pulse of the code. At the same time the main lobe amplitude of the received signal, represents the energy of the composite pulse. In this way a target detectability is obtained which otherwise would correspond to a transmitted pulse power much higher than that actually radiated.

In ionospheric applications the variability of the scatter medium may cause a problem. The matched filter technique will only work when the correlation time of the scatter medium is larger than the time spanned by the coded sequence. In practice little degradation is introduced if the composite pulse length is less than the time to the first zero crossing of the autocorrelation function of the scatter medium (Grey and Farley, 1973).

Another problem is that the sidelobes of the MF autocorrelation function may mask weak echoes which may occur in a situation with multiple scattering.

In the receiver chain the matched filter may work at the IF-stage or at the baseband level. The latter application has been adopted in the EISCAT system, as schematically indicated in the figure below.

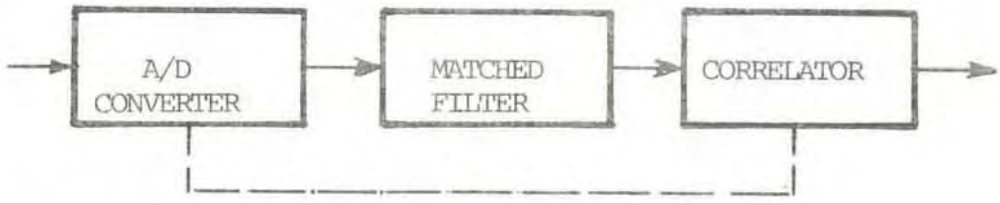


Figure 11.19 Matched filter position in the EISCAT receiver system

For more details on the technical realization of the EISCAT filter we refer to the proceedings of the annual review meeting at Abisco, March 1979 (volume II).

11.6.2 Optimal linear filter

In applications with matched filters the form of the uncorrupted signal is known. This is not the case with the filter we consider in this section. The objective of the optimal linear filter, as we here understand it, is to provide the best fit or the best estimate of a random (stokastic) signal, "best fit" or "best estimate" being understood in terms of minimum mean-square error. It is noted that the problem dealt with in this section belongs to the area of statistical filtering where probability theory provides the mathematical apparatus for formulating the solutions.

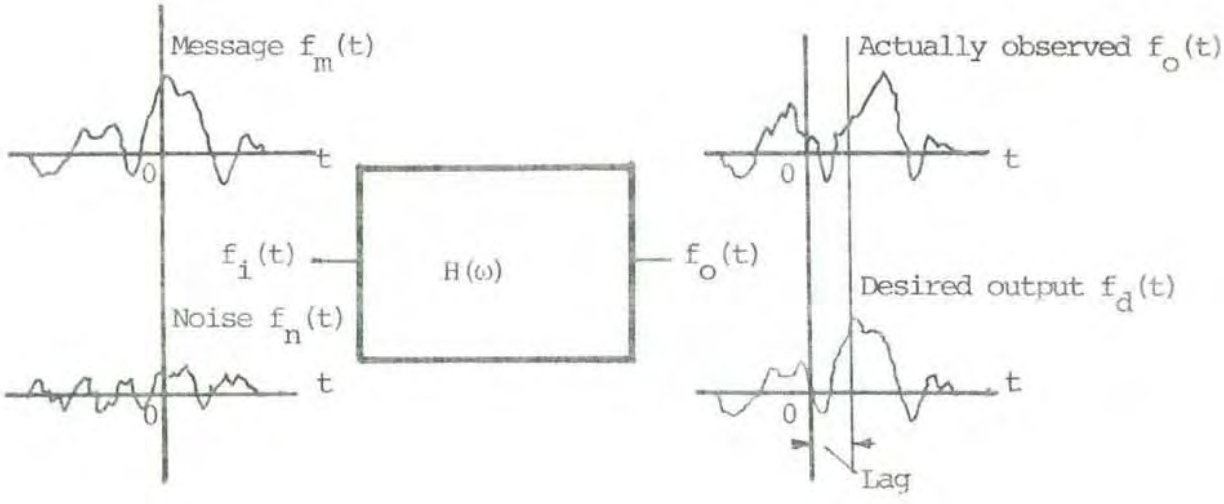


Figure 11.20 Optimum filtering

The input signal, $f_i(t)$, generally consists of a message $f_m(t)$ and noise $f_n(t)$. At the output one measures a signal $f_o(t)$. The desired output is $f_d(t)$. In general the actually measured output differs from the desired output. The problem is to find a linear system with an impulse response $h(t)$ which is such that the mean-square error

$$\overline{\varepsilon^2(t)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} (f_o(t) - f_d(\tau))^2 dt \qquad (11.53)$$

is minimized. We shall not go into discussion of the mathematics of the problem. It may be shown (see f.inst. Lee, 1960) that the problem amounts to finding the impulse response which satisfies the so-called Wiener-Hopf equation:

$$\varphi_{id}(\tau) = \int_{-\infty}^{\infty} h_{opt}(\sigma)\varphi_{ii}(\tau - \sigma)d\sigma \qquad (11.54)$$

Here $\varphi_{ii}(\tau)$ is the autocorrelation function of the input signal; $\varphi_{id}(\tau)$ denotes the crosscorrelation between the input and the desired output. In an alternative formulation the Wiener-Hopf condition states that the impulse response is optimal if, for all time displacements, the error is uncorrelated with the input signal.

For a non-causal filter ($h(\sigma) \neq 0$ for $\sigma < 0$) equation (11.54) can be Fourier transformed and we obtain:

$$H_{opt}(\omega) = \frac{\Phi_{id}(\omega)}{\Phi_{ii}(\omega)} \qquad (11.55)$$

In this case the filter utilizes both past and future input values. It may have a practical significance when the filtering is carried out off-line, with stored signals.

For a causal filter we have:

$$\varphi_{id}(\tau) = \int_0^\infty h_{opt}(\sigma)\varphi_{ii}(\tau - \sigma)d\sigma \qquad \tau \geq 0 \qquad (11.56)$$

If $f_i(t)$ is white (implying a constant effect spectrum), $\varphi_{ii}(\tau) = \delta(\tau)$:

$$\varphi_{id}(\tau) = \int_0^\infty h_{opt}(\sigma)\delta_{ii}(\tau - \sigma)d\sigma \qquad (11.57)$$

$$\varphi_{id}(\tau) = h_{opt}(\tau) \qquad (11.58)$$

Generally $f_i(t)$ is not white. It may be shown, however, that for rational spectra a filter can be found, making the output white when $f_i(t)$ is fed in. The solution to the filter problem in this case can be envisaged as a cascade arrangement of the whitening circuitry, with output $g(t)$, followed by a network whose impulse response is given by the equation:

$$h_{opt}(\tau) = \begin{cases} \varphi_{gd}(\tau) & \text{for } \tau \geq 0 \\ 0 & \text{for } \tau < 0 \end{cases} \qquad (11.59)$$

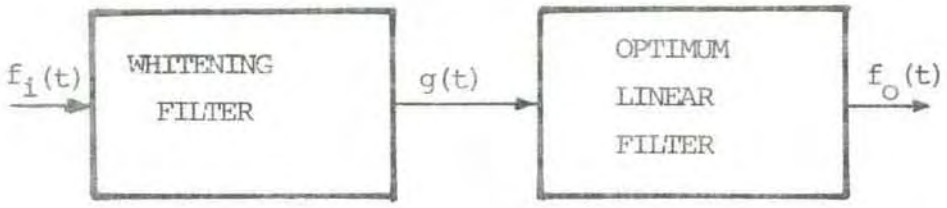


Figure 11.21 Optimal filter preceded by a whitening network

11.7 Examples and comments

(i) Digital feedback system

Consider the following arrangement:

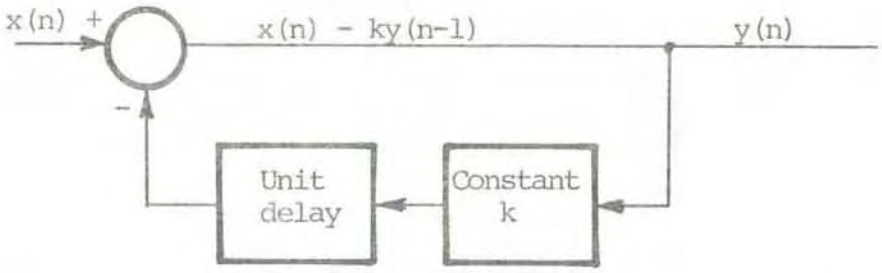


Figure 11.22 Schematic of digital feedback system

The impulse responses (input 1), with the corresponding z-transforms, for the constant- and unit delay - boxes are:

	Impulse response	z -transform
constant box	k, 0, 0, ...	$k + 0 \times z^{-1} + 0 \times z^{-2} \ldots = k$
unit delay box	0, 1, 0, ...	$0 \times z^{0} + 1 \times z^{-1} + 0 \times z^{-2} \ldots = z^{-1}$

The fundamental equation becomes:

$$y(n) = x(n) - k\, y(n-1) \qquad (11.60)$$

z-transform:

$$Y(z) = X(z) - k\, z^{-1}\, Y(z) \qquad (11.61)$$

Transfer function:

$$F(z) = \frac{Y(z)}{X(z)} = \frac{1}{1+k\, z^{-1}} \qquad (11.62)$$

y(n-1) in equation (11.60) may be expressed in terms of x(n-1) and y(n-2):

$$y(n) = x(n) - k\,[x(n-1) - k\,y(n-2)] = x(n) - k\,x(n-1) + k^2\,y(n-2) \qquad (11.63)$$

By continuing the substitution:

$$y(n) = x(n) - k\,x(n-1) + k^2 x(n-2) - k^3 x(n-3) + \ldots. \qquad (11.64)$$

If $|k| < 1$ the sequence

$1, -k, k^2, -k^3, \ldots\ldots$

tends to zero and the closed loop system is stable.

Let us as a numerical example select $k = 0.5$. At time index 2 we have:

$$y(2) = x(2) - 0.5x(1) + 0.25x(0) - 0.125x(-1) + \ldots.$$

On the other hand, if $|k| > 1$, the system is unstable and the feedback grows with respect to time.

If for some reason our system feedback amplification must remain > 1, then we can only obtain stability with respect to future time. In practical terms this implies that we have to delay our system output.

Going back to (11.60) and considering future input values:

$$y(n-1) = \frac{1}{k}\,x(n) - \frac{1}{k^2}\,x(n+1) + \frac{1}{k^3}\,x(n+3) + \ldots. \qquad (11.65)$$

Selecting $k = 2$ as a numerical example:

$$y(n-1) = 0.5x(n) - 0.25x(n+1) + 0.125x(n+3) + \ldots.$$

This series is convergent.

(ii) Memory and anticipation

We know that an output sequence of a network is given by the convolution of the input- and pulse response sequences (expression 11.29):

$$g(n) = \sum_{m=-\infty}^{\infty} h(m)\, x(n - m) \qquad (11.66)$$

With $h(m) = 0$ for $m < 0$ we find that the operator coefficients only operate on present and past values of the input. The impulse response function in this case may therefore be termed the memory function of the filter.

When $h(m) = 0$ for $m \geq 0$ the convolution formula becomes:

$$y(n) = \sum_{m=-\infty}^{-1} h(m)\, x(n - m) \qquad (11.67)$$

In this case the non-vanishing values of the operator coefficients only operate on future values of the input and the impulse function may be termed the anticipation function of the digital filter. As explained in the preceding example, the operation represented by the anticipation function presupposes that the computation has been delayed until the needed values of the future inputs are available.

(iii) Deconvolution

A common practical problem is the following: Given a wave sequence which has undergone convolution. Find the original signal.

We first discuss the problem for a 2-length memory system b(0), b(1), and thereafter indicate the solution of the problem of deconvolution of a waveform of arbitrary length.

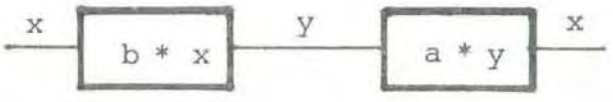


Figure 11.23 Schematic display of convolution and deconvolution process

Without loss of generality we assume: b(0)=1 and b(1)=k.

From the tandem arrangement shown in figure 11.23 we have:

$$x = a * b * x = \delta * x \qquad (11.68)$$

The problem consists of finding the sequence a(0), a(1), a(2) .... such that the convolution

$$(a(0), a(1), a(2) ....) * (1, k) = (1, 0, 0, ...)$$

In terms of z-transforms:

$$a(0) + a(1)z^{-1} + a(2)z^{-2} .... = \frac{1}{1+kz^{-1}} \qquad (11.69)$$

By a series development:

$$\frac{1}{1+1z^{-1}} = 1-kz^{-1} + k^2z^{-2} - k^3z^{-3} + ... \qquad (11.70)$$

which gives:

$$a(n) = \begin{cases} 0 & \text{for } n = -1,-2, ... \\ (-k)^n & \text{for } n = 0, 1, 2, ... \end{cases} \qquad (11.71)$$

To retrieve the original signal the inverse memory function, a, must be infinitely long. The truncation error introduced in practical applications will depend upon the value of k. For $|k| > 1$ stability problems arise.

In the more general case the original signal has been convolved by a response function having an arbitrary length:

$$b = \{ b(0), b(1), \ldots . b(n) \} \qquad (11.72)$$

with z-transform:

$$B(z) = b(0) + b(1)\, z^{-1} + \ldots . + b(n)\, z^{-n} \qquad (11.73)$$

By factorization:

$$B(z) = (\beta_{oo} + \beta_{o1} z^{-1})(\beta_{1o} + \beta_{11} z^{-1}) \ldots . (\beta_{no} + \beta_{n1} z^{-1}) \qquad (11.74)$$

Since multiplying polynomials amounts to convolving their coefficients:

$$b = (\beta_{oo}, \beta_{o1}) * (\beta_{1o}, \beta_{11}) * \ldots\ldots\ldots\ldots * (\beta_{no}, \beta_{n1}) \qquad (11.75)$$

A box with a long memory function is equivalent to a sequence of boxes with 2-length memory functions.

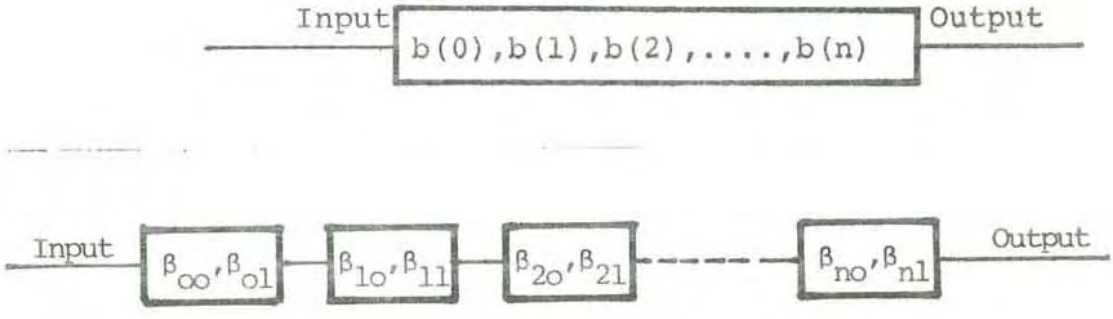


Figure 11.24 Equivalent systems

A numerical example: The memory function 4, 12, 5 has the z-transform

$$4 + 12\ z^{-1} + 5\ z^{-2} \tag{11.75}$$

which may be factored as

$$(2 + z^{-1})\ (2 + 5\ z^{-1}) \tag{11.76}$$

By the polynomial division:

$$\frac{1}{b(0) + b(1)z^{-1} + -- + b(n)\ z^{-n}} \tag{11.77}$$

the coefficient of the sequence being the inverse of b are found. Alternatively, the inverse of each 2-length wavelet is derived. By convolving these together the desired inverse response coefficients may be obtained.

