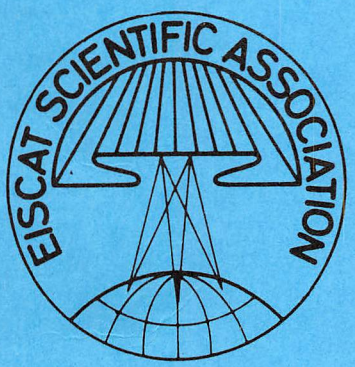
86/45



EISCAT
TECHNICAL
NOTE

STATISTICAL THEORY OF INCOHERENT SCATTER RADAR MEASUREMENTS

by

Markku S. Lehtinen

Finnish Meteorological Institute

Box 503, SF-00101 Helsinki, Finland

Ph.D. Thesis

University of Helsinki, 1986

KIRUNA
Sweden

STATISTICAL THEORY OF INCOHERENT SCATTER RADAR MEASUREMENTS

by

Markku S. Lehtinen

Finnish Meteorological Institute
Box 503, SF-00101 Helsinki, Finland

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in Auditorium XII on August 25th, 1986, at 12 o'clock noon

EISCAT Scientific Association
S-98127 Kiruna, Sweden
June 1986
EISCAT Technical Note 86/45
Printed in Sweden
ISSN 0349-2710
ISBN 951-99743-2-6 (diss.)

EISCAT Scientific Association
S-98127 Kiruna, Sweden
June 1986
EISCAT Technical Note 86/45
Printed in Sweden
ISSN 0349-2710
ISBN 951-99743-2-6 (diss.)

PREFACE

The idea about a mathematical formalism for describing the incoherent scatter measurements arose while I was working at the EISCAT Sodankylä site 1979-84. It seemed that a very miscellaneous set of tools was used in discussions about the coding principles. The ideas were inherited from the well developed formalisms for communications theory, the theory of spectral estimation of stationary processes and other related areas, almost but not exactly similar to the problems encountered in probing a continuous medium with radar pulses.

I hope this work will show that it is possible to solve the problems arising in the design of incoherent scatter measurements by exact calculations. A thorough mathematical understanding of the methods used will lead to better experiments than a study based on computer simulations of the hardware or a well developed set of rules of thumb. In this work the basic theory of the incoherent scatter radar responses and their statistical behaviour is developed and some practical applications are shown. I wish to be able to continue this work in two directions: 1) to analyze more explicitly all the different classical methods used so that it could be better used as an introductory textbook to the subject and 2) to develop further the possibilities offered by the statistical inversion methods to optimize the measurements and data analysis.

I have to thank several people for help in the preparation of this thesis. Markku Mäkelä has drawn the diagrams in chapter 2. Alan Farmer and Matti Vallinkoski have read through parts of the manuscript and picked away the most horrible errors in language. Tuomo Rantanen has been the leading TEXnician in typesetting the text. Tuomo Nygrén has found a significant amount of errors in the formulae so that he has to be thanked by anybody needing them to be correct. Asko Huuskonen has computed the curves in fig. 7.2.

The work has been supported financially by the foundations *Emil Aaltosen Säätiö* and *Jenny ja Antti Wihurin Rahasto* during a time period of three years and recently also by the Academy of Finland. Niilo Kaartinen has borrowed the computer used (and ruined) in the search for the alternating codes.

The most important thanks go to the people whose serious scientific attitude has provided inspiration and many times also practical help in solving the problems. I have to mention especially the inversion mathematicians Lassi Päivärinta, Erkki Somersalo and Heikki Haario in Helsinki, Tauno Turunen working in EISCAT and the whole EISCAT Sodankylä site staff.

Helsinki, June 1986

Markku S. Lehtinen

CONTENTS

Lehtinen, Markku S.: *Statistical theory of incoherent scatter measurements.*
University of Helsinki, Department of Theoretical Physics,
Helsinki 1986, Finland (SF). 97 pp.
(EISCAT Scientific Association, EISCAT Technical Note 86/45,
ISSN 0349-2710.)
ISBN 951-99743-2-6 (diss.)

ABSTRACT

High-resolution measurements of the ionosphere with an incoherent scatter radar using various pulse coding techniques are discussed with the aim to set up the theoretical tools necessary to optimize the accuracy of the measurements.

The statistical concepts needed are defined in the framework of statistical inversion theory.

A general formalism explaining the relationship between the expectation of a signal lagged product and the plasma scattering properties for an arbitrary coding scheme is derived. It is shown that this expectation is given as an average of the plasma scattering cross section as a function of both the range and lag variables. These averages are given in terms of ambiguity functions. Ambiguity functions are defined in the two-dimensional lag-range space. Reduced ambiguity functions are also discussed for those cases where the behaviour of the ambiguity function itself with respect to one of its parameters is not interesting.

The fourth moments theorem for gaussian processes is used to derive error fluctuations for the crossed product estimates of the signal in a general situation. Range-gating is also defined and considered here.

The concept of speed is defined for different coding strategies. The definition is based on the fact that the integration time necessary to attain a specified statistical accuracy is inversely proportional to the error variances of the estimates, and the Fisher information matrix is used as a basis for the exact definition. The speeds of the most generally used coding schemes are discussed and compared to each other in the white noise -dominated situation. Different practical post-detection filters are compared in terms of speeds and ambiguities.

It is proven that, given a range and lag resolution, there is a theoretical upper limit to the speed of a measurement if the design of the experiment is constrained by a maximum available modulation time and radar power. The noise is supposed to be dominated by white noise.

A new principle of radar coding, called the alternating code, is discussed. This principle differs from the previous ones in that the ambiguity functions for the codes used may be much wider than the final resolution, but different codes are used in different scans, so that when the data from different scans are suitably summed/subtracted from each other, the resulting ambiguity function may have one narrow peak only. This method comes close to the theoretical speed limit derived.

The new possibilities of a full inversion approach to the analysis of IS (incoherent scatter) data are discussed, including a way to use the ambiguous multipulse zero lag data to improve power profile accuracy by a significant factor.

KEY WORDS: ambiguity functions, optimization of measurements, inversion theory, deconvolution, alternating codes, speed of measurements, upper limits for accuracy.

0. INTRODUCTION

0.1 Physical Principles of Incoherent Scatter Measurements

In incoherent scatter (IS) measurements, the properties of ionospheric plasmas are measured by recording the correlation properties of backscattered radar signals. The radar power is scattered from the free electrons in the plasma by the mechanism of Thomson scattering. As the electrons have thermal velocity, the elementary scatters suffer a Doppler shift depending on the scatterers' velocities. Thus, it might be expected that a scattered monochromatic radio wave should have a spectrum whose form depends on the projection of the electron velocity distribution on the direction of the scattering vector.

In the early days of IS measurements, it was quickly found out that the spectrum looked quite different, however. The spectrum width was such that it more closely corresponded to the Doppler shifts due to the ion thermal velocities. Moreover, it was found not to be gaussian, but instead composed of two broad peaks at either side of the centre frequency. These two peaks not clearly separated, but had only a small dip in power between them. In active situations, the peaks may grow separate from each other.

The phenomenon was explained by many different authors (see *e.g.* Dougherty and Farley 1960,1961,1963, Fejer 1960,1961, Hagfors 1961). The reason was found to be that while the basic scattering mechanism is Thomson scattering from different free electrons, the electrons cannot be considered independently distributed in space. Rather, the electrons are concentrated in Debye spheres around ions moving with their thermal velocities. These interdependencies between different electrons cause the elementary scatterings to interfere rather analogously with the Bragg scattering from crystal structures. In this case the effect of constructive and destructive interference is seen as a sharp structure in the frequency space instead of a sharp structure in the angular dependence of the cross section. The two peaks composing the spectrum can be explained to originate from severely damped ion acoustic waves, one travelling towards the receiver and the other away from it. When the temperature of the electron population is higher than the ion temperature, the damping becomes smaller and the two peaks separate.

The exact derivation of the scattering spectrum from a uniform plasma volume (supposed to be small with respect to the geometry of the antennas and large with respect to the characteristic dimensions of the plasma correlations) can be

found in the papers mentioned above. In the one-ion case with neither magnetic field nor collisions taken into account it is expressed by the formulae

$$\langle e\left(t; d^3\vec{x}\right) \overline{e\left(t'; d^3\vec{x}\right)}\rangle / R = P_0\left(\vec{x}\right) \sigma_{\text{eff}}\left(t - t'; \vec{x}\right) d^3\vec{x}, \tag{0.1}$$

where $P_0(\vec{x})$ is the single electron scattering power, $d^3\vec{x}$ is the size of the elementary volume studied and σ_{eff} is the plasma effective cross section for lag $t - t'$ at point $\vec{x}$ related to the plasma spectrum $S_{\text{eff}}(\omega)$ by the formulae

$$\begin{aligned} \sigma_{\text{eff}}\left(\tau; \vec{x}\right) &= \int d\omega e^{i\omega\tau} S_{\text{eff}}\left(\omega; N_e(\vec{x}), T_e(\vec{x}), T_i(\vec{x})\right), \quad \text{where} \\ S_{\text{eff}}(\omega) &= \frac{2\pi^{1/2}}{\omega_{th}} \frac{e^{-\omega/\omega_{th}} N_e}{\left|1 + (kD)^2 + (T_e/T_i) W\left(\omega/\omega_{th}\right)\right|^2}. \end{aligned} \tag{0.2}$$

Here N_e is the electron density, T_e the electron temperature and T_i the ion temperature. D is the Debye length given by $D = (\varepsilon_0 \kappa T_e / e^2 N_e)^{1/2}$ and ω_{th} is the Doppler shift corresponding to the ion thermal velocity given by $\omega_{th} = k(2\kappa T_i/m_i)^{1/2}$. Ion and electron mass is denoted by m_i and m_e, respectively, κ is the Boltzmann constant and ε_0 is the dielectric constant of the vacuum. The length of the scattering wave vector is denoted by k with $k = 2\omega_0 / \sin(\theta/2)/c$, where ω_0 is the radar angular frequency, θ is the angle between the incident and scattered wave vectors and c is the velocity of light. For monostatic measurements $\theta = 180°$ and k is thus given by $k = 2\omega_0/c$. W is the Fried–Conte function or plasma dispersion function. The effective scattering power is

$$\sigma_{\text{eff}}\left(0, \vec{x}\right) = \frac{N_e(\vec{x})}{\left(1 + (kD)^2\right)\left(1 + (kD)^2 + T_e/T_i\right)}. \tag{0.3}$$

If the spectrum is measured by an experiment, the plasma parameters appearing in (0.2) can be found by using statistical estimation methods such as least squares fitting of the theory to the data. In practice, the numbers produced by a measurement often represent some averages of the autocorrelation function rather than the spectrum itself. Consequently, one has to fit these theoretically calculated averages to the data measured.