(iv) Some practical considerations in the design of digital filters

As mentioned before in this chapter it is a commonly used design procedure to derive digital filter coefficients by a transformation from analog filter devices. Section 11.4.1 lists four specific transformation techniques.

The method of mapping differentials, transforms a stable continuous system into a stable discrete system. It does not, however, have the desired property that the continuous frequency axis (the imaginary axis in the s-plane) maps into the unit circle in the z-plane. The technique does not appear to be much used in digitizing analog filters.

The impulse invariant method and the matched z-transform technique are quite similar. For an analog filter response, h(t), the z-transform, H(z), with the impulse invariant method, is obtained by taking the z-transform of the sampled version, h(nT), of the impulse response. Mathematically the replacement is described by (11.36). The matched

z-transform gives the same poles as the impulse invariant technique, but the zeros are different. Both methods are subject to aliasing, which may seriously distort the transfer function when the critical frequency is less than half the sampling frequency.

With the bilinear transformation the replacement relation is given by the conformal mapping

$$s \longrightarrow \frac{1 - z^{-1}}{1 + z^{-1}} \tag{11.78}$$

In this case the imaginary axis in the s-plane maps into the unit circle in such a way that the entire analog frequency range corresponds to one rotation in the z-plane. This implies that aliasing does not appear by use of the bilinear mapping technique. However, the analog and the digital frequency variables, ω_A and ω_1, are interrelated by the following equation

$$\omega_A = \tan(\omega T/2) \tag{11.79}$$

which decribes a highly nonlinear relationship. In compensating for the frequency distortion, applying prewarping, the critical set of digital cutoff frequencies are first specified. Subsequently these frequencies are transformed to analog critical frequencies, using (11.79). These latter frequencies specify the analog filter which in turn is applied to derive the digital filter coefficients.

As an example we consider the n-th order Butterworth filter with squared amplitude response (11.38):

$$\left|H(\omega)\right|^2 = \frac{1}{1 + (\omega/\omega_o)^{2n}} \tag{11.80}$$

Introducing $\omega = s/j$ and considering only the poles in the left-hand half plane (Tretten, 1976), we obtain:

$$H(s) = \frac{\omega_o^n}{\prod_{k=0}^{n-1} (s - s_k)} \tag{11.81}$$

Selecting the 3 dB cutoff frequency $\omega_o = \omega_s/4$ we obtain for the warped frequency (11.79):

$$\omega_o = \text{tg}\left(\frac{2\pi T}{4T2}\right) = \text{tg}\,\frac{\pi}{4} = 1 \qquad (11.82)$$

Specifying the Butterworth filter to be of 3rd order the poles are at -1 and at $e^{j2\pi/3}$ and $e^{-j2\pi/3}$.

The transfer function becomes:

$$H(s) = \frac{1}{s^3 + 2s^2 + 2s + 1} \qquad (11.83)$$

and by the substitution $s = (1 - z^{-1})/(1 + z^{-1})$

$$H(z) = \frac{1}{6}\,\frac{1 + 3z^{-1} + 3z^{-2} + z^{-3}}{1 + \frac{1}{3}z^{-2}} \qquad (11.84)$$

which can be realized by the difference equation:

$$g(n) = \frac{1}{6}\,f(n) + \frac{1}{2}\,f(n-1) + \frac{1}{2}\,f(n-2) + \frac{1}{6}\,f(n-3) - \frac{1}{3}\,g(n-2) \qquad (11.85)$$

References and general reading:

Grey R W and D T Farley - Theory of incoherent-scatter measurements using compressed pulses, Radio Sci 8, 123 (1973)

Lee Y W - Statistical Theory of Communication, John Wiley & Sons, Inc. (1960)

Oppenheim A V and R W Schafer - Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliffs, New Jersy (1975)

Rabiner L R and B Gold - Theory and Application of Digital Signal Processing, Prentice-Hall, Inc. (1975)

Robinson E A - Statistical communication and detection, Griffin, London (1967)

Stark H and F B Tuteur - Modern Electrical Communications. Theory and Systems, Prentice-Hall, Inc., Englewood, Cliffs, New Jersey (1979)

Tretter S A - Introduction to Discrete-Time Signal Processing, John Wiley & Sons, Inc. (1976)

12 ANTENNAE

In this chapter we intend to describe some characteristics of the EISCAT antennae. Parts of the antenna systems, notably the UHF polarizer and the switchyard used in the VHF system, are already dealt with in chapter 6 and 7 respectively.

For completeness and for turorial purposes we start with a few basic definitions and some excerpts of general antenna theory. Thereafter we go on to describe specific features of the EISCAT antennae, including a set of geometrical relationships pertinent ot the phase steering of the VHF antenna. Towards the end of the chapter we briefly treat the problem of radiation hazard as related to the conditions near the transmitting antennae at Ramfjordmoen. The final section presents the results of mapping the horizon profile at the Ramfjordmoen site by passive and active measurements carried out using the EISCAT UHF-antenna.

12.1 Definitions

A completely isotropic radiator, having uniform distribution of radiation intensity in all directions, cannot be realized. However, the concept serves a worthwhile purpose as a principal reference in antenna theory.

For any antenna the averaged total radiated power is represented by the integral:

$$W_t = \int_S P_r \, dS \qquad (12.1)$$

where P_r denotes the averaged power density (Watts/m^2) on the spherical surface S surrounding the antenna. For an isotropic radiator:

$$P_r = \frac{W_t}{4 \pi r_o^2} \qquad (12.2)$$

where r_o is the radius of the sphere.

The averaged isotropic radiation intensity, defined as the average power per solid angle, is given as:

$$U_o = \frac{W_t}{4 \pi} \qquad (12.3)$$

The antenna gain function, $G(\Omega)$, generally depends upon the direction Ω (specified by aximuth Φ and elevation Θ). It is defined as:

$$G(\Omega) = \frac{U(\Omega)}{U_o} \qquad (12.4)$$

where $U(\Omega)$ is the radiation intensity in the direction considered.

The antenna gain is represented by the maximum value attained by expression (12.4):

$$G_a = \frac{U_{max}}{U_o} \qquad (12.5)$$

which is the ratio of maximum radiation intensity to the radiation intensity of a lossless isotrope fed by the same input power. The capability of the antenna in directing

radiated power is given by the radiation pattern:

$$r(\Omega) = \frac{G(\Omega)}{G_a} = \frac{U(\Omega)}{U_{max}} \qquad (12.6)$$

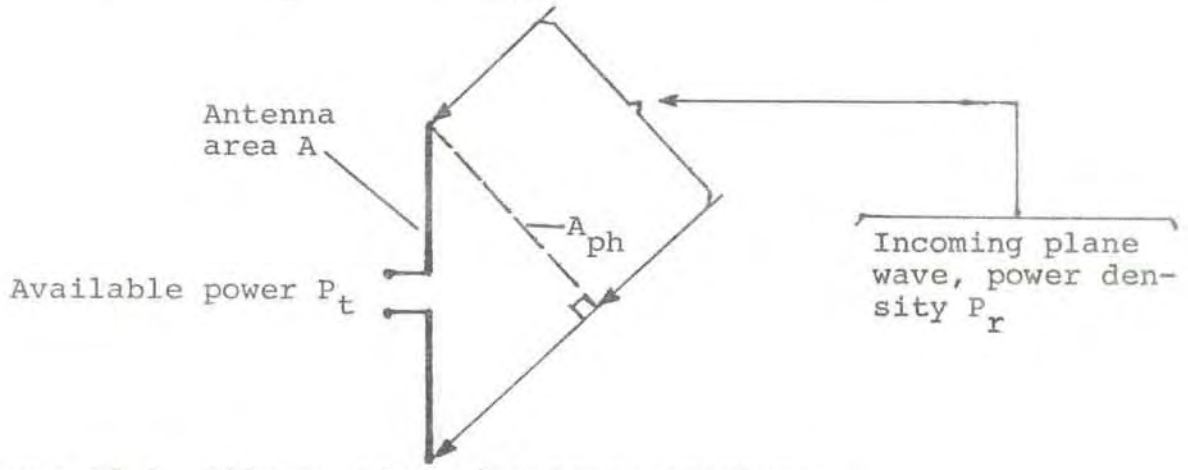


Figure 12.1 Illustration of antenna aperture.

In figure 12.1 A signifies the geometrical area of an antenna. A_{ph} is the cross-sectional area seen by an incoming wave. The antenna's capture area or effective area describes the efficiency of the antenna in extracting energy from an incoming wave and making this energy available at the terminal of a matched receiver. If P_t is the power availbable at the terminal, we have:

$$P_t = A_{eff} P_m \qquad (12.7)$$

P_m is the component of the source radiation matching the polarization of the antenna. Note that the power density, P_r, of the incoming wave is larger, or equal to, the matched radiation, $P_m \leq P_r$. The proportionality factor, A_{eff}, in (12.7) is by definition the capture area. We should emphasize that A_{eff} depends upon direction as well as upon the losses incurred in transforming the intercepted wave energy into terminal energy at the receiver input.

An antenna aperture efficiency may be defined as:

$$\eta = \frac{A_{eff}}{A_{ph}} \qquad (12.8)$$

Generally η is between 0 and 1. However, theoretically, in supergain antennae, η may exceed 1.

The effective area of an antenna and the gain are related as follows (Schelkunoff, 1943):

$$A_{eff} = \frac{G_a \Lambda^2}{4 \pi} \qquad (12.9)$$

It follows from the reciprocity theorem in electromagnetic theory that characteristics like antenna gain, radiation pattern and impedance remain the same irrespective of whether the antenna is used for transmission or reception.

12.2 Diffraction integral and radiation zones

A transmitter antenna serves the function of converting guided electromagnetic energy into freely propagating waves, whereas the receiving antenna has the opposite function. The mathematical apparatus for describing the wave fields and the antenna current distribution, be it the source of the waves or induced by incoming radiation, is provided by Maxwell's equations. In the source-free region the wave fields, the electric or magnetic field intesities, are described by the equation:

$$\nabla^2 \overline{F} + k^2 \overline{F} = 0 \qquad (12.10)$$

where $k = 2\pi/\Lambda$, $\overline{F}$ denotes the vector field and ∇^2 is the Laplace operator:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^3}{\partial z^2} \qquad (12.11)$$

Each rectangular component of the electric field (E_x, E_y, E_z) or the magnetic field (H_x, H_y, H_z) obeys the scalar Helmholtz equation:

$$\nabla^2 \Psi + k^2 \Psi = 0 \qquad (12.12)$$

The Kirchoff's diffraction integral follows from equation (12.12) and Green's theorem (Rusch and Potter, 1970):

$$\Psi(x, y, z) = \frac{1}{4\pi} \int_{S_1} \left(\Psi \frac{\partial}{\partial \eta}\left(\frac{e^{-jkr}}{r}\right) - \frac{e^{-jkr}}{r} \frac{\partial \Psi}{\partial \eta}\right) dS \qquad (12.13)$$

Here $\Psi(x, y, z)$ signifies the field at the point of observation (x, y, z). S_1 is a bounded surface withing the medium, which is assumed to be homogeneous and isotropic. $\partial\Psi/\partial\eta$ is the normal derivative of Ψ at the surface. r denotes the distance from the incremental element $d\xi d\eta$ on S_1 to the field point.

In applying relation (12.13) one must know the value of Ψ and $\partial\Psi/\partial\eta$ on the surface. Kirchoff's approximations assume that:

(i) Ψ and $\partial\Psi/\partial\eta$ are zero except in the openings of S_1.

(ii) The values of Ψ and $\partial\Psi/\partial\eta$ in the openings are those given by an incident wave without the presence of screens or obstacles.

It may be shown that assumptions inherent in Kirchoff's theory lead to mathematical inconsistencies (Jackson, 1975). Moreover, the scalar nature of the approach limits the value

of the diffraction integral in describing the vectorial structure of the electromagnetic fields. However, in many practical circumstances, with apertures being large compared to the wavelength, the Kirchoff theory represents a good approximation to the results obtained by more rigorous treatments. It serves as a basis for classical works within the field of geometrical optics and we will use it in the following in a brief discussion of the traditional subdivision of the field of a radiating structure.

We assume that the radiation is from a planar aperture as shown in figure 12.2

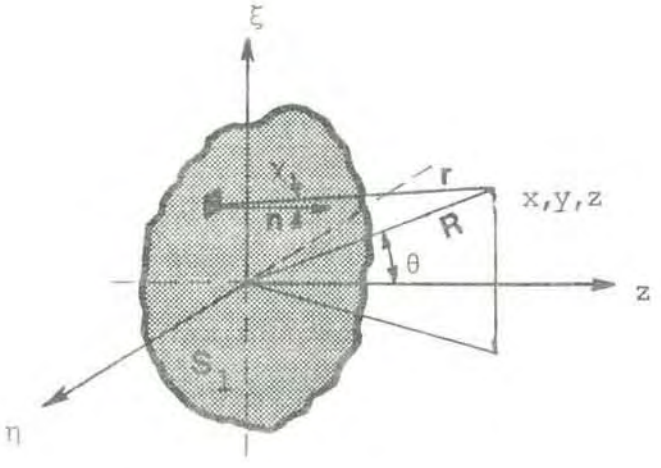


Figure 12.2 Geometry of diffraction field from a planar aperture.

Assuming an illumination function

$$F(\xi,\eta) = F_o(\xi,\eta)\, e^{-jk\,\Phi(\xi,\eta)} \qquad (12.14)$$

(12.13) may be transformed into:

$$\Psi(x, y, z) = \frac{1}{4\pi} \int_{S_1} \left[F(\xi,\eta)\, \frac{e^{-jkr}}{r} \left(jk + \frac{1}{r}\right)\cos\chi + jk\, \frac{\partial \Phi}{\partial \eta} - \frac{1}{F_o} \frac{\partial F_o}{\partial \eta} \right] d\xi d\eta \qquad (12.15)$$

The traditional subdivision of the antenna field into three regions stems from the mathematical simplifications introduced to solve (12.15):

- Near zone: this is the region in the immediate neighbourhood ($r < \Lambda$) of the antenna. No approximations are allowed in solving the integral in this domain, in fact, the integral is itself an approximation. Here the field is almost equal to the aperture field, propagating along geometrical rays from the surface.

- Fresnel zone:

At larger distances phase differences along different paths cause fluctuations in the patterns. At the Rayleigh distance, $0.5D^2 / \Lambda$ (D being the largest dimension of the aperture), the pattern changes from essentially parallel- to angular-beam characteristic.

In the Fresnel region the following approximations are considered adequate:

$$\begin{aligned} 1/r &\ll k \\ 1/r &\approx 1/R \\ \cos\chi &\approx \cos\Theta \end{aligned}$$

Differences in distance from the observation point to the radiating elemental areas are disregarded in assessing amplitudes, but retained in the pahse terms. The diffraction integral becomes.

$$\Psi(x, y, z) = \frac{je^{-jkr}}{2\Lambda R} \int_{S_1} F(\xi,\eta) \exp\{-jk\left[\frac{(x-\xi)^2}{2z} + \frac{(y-\eta)^2}{2z}\right]\}(\cos\Theta + \frac{\partial\Phi}{\partial\eta})\, d\xi d\eta \qquad (12.16)$$

Note that with a constant-phase distribution over the aperture $\partial\Phi/\partial\eta = 1$.

An evaluation of (12.16) for points along the axis of an axisymmetrical aperture yields a field which oscillates as the point recedes outwards. Figure 12.3 gives the relative axial power density of a uniformly illuminated circular aperture. (Bickmore and Hansen, 1959).

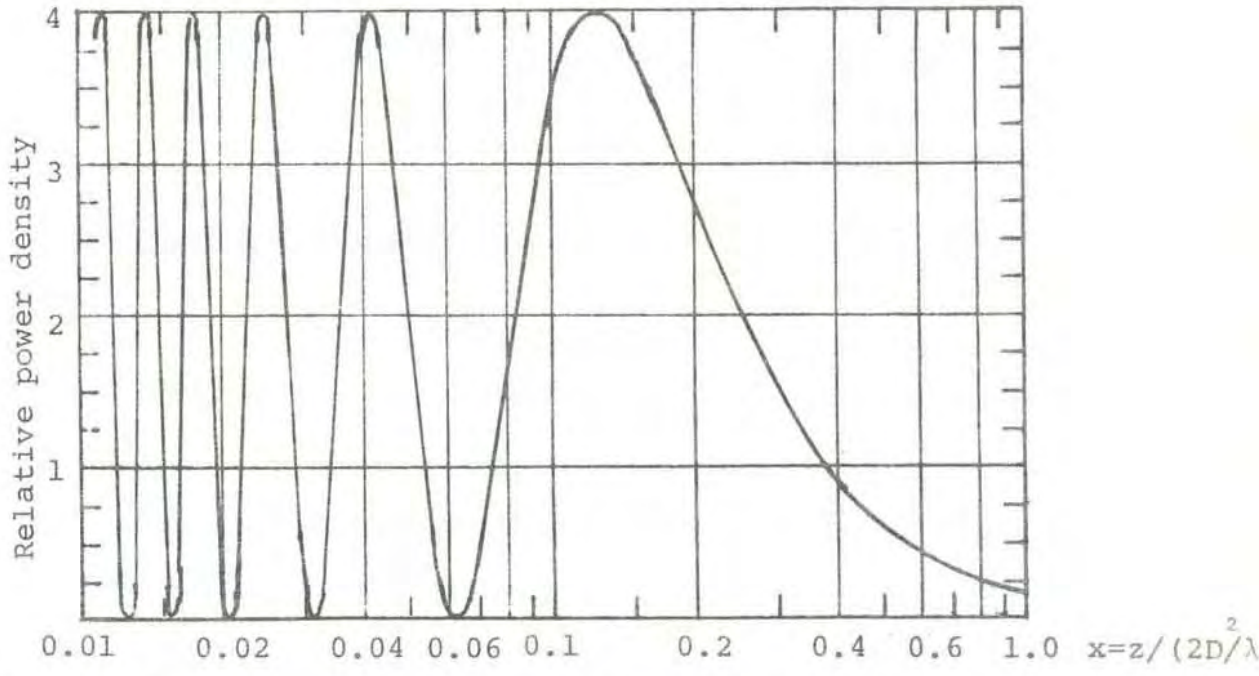


Figure 12.3 Strength of axial power density relative to aperture density for a circular, uniformly illuminated, aperture

- Fraunhofer region:

The boundary between the Fraunhofer region, or the radiation zone, and the Fresnel domain is determined by the requirement that in this most remote region no phase differences shall exceed $\Lambda/16$ for a constant-phase illumination across an aperture of maximum dimension D.

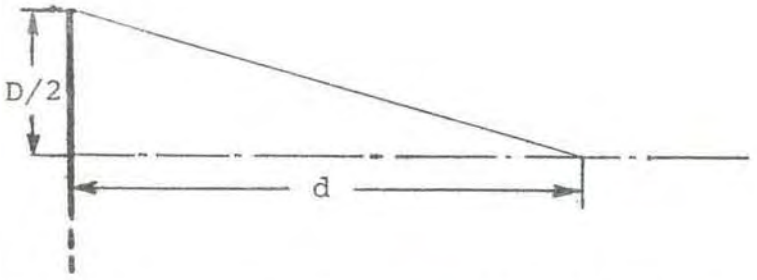


Figure 12.4 Geometry for computing the boundary of the Fraunhofer region.

According to figure 12.4:

$$\sqrt{(D/2)^2 + d^2} - d < \Lambda/16 \qquad (12.17)$$

Neglecting the term $(\Lambda/16)^2$ yields for the boundary:

$$d > 2D^2/\Lambda \qquad (12.18)$$

In the Fraunhofer region only points in the vicinity of the antenna axis are considered. Hence, $1 + \cos\Theta \approx 2$. Furthermore, all terms beyond the first order in the expansion of r (distance) in the phase term are neglected. This gives for a constant-phase aperture distribution, in spherical coordinates:

$$\Psi(x, y, z) = \frac{j\, e^{-jKR}}{\Lambda R} \int_{S_1} F(\xi,\eta) \exp\{jk \sin\Theta\, (\xi\cos\phi + \eta \sin\phi)\}\, d\xi d\eta \qquad (12.19)$$

<u>which is the Fourier transform of the aperture distribution</u>.

If the amplitude of the illumination function remains constant the half-power beamwidth is found to be $\Theta_B = \arcsin(1.029\,\Lambda/D)$.

The medium to be explored by the EISCAT facility will almost exclusively be in the Fraunhofer regions of the antennae (the

conceivable exception might be observations below an altitude of 20 km with the VHF-system). The Fresnel diffraction relations are important, though, for an assessment of possible radiation hazards on the ground in the neighbourhood of the radiators. In general, mathematical analyses tend to become intractable where the electric field changes abruptly over short distances, for instance, at the edges of the reflectors, at shadow boundaries and in areas where the rays are focussed. This has a practical significance in that it complicates attempts to reliably estimate the spillover from primary or secondary antenna feeds. We will return to this subject in a later section.

12.3 Measuring antenna gain by reception of signals from radio stars

According to formula (12.9) the antenna gain is given as

$$G_a = \frac{4\pi A_{eff}}{\lambda^2} \qquad (12.20)$$

If S_m is the flux of the standard source (in W/Hz), at the polarization which matches that of the antenna, the measured antenna temperature is:

$$T_a = \frac{S_m A_{eff}}{gk} \qquad (12.21)$$

where k is Boltzman's constant and g is a factor specifying the coupling between the source distribution and the antenna radiation pattern (Kuzmin and Salomonovich, 1966).

$$g = \frac{\int_{\Omega \text{ source}} T_b\ (\Omega)\ d\Omega}{\int_{\Omega \text{ source}} T_b\ (\Omega)\ r\ (\Omega)\ d\Omega} \qquad (12.22)$$

Here T_b (Ω) denotes the brightness temperature of the source, defines as (page 39, volume I):

$$B_m (\Omega) = kT_b (\Omega) \qquad (12.23)$$

If the radio star is a point source and the measurement is made with the star on the main lobe axis, g = 1 and (12.21) reduces to:

$$T_a = \frac{S_m A_{eff}}{k} \qquad (12.24)$$

In general the temperature measurement with the star in the main lobe includes also the contribution from the background sky. This must be eliminated before (12.24) can be applied. The practical arrangements will depend upon the characteristics of the receiver system. We will briefly consider two cases, assuming that the antenna is prepositioned so as to drift through the polar pattern of the source with the rotation of the earth.

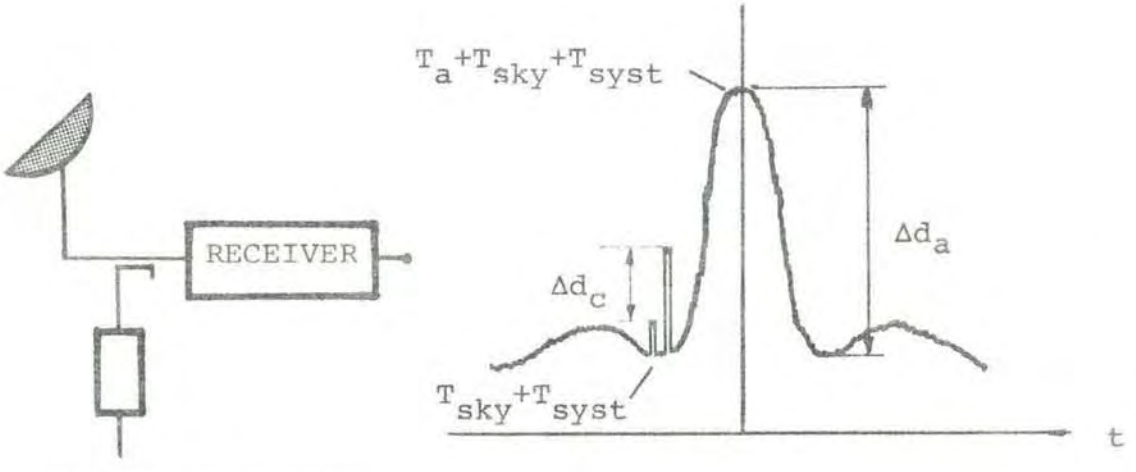


Figure 12.5 Antenna calibration setup and typical drift scan.

(i) The receiver (radiometer) is linear with respect to power (square-law detector)

Subtracting the output reading at the minimum between the main lobe and the first side-lobe (assuming zero contribution from the source at this point) from the reading at the drift maximum yields.

$$T_a k \Delta B G_1 = \Delta d_a / p_o \qquad (12.25)$$

Δd_a is here the pen deflection and p_o (mm/dB) a scaling factor found by measuring the pen deflection caused by a controlled change of the receiver gain.

By calibrating the readout with a hot/cold source we obtain:

$$\Delta T k \Delta B G_2 = \Delta d_c / p_o \qquad (12.26)$$

where $T_{hot} - T_{cold} = \Delta T$ (note that T_{cold} can, for instance, be T_{sky}). Combining (12.25) and (12.26) gives:

$$T_a = \frac{\Delta T}{a} \frac{\Delta d_a}{\Delta d_c} \qquad (12.27)$$

If the gain is the same during the two sets of readings, $a = 1$.

(ii) Logarithmic receiver

In this case we derive the following expression:

$$T_a = T_c \frac{10^{\Delta d_a/(10 p_o)} - 1}{10^{\Delta d_c/(10 p_o)} - 1} \qquad (12.28)$$

T_c is a known noise impulse injected at the first interlobe minimum. Δd_c is the corresponding pen deflection.

If the receiver is not strictly linear in power or in the logarithmic sense over a wide dynamic range, it may be desirable to perform the calibration both at the signal minimum and at the maximum in the drift scan (Westerlund, private communication).

The most commonly used radio star for antenna-calibration and measurements is Cassiopeia A. The actual strength of this source seems to have been a matter of some conjecture. More recent publications (see Kildal, 1980) gives the following relationship for the flux density, with numerical constants:

$$F = F_o \left(\frac{f}{f_o}\right)^k [\, 1 + B(f) \,]^n \qquad (12.29)$$

$F_o = 3181 \pm 25$ Jy at $f_o = 1$ GHz

(1 Jy = 10^{-26} Wm^{-2} Hz^{-1})

$k = -0.792 \pm 0.006$

$B(f) = -0.0097 \pm 0.0005 + (0.00126 \pm 0.00023) \ln (f)$

Epoch $t_o = 1965$, frequency f is in GHz

12.4 The EISCAT UHF-antenna

12.4.1 Some geometrical considerations

The EISCAT UHF-antenna is of the Cassegrain type, comprising a feed horn and two reflectors as shown in figure 12.6.

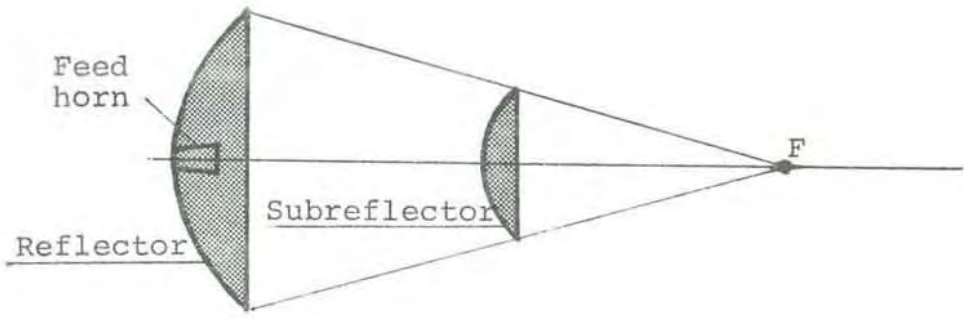


Figure 12.6 Schematic presentation of the Cassegrain system.

Since the reflectors are surfaces of revolution their geometry may be studied in a two-dimensional coordinate system. We will derive the geometrical form of the two reflector surfaces. We divide the problem in two:

(i) First we seek the form which the main reflector must have to convert a spherical incoming wave to a plane wave propagating along the symmetry axis.

(ii) Thereafter we address the problem of assessing the form of the subreflector, requiring that it shall transform a spherical incoming wave, emanating from a point on the symmetry axis into a spherical wave propagating towards the main reflector.

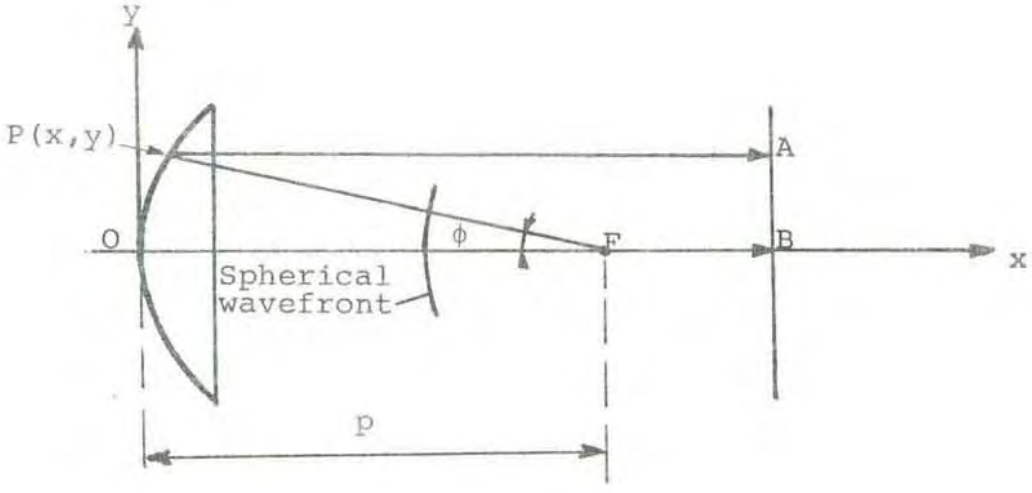


Figure 12.7 Geometry to determine form of main reflector.

Assuming that the spherical wave is emitted from a source at F and that A and B are points on the plane wave front, we have:

$$FP + PA = FO + OB \tag{12.30}$$

Introducing xy-coordinates for the arbitrary point P (x, y) on the reflector, and using the notations in the figure, we easily find:

$$\frac{p - x}{p + x} = \cos \phi \tag{12.31}$$

$$y = (p - x) \operatorname{tg} \phi \tag{12.32}$$

(12.31) and (12.32) combine to give

$$y^2 = 4px \tag{12.33}$$

Equation (12.33) represents a <u>parabola</u> with focal length p.

We next assume that the subreflector is inserted between the focal point and the main reflector, shaped in such a way that it transforms a spherically shaped wave, generated at a point on the axis at the mouth of the horn, into a spherical wave travelling towards the main reflector.