In addition to the ion spectrum expressed in (0.2), two other peaks may appear at the electron plasma frequencies corresponding to the electron plasma oscillations. While the methods developed in this work are completely general, we will be mainly interested in their applications to measuring the ion line only.

0.2 The Basic Requirements Posed on an Experiment

The main problem in conducting incoherent scatter measurements is the fact that, as the Thomson cross section is small and the electron densities are not very high, the signal received is weak. It is often smaller than the background noise from the sky and the receivers. Moreover, if the signal is received with the same antenna as is used for transmission, it is necessary to use small pulses instead of a continuous wave transmission, so that the responses from different altitudes would not be mixed with each other. This makes the situation still worse.

There are four resolution requirements that have to be taken into account in a design of a measurement. The first one is *spatial resolution.* This resolution determines the basic pulse widths used. The second one is the *lag resolution,* by which we mean the time in which the phase and amplitude of the scattered signal itself does not change significantly. Thus, it is the typical scale of the autocorrelation functions. The third one is the *time resolution,* by which we mean the time scale for the changes of the plasma parameters themselves. During this length of time it is possible to reduce measurement errors by repeating the experiment as many times as the radar duty cycles and other relevant factors allow. This is called integration, and consequently, the time resolution is often called integration time. The last one is then the required *accuracy* of the autocorrelation function estimates calculated.

In addition there are two basic extent requirements. The first one is that the experiment produces data from a long enough altitude interval to be useful. This poses a limit to the repetition frequency of the experiment. The second one is that the experiment must provide data over a broad enough lag interval. These intervals are called *range extent* and *lag extent*, respectively.

0.3 Classical Methods

The requirements of having all of the resolutions accurate are contradictory with each other when the design of the experiment is restricted with technical factors, such as radar peak power, radar duty cycle etc. As it is often impossible to meet all the requirements with straightforward techniques, a number of different methods of coding the transmission, perhaps using a series of coded groups with different frequencies, are being employed to use the radar equipment as effectively as possible. These methods have in common the property that they produce data with specified resolutions, but with better accuracy than a simple pulse or a pair of pulses would do.

The method used is basically determined by the first two resolution requirements and the extent requirements, as the form and duration of the transmitted pulses determines them. Time resolution is determined simply by the number of repeats

and the repeat frequency of the experiment. The classical methods may be divided in several categories:

1) Long pulse methods. These methods are used when the spatial resolution required is longer than the lag extent.

2) Pulse codes. These methods are best when the spatial resolution is shorter than the lag extent but of the same order than the lag resolution.

3) Phase coded pulse codes. These methods are best when the spatial resolution is shorter than the lag resolution. The knowledge that the amplitude or phase of the scattered signal does not change very much during several basic pulse lengths is used here to improve data accuracy. This is accomplished by special pulse compression techniques where sequences of the basic pulses are sent with possibly differring phases.

4) Pulse-to-pulse methods. These methods are used when the lag resolution is longer than the range extent. Usually these methods are used in D-layer work, where the spatial resolution is very small and thus the pulses can be also phase coded.

In addition, there are power profile experiments, in which one is not interested in the structure of the autocorrelation function at all, but only wants to get an estimate of the backscatter power. Here the total pulse length should be so small that one can be sure that the ACF stays almost constant. If the spatial resolution is much smaller than the possible total pulse length, phase coding can be used to improve data accuracy.

The data gathered by these methods never represent point values of the plasma properties in the ionosphere, but rather some averages of the plasma effective autocorrelation function in both the lag variable and the range variable. The functions specifying the form of the averages are called weighting or ambiguity functions. In most general form, the ambiguity functions depend on both the range and the lag variables. If the plasma effective correlation can be supposed to be approximately constant in either variable, the reduced ambiguity functions can be used. These are integrals of the two-dimensional ambiguity function over the variable in the direction of which the plasma effective correlation function does not change too much.

The theory of ambiguity functions for pulse coding and phase coding methods was pioneered by Woodman and Hagfors (1969), Farley (1969,1972) and Gray and Farley (1973), where the reduced ambiguity functions for the lag variable are shown. The study has been developed by Rino (1978) to include the reduced ambiguity functions for range. The error variances of the measured ACF estimates have also been calculated in various situations.

In the previous works, the two-dimensional ambiguity functions have not been clearly defined. Moreover, the effects of the pulse forms have not been related to the other factors determining the signal strength. Complete formulae relating the averages of signal crossed products to the plasma correlation function $\sigma_{\text{eff}}(\tau; S)$, including also the antenna beam shapes, the distance effects of the scattering point and the Thomson cross section, have not appeared in the literature. The study of the error variances of the estimators has also usually been limited to some special cases.

The lack of an exact mathematical formalism has had the consequence that the principles have mainly been understood intuitively, often using graphical devices such as time-range diagrams. These are obviously very useful devices, but it is impossible to base any advanced estimation methods, such as deconvolutions of the ambiguity functions or general inversion solutions of the estimation problems to them, unless exact and complete mathematical formulae describing the situation and the diagrams are simultaneously displayed.

0.4 Aims of the Present Study

The aims of this thesis are the following:

1) To develop a theory that fully explains the relationship of the measured crossed products of the signal to the plasma properties. This theory can be applied as well to monostatic and bistatic cases. The transmitted waveform and the receiver filter impulse response can be arbitrary; the effects of the beam shapes and distance factors are also included. The formulae are complete with all constants such as the Thomson cross section, the antenna gains etc. included, so that real signal strengths can be calculated, not just values up to an unknown constant. It is shown that the expectations of the crossed products are averages of the plasma effective scattering cross section σ_{eff} over both the lag variable and the range variable. The reduced ambiguity functions are also calculated leading to results similar to the previous ones.

2) To develop a general formalism by which it is possible to calculate the variances of any kinds of crossed product estimates and their sums.

3) To define the concept of measurement speed and to compare the classical methods in terms of it.

4) To state the problem of optimization of measurements as a precise mathematical problem and to derive theoretical limits to the possible accuracy of a measurement, given prescribed resolution requirements.

5) To describe a new principle of radar coding, by which it is possible to come close to the theoretical limits of accuracy.

6) To discuss the possibilities of a statistical inversion approach to develope incoherent scatter radar signal analysis and to show some examples of it.

The direct formalism for calculating the responses and error variances can be applied as well in the high and low SNR (signal-to-noise ratio) situations. The discussion of optimization and the comparison of the accuracies of various methods is, however, restricted to the low SNR case only.

0.5 The Structure of the Thesis

This thesis is divided into seven chapters in addition to this introduction. In the first one, the necessary statistical framework is defined. This is done in the spirit of statistical inversion theory. In the second chapter, the direct theory of the responses is developed in terms of the ambiguity functions. In the third chapter, the variances of arbitrary measurements are calculated. The fourth chapter is a comparison of the most important classical methods. In the fifth chapter, the problem of optimization of measurements is stated and a theoretical limit to the possible accuracy is found. The sixth chapter describes a new principle of radar coding and relates it to the theoretical accuracy limit. In the seventh chapter, the possibilities of a full inversion approach to the analysis are discussed, including a method to include multipulse zero lag data for making power profile estimates more accurate by a significant factor.

1. ELEMENTARY INVERSION THEORY

1.1 Introduction

In this chapter we derive the necessary statistical formalism referred to in the later chapters. The exposition is a somewhat modified approach from that of Tarantola and Valette (1982a,b). The language we use will be closer to that traditionally used by mathematicians studying random variables and their expectional distributions.

1.2 Basic Definitions

Let Ω be a probability space and let $m_i : \Omega \to \mathbf{R}^{n_i} = M_i, i = 0..N$ be random variables. We suppose that their distribution has a joint probability density

$$D(m_0, ..., m_N) : \prod_{i=0}^{N} M_i \to \mathbf{R} \ . \tag{1.1}$$

We call the variables $m_1, ..., m_N$ independent measurements of m_0, if the joint density can be represented in the form

$$D(m_0, ..., m_N) = D_{pr}(m_0) \prod_{i=1}^{N} D(m_i|m_0) \ , \tag{1.2}$$

where D_{pr} is called the a priori density of m_0 and where the $D(m_i|m_0)$ are called transition densities of the measurements. The notation shows the fact that the transition densities are conditional densities of the measurements where the desired variable m_0 is taken to be known.

Let us suppose that, in a measurement, we actually observe the values $\overline{m_1}, ..., \overline{m_N}$ of the measured variables. Then the conditional density

$$D_p(m_0) = D(m_0|\overline{m_1},...,\overline{m_N}) = \frac{D(m_0,\overline{m_1},...,\overline{m_N})}{\int\limits_{M_0} D(m_0,\overline{m_1},...,\overline{m_N})dm_0} \sim D_{pr}(m_0)\prod_{i=1}^{N} D(\overline{m_i}|m_0) \tag{1.3}$$

is called the a posteriori density of the desired variable m_0. By the single approximate equal sign we mean that the equation concerning probability densities is true except for normalization.