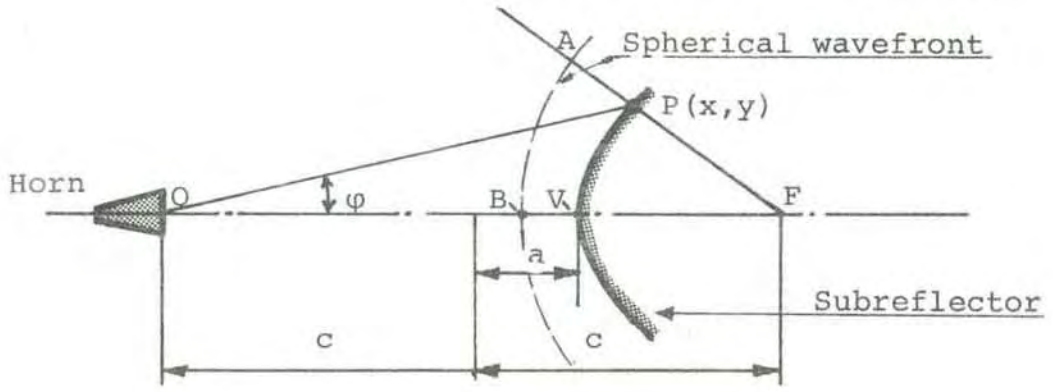


Figure 12.8 Subreflector geometry.

From the figure we have

$$OP + PA = OV + VB = c + a + VB \qquad (12.34)$$

and further, since the wave scattered from the subreflector is spherical with centre at F:

$$PA = FA - FP = FB - FP \qquad (12.35)$$

Combining the two equations above gives:

$$OP - FP = 2a \qquad (12.36)$$

a relation which represents a hyperbola with foci at O and F

In modern computer designs of Cassegrain antennae the main and/or subreflectors are often shaped to minimize losses and allow for an optimization of the gain. As a result the final shape of the reflector reflectors may deviate somewhat from the conical surfaces indicated by (12.33) and (12.36).

For the EISCAT UHF-antenna the subreflector has been shaped to yield a main aperture diffraction field as shown in figure 12.9 (Antenna Manual, Volume IV). The computed overall efficiency, taking account of spillover losses, phase errors, amplitude ripples and ohmic losses, amounts to 0.651 giving an efficient antenna area of 523.6 m^2 and a gain of 41.8 dB.

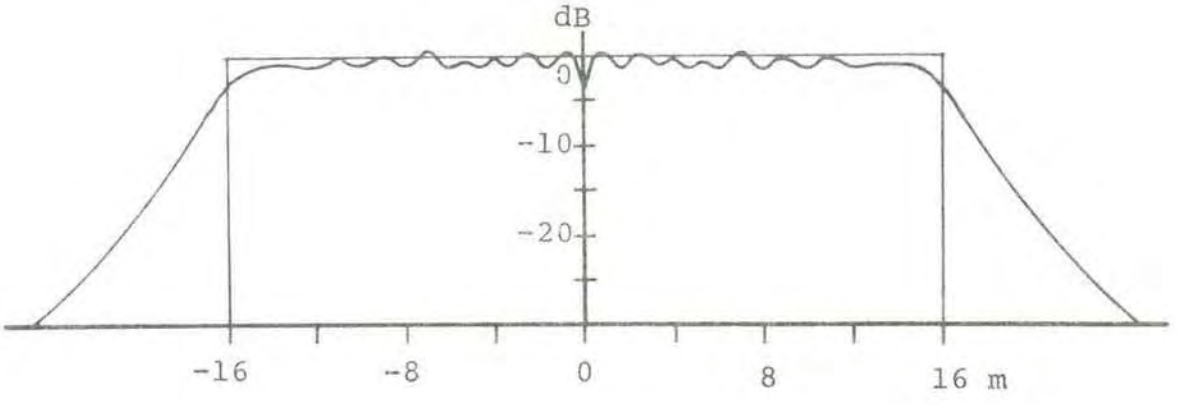


Figure 12.9 Main aperture illumination.

12.4.2 Antenna characteristics

Location and height above sea level (MSL), (Datum ED-1950):

Ramfjordmoen site:	69°35'11.239'' N 19°13'38.746'' E H = 86.5 m
Kiruna site:	67°51'38.373'' N 20°26'06.806'' E H = 417.9 m
Sodankylä site:	67°21'49.266'' N 26°37'37.497'' E H = 197.3 m

Diameter of main reflector	32.00 m
Diameter of subreflector:	4.58 m
Focal length:	11.01 m
Gain, calculated:	48.1 dB

Efficiency, calculated:	0.651
Noise contribution, measured:	20-21 K
Half-power beamwidth:	0.6^{o}
Speed of mechanical movement, azimuth & elevation	80^{o}/min
Fresnel zone limit:	6.37 km
Rayleigh distance:	1.59 km

For more details we refer to the manual, volume IV.

An inspection of the UHF Antenna Test Report (78/7) shows that the antenna gains at the three sites, as measured in November 1978, are close to the computed values.

The figures given for the antenna positions result a from a geodetic survey, carried out under the auspices of the Nordic Geodetic Commission. The quoted coordinates are extrapolated from reference points differently positioned in the antenna structures at the three stations. For Ramfjordmoen and Sodankylä the tabled figures refer to a point on the antenna azimuth axis, at the height of the elevation axis (the two axes are not coplanar). For Kiruna the geodetically determined reference is a point outside the pintle bearing hub. The figures quoted again refer to the level of the elevation axis, but define a point about 2.2 m from the azimuth axis. For ionospheric measurements the inaccuracies of the quoted positions hardly matter. For possible VLBI observations one should confer with the geodetic survey documentation.

12.5 EISCAT VHF-antenna

The antenna consists of four identical elements, each shaped as a parabolic cylindre with dimensions 30m x 46 m. A layout of the antenna is shown in figure 12.10.

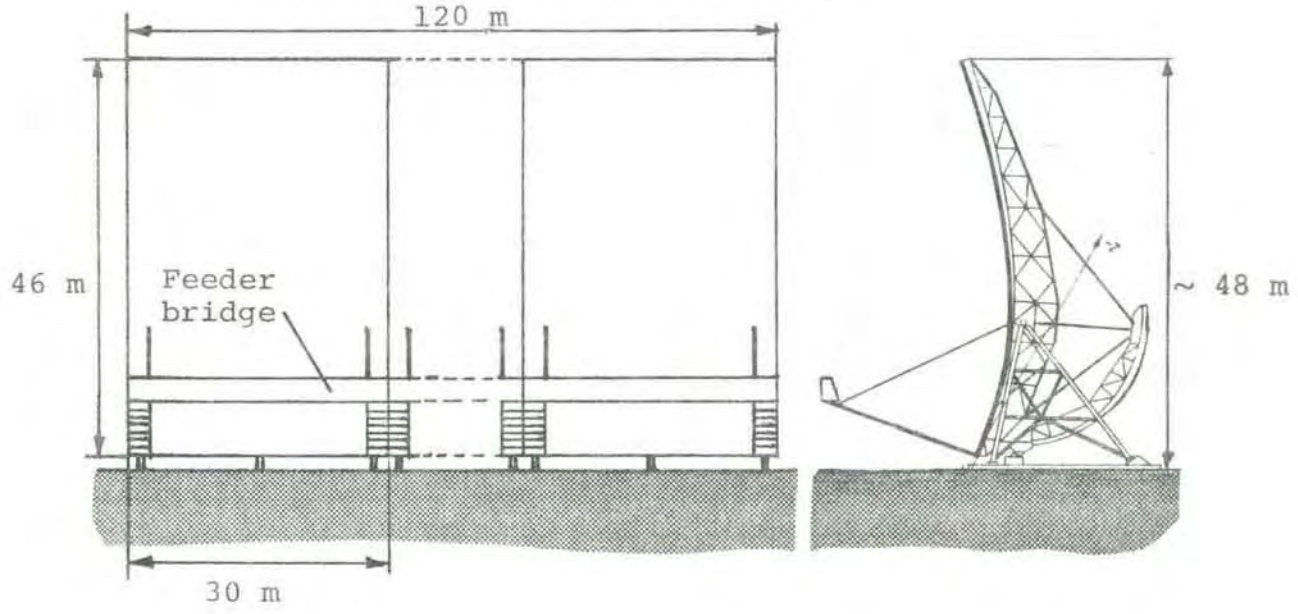


Figure 12.10 VHF-antenna structure.

The feed antenna consists of an array of 128 crossed dipoles positioned along the focal line of the main reflector. The design of the feed is essential to the operation of the antenna. Construction details have been described by Kildal (1979) and by Hagfors (1977) and we shall not repeat these treatments here. However, in line with the tutorial aims of these lectures, in the following subsection we digress to expound some general features of antenna arrays, before we go on to a more specific treatment of the EISCAT array feed. We shall restrict our considerations to linear arrays only.

12.5.1 Array concept

Let us first consider an array of N isotropic, equally spaced radiators, all fed with signals having identical amplitudes and (zero) phases. We assume that unit total power is fed to the array.

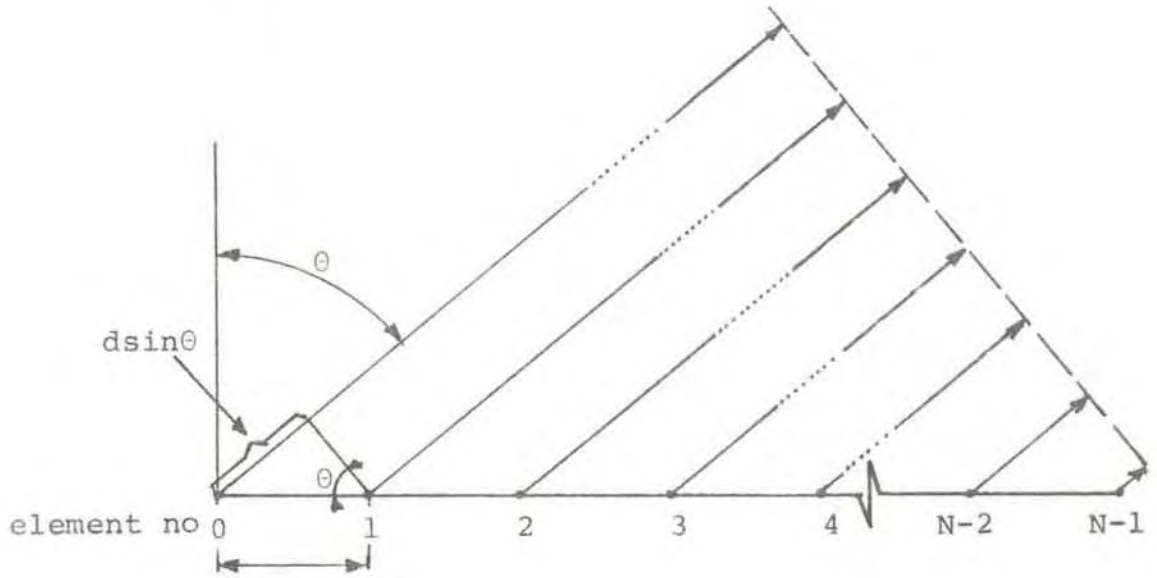


Figure 12.11 Array of isotropic radiators.

For the far field (Fraunhofer) zone the field radiated in direction θ is established by the vectorial addition of the contributions from all N elements:

$$E_a(\theta) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N} e^{-j \frac{2\pi}{\Lambda} nd \sin \theta} \qquad (12.37)$$

Here d denotes the distance between the radiators. Since the power radiated by each element is 1/N, the amplitude of the resulting signal is $1/\sqrt{N}$, accounting for the term ahead of the summation sign. In the subsequent treatment we introduce the common abridgement S = sin θ.

Carrying out the summation and squaring the field we obtain for the power:

$$P_a(\theta) = E_a(\theta)^2 = \frac{1}{N}\left|\frac{\sin(\frac{\pi}{\lambda} dN\ S)}{\sin(\frac{\pi}{\lambda} d\ S)}\right|^2 \qquad (12.38)$$

If the individual radiators were not isotropic, but having a radiation pattern $P_e(\theta)$, the array polar diagram would become:

$$P(\theta) = P_e(\theta)\, P_a(\theta) \qquad (12.39)$$

It is common to term $P_a(\theta)$ the array factor or space factor (sometimes the same notations are used with reference to $E_a(\theta)$) and $P_e(\theta)$ the element factor. Equation (12.39) embodies the principle of pattern multiplication, stating that the radiation pattern of an array is given as the product of the element pattern and the array factor.

The array factor is essential in providing information on the location of the principal and subsidiary beams and the nulls in the polar diagram. Expression (12.38), without the scaling factor 1/n, has been plotted in figure 12.12 for N = 10 and $d/\Lambda = 0.5$ and 1.0.

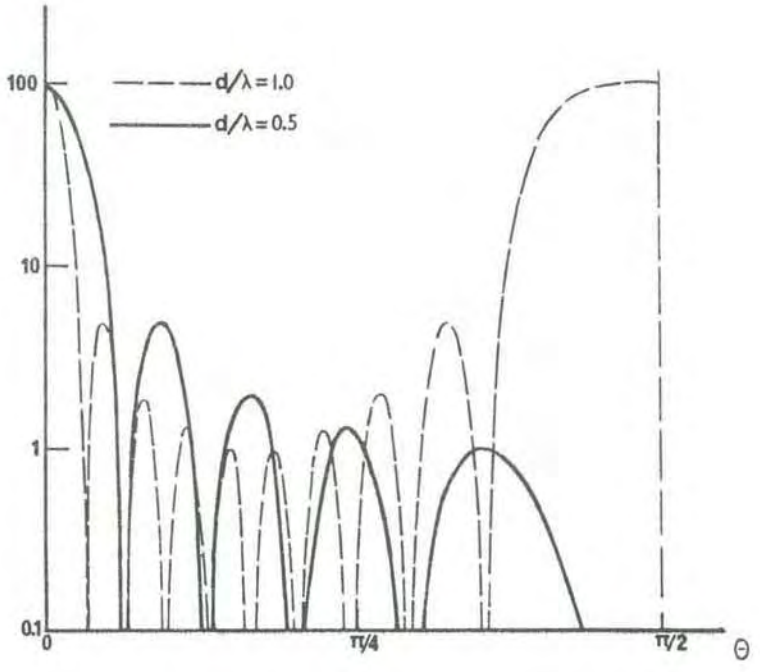


Figure 12.12 Radiation pattern for N = 10 and d/Λ = 0.5 and 1.0

Subsidiary beams (sidelobes) having the same magnitude as the main beam are termed grating lobes. We note that a grating lobe appears in figure 12.12 for $d/\Lambda = 1.0$ at $\Theta = \pi/2$. An inspection of (12.38) shows that grating lobes occur when the argument of the sinusoidal expression in the denominator assumes integral values of π:

$$\frac{\pi\, d}{\Lambda} \sin \Theta = m\, \pi \qquad (12.40)$$

To avoid grating lobes for $-\pi/2 < \Theta < \pi/2$ we find that we must have $d < \Lambda$, the spacing between adjacent elements must be less than the wavelength. It may be noted here that even with spacings exceeding the wavelength, grating lobes may be eliminated

by the directivity of the element pattern $P_e(\Theta)$.

So far we have assumed that all radiating elements were fed with equally phased signals. In this case the main beam will be normal to the array, which is said to be of the broadside type. The array may be steered to an angle Θ_o by applying a linearly progressive phase increment, $2\pi\ (d/\Lambda)\ \sin\Theta_o$, from element to element, yielding:

$$P(\Theta) = P_e(\Theta) \left| \frac{\sin\left(N \frac{d}{\lambda} (S - S_o)\ \pi\right)}{N \sin\left(\frac{d}{\lambda} (S - S_o)\ \pi\right)} \right|^2 \qquad (12.41)$$

If $\Theta_o = \pi/2$ the main beam is parallel to the line of elements. It is customary in this case to talk about an endfire array.

A note of warning may be in order here. Since mutual coupling will generally exist between neighbouring elements in an array, and since radiation from one element may be scattered from near-by radiating structures, the element factor $P_e(\Theta)$ as used in (12.41), may be different from that describing the radiation from a singel element in free air.

For a narrow beam in the broadside direction the array factor of equation (12.38) simplifies to:

$$P_a(\Theta) = \left| \frac{\sin\left(\frac{\pi}{\lambda} dN\ S\right)}{N \frac{\pi}{\lambda} dS} \right|^2 \qquad (12.42)$$

Equation (12.42) is of the well-known form $\sin x/x^{\,2}$ and we obtain for the half-power beamwidth:

$$\Theta_B = \frac{0.886}{Nd/\lambda} \qquad (12.43)$$

We note that the beam becomes narrower with increasing number of elements and with increasing frequency. The beam broadens with the steering angle. It may be shown (Skolnik) that for moderate steering angles (from broadside) we have:

$$\Theta_B = \frac{0.886}{\frac{Nd}{\Lambda}\cos\Theta_o} \qquad (12.44)$$

12.5.2 Band width restrictions in phase steering

Let us consider the feed system shown in figure 12.13 where a uniform phase progression along the array is obtained by increasing the cable length by a fixed amount from element to element.

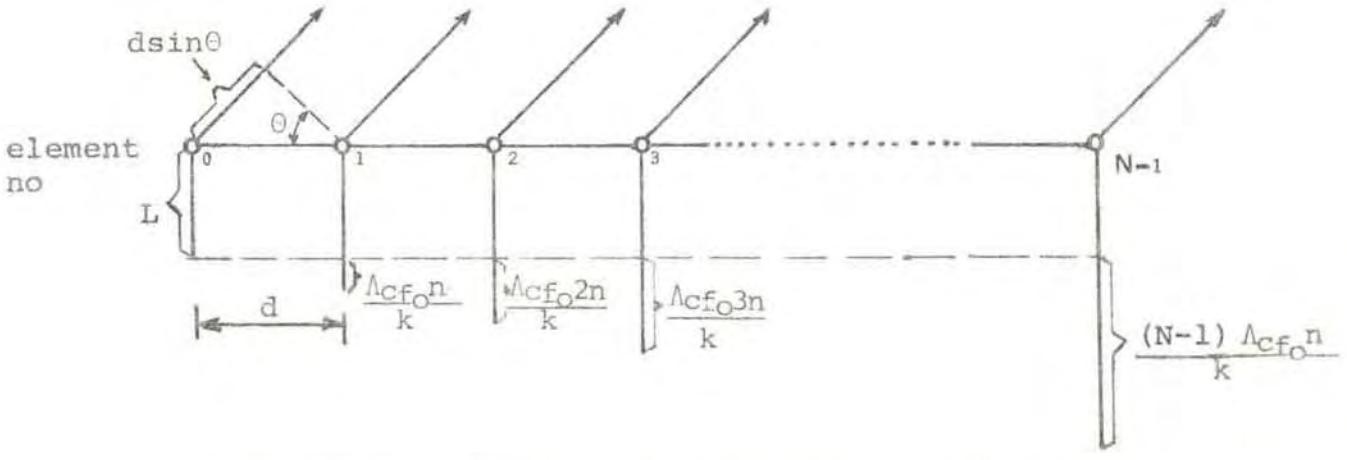


Figure 12.13 Steering by incrementing cable lengths.

The phase excursions caused by the constant cable length , L , act equally on all array elements and can be disregarded in the following. Λ_{cf_o} is the wavelength in the cable at frequency f_o. Λ_{cf} is the corresponding wavelength at an arbitrary

frequency f. Λ_{f_o} and Λ_f are free-air wavelengths at the two frequencies. d denotes the distance between adjacent radiators. n and k are integers. Assuming unit amplitude signals and zero phase at terminal 0, we have the following signal sequence, proceeding from left to right:

$$1,\ e^{-j2\pi\Lambda_{cf_o}n/(k\Lambda_{cf})},\ldots\ldots,e^{-j(N-1)2\pi\Lambda_{cf_o}n/(k\Lambda_{cf})}$$

For the centre frequency f_o we find for the direction, ϕ_o, of the principal beam:

$$2\pi\,\frac{dS_o}{\Lambda_{f_o}} = 2\pi\,\frac{n}{k} \qquad (12.45)$$

$$S_o = \frac{n\,\Lambda_{f_o}}{dk} = \frac{n\Lambda_{f_o} f_o}{dk\, f_o} = \frac{nc}{dkf_o} \qquad (12.46)$$

where c is the velocity of light.

At the arbitrary frequency f we have in a similar way:

$$S = \frac{\Lambda_{cf_o}\Lambda_f\, n}{k\,\Lambda_{cf}\,d} = \frac{\Lambda_{cf_o}\,\Lambda_f\, ff_o\, n}{k\,\Lambda_{cf}\,d\,ff_o} \qquad (12.47)$$

$$S = \frac{\upsilon_{f_o}\, c\, n}{k\,d\,\upsilon_f\, f_o} \qquad (12.48)$$

υ_{f_o} and υ_f are phase velocities in the cable at f_o and f. Assuming the phase velocity to remain constant over the actual frequency range yields:

$$S = \frac{cn}{dkf_o} = S_o \qquad (12.49)$$

Relations (12.47) to (12.49) could have been derived by considering any pair of neighbouring element. Under the given conditions the main bean direction does not change for a change of operating frequency.

For practical reasons (including economy) one is often interested in phasing modulo 2π. In the EISCAT VHF feed system two identical sets of cables exist (for each polarization), each set comprising 68 cable lengths:

$$l_i = \frac{\Lambda_{cf_o} i}{67} \qquad (i = 0,1,2,.....67)$$

(Note that i refers to cable number, not to array element). As Kildal (1978) has shown, this cabling arrangement allows beam steering up to 21.3^o in steps of approximately 1.2^o by appropriate permutations of the cables (in fact the beam can be steered to 25.2^o, but the range 21.3^o + 25.2^o should not be used due to impairment of the VSWR).

Contradictory to the conditions obtained with the phasing scheme of figure 12.13, in phasing modulo 2π the beam direction depends upon frequency. The underlying principle is outlined in the following, with reference to figure 12.14.

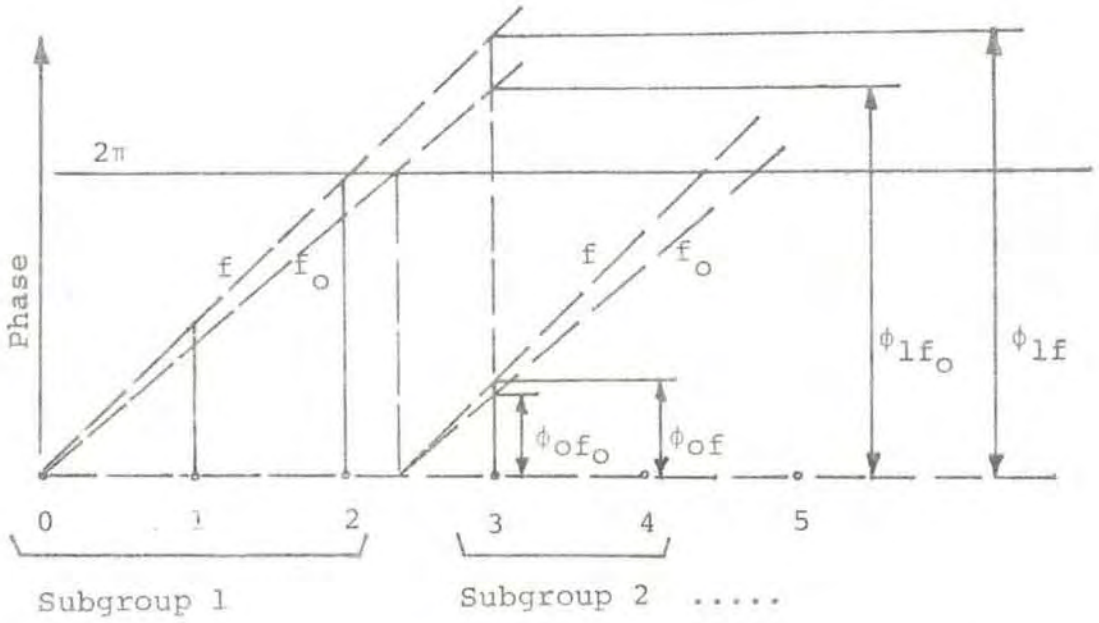


Figure 12.14 Phasing modulo 2π for frequencies f_o and f

For the three first elements (numbered 0 to 2) in figure 12.14 the phase is progressed by 2π n/k from element to element for the frequency f_o. For element number 3 the phase is

$$2\pi \left(\frac{4n}{k} - 1\right)$$

Since we have the phasor relationship:

$$e^{-j\phi_{1f_o}} = e^{-j\,(2\pi + \phi_{of_o})} = e^{-j\phi_{of_o}} \qquad (12.50)$$

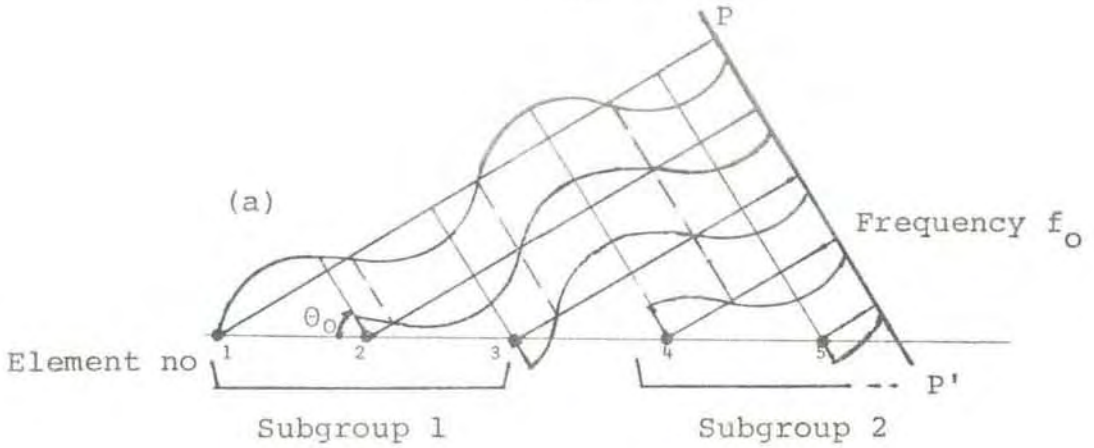
the field contributions sum as if there were a constant phase progression along the array. The main beam direction is still given by (12.46).

The phase variation for frequency f is indicated by the broken lines in figure 12.14. We readily se that

$$e^{-j(\phi_{1f} - 2\pi)} \neq e^{-j\phi_{of}} \qquad (12.51)$$

Within each subgroup the phase increment remains constant, but phase jumps occur in proceeding from one group to the next.

As indicated in figure 12.15 the beams formed by the radiation within each subgroup all point in the same direction, Θ_o. However, due to the phase jumps the phasors are not completely aligned in this direction. The result is a swinging of the main beam direction, dependent upon the deviation from the centre frequency.



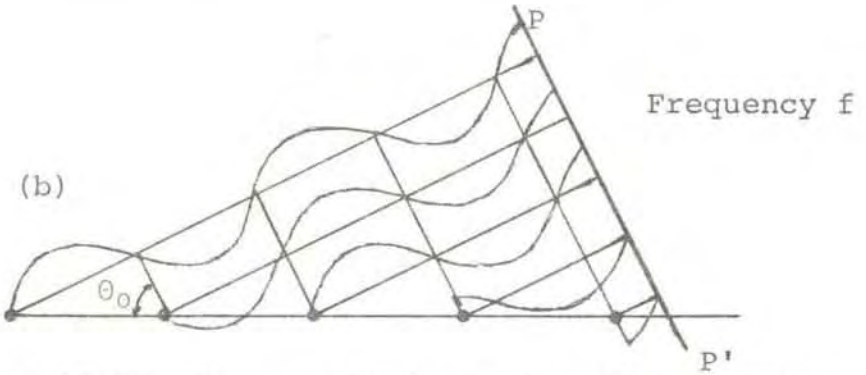


Figure 12.15 Phase relations of radiated waves at frequencies f_o and f. Note that in case (a) all waves are in phase along PP'. This is not so in case (b).

Figure 12.16 and 12.17 illustrate how the beam directions and bandwidths of the VHF-array feed antenna depends upon frequency.

The formulae used in deriving the graphs are:

$$P_a(\Theta) = \{ \sum_{I=1}^{I_{max}} \cos A(I) \}^2 + \{ \sum_{I=1}^{I_{max}} \sin A(I) \}^2 \qquad (12.52)$$

$$A(I) = 2\pi \left[(I-1)0.7\frac{f}{f_o} S + \frac{K(I)f}{67f_o} \right] \qquad (12.53)$$

$S = \sin \Theta$

$f_o = 224$ MHz

$K(I) = M - (I-1)N$

$P_a(\Theta)$: Array factor

I : No of array element $(1,2,.....,I_{max})$

I_{max} : 128 for full array, 64 for half array

N : Cable no ($0 \leq N \leq 67$), defined by the relation $\sin \theta_o = N/(67 \cdot 0.7)$

M : An integer initially given the value 67, augmented by 67 each time K(I) falls below zero

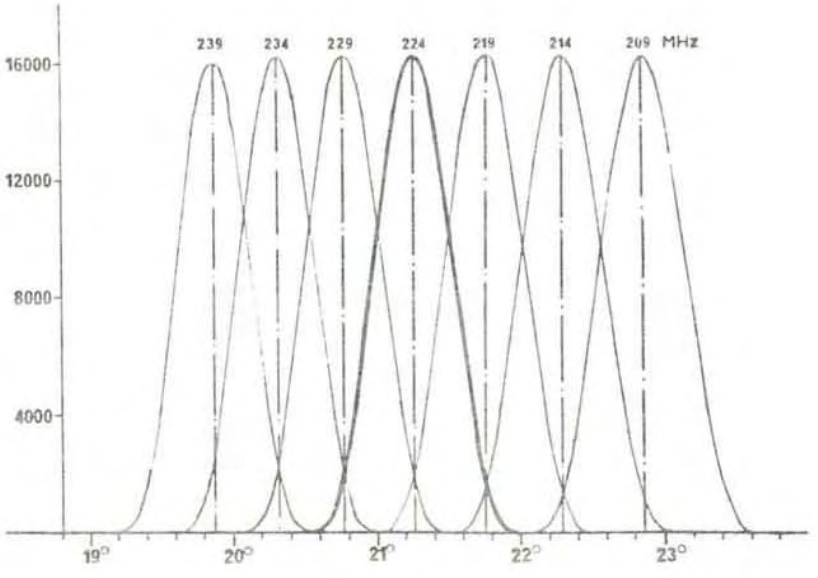


Figure 12.16 Beam off-sets as function of frequency for ultimate steering angle (21.3^{o})