By inversion theory we understand here the study of a posteriori densities. More generally, a study would be based on probability measures in some function spaces, instead of probability densities in finite-dimensional Euclidean spaces, but for our purposes here, this simpler setting will be sufficient.

The a posteriori density is an objective way to describe measurement results since it gives the probability of different values of the desired variable, given the available information, *i.e.* the measurement values and the a priori density. To be able to calculate it, one has to know the transition densities and the a priori density. As the formula for the a posteriori density is so simple, it may be difficult to understand why it is used so little. The reason is that the desired variable m_0 is often a vector with many components, and so its density is a function of a great number of parameters. As functions of more than two parameters are practically impossible to plot, it is necessary to describe the a posteriori distribution in a simpler manner.

A good way is to give the the marginal densities of $D_p(m_0)$ with respect to the vector components of m_0. The interdependences of the uncertainties of the different components of m_0 are then lost. Another way is to describe a set in M_0 which contains a significant amount, say 95%, of the total probability mass. Such sets are called confidence sets. However, in most cases one just gives the point m_0 where $D_p(m_0)$ attains its largest value and, in addition, some error estimates describing the width of the a posteriori distribution in the directions of the different vector components of m_0. The exact meaning of the error estimates is often not specified clearly. The method of finding the maximum point of $D_p(m_0)$ is called the maximum likelihood estimation.

A basic difficulty of this kind of methods is that one has to specify an a priori distribution for the unknown variable m_0. Often the experimenter does not know the probability law that governs the behaviour of the desired variables in nature, and some guesswork is needed in choosing the a priori distribution. This difficulty divides statisticians in the classical and Bayesian schools. From the Bayesian viewpoint, the basic difference between these is that the classical school wants to use only such methods as are independent of the a priori distribution,

while in the Bayesian approach the a priori distribution may well appear in the final formulae, too.

There are many cases, especially in the measurements in natural sciences, where the transition densities have such sharp distributions that the exact form of the a priori density does not have any effect in (1.3). In this case one may as well suppose that $D_{pr} = 1$. For example, in a typical situation in measuring electron densities by the incoherent scatter radar method, the relative widths of the transition densities may be of the order of a few percent of the correct value, while the experimenter's a priori knowledge is limited to knowing the correct order of magnitude. The conceptually simple definition of the a posteriori distribution as a conditional distribution makes the Bayesian approach attractive and it certainly has many applications in interpreting incoherent scatter measurements particularly.

1.3 Example of a Non-linear Theory with Gaussian Errors

A typical example of a measurement can be described as follows. Let us suppose that $x = m_0$ is a random variable with distribution $D_{pr}(x)$. In addition, let us suppose that the measurement $m = m_1$ depends on x through the equation

$$m = f(x) + \varepsilon, \tag{1.4}$$

where f is some measurable function $X \to M$. By X we denote here the space of the x values and by M the space of m values. We have denoted the noise by ε, which we suppose is a zero mean Gaussian M-valued random vector independent of x. In this situation the function f is often called the direct theory.

The transition density can now be expressed as

$$D(m|x) = \frac{1}{(2\pi)^{n_M/2}|\Gamma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(m - f(x))^{\mathrm{T}}\Gamma^{-1}(m - f(x))\right) , \tag{1.5}$$

where Γ is the covariance matrix of ε. The total density of m and x is then $D_{pr}(x)D(m|x)$ and the un-normalized a posteriori density given a measured value $\overline{m}$ is

$$D_p(x) = D(x|\overline{m}) \sim D_{pr}(x) \exp\left(-\frac{1}{2}(f(x) - \overline{m})^{\mathrm{T}}\Gamma^{-1}(f(x) - \overline{m})\right) . \tag{1.6}$$

Thus, we see that it is very easy to express the a posteriori density as a function of x. Except for normalization, it is simply the transition density multiplied by the a priori density. The most important difference is that the transition density is the density of m, with x playing the role of a parameter, and the a posteriori density is the density of x, with $\overline{m}$ being a parameter.

Let us suppose that the a priori density is approximately constant in the region where the second term in (1.6) is significantly different from zero. The maximum likelihood estimate is then equivalent to minimizing the quadratic form in the exponent function of (1.6). If the matrix Γ is diagonal, this quadratic form is just the sum of the squares of the differences between the components of the theoretical mapping function, $f(x)$, and the components of the measurement result $\overline{m}$ weighted by the inverses of the diagonal elements of Γ. The maximum likelihood estimate is then equivalent to the correctly weighted least squares minimization problem. If the matrix Γ is not diagonal, the minimization of the quadratic form in (1.6) defines a generalized least squares estimation method with a non-diagonal weighting matrix Γ^{-1}.

The problem of fitting theoretically calculated ACF values to incoherent scatter measurements is of the type described here. The number of parameters fitted is usually of the order 2...5. It is often seen that the fitting of composition, for example, seems to lead to very odd behaviour of the minimization programs. Methods of alternately fixing some of the parameters while varying others are also used to make things work. This is probably an indication that the a posteriori distribution is rather complicated. As the error estimates provided by the standard least-squares approach are usually erroneous and the underlying assumption of approximate linearity of the theory is also suspect, it seems that the plotting of the whole a posteriori distributions would be necessary to justify the statistical significance of the fitted results. The number of fitted parameters is so small that the plotting of a number of projections (=marginal distributions) of the a posteriori distribution is practical.

1.4 A Set of Linear Measurements with Gaussian Errors

Linear measurements with Gaussian errors is a very interesting example because the a posteriori distribution can be described in analytical form. Let us suppose that the a priori distribution of m_0 is Gaussian with centre point $\overline{m_0}$:

$$D_{pr}(m_0) = \frac{1}{(2\pi)^{n_0/2}|\Gamma_0|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(m_0 - \overline{m_0})^T \Gamma_0^{-1}(m_0 - \overline{m_0})\right) . \qquad (1.7)$$

We suppose that the measurements $m_i : \Omega \to M_i$ are related to m_0 by the random variable equations

$$m = A_i m_0 + \varepsilon_i, \ i = 1...N,$$

where the noises $\varepsilon_i : \Omega \to M_i$ are independent Gaussian random variables with zero mean and covariance Γ_i. The transition densities are then given by

$$D(m_i|m_0) = \frac{1}{(2\pi)^{n_i/2}|\Gamma_i|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(m_i - A_i m_0)^{\mathrm{T}} \Gamma_i^{-1}(m_i - A_i m_0)\right) . \quad (1.8)$$

Let us denote $A_0 = \mathbf{1}$ to simplify the notation. The unnormalized a posteriori density, given measurements $\overline{m_i}$, can then be calculated by

$$\begin{aligned} D_p(m_0) &\sim D_{pr}(m_0) \prod_{i=1}^{N} D(\overline{m_i}|m_0) \\ &= \prod_{i=0}^{N} \exp\left(-\frac{1}{2}(A_i m_0 - \overline{m_i})^{\mathrm{T}} \Gamma_i^{-1}(A_i m_0 - \overline{m_i})\right) \quad (1.9) \\ &= \exp\left(-\frac{1}{2}\sum_{i=0}^{N}(A_i m_0 - \overline{m_i})^{\mathrm{T}} \Gamma_i^{-1}(A_i m_0 - \overline{m_i})\right) . \end{aligned}$$

By algebraic manipulations, the quadratic form in (1.9) can be rewritten in the standard form

$$D_p(m_0) = \frac{1}{(2\pi)^{n_0/2}|Q|^{-\frac{1}{2}}} \exp\left(-\frac{1}{2}(m_0 - \widetilde{m_0})^{\mathrm{T}} Q (m_0 - \widetilde{m_0})\right) \quad (1.10)$$

with

$$Q = \sum_{i=0}^{N} A_i^{\mathrm{T}} \Gamma_i^{-1} A_i \quad (1.11)$$

and with $\widetilde{m_0}$ any vector in M_0 satisfying the equation

$$Q\widetilde{m_0} = \sum_{i=0}^{N} A_i^{\mathrm{T}} \Gamma_i^{-1} \overline{m_i} . \quad (1.12)$$

If one can invert Q, the formula for the centre point of the a posteriori distribution becomes

$$\widetilde{m_0} = \left(\sum_{i=0}^{N} A_i^{\mathrm{T}} \Gamma_i^{-1} A_i \right)^{-1} \sum_{i=0}^{N} A_i^{\mathrm{T}} \Gamma_i^{-1} \overline{m_i} \,. \tag{1.13}$$

We have thus found a complete, analytical description of the a posteriori distribution for any linear, finite-dimensional estimation problem with Gaussian additive errors. The solution is given by the centre point of the resulting Gaussian distribution and by the a posteriori covariance matrix.

The matrix Q is usually called the Fisher information matrix if the a priori density is uniform. We will call it the Fisher information matrix also in the case when the a priori density is arbitrarily Gaussian. It is shown in (1.11) that each independent measurement adds its own contribution to the information matrix.

The formula (1.13) may be understood as an estimator of the desired variable m_0 in terms of the measurements m_i. It can be proven that this estimator is better than any other estimator of m_0:

Theorem 1.1 Let $f : M_1 \times ... \times M_N \to M_o$ be any measurable function (that is, any estimator of m_0). If the estimation error variance is defined by

$$\mathrm{error}(f) = \mathbf{E}(f(m_1, ..., m_N) - m_0)(f(m_1, ..., m_N) - m_0)^{\mathrm{T}},$$

then $\mathrm{error}(f) - \mathrm{error}(\widetilde{m_0})$ is a positive semidefinite matrix.

Proof: For the proof let us denote m_0 by x and the sequence $m_1, ..., m_N$ by m. It is clear that $\widetilde{m_0} = \mathbf{E}(x|m)$ and thus we can write

$$\mathrm{error}(f) - \mathrm{error}(\widetilde{m_0})$$

$$= \mathbf{E}(f - x)(f - x)^{\mathrm{T}} - \mathbf{E}\Big(\big(\mathbf{E}(x|m) - x\big) \big(\mathbf{E}(x|m) - x\big)^{\mathrm{T}} \Big)$$

$$= \mathbf{E}\Big(\mathbf{E}\big((f - x)(f - x)^{\mathrm{T}}|m\big) - \big(\mathbf{E}(x|m) - x\big) \big(\mathbf{E}(x|m) - x\big)^{\mathrm{T}} \Big)$$

$$= \mathbf{E}\Big(ff^{\mathrm{T}} - f\mathbf{E}(x^{\mathrm{T}}|m) - \mathbf{E}(x|m)f^{\mathrm{T}} + \mathbf{E}(x|m)\mathbf{E}(x|m)^{\mathrm{T}}$$
$$-2\mathbf{E}(x|m)\mathbf{E}(x|m)^{\mathrm{T}} + \mathbf{E}(x|m)x^{\mathrm{T}} + x\mathbf{E}(x^{\mathrm{T}}|m) + \mathbf{E}(xx^{\mathrm{T}}|m) - xx^{\mathrm{T}}\Big)$$

$$= \mathbf{E}\Big(\big(f - \mathbf{E}(x|m)\big)\big(f - \mathbf{E}(x|m)\big)^{\mathrm{T}}\Big),$$

which is clearly a positive semidefinite matrix. We have used the fact that $\mathbf{E}(f(m)\xi|m) = f(m)\mathbf{E}(\xi|m)$ for any random variable ξ. ■

It follows from theorem 1 that for any $a \in M_0$

$$\mathbf{E}\left(a^{\mathrm{T}} f(m_1, ..., m_N) - a^{\mathrm{T}} m_0\right)^2 \geq \mathbf{E}\left(a^{\mathrm{T}} \widetilde{m_0} - a^{\mathrm{T}} m_0\right)^2,$$

that is, the error of any linear combination of the components of m_0 will have a bigger variance if m_0 is estimated by any other estimator than $\widetilde{m_0}$.

It is possible to generalize theorem 1.1 for arbitrary distributions if one restricts oneself to affine estimators.

Theorem 1.2 Let m_0 be a random vector with mean $\overline{m_0}$ and correlation matrix Γ_0 and let

$$m = A_i m_0 + \varepsilon_i,$$

where the ε_i are random variables $\Omega \to M_i$ with zero mean which do not correlate with each other or with m_i. Let us suppose that the correlation matrices of ε_i are Γ_i. If $f : M_1, ..., M_N \to M_0$ is any affine mapping, and $\widetilde{m_0}$ is as in (1.13), then error(f) − error$(\widetilde{m_0})$ is a positive semidefinite matrix.

Proof: As f and $\widetilde{m_0}$ are affine mappings of $m_0 \times m_1 \times, ..., \times m_N$, it is clear that error$(f)$ − error$(\widetilde{m_0})$ depends only on the first and second moments of $m_0, ..., m_N$. Thus, instead of $m_0, ..., m_N$, one can consider a set of Gaussian variables with the same first and second moments and use theorem 1.1 to complete the proof. ■

2. DIRECT THEORY OF PULSED RADAR MEASUREMENTS

2.1 Introduction

In this chapter, we study the general theory behind all the different coding principles. Since the pulse duration is finite, the measured ACF values always represent the plasma fluctuation correlations averaged over a certain range interval determined by the transmitter pulse duration and form. In the receivers, the signal sampling is not instantaneous, but occurs only after some filtering. This means that the estimated lagged products, in addition to being averages of the plasma fluctuation correlations over certain range intervals, are also averages of the plasma fluctuations over certain lag intervals.

The tool we use to specify exactly the nature of these averages are the ambiguity functions, sometimes also called weighting functions. We define the ambiguity functions in terms of two variables: the lag τ and range. To make the results applicable also in the case of bistatic measurements, we use the total travel time from the transmitter to the scattering point and then back to the receiver as the range variable.

In some cases one can suppose that the properties of the plasma stay constant in the lag-range space along one of the variables in the region where the ambiguity function is nonzero. In these cases, one does not need the two-dimensional ambiguity functions, and for these cases the reduced ambiguity functions which depend on only one of the two parameters are defined. It is shown that the reduced ambiguity functions are just the projections of the two-dimensional ambiguity functions on one of the axes.

We will derive the mathematical expressions both for the two-dimensional ambiguity functions and for the reduced ambiguity functions in the general case where the transmitted pulse envelope and the receiver filter impulse response are given as arbitrary complex functions. For this purpose we will first show how the filtered signal amplitude is composed of contributions from elementary volumes with different ranges scattered at different times. To describe this average, we define the amplitude ambiguity functions as functions of scatter time and range. The final ambiguity functions are then derived from the amplitude ambiguity functions by correlations in the time variable direction.

We need the concept of stochastic integral in describing the composition of the received signal as an average over scatterings from elementary volumes. We feel that the introduction of such rather poorly known mathematical machinery is justified by the simplicity of its physical interpretation and by the fact that it alone makes possible the transition from the world of amplitudes, where the pulse phases are most easily understood, to the world of lagged products, which are the final results of the measurements. It would be possible to circumvent the use of this concept, but it would lead to a loss of much of the visuality due to the ability to derive the results stepwise.

The ambiguity functions provide us the means to estimate the signal strength in various situations. In addition to this, we show how to estimate the contribution of the background noise in our measurements using an arbitrary receiver filter, which makes it possible to calculate the signal-to-noise ratios for arbitrary measurements.

2.2 The Contribution of a Single Electron

Let us adopt the following notation: by $\vec{x}$ we denote the scattering point and by $d^3\vec{x}$ we denote an infinitesimal scattering volume. $G_0(\vec{x})$ is used to denote the transmitter antenna gain in the direction to $\vec{x}$, and $R_0(\vec{x})$ is used to denote the distance from the transmitter antenna to $\vec{x}$. $G_1(\vec{x})$ and $R_1(\vec{x})$ are used to denote the corresponding quantities of the receiver antenna. The receiver input impedance is denoted by R. The basic variable we use to represent range is defined by

$$S(\vec{x}) = (R_0(\vec{x}) + R_1(\vec{x}))/c. \tag{2.1}$$

We suppose that a complex hybrid detection scheme is used so that the receiver detects both the phase and amplitude of the signal using a reference signal pair of transmission frequency. The receiver thus generates a complex signal containing phase and amplitude information of the received narrow band process. The complex envelope of the voltage at the receiver input due to scattering of a continuous wave from the elementary volume $d^3\vec{x}$ is then denoted by $e(t; d^3\vec{x})$. Radar wavelength is denoted by λ.

Let us first suppose that there is only one electron in volume $d^3\vec{x}$, and that the influence of other possible electrons elsewhere need not to be taken into account. By using the formula for classical Thomson scattering cross-section and the antenna reciprocity relation for the receiver antenna effective area A_e

$$A_e = \frac{G_1(\vec{x})\lambda^2}{4\pi}\,, \tag{2.2}$$

it is an easy exercise to show that the power at the receiver input is given by

$$P_0(\vec{x}) = e\left(t; d^3\vec{x}\right)\overline{e\left(t; d^3\vec{x}\right)}/R = 4\pi r_0^2 \sin^2\chi \frac{P_t G_0(\vec{x})}{4\pi R_0^2(\vec{x})} \frac{G_1(\vec{x})}{4\pi R_1^2(\vec{x})} \frac{\lambda^2}{4\pi}, \tag{2.3}$$

where P_t is the transmitter power and $4\pi r_0^2 \sim 10^{-28} m^2$ is the classical electron cross section. The receiver input impedance is denoted by R. The polarization angle is denoted by χ.

2.3 The Scattering from an Elementary Plasma Volume

Actually, the electrons in the ionosphere are not stationary. Moreover, the electromagnetic forces in the plasma cause some regularity in the plasma fluctuations, which is seen as an interference between the scattered waves of the different electrons. Then the scattering from the elementary volumes may be described by the effective plasma correlation function σ_{eff} as derived in plasma physics. If A is a small volume containing $\vec{x}$ with uniform plasma parameters, it can be shown that the complex envelope of the signal scattered from A is a zero-mean complex Gaussian process with covariance

$$\langle e(t; A)\overline{e(t'; A)}\rangle/R = P_0(\vec{x})\,\sigma_{\text{eff}}(t - t'; \vec{x})\,\mu(A), \tag{2.4}$$

where all the geometrical effects have been included in the single electron scattering power as given in (2.3), and $\mu(A)$ is the volume of A. We suppose that the elementary volumes can be considered large enough that the scatterings from different volumes will be independent:

$$\langle e(t; A)\overline{e(t'; B)}\rangle/R = 0, \tag{2.5}$$

if $A \cap B = \emptyset$. Thus the scatters $e(t; A)$ as a set function form a family of stochastic orthogonal measures with $\sigma_{\text{eff}}(t - t'; \vec{x})$ as their (cross) structure function (see Gihman and Skorohod 1980 or Priestley 1981).

Formally, the formulae (2.4) and (2.5) can be combined to give

$$\langle e\left(t; d^3\vec{x}\right)\overline{e\left(t'; d^3\vec{x}'\right)}\rangle/R = \delta\left(\vec{x} - \vec{x'}\right) P_0(\vec{x})\,\sigma_{\text{eff}}(t - t'; \vec{x})\,d^3\vec{x}d^3\vec{x}', \tag{2.6}$$

where δ is the Dirac delta function and σ_{eff} is shown in (0.2)–(0.3).