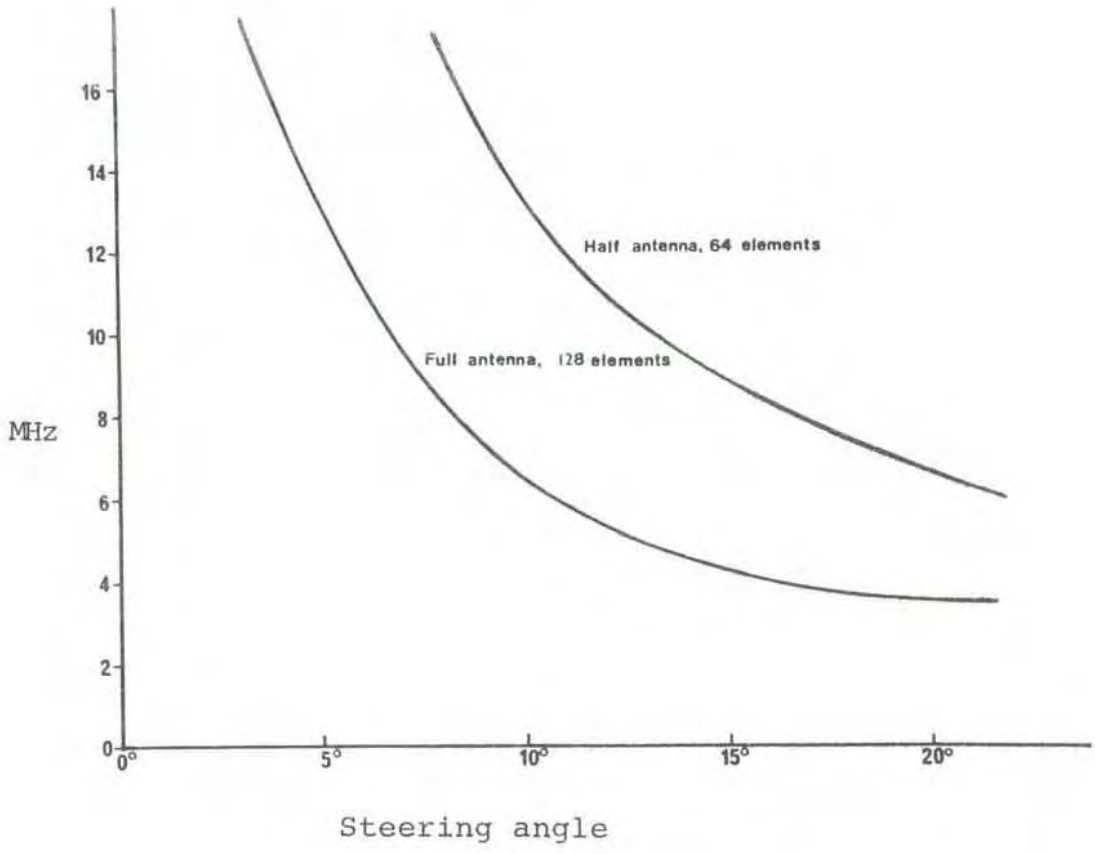


Figure 12.17 Bandwidths as function of steering angle for full and half antenna

Figure 12.16 demonstrates how the main beams are offset by varying the frequency, the array being phased for the ultimate steering angle (21.3^{o} at f_o). The bandwidth relations shown in figure 12.17 have been determined in the following way:

(i) For the centre frequency lobe in figure 12.16 (f_o = 224 MHz) the angle corresponding to the lower half-power point has been noted.

(ii) With the angle determined as indicated under (i) formula 12.52 has been repeatedly applied to find the frequency yielding maximum response.

(iii) The bandwidth has been defined as the difference between the maximum response frequency and the centre frequency.

(iv) By repeating the procedure above for other steering angles the curves depicted in figure 12.17 have been derived.

It may be noted that the criterion used here to defined the bandwidth is essentially the same as introduced by Hagfors (1978) in an analytical treatment of an array feed with continuous field distribution. The present analysis only takes account of offsets in one direction. It is a matter of convention whether the bandwidth should be defined as twice the values given here, noting the offsets in both directions.

In the EISCAT context the more serious offsets will be those associated with the plasma lines, being displaced from the sentre frequency by the values of the ionospheric plasma frequencies. We know that these frequencies are both height- and time-dependent quantities, occasionally reaching values in the range 10 to 15 MHz. A glance at figure 12.17 indicates that there are no chances of observing such offset frequencies for steering angles larger than about 8-10^{o} with the full antenna. The chances are somewhat better with in the split beam option, but here it has to be borne in mind that the reduced antenna area implies a very notable gain reduction.

As stated before, the polar patterns considered in this chapter for array feed antennae are all referred to the Fraunhofer approximation. In the EISCAT VHF antenna the array radiates into the main reflector. While the Fraunhofer assumption holds good for the cross sectional radiation pattern, the reflector is in the near zone with respect to the longitudinal antenna dimension. Kildal (1978) has treated the feed- and polar pattern problems in fair detail, including optimizing features like beam- and impedance matching, efficiency computations and polarization behaviour. We

shall be content here with referring to his report for a study of all these topics.

12.5.3 Some useful geometrical relations related to phase steering of the VHF-antenna.

Let us define a cartesian coordinate system with the z-axis pointing vertically upwards and the x-axis directed geographical north as shown in figure 12.18.

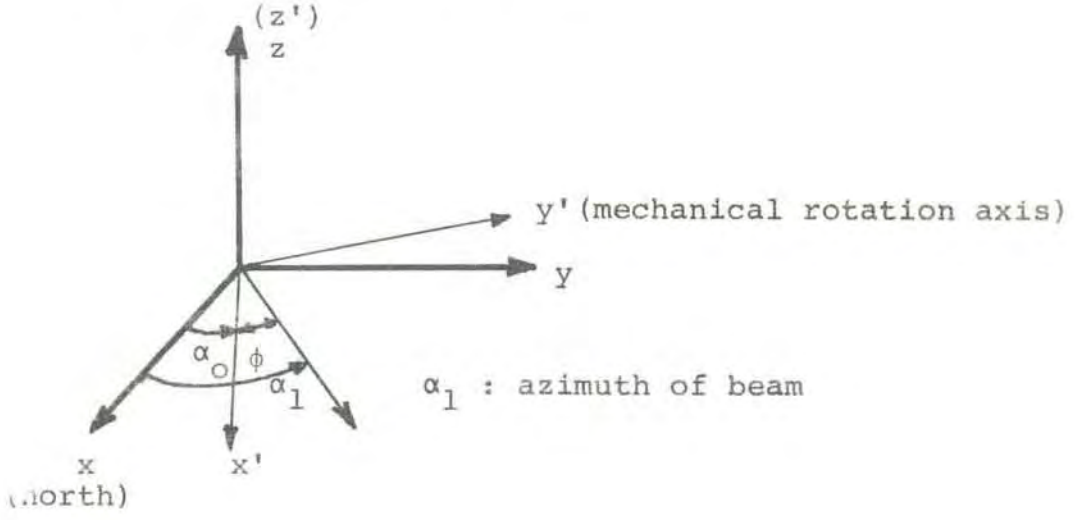


Figure 12.18 Coordinate system used for studies of phasing of VHF-antenna

To describe the phasing of the VHF-antenna we introduce the system x'y'z' with the z'-axis coinciding with the z-axis, but with the x'- and y'-axes rotated by the angle α_o with respect to the x- and y-axes respectively. The antenna rotates mechanically around the y'-axis. With no phasing the main lobe is in the x'z'-plane, and the antenna is said to radiate in the broadside direction. An arbitrary great circle cuts the horizontal plane along the line defined by the angle $\phi = \alpha_1 - \alpha_o$ using the notations of figure 12.18.

When the antenna moves, the phased beam describes a cone with the symmetry axis along y', as shown in figure 12.19.

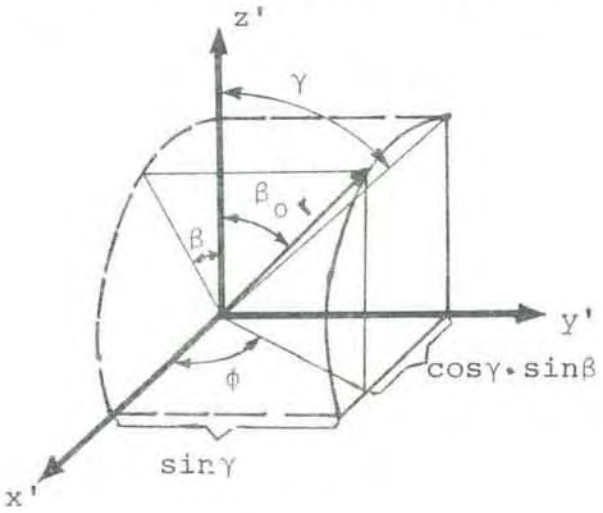


Figure 12.19 Beam pointing geometry.

At an arbitrary position the antenna beam axis is specified by the vector $\bar{r}$. γ is the phasing angle. Taking $\bar{r}$ as a unity vector we may easily derive the following relation from the figure.

$$\operatorname{tg}(\alpha_1 - \alpha_o) = \frac{\sin\gamma}{\cos\gamma\,\sin\beta} = \frac{\operatorname{tg}\gamma}{\sin\beta} \qquad (12.54)$$

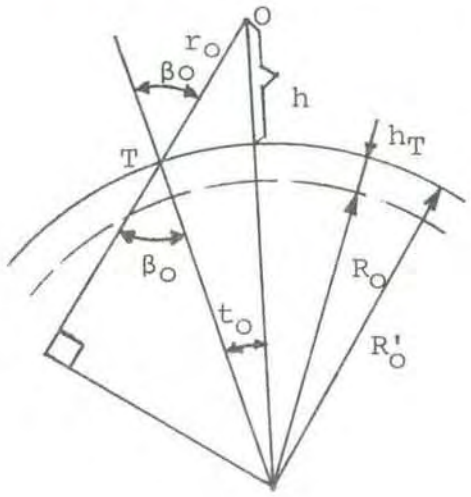
where β is the rotation angle of the antenna referred to the vertical (the complement of the angle displayed on the antenna control panel).

Since the length of the projection of the $\bar{r}$-vector in the x'y'-plane remains fixed, equal to $\cos\gamma$, during the rotation of the antenna, we can express the angle β_o by the following relationships:

$$\sin\beta_o \cdot \cos\phi = \cos\gamma \cdot \sin\beta \qquad (12.55)$$

$$\sin\beta_o = \frac{\cos\gamma \cdot \sin\beta}{\cos(\alpha_1 - \alpha_o)} = \frac{\sin\gamma}{\sin(\alpha_1 - 0.5^{\circ})} \qquad (12.56)$$

($\alpha_o = 0.5^o$). In figure 12.20 we consider the plane of the great circle through $\bar{r}$.



h_T: height of transmitter above ocean level

R_o: radius of the earth

R'_o: adjusted radius, including h_T

O : point of observation

Figure 12.20 Great circle plane through main beam.

The height of the transmitter above sea level is only 85 m and can be neglected for most practical purposes (this is done in the subsequent treatment). t_o denotes the angular distance to the point of observation.

Simple geometrical considerations give:

$$r_o = \sqrt{(R_o \sin t_o)^2 + \{(R_o + h) - R_o \cos t_o\}^2}$$
$$= \sqrt{R_o^2 + (R_o + h)^2 - 2R_o (R_o + h) \cos t_o} \qquad (12.57)$$

$$r_o = \sqrt{R_o^2 \cos^2 \beta_o + 2R_o h + h^2} - R_o \cos \beta_o \qquad (12.58)$$

$$\sin \beta_o = \frac{R_o + h}{r} \sin t_o \qquad (12.59)$$

$$\text{tg } t_o = \frac{r_o \sin \beta_o}{R_o + r_o \cos \beta_o} \qquad (12.60)$$

By applying standard formulae in spherical geometry:

$$\cos t_o = \sin LA_T \sin LA_O + \cos LA_T \cos LA_O \cos (LO_O - LO_T) \quad (12.61)$$

$$\sin \alpha_1 = \cos LA_O \sin (LO_O - LO_T) / \sin t_o \quad (12.62)$$

$$\sin LA_O = \cos t_o \sin LA_T + \sin t_o \cos LA_T \cos \alpha_1 \quad (12.63)$$

(Note that the last equation may be derived from the two preceeding ones).

Notations:

α_1 : Antenna bearing

LA_T: Latitude of transmitter antenna

LO_T: Longitude of transmitter antenna

LA_O: Latitude of point observation

LO_O: Longitude of point observation

Introducing numerical values for LA_T and LO_T (see specification in next section:

$$\cos t_o = 0.93720 \sin LA_O + 0.34879 \cos LA_O \cos(LO_O + 19.22034) \quad (12.64)$$

$$\sin \alpha_1 = \cos LA_O \sin (LO_O + 19.22034) / \sin t_o \quad (12.65)$$

$$\sin LA_O = 0.93720 \cos t_o + 0.34879 \sin t_o \cos \alpha_1 \quad (12.66)$$

Here we have taken eastern longitudes as being negative.

The formulae may be used in various combinations under different circumstances. Let us consider two possibilities:

1. The observing point is defined by its geographical coordinates and its height above ground. What are the phasing angle and mechanical antenna setting?

   Formulae (12.64) and (12.65) give immediately the angular distance and the bearing of the beam. Expression (12.57) specifies range and equation (12.59) the beam zenith angle. The phasing angle is computed from (12.56) and the mechanical antenna setting by (12.54).

2. The antenna position β and the phasing angle γ are given. What are the geographical coordinates and the height above ground of the point corresponding to a given range r_{ol}?

   In this case equation (12.54) gives the bearing of the beam and (12.56) the zenith angle. Formula (12.58) provides the altitude of the observing point and (12.60) the angular distance. The desired latitude is derived from (12.66); the longitude from (12.65).

In figure 12.21 the set of formulae described have been applied to find the areas covered by the antenna at the altitudes of 1000, 2000 and 3000 km.

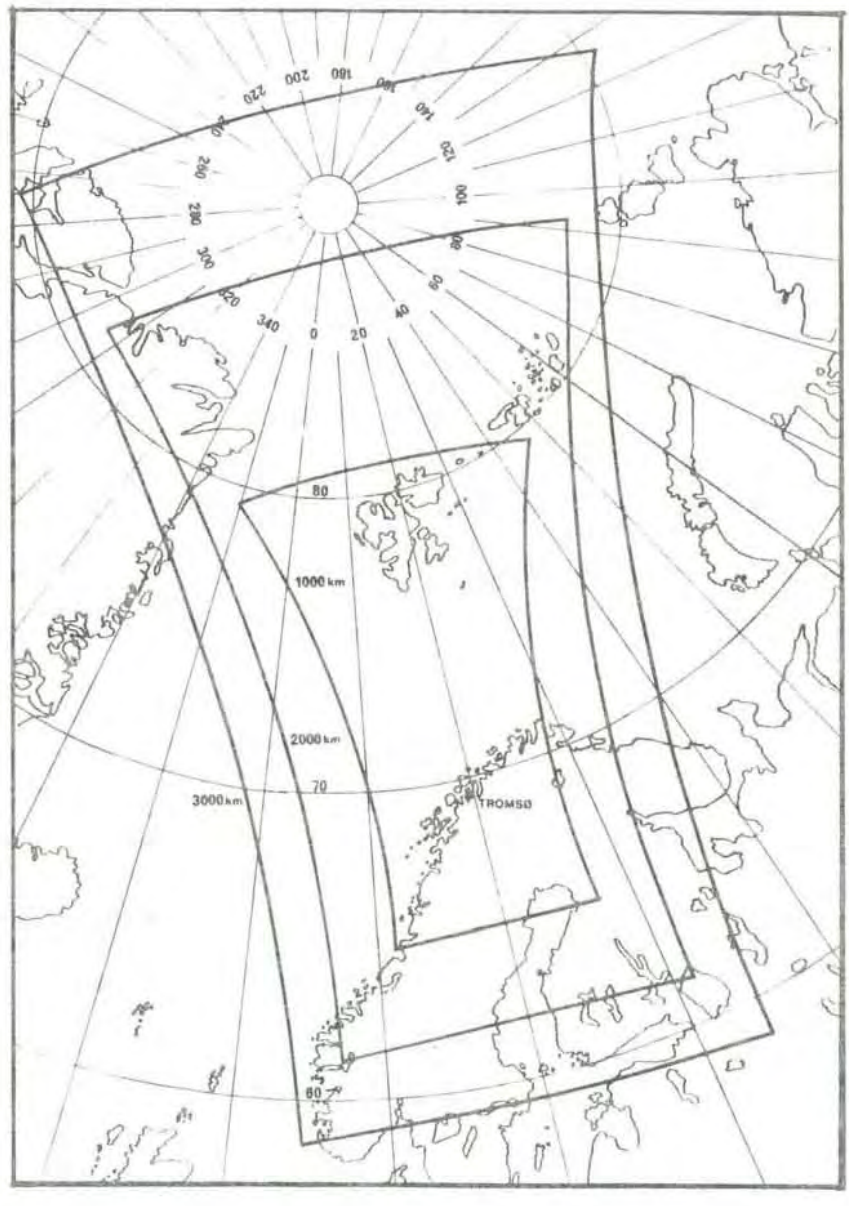


Figure 12.21 Areas covered by the VHF-antenna at the heights of 1000, 2000 and 3000 km. Boundaries correspond to scan limits 30° south and 60° north of zenith and phasing angles 21.3° east and west of transit plane.