2.4 The Amplitude Ambiguity Functions

We have derived the correlation properties of a continuous wave signal scattered from a single elementary volume and received at times t and t' in formula (2.6). The complex envelope of the total signal $e(t)$ at the receiver input at time t can be obtained by summing together all the elementary signals from the different scattering volumes. If we suppose that the transmission is pulsed with a complex envelope $\mathrm{env}(t)$, we arrive at the integral

$$e(t) = \int_{\vec{x}} \mathrm{env}(t - S(\vec{x}))e(t; d^3\vec{x}), \tag{2.7}$$

where S is the total travel time through $\vec{x}$ from the transmitter to the receiver as given in (2.1).

The integral in (2.7) is a stochastic integral with respect to the stochastic measure $e(t; d^3\vec{x})$. The receiver detects the complex envelope $e(t)$ of the signal. This envelope is filtered to improve noise performance before making the lagged product estimates in a correlator device. If we suppose that the impulse response function of the receiver filters is $p(t)$, we can represent the filtered signal $z(t)$ by the convolution

$$z(t) = (p * e)(t) = \int_{-\infty}^{\infty} p(t - \tau)e(\tau)d\tau, \tag{2.8}$$

and taking (2.7) into account, we arrive at

$$z(t) = \int_{-\infty}^{\infty} d\tau \int_{\vec{x}} e(\tau; d^3x)W_t^A(\tau; \vec{x}), \tag{2.9}$$

where

$$W_t^A(\tau; \vec{x}) = p(t - \tau)\mathrm{env}(\tau - S(\vec{x})) \tag{2.10}$$

is called the *amplitude ambiguity function* for sampling time t. Formula (2.9) tells us how the signal sampled at time t is composed of the elementary signals $e(\tau; d^3\vec{x})$ scattered from elementary volumes at $\vec{x}$ and received at times τ.

2.5 The Two-dimensional Ambiguity Functions

We can now study the lagged products of the filtered complex envelope of the backscattered signal. By using (2.9), we see that

$$\begin{aligned}\langle z(t)\overline{z(t')}\rangle/R &= \int_{-\infty}^{\infty} d\tau \int_{\vec{x}} \int_{-\infty}^{\infty} d\tau' \int_{\vec{x'}} W_t^A(\tau;\vec{x})\overline{W_{t'}^A(\tau';\vec{x'})}\langle e(\tau;d^3\vec{x})\overline{e(\tau';d^3\vec{x'})}\rangle \\ &= \int d^3\vec{x} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\tau d\tau' W_t^A(\tau;\vec{x})\overline{W_{t'}^A(\tau';\vec{x'})}P_0(\vec{x})\,\sigma_{\text{eff}}(\tau-\tau';\vec{x}),\end{aligned} \tag{2.11}$$

where we have used the basic equation (2.6) for the elementary signal contributions.

Let us define the two-dimensional ambiguity functions for the lagged products by the formula

$$W_{t,t'}(\tau;x) = \int_{-\infty}^{\infty} d\nu W_t^A(\nu;\vec{x})\,\overline{W_{t'}^A(\nu-\tau;\vec{x})}\,, \tag{2.12}$$

which we can use in representing the lagged product as an average of the effective plasma correlation over both space and lag variables as

$$\langle z(t)\overline{z(t')}\rangle/R = \int_{\vec{x}} d^3\vec{x} \int_{-\infty}^{\infty} d\nu P_0(\vec{x})\,W_{t,t'}(\nu;\vec{x})\sigma_{\text{eff}}(\nu;\vec{x})\,. \tag{2.13}$$

One should note that all geometrical factors depending on beam geometry, range variation, polarization, etc. are taken into account in the single electron scattering power $P_0(\vec{x})$, while the factors depending on transmitter modulation and receiver filtering are taken into account in the two-dimensional ambiguity function $W_{t,t'}(\nu;\vec{x})$.

2.6 The Reduced Ambiguity Functions

Formula (2.13) gives the exact value of the estimation average of the lagged product $z(t)\overline{z(t')}$. However, in some cases the two-dimensional representation of (2.13) may be unnecessarily complicated. This situation arises when the effective

plasma correlation function $\sigma_{\text{eff}}(\nu;\vec{x})$ may be assumed to be constant in one of the variables ν or $\vec{x}$ in the region where $W_{t,t'}(\nu;\vec{x}) \neq 0$. For example, let us assume that

$$\sigma_{\text{eff}}(\tau;\vec{x}) \approx \sigma_{\text{eff}}\,(t-t';\vec{x}) \quad \text{if} \quad W_{t,t'}(\tau;\vec{x}) \neq 0. \tag{2.14}$$

In this case we may perform the integration over the lag variable in (2.13), leading to

$$\langle z(t)\overline{z(t')}\rangle/R = \int_{\vec{x}} d^3\vec{x} P_0\,(\vec{x})\, W_{t,t'}(\vec{x})\sigma_{\text{eff}}\,(t-t';\vec{x})\,. \tag{2.15}$$

The reduced *range ambiguity function* $W_{t,t'}(\vec{x})$ is given by

$$W_{t,t'}(\vec{x}) = \int_{-\infty}^{\infty} d\tau\, W_{t,t'}(\tau,\vec{x}), \tag{2.16}$$

and by using formulae (2.12) and (2.10), we see that

$$W_{t,t'}(\vec{x}) = (p * \text{env})(t - S(\vec{x}))\,\overline{(p * \text{env})(t' - S(\vec{x}))}. \tag{2.17}$$

On the other hand, if we suppose that $\sigma_{\text{eff}}(\tau;\vec{x})$ is approximately constant with respect to $\vec{x}$ in the region where $P_0(\vec{x})W_{t,t'}(\tau;\vec{x})$ does not vanish, we can write

$$\langle z(t)\overline{z(t')}\rangle/R = P_0 \int_{-\infty}^{\infty} d\tau\, W_{t,t'}(\tau)\sigma_{\text{eff}}\,(\tau;\vec{x_0})\,, \tag{2.18}$$

where $\vec{x}_0$ is some fixed point in the region which contributes, and $W_{t,t'}(\tau)$ is the reduced *lag ambiguity function*, defined by

$$W_{t,t'}(\tau) = \int_{\vec{x}} d^3\vec{x}\,\text{SV}(\vec{x}) W_{t,t'}(\tau; S(\vec{x}))\,. \tag{2.19}$$

Here, the form of the beam intersection is given by

$$\mathrm{SV}(\vec{x}) = \frac{G_0(\vec{x})}{G_0(0)} \frac{G_1(\vec{x})}{G_1(0)} , \tag{2.20}$$

and the single electron power at the intersection of the beam axes is given by

$$P_0 = 4\pi r_0^2 \sin^2 \chi \frac{P_t G_0(0)}{4\pi R_0^2} \frac{G_1(0)}{4\pi R_1^2} \frac{\lambda^2}{4\pi}. \tag{2.21}$$

2.7 Physical Significance of the Formalism

It is easy to visualize the amplitude ambiguity function W_t^A as given by (2.10) in the $\tau - S(\vec{x})$ -plane. It is simply a product of two functions, one of which is constant along the lines τ=constant and the other constant along the lines $\tau - S(\vec{x})$=constant.

For illustration, let us suppose that the receiver impulse response is a boxcar function of length τ_0 and that the transmitter modulation envelope is a (poorly designed) pulse code consisting of three pulses as shown in figure 2.1.

In figure 2.2 we have chosen two samples taken at times t and t' and drawn the functions $p(t-\tau)$, $p(t'-\tau)$ and $\mathrm{env}(\tau - S(\vec{x}))$ in the $\tau - S(\vec{x})$ plane. The function $p(t'-\tau)$ is nonzero in the first vertical strip, the function $p(t-\tau)$ is nonzero in the second vertical strip, while the function $\mathrm{env}(\tau - S(\vec{x}))$ is nonzero in the area under the slanted strips. According to formula (2.10), the amplitude ambiguity function $W_{t'}^A(\tau; S(\vec{x}))$ is nonzero in the intersection of the slanted strips and the first vertical strip (shaded parallelograms), while the amplitude ambiguity function W_t^A is nonzero in the intersection of the slanted strips and the second vertical strip (shaded and hatched parallelograms).

A physical interpretation of the formalism may be arrived at by considering a fixed vertical line in figure 2.2 going through some time instant, say t, in the τ-axis. The intersection of that line with the slanted strips gives the intervals in range $S(\vec{x})$ from which the signal received at time t has been scattered. If we move the vertical line to the right (increase t), we see that the corresponding scattering region preserves its form, but moves higher, and if we decrease t, it moves lower. This is true for the signal at the receiver input. When we take the receiver filtering into account, we must recall that the filtered signal at time t is composed of the unfiltered signal averaged over past times as specified by the impulse response function $p(t-\tau)$. This time interval is represented by the first vertical strip in figure 2.2. Thus, we arrive at the conclusion that the filtered signal $z(t)$ sampled at time t must indeed be represented by the hatched parallelograms in figure 2.2. Correspondingly, the filtered signal $z(t')$ sampled at t' is represented by the shaded parallelograms.

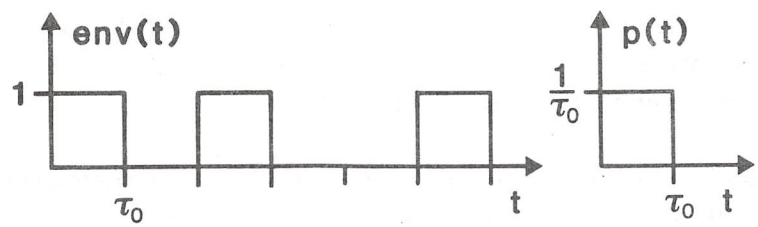


Figure 2.1 The transmitter envelope env(t) and receiver impulse response $p(t)$ used in the illustrations.