2.5.4 Some characteristics of the EISCAT VHF-antenna

Location	:	69^{o} 35' 11.9408'' N 19^{o} 13' 13.230'' E H = 85.3 m (reference point centre of elevation axis)
Orientation of plane of mechanical movement	:	0.5^{o} (measured $0.55^{o} \pm 0.05^{o}$ Kildal (1980)) west of north
Operational modes	:	I. Combined mode, all elements aligned, physical area 5520 m^2. II. Split beam mode, element 1 & 2 and 3 & 4 pairwise aligned, structure behaves as two independent antennae, each of physical area 2400 m^2.
Effective area, broadside, mode I	:	Calculated for circular polarization 3250 m^2. Measured (Kildal (1980)): Horizontal polarization 3330 $\pm$ 240 m^2 Vertical polarization 2890 $\pm$ 210 m^2
Polarization	:	Mode I: (i) Right- and lefthand circular, possibility for polarization flipping from pulse to pulse (ii) Linear at $\pm 45^{o}$ with respect to vertical Mode II: Lefthand circular (transmission)
Beamwidths, broadside, calculated values	:	Mode I: 0.6^{o} east/west 1.7^{o} north/south Mode II: 1.2^{o} east/west 1.7^{o} north/south
Phasing range	:	21.3^{o} east and west of transit plane, steps approximately 1.2^{o}
Range of mechanical movement	:	30^{o} south to 60^{o} north of zenith

Speed of mechanical movement	:	5^{o}/min
Frequencies:	:	224 $\pm$ 1.25 MHz, steps 0.25 MHz
Peak pulse power	:	5 MW (specification)
Average	:	675 kW (specification)

12.6 RF and microwave radiation hazard

12.6.1 General

The recent years have seen an increasing interest in bioeffects of microwave radiation. Along with the growing concern for the potential danger which the radiation may represent for living organisms there has been, and continues to be, a discussion of which radiation levels to impose as safety limits. Topics conjectured are the relative importance of thermal and non-thermal effects, scaling problems related to the difficulties which arises in attempts to extrapolate from observations on small mammals to the exposure conditions of human beings, CW versus pulsed radiation, the frequency dependence of the radiation hazard and the significance of the exposure length. As yet there is no internationally agreed standard.

In Western countries the medical emphasis seems to have been on the heating of local tissues in the human body. For several years a radiation level of 10 mW/cm^2 was considered to be relatively safe limit for an exposure time limited to an ordinary working day, or 8 hours a day. The trend now is to lower this value by an order of magnitude and take

1 mW/cm^2

as the valid limit, somewhat higher doses being permitted for shorter exposure intervals.

In Eastern Europa investigators, evidently focussing their interest on non-thermal effects, have claimed that even radiation intensities more than two order of magnitude lower may have adverse effects on the central nervous system (CNS). These results have so far not been generally corroborated by Western studies.

12.6.2 Radiation near the EISCAT antennae at Ramfjordmoen

It will be convenient to divide the environment into three areas:

(i) Antenna reflectors/subreflectors (UHF-antenna only)

UHF:

As presently configured, with one output tube only, the transmitter is specified to yield an average power of 250 kW. Assuming a uniformly illuminated reflector the radiation intensity at the antenna becomes 31 mW/cm^2. With a second klystron installed this value should be doubled. The subreflector will be exposed to a much higher radiation intensity.

VHF:

Assuming an average power of 750 kW Kildal (1978) quotes intensities of 22 mW/cm^2 along the reflector at the shortest distance from the focal line, and 7 and 4 mW/cm^2 at the northern and southern edges of the reflector respectively.

The evident conclusion is that under no circumstances should personell be in the reflector or subreflector areas of the antennae during transmissions.

(ii) Near field of main beams

UHF:

Figure 12.3 displays how the axial power density from a uniformly illuminated circular antenna varies with the distance from the antenna, from zero to a value four times the density on the aperture. The maximum radiation density corresponds to 125 mW/cm^2 for an average radiated power of 250 kW. With a second klystron inserted this value should be doubled.

From the last maximum, occurring at a distance of 800 m from the antenna, the axial intensity falls off to about 4.8 mW/cm^2 at the Fresnel limit of 6.4 km.

VHF:

As stated under (i) the maximum reflector intensity amounts to 22 mW/cm^2. Hansen and Bickmore (1959) have shown that for a quadratic antenna the axial field oscillates with increasing amplitude, reaching a maximum 13 times the intensity at the Fresnel limit at a distance a little less than 0.2 times the Fresnel distance. Although it is difficult to extrapolate from these results to the conditions of the VHF-antenna, we might expect that fields considerably higher than the maximum surface field may be present up to 2-3 kilometers from the reflector.

For the VHF-antenna, with scanning limits at zenith angles 30° south and 60° north, the main beam cannot hit the ground around the antenna or the nearby mountain ranges. The UHF antenna will be equipped with a

"radiation inhibit" interlock, preventing the beam region from intersecting the ground or buildings in the neighbourhood at low elevation angles.

Planes which happen to fly through the axial part of the near zone radiation at distances less than 1-2 km may experience high field intensities. However, in almost all cases nowadays are the aircraft bodies covered by metallic material and crew and passengers are screened from the radiation. Moreover, vehicles ordinarily move with velocities at 100 m/s, or higher, and will therefore travers the central portion of the beam cross section in fractions of a second. Commercial planes moving at heights more than 5 km above ground at speed 8-900 km/hour should pose no problem. The most serious case might be helicopters with relatively large windows which for some reason might linger in the air at close distance above the UHF-antenna.

(iii) Ground immediately around the antenna structures

Kildal (1978) has made some calculations of the radiation intensities to expect and we shall confine ourselves to a presentation and discussion of his results. His computations are based on geometrical optics. As indicated in previous sections a thorough analytical treatment of the field distribution in the penumbral regions at the reflector boundaries tends to be impractical due to mathematical complications. Therefore, we regards the figures to be presented more as guide lines than an accurate decription of the real field distributions.

UHF:

There are essentially two effects to consider: (i) the spillover from the horn and (ii) the spillover from the subreflector scattering. As derived by Kildal (1978) the two contributions are shown in figure 12.22 for an assumed average radiated power of 500 kW. For the present average radiated power of 250 kW the values in the graph should be halved. We see from the figure that with the "radiation inhibit" preventing radiation below 20^{o} ($\theta = 70^{o}$) the computed radiation levels caused by the feed horn will not for any distance exceed 3-4 mW/cm^2 for 500 kW average and 1.5-2 mW/cm^2 for 250 kW. On the other hand at the shadow boundary for the main reflector the scattering radiation from the subreflector may reach values as high as 18 mW/cm^2 for 500 kW average and 9 mW/cm^2 for 250 kW about 13 m from the antenna centre. At the distance of the closest point of the fence which surrounds the antenna, the intensity is decreased to 7 and 3.5 mW/cm^2 for the quoted average powers.

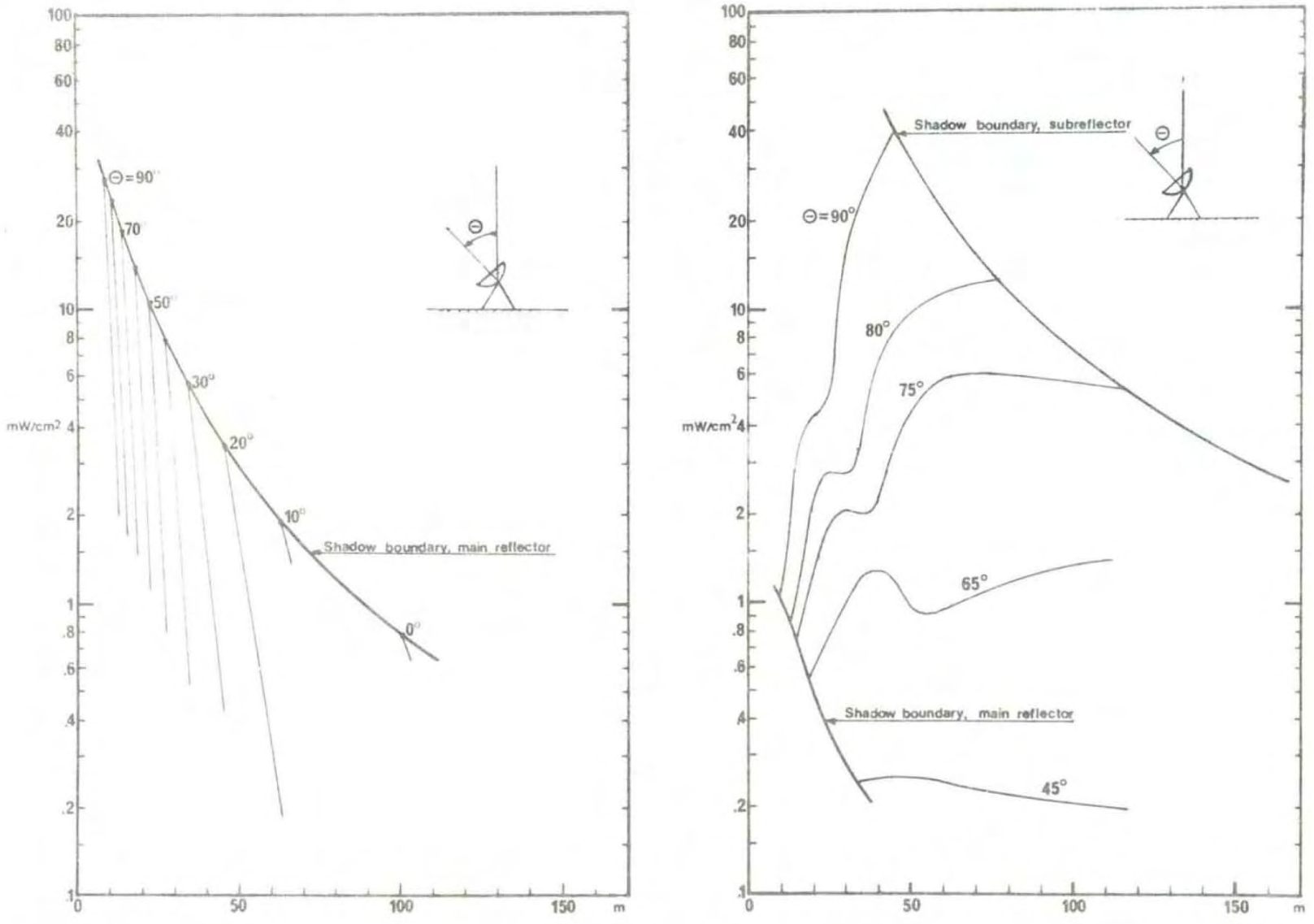


Figure 12.22 Computed radiation from feed horn and subreflector in the vicinity of the UHF-antenna, assumed average radiated power 500 kW. Note that with present transmitter the average radiated power is only half of assumed value (Kildal (1978)).

VHF:

Calculated ground intensities for the VHF-antenna are presented in figure 12.23. For broadside radiation the highest field experienced will be at the northern edge of the reflector. As quoted under (i) the intensity here may reach 7 mW/cm^2. As indicated in the figure the highest intensities are observed at the end areas of the antenna in case of phased beam operation, being as high as 12 mW/cm^2 at a phasing angle of 25^o. Some reduction of these values is expected since the maximal operating phasing angle now is 21.3^o.

We note that at the position of the fence the highest intensities are found toward the south. Here 3-4 mW/cm^2 may be reached.

To conclude: inside the fenced areas, close to the antennae we expect radiation intensities around 10 mW/cm^2 (with 250 kW average from the UHF-antenna). Outside the fence the maximum levels are anticipated to be less by a factor from 2 to 3. As a general precaution people should not stay within the fenced area during transmissions, neither should work be carried out close to the fence as long as the antennae are radiating.

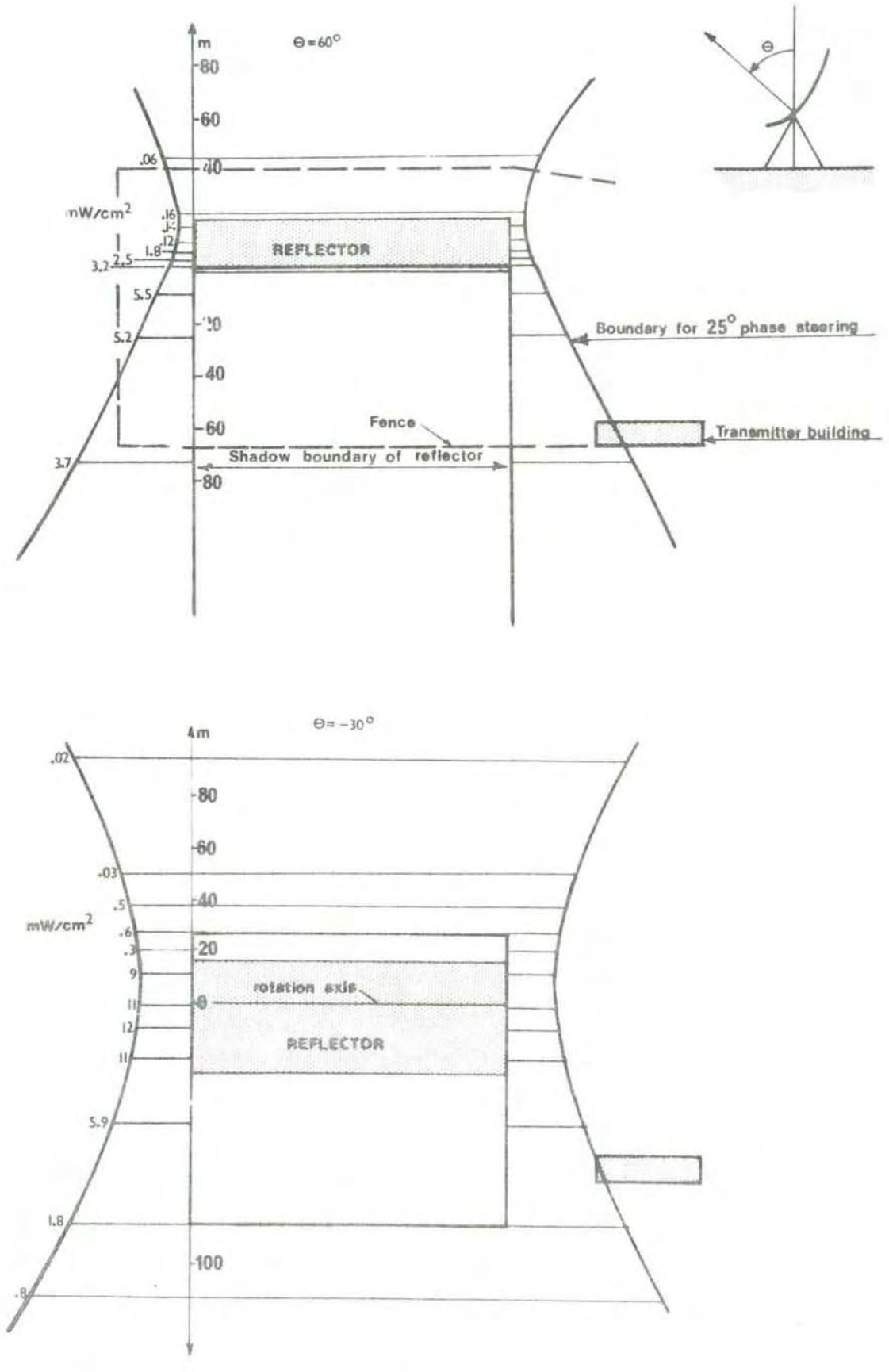


Figure 12.23 Radiation level at the ground around the VHF-antenna computed for an assumed average radiated power of 750 kW (Kildal 1978).