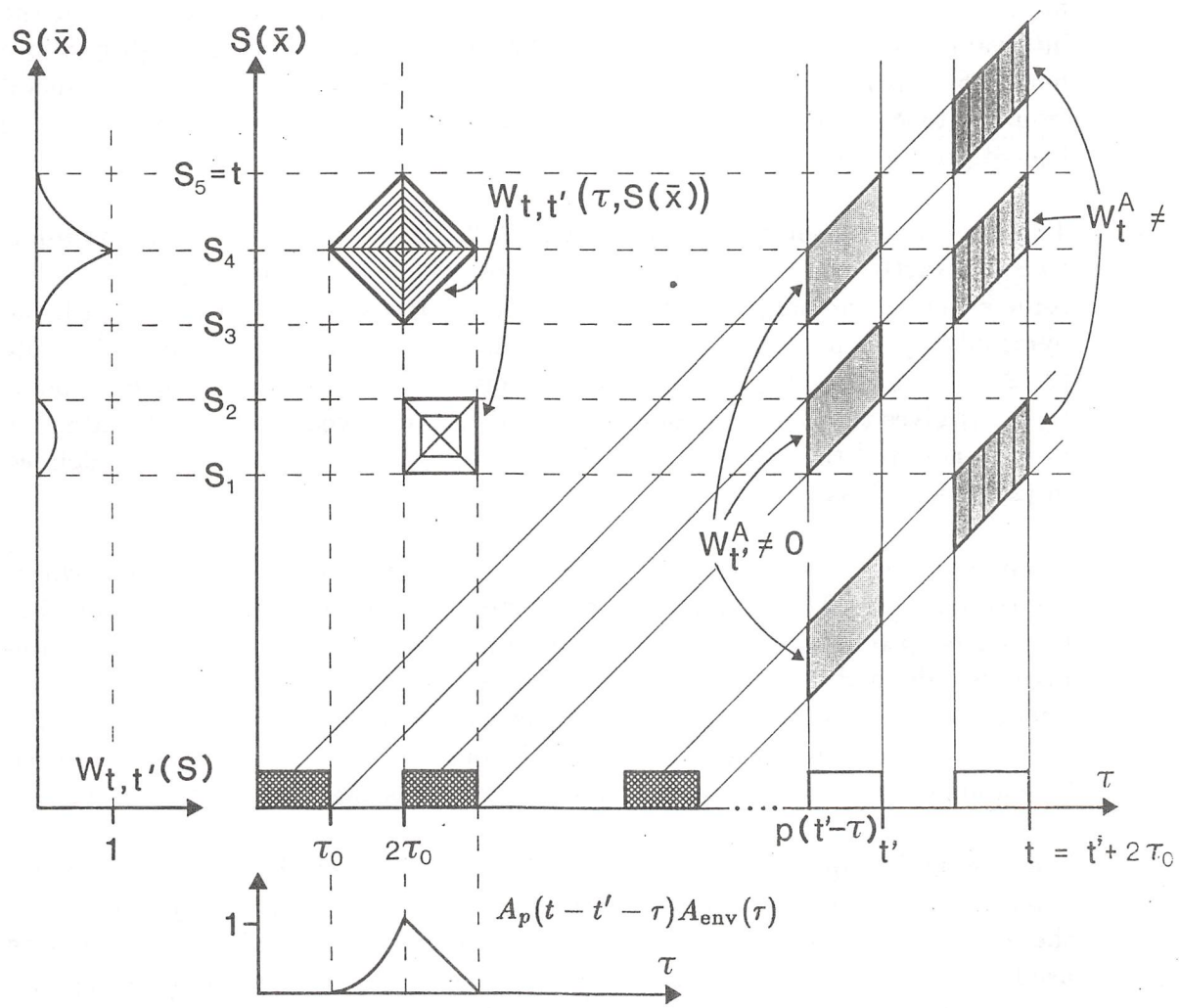


Figure 2.2 An illustration of the two-dimensional and reduced ambiguity functions for a three-pulse code.

In the correlator device the product $z(t)\overline{z(t')}$ is calculated. Since the signals scattered from different heights are independent of each other, contributions to this estimate arise only from those heights where both of the amplitude ambiguity functions W_t^A and $W_{t'}^A$ are nonzero. This occurs where the shaded and hatched parallelograms in figure 2.2 overlap. In figure 2.2, there are two distinct regions where this occurs: the height range from S_1 to S_2 and the height range from S_3 to S_5. Thus, we know that contributions to the average of the lagged product $\langle z(t)\overline{z(t')}\rangle$ will come from these height intervals.

Let us next study which lag intervals of the plasma correlation function σ_{eff} appear in the lagged product estimate. Let us take height S_4 as an example. In $z(t)$ signals received at the time interval $(t-\tau_0, t)$ are present, and in $z(t')$, signals received at the time interval $(t'-\tau_0, t')$ appear. Thus, the shortest lag in the correlation of $z(t)$ and $z(t')$ is $t-t'-\tau_0$ and the longest lag is $t-t'+\tau_0$. This means that from the height S_4 the lagged product of $z(t)$ and $z(t')$ will have some contribution from the plasma correlation $\sigma_{\text{eff}}(\tau; S_4)$ for all τ in the interval $(t-t'-\tau_0, t-t'+\tau_0)$. Slightly below height S_5 as well as slightly above height S_3, only the lag value $\tau = t-t'$ will contribute. At the height interval from S_1 to S_2, the values of $\sigma_{\text{eff}}(\tau; \vec{x})$ for $\tau \in (t-t', t-t'+\tau_0)$ will affect the lagged product estimate.

There are two pyramids drawn on figure 2.2. The areas under these pyramids consist exactly of those $(\tau, S(\vec{x}))$ pairs for which the value $\sigma_{\text{eff}}(\tau; S(\vec{x}))$ has some effect on the lagged product estimate $\langle z(t)\overline{z(t')}\rangle$. As is easily verified, the pyramids represent the two-dimensional ambiguity function shown in the formula (2.12). Formula (2.13) shows that the height of these pyramids over a point $(\tau, S(\vec{x}))$ gives the weighting of $\sigma_{\text{eff}}(\tau; S(\vec{x}))$ in the lagged product estimate. We have thus found the interpretation for the ambiguity functions formally defined in (2.10) and (2.12).

It should be noted that the presence of two separate regions where $W_{t,t'}(\tau; S(\vec{x})) \neq 0$ is usually an unsatisfactory situation, since it means that the lagged products estimated will not contain information concerning unambiguously defined height ranges. The smaller pyramid in figure 2.2 is called an *ambiguity*. The volumes of the pyramids are $(2/3)\tau_0$ and $(1/6)\tau_0$. Thus, the ambiguous contribution will in this case be one fourth of the main contribution. The ambiguity is not centred in time at $t'-t$, as is the main contribution.

The ambiguity appears here because the separation of the second and third subpulses in the transmitter modulation pattern is an odd multiple of τ_0. In these kinds of multiple pulse measurements, only even multiples of τ_0 should be used as separations between subpulses. However, if the receiver impulse response does not have the same length as the basic transmitted pulse, this rule does not apply, and the situation must be studied separately. Our example could be corrected by increasing the separation between the second and third subpulses by τ_0 so that the smaller pyramid disappears.

One should note that although boxcar functions have been used here for illustration purposes, the formulae given apply perfectly well to arbitrary functions. Figure 2.2 is also easily interpreted in the case of a more general receiver response function and perhaps a more general form of the envelope. The boxcar functions, however, will in many cases give a good overall picture of the possible ambiguities of the system and of the regions which contribute to the lagged product estimates.

The reduced ambiguity functions are integrals of the two-dimensional ambiguity function with respect to either of the variables. Both of these functions are shown in figure 2.2, the reduced ambiguity for range on the left and the reduced ambiguity for lag at the bottom. The ambiguity appears clearly here only on the reduced ambiguity function for range, although its contribution is slightly displaced from the middle point of the main contribution ($2\tau_0$) for lag, too.