With the inherent uncertainties of the computed radiation intensities it is mandatory to map the radiation environment by performing systematic measurements once the transmitters become operable. Measurements of possible leakages from flanges and joints of the feed system should also be included. Further, an experimental test series should be carried out to determine the indoor radiation environment. The metal in the roof and the walls of the building is likely to reduce the radiation levels to acceptable values within any reasonable standard. The only exception might be the areas close to the windows. These should be carefully checked. Should the levels here proove to be higher than acceptable, the windows may be screened in a suitable way. The author is aware of at least one station abroad where meshed window panes have been applied.

12.7 Horizon profile at Ramfjordmoen

Figure 12.24 gives an experimentally determined horizon profile for the UHF-antenna at Ramfjordmoen. In carrying out the measurements 75 Watt pulses were fed to the antenna by the transmitter exciter. The displayed graph has been obtained by successively stepping the antenna in azimuth and noting the highest elevation angles from which perceptible ground echoes were observed.

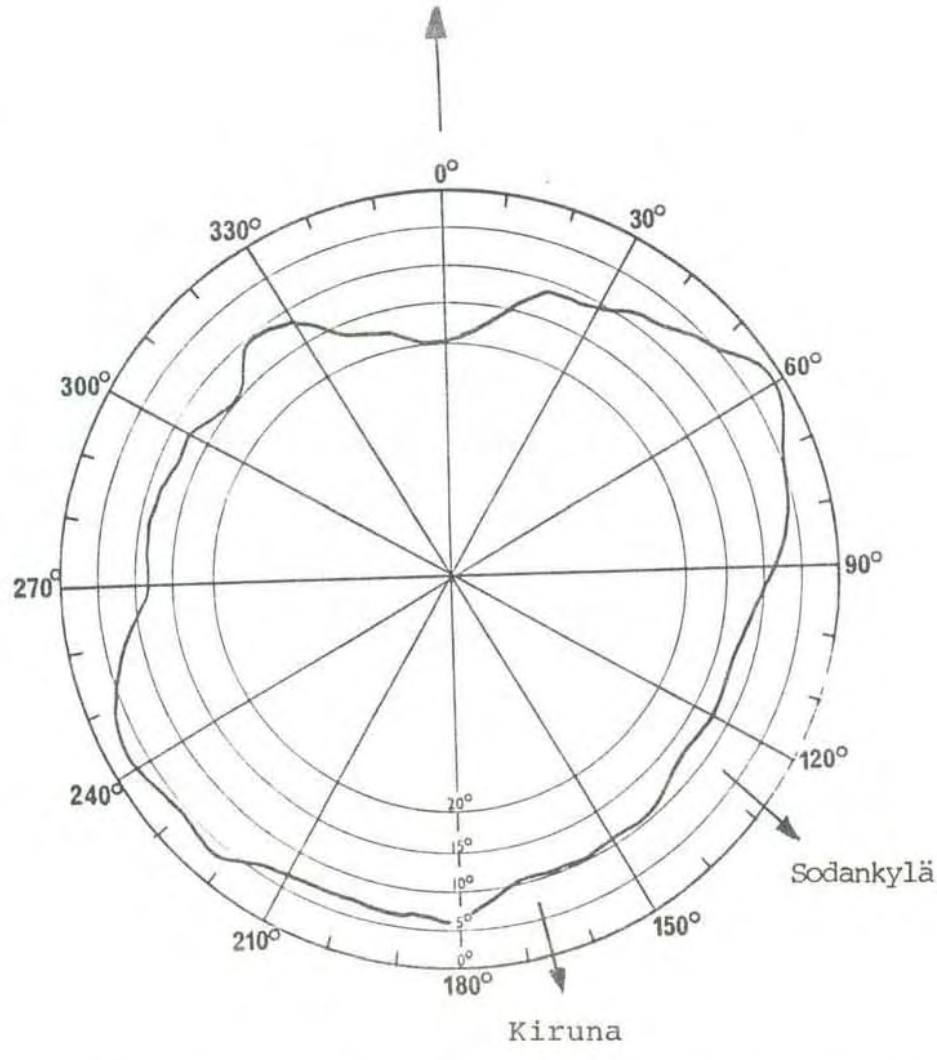


Figure 12.24 Horizon profile for the UHF-antenna, determined as maximum angles for ground reflections. Elevations scale extends from 0° to 20°.

In figure 12.25 the distances observed for the echoes yielding the horizon profile are plotted.

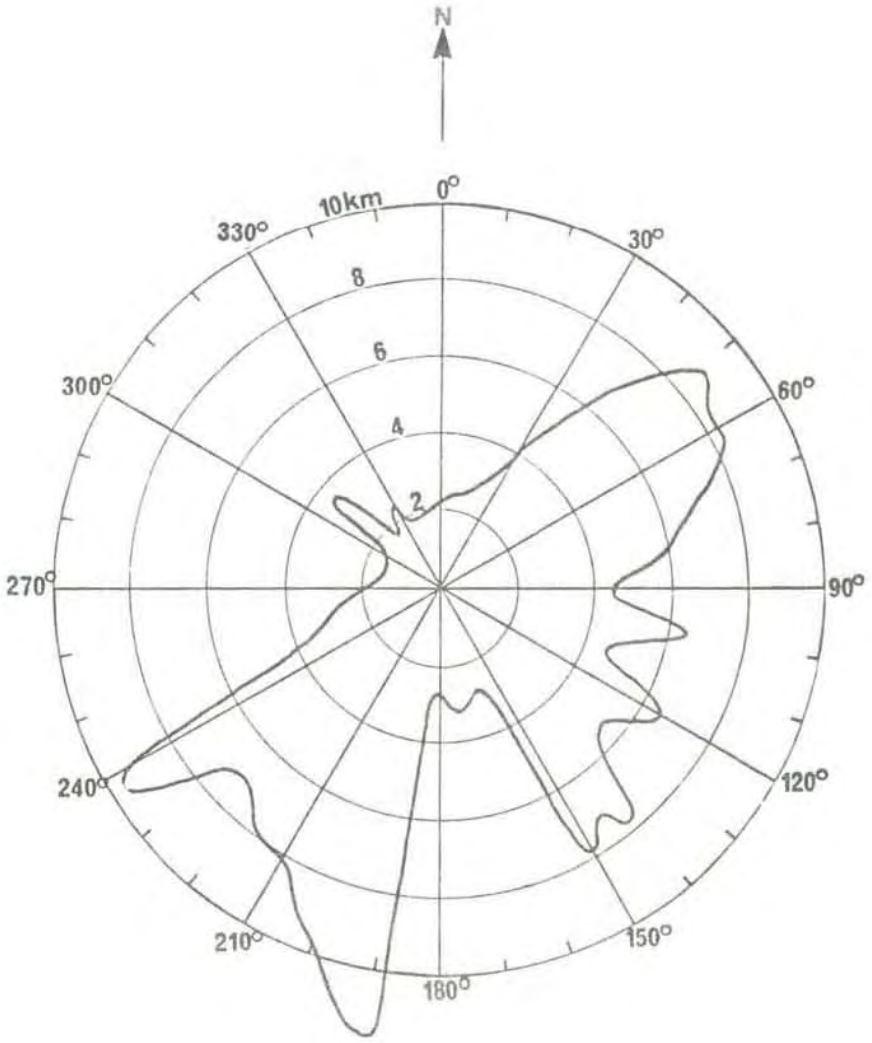


Figure 12.25 Distance, in km, of ground echoes determining horizon profile.

References and reading

Bickmore, R W
R C Hansen - Antenna Power Densities in the Fresnel Region, Proc IRE 47, 2119 (1959)

Hagfors, T - Phased Linear Array - Fixed Reflector Concept for the EISCAT VHF Antenna, Design Study (1977)

Hagfors, T - The Bandwidth of a Linear Phased Array with Stepped Delay Corrections, EISCAT Techn Note 78/4 (1978)

Jackson, J D - Classical Electrodynamics, John Wiley & Sons, Inc. (1975)

Kildal, P-S - Beregning av strålingsintensitet rundt EISCATs antenneanlegg på Ramfjordmoen, informal report (1978)

Kildal, P-S - Discrete Phase Steering by Permuting Precut Phase Cables, EISCAT Techn Note 78/6 (1978)

Kildal, P-S - Feeder Elements for the EISCAT VHF Parabolic Cylinder Antenna, EISCAT Techn Note 78/8 (1979)

Kildal, P-S - EISCAT VHF Antenna Tests, EISCAT Techn Note 80/21 (1980)

Kuzmin, A D
A E Salomonovich - Methods of Antenna Measurements, Academic Press (1966)

Kärcher, H J - Die Stahlkonstruktion der EISCAT-VHF-Zylinderparabol-Antenne in Tromsø, Der Stahlbau 49, 269 (1980)

Rusch, W V T
P D Potter - Analysis of Reflector Antennas, Academic Press (1970)

Schelkunoff, S A - Electromagnetic Waves, D Van Nostrand Company, Inc. (1943)

Skolnik, M I - Introduction to Radar Systems, Mc Graw-Hill

EISCAT publications

F. du Castel, O. Holt, B. Hultqvist, H. Kohl and M. Tiuri:
A European Incoherent Scatter Facility in the Auroral Zone (EISCAT).
A Feasibility Study ("The Green Report") June 1971. (Out of print).

O. Bratteng and A. Haug:
Model Ionosphere at High Latitude, EISCAT Feasibility Study, Report No. 9.
The Auroral Observatory, Tromsö July 1971. (Out of print).

A European Incoherent Scatter Facility in the Auroral Zone, UHF System and Organization ("The Yellow Report"), June 1974.

EISCAT Annual Report 1976.(Out of print).

P.S. Kildal and T. Hagfors:
Balance between investment in reflector and feed in the VHF cylindrical antenna.
EISCAT Technical Notes No. 77/1, 1977.

T. Hagfors:
Least mean square fitting of data to physical models.
EISCAT Technical Notes No. 78/2, 1978.

T. Hagfors:
The effect of ice on an antenna reflector.
EISCAT Technical Notes No. 78/3, 1978.

T. Hagfors:
The bandwidth of a linear phased array with stepped delay corrections.
EISCAT Technical Notes No. 78/4, 1978.

Data Group meeting in Kiruna, Sweden, 18-20 Jan. 1978
EISCAT Meetings No. 78/1, 1978

EISCAT Annual Report 1977

H-J. Alker:
Measurement principles in the EISCAT system
EISCAT Technical Notes No. 78/5, 1978

EISCAT Data Group meeting in Tromsö, Norway 30-31 May, 1978
EISCAT Meetings No. 78/2, 1978.

P-S. Kildal:
Discrete phase steering by permuting precut phase cables.
EISCAT Technical Notes No. 78/6, 1978

EISCAT UHF antenna acceptance test.
EISCAT Technical Notes No. 78/7, 1978.

P-S. Kildal:
Feeder elements for the EISCAT VHF parabolic cylinder antenna.
EISCAT Technical Notes No. 78/8, 1978.

H-J. Alker:
Program CORRSIM: System for program development and software simulation of EISCAT digital correlator, User's Manual.
EISCAT Technical Notes No. 79/9, 1979.

H-J. Alker:
Instruction manual for EISCAT digital correlator.
EISCAT Technical Notes No. 79/10, 1979

H-J. Alker:
A programmable correlator module for the EISCAT radar system.
EISCAT Technical Notes No. 79/11, 1979.

T. Ho and H-J. Alker:
Scientific programming of the EISCAT digital correlator.
EISCAT Technical Notes No. 79/12, 1979.

S. Westerlund (editor):
Proceedings EISCAT Annual Review Meeting 1969. Part I and II,
Abisko, Sweden, 12-16 March 1979.
EISCAT Meetings No. 79/3, 1979.

J. Murdin:
EISCAT UHF Geometry.
EISCAT Technical Notes No. 79/13, 1979.

T. Hagfors:
Transmitter Polarization Control in the EISCAT UHF System.
EISCAT Technical Notes No. 79/14, 1979.

B. Törustad:
A description of the assembly language for the EISCAT digital correlator.
EISCAT Technical Notes No. 79/15, 1979.

J. Murdin:
Errors in incoherent scatter radar measurements.
EISCAT Technical Notes No. 79/16, 1979.

EISCAT Digital Correlator. TEST MANUAL.
EISCAT Technical Notes No. 79/17, 1979.

G. Lejeune:
A program library for incoherent scatter calculation.
EISCAT Technical Notes No. 79/18, 1979.

K. Folkestad:
Lectures for EISCAT Personnel, Volume I
EISCAT Technical Notes No. 79/19, 1979.

Svein A. Kvalvik:
Correlator Buffer-Memory for the EISCAT Radar system
EISCAT Technical Notes. No. 80/20.

P-S. Kildal:
EISCAT VHF Antenna Tests
EISCAT Technical Notes No. 80/21

J. Armstrong:
EISCAT Experiment Preparation Manual
EISCAT Technical Notes No. 80/22

A. Farmer:
EISCAT Data Gathering and Dissemination
EISCAT Technical Note 80/23

Terrance Ho and Hans-Jørgen Alker:
Scientific Programming of the EISCAT Digital Correlator (Revised)
EISCAT Technical Note 81/24

Terrance Ho:
Programs Corrsim, Corrtest: System for Program Development and
Software Simulation of EISCAT Digital Correlator. User´s manual.
EISCAT Technical Note 81/25

Terrance Ho:
Instruction Manual for EISCAT Digital Correlator (Revised),
EISCAT Technical Note 81/26

Terrance Ho:
Standard Subroutines and Programs for EISCAT Digital Correlator.
EISCAT Technical Note 81/27

Terrance Ho:
Pocket Manual for Programming the EISCAT Digital Correlator.
EISCAT Technical Note 81/28