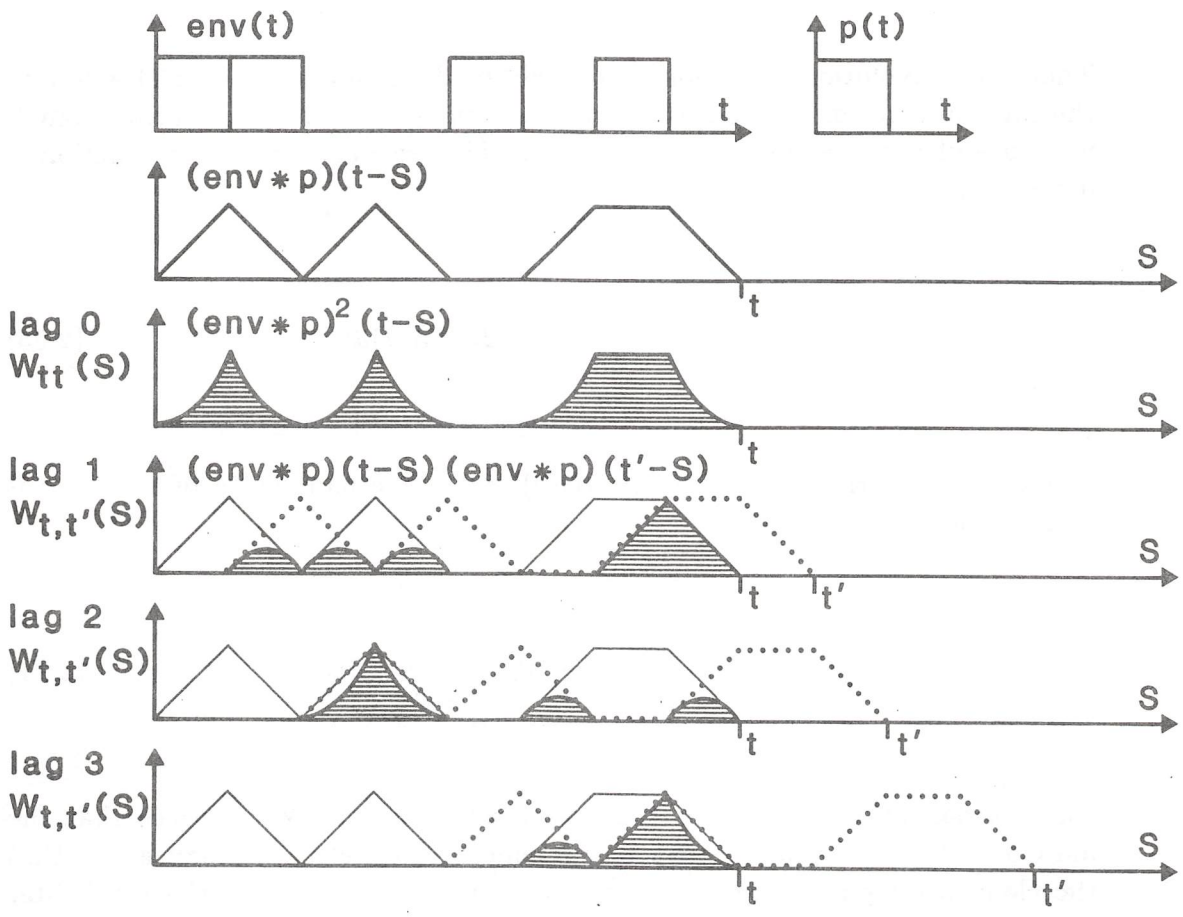


Figure 2.3 An illustration of the reduced range ambiguity functions for an ambiguous four-pulse code.

In figure 2.3, the construction of the reduced ambiguity function for range is illustrated for a four-pulse code. The interpulse separation is not an even multiple of the pulse length, resulting in several ambiguities. All the terms appearing in (2.17) are shown in figure 2.3 for a choice of (t, t') pairs. The two factors appearing at the right hand side of (2.17) are shown in the panels for different lags by solid and dotted lines, respectively, and their product is shown by the hatched areas. One should note that this kind of construction is possible by using ruled paper and graphical methods only. It is of course possible to use arbitrary forms for env and p, but then the calculations have to be made on a computer.

2.8 Effective Pulse Length and Scattering Volume

It is often useful to normalize the receiver filter response functions $p(t)$ so that

$$\int_{-\infty}^{\infty} p(t)dt = 1. \tag{2.22}$$

Then the convolution operation with p will be dimensionless, and $z(t)$ will have the same dimensions as $e(t)$. Moreover, the convolution of low frequency signals with p will not change the signal at all. The frequency response function is defined by

$$B_p(f) = \int_{-\infty}^{\infty} \exp(-i2\pi ft)p(t)dt, \tag{2.23}$$

and with the normalization (2.22), $B_p(0) = 1$. The bandwidth of the filter is then defined by

$$\mathrm{BW}_p = \frac{1}{2}\int_{-\infty}^{\infty} |B_p(f)|^2 df = \frac{1}{2}\int_{-\infty}^{\infty} |p(t)|^2 dt. \tag{2.24}$$

The normalization (2.22) is not always used, however. We will see that the matched filters for phase coded measurements are better normalized so that the elementary pulses are normalized according to (2.22), not the total filter response.

The reduced ambiguity function for range

$$W_{t,t'}(S) = (p * \mathrm{env})(t - S)\overline{(p * \mathrm{env})(t' - S)} \tag{2.25}$$

when considered as a function of travel time S can be understood as the *effective pulse form.* As shown in formula (2.15), it determines, together with the geometrical factors included in $P_0(\vec{x})$, the form of the volume that contributes to the backscattered signal strength. It is then natural to define the *effective pulse length* by the formula

$$l_{t,t'} = \int_{-\infty}^{\infty} W_{t,t'}(S)dS. \tag{2.26}$$

The *effective scattering volume* is defined analogously, by integrating with respect to $\vec{x}$ instead of s:

$$\begin{aligned} V_{t,t'}^{\mathrm{eff}} &= \int_{\vec{x}} d^3\vec{x} \int_{-\infty}^{\infty} d\tau\, \mathrm{SV}(\vec{x}) W_{t,t'}(\tau; S(\vec{x})) \\ &= \int_{\vec{x}} d^3\vec{x}\, \mathrm{SV}(\vec{x}) W_{t,t'}(S(\vec{x}))\ , \end{aligned} \tag{2.27}$$

If one can suppose that the functions $R_0(\vec{x})$, $R_1(\vec{x})$ and $\sigma_{\mathrm{eff}}(\tau;\vec{x})$ stay constant in the scattering region defined by the product $SV(\vec{x})W_{t,t'}(\tau; S(\vec{x}))$, the signal strength is given by

$$\langle z(t)\overline{z(t')}\rangle / R = 4\pi r_0^2 \sin^2\chi \frac{P_t G_0(0)}{4\pi R_0^2} \frac{G_1(0)}{4\pi R_1^2} \frac{\lambda^2}{4\pi} V_{t,t'}^{\mathrm{eff}} \sigma_{\mathrm{eff}}\,(t - t'; \vec{x}_0)\,. \tag{2.28}$$

where $\vec{x}_0$ is some point in the scattering volume and

$$R_0 = R_0(\vec{x}_0), \quad R_1 = R_1(\vec{x}_0). \tag{2.29}$$

2.9 The Signal Power for Monostatic Measurements

Let us now suppose that the transmitter antenna is used for receiving also. Then $S(\vec{x}) = 2r/c$ and $R_0 = R_1 = r$, where r is the range from the antenna to the scattering point. Moreover, let us suppose that the effective pulse length function is short enough so that $\sigma_{\text{eff}}(\tau;\vec{x})$ can be assumed constant in the scattering region. The reduced lag ambiguity function (2.19) can be simplified

$$W_{t,t'}(\tau) = G_0^{-2}(0) \int_\Omega G^2(\Omega) R_0^2 d\Omega \int_0^\infty W_{t,t'}(\tau; 2r/c) dr. \tag{2.30}$$

The integral of the antenna gain over the angular variables Ω can be written

$$\int G^2(\Omega) d\Omega = C_{beam} G(0) \int G(\Omega) d\Omega = C_{beam} 4\pi G(0), \tag{2.31}$$

where C_{beam} is a constant depending on the form of the beam geometry. (It does not depend on the beam width, if the form stays constant.) This constant can be numerically evaluated for different beam forms. If a Gaussian form is used, then $C_{beam} = 0.500$, and if the diffraction pattern for a circular aperture

$$G(\theta, \phi) = (J_1(a\theta)/a\theta)^2 \tag{2.32}$$

is used, $C_{beam} = 0.460$. In general C_{beam} is not a strong function of the beam form, and the value 0.460 for circular antennae will probably be a good approximation, even if the gain does not exactly obey formula (2.32).

The second integral in (2.30) can be written

$$\begin{aligned} \int_0^\infty dr\, W_{t,t'}(\tau; 2r/c) &= \frac{c}{2} \int_0^\infty dS\, W_{t,t'}(\tau; S) \\ &= \frac{c}{2} A_p(t - t' - \tau) A_{\text{env}}(\tau). \end{aligned} \tag{2.33}$$

Here, A_p and A_{env} denote the autocorrelation functions of p and env, given by

$$\begin{aligned} A_p(\tau) &= \int_{-\infty}^{\infty} dt\, p(t+\tau)\overline{p(t)} \quad \text{and} \\ A_{\mathrm{env}}(\tau) &= \int_{-\infty}^{\infty} dt\, \mathrm{env}(t+\tau)\overline{\mathrm{env}(t)}. \end{aligned} \tag{2.34}$$

Thus, we arrive at

$$W_{t,t'}(\tau) = A_{beam}^{\mathrm{eff}} \frac{c}{2} A_p(t-t'-\tau) A_{\mathrm{env}}(\tau) \ , \tag{2.35}$$

where

$$A_{beam}^{\mathrm{eff}} = C_{beam} \frac{4\pi R_0^2}{G(0)} \ . \tag{2.36}$$

The factor A_{beam}^{eff} can be considered here as the *effective beam cross section*, while the same term without the R_0^2 contribution can be understood as the effective beam cross section in steradians.

The effective scattering volume is obtained by an integration of $W_{t,t'}(\tau)$ over the lag variable τ, leading to

$$\begin{aligned} V_{t,t'}^{\mathrm{eff}} &= A_{beam}^{\mathrm{eff}} \frac{c}{2} l_{t,t'} \quad \text{with} \\ l_{t,t'} &= \int_{-\infty}^{\infty} dS\, (p * \mathrm{env})(t-S)\, \overline{(p * \mathrm{env})(t'-S)} \\ &= \int_{-\infty}^{\infty} d\tau\, A_p(t-t'-\tau) A_{\mathrm{env}}(\tau) \ . \end{aligned} \tag{2.37}$$

2.10 Noise

In all radar receiver systems there is some noise, partially due to the thermal noise in the receivers and partially due to the background noise from the universe. The bandwidth of this noise is very broad compared with the bandwidth of the receivers, and for this reason the background noise $e_n(t)$ is usually modelled as white noise, characterized by

$$\langle e_n(t)\overline{e_n(t')}\rangle/R = \kappa T\delta(t-t'), \tag{2.38}$$

where κ is the Boltzmann constant and T is called the system noise temperature.

The reason for this kind of representation is that the noise would be given by formula (2.38) if it were measured from a resistance at temperature T connected to an antenna with corresponding radiation resistance looking at a cold universe. Even if the real situation is more complicated, this is a convenient model to use and quite accurate, if the correct factor T is found by calibration measurements.

It is easy to see that the filtered noise signal $z_n = p*e_n$ obeys the formula

$$\langle z_n(t)\overline{z_n(t')}\rangle/R = \kappa T A_p(t-t'). \tag{2.39}$$

The filtered noise power is then given by

$$\langle |z_n(t)|^2\rangle/R = \kappa T A_p(0). \tag{2.40}$$

If the total signal is denoted by

$$Z(t) = z(t) + z_n(t), \tag{2.41}$$

its expectations are given by

$$\langle Z(t)Z(t')\rangle/R = \langle z(t)z(t')\rangle/R + \langle z_n(t)z_n(t')\rangle/R \tag{2.42}$$

because of the independence of z and z_n.

3. VARIANCES OF THE ACF ESTIMATES

3.1 Correlation Functions of the Signal and Noise

An incoherent scatter radar signal $Z(t)$ is a sum of two signals, the filtered radar echo $z(t)$ and the filtered noise $z_n(t)$. Because the radar echo is a sum of a large number of independent elementary contributions from different parts of the scattering volume, it is a Gaussian stochastic process by the central limit theorem. The filtered noise is also a Gaussian stochastic process, the basic reason for the Gaussian property being that the filtering operation sums together independent portions of the basic wideband noise process.

The filtered radar echo is a non-stationary process because of variations of the plasma scattering properties along the radar beam. The noise process is a stationary process, which is independent of the radar echo itself. Since both processes have zero mean and are Gaussian, they are specified by their correlation functions. Let us denote these correlation functions by

$$\begin{aligned} k(t,t') &= \langle z(t)\overline{z(t')}\rangle/R, \\ k_n(t,t') &= \langle z_n(t)\overline{z_n(t')}\rangle/R, \\ K(t,t') &= \langle (z(t)+z_n(t))\,\overline{(z(t')+z_n(t'))}\rangle/R = k(t,t')+k_n(t,t'). \end{aligned} \tag{3.1}$$

It has been shown in the previous chapter that the theoretical dependence of the correlation function k on the plasma correlation properties $\sigma_{\text{eff}}(\tau,\vec{x})$ is given by

$$\begin{aligned} k(t,t') &= P_0 \int_{\vec{x}} d^3\vec{x} \int_{-\infty}^{\infty} d\tau\, \mathrm{SV}(\vec{x}) W_{t,t'}(\tau; S(\vec{x}))\sigma_{\text{eff}}(\tau;\vec{x}) \\ &\approx P_0 \int_{\vec{x}} d^3\vec{x}\, \mathrm{SV}(\vec{x}) W_{t,t'}(S(\vec{x}))\sigma_{\text{eff}}(t-t';\vec{x}) \\ &\approx P_0 \int_{-\infty}^{\infty} d\tau\, W_{t,t'}(\tau)\sigma_{\text{eff}}(\tau;\vec{x_0}) \\ &\approx P_0 V^{\text{eff}}_{t,t'} \sigma_{\text{eff}}(t-t';\vec{x_0}). \end{aligned} \tag{3.2}$$

where x_0 is the centre point of the scattering volume and P_0 is the single electron backscattering power for point $\vec{x}_0$, given by

$$P_0 = 4\pi r_0^2 \sin^2 \chi \frac{P_t G_0(0)}{4\pi R_0^2} \frac{G_1(0)}{4\pi R_1^2} \frac{\lambda^2}{4\pi}. \tag{3.3}$$

The form of the beam intersection is given by

$$\mathrm{SV}(\vec{x}) = \frac{G_0(\vec{x}) G_1(\vec{x})}{G_0(0) G_1(0)}. \tag{3.4}$$

The different versions of the ambiguity functions are given by

$$W_{t,t'}(\tau; S(\vec{x})) = \int_{-\infty}^{\infty} d\nu \, p(t-\nu) env(\nu - S(\vec{x})) \, \overline{p(t'-\nu+\tau)} \, \overline{env(\nu - \tau - S(\vec{x}))} \tag{3.5}$$

and

$$\begin{aligned} W_{t,t'}(\tau; S(\vec{x})) &= \int_{-\infty}^{\infty} d\tau \, W_{t,t'}(\tau; S(\vec{x})) = (p*env)(t - S(\vec{x})) \cdot \overline{(p*env)(t' - S(\vec{x}))}, \\ W_{t,t'}(\tau) &= \int_{-\infty}^{\infty} d\vec{x} \, W_{t,t'}(\tau; S(\vec{x}) = A_{beam}^{\mathrm{eff}} \frac{c}{2} A_p(t - t' - \tau) \cdot A_{env}(\tau) \end{aligned} \tag{3.6}$$

The effective scattering volume is

$$V_{t,t'}^{\mathrm{eff}} = \int_{\vec{x}} d^3\vec{x} \, \mathrm{SV}(\vec{x}) W_{t,t'}(S(\vec{x})), \tag{3.7}$$

which reduces to

$$V_{t,t'}^{\text{eff}} = A_{beam}^{\text{eff}} \frac{c}{2} l_{t,t'} = C_{beam} \frac{4\pi R_0^2}{G(0)} \frac{c}{2} l_{t,t'} \ ,$$
$$l_{t,t'} = \int_{-\infty}^{\infty} dS \, W_{t,t'}(S) \tag{3.8}$$

in the case of monostatic operation.

It has also been shown that the covariance function for the noise signal is given by

$$k_n(t,t') = \kappa T A_p(t-t'), \tag{3.9}$$

where κ is the Boltzmann constant, T is the system noise temperature, and A_p is the autocorrelation function of the receiver filter impulse response.

3.2 ACF Estimation

Since the covariance function $k(t,t')$ carries the information from the plasma properties σ_{eff}, it is the purpose of incoherent scatter measurements to estimate this function from the measured backscatter signal. For this purpose, the matrices

$$M(t,t') = \text{ND}^{-1} \sum_{sc=1}^{\text{ND}} (z(t) + z_n(t)) \, \overline{(z(t') + z_n(t'))}/R \quad \text{and}$$
$$m_n(t,t') = \text{NB}^{-1} \sum_{sc=1}^{\text{NB}} z_n(t) \overline{z_n(t')}/R \tag{3.10}$$

are calculated in a special correlator device. Moreover, the matrix m is defined by

$$m(t,t') = M(t,t') - m_n(t,t') \ . \tag{3.11}$$

These estimates are calculated for a certain set of time pairs (t,t'), determined by the experimenter. The sums in (3.10) mean sums over different scans when the same radar pulse train is sent over and over again, and the calculated products are summed together to improve the statistical accuracy of the estimates. This

summation is called *integration.* Here, ND is called the *scan count* for data and NB is called the *scan count* for the background estimate.

Since it is impossible to separate z_n from the sum $z + z_n$, the second estimate (3.11) is actually calculated as

$$m_n(t,t') = \mathrm{NB}^{-1} \sum_{sc=1}^{\mathrm{NB}} z_n(t+t_n)\overline{z_n(t'+t_n)}/R \tag{3.12}$$

where t_n is chosen so that no radar echoes z are received at the times indicated in (3.12). This can be done because of the supposed stationarity of the noise signal.

In addition to the estimates M and m_n, one additional estimate is usually calculated for calibration purposes. Since the system gains as well as the system and background noise can change, especially when the antenna is moved, a known noise source is switched on to increase the system noise temperature by an amount T_c at some suitable instance where no echoes are received. The ACF estimates for calibration

$$m_c(t,t') = \mathrm{NC}^{-1} \sum_{sc=1}^{\mathrm{NC}} z_n(t+t_c)\overline{z_n(t'+t_c)}/R \tag{3.13}$$

are formed. By comparing $m_c(t,t)$ to $m_n(t,t)$, the system noise temperature can be found by a simple scaling using the known value T_c for the noise temperature increase and the formula

$$\langle m_c(t,t)\rangle / \langle m_n(t,t)\rangle = (T + T_c)/T. \tag{3.14}$$

If we suppose that the different scans give results independent of each other, it follows that

$$\begin{aligned} \langle M(t,t')\rangle &= K(t,t'), \\ \langle m_n(t,t')\rangle &= k_n(t,t') \text{ and} \\ \langle m(t,t')\rangle &= k(t,t'). \end{aligned} \tag{3.15}$$

3.3 Variances of the Estimates

Let us define the fluctuations of the estimated covariances by

$$\begin{aligned} \Delta M(t,t') &= M(t,t') - \langle M(t,t')\rangle, \\ \Delta m_n(t,t') &= m_n(t,t') - \langle m_n(t,t')\rangle \text{ and} \\ \Delta m(t,t') &= m(t,t') - \langle m(t,t')\rangle. \end{aligned} \tag{3.16}$$

The fourth moments theorem for complex Gaussian processes can be used to derive the formulae

$$\begin{aligned} \langle \Delta M(t,\tau)\Delta M(t',\tau')\rangle &= \mathrm{ND}^{-1}K(t,\tau')K(t',\tau), \\ \langle \Delta M(t,\tau)\overline{\Delta M(t',\tau')}\rangle &= \mathrm{ND}^{-1}K(t,t')K(\tau',\tau). \end{aligned} \tag{3.17}$$

The corresponding formulae for m_n can be obtained from (3.17) by replacing M with m_n and K with k_n. The corresponding covariances for $m = M - m_n$ can be obtained by summing together the covariances for M and m_n.

As such, the formula (3.17) does not seem very visual. By some algebraic manipulations, it is possible to derive formulae for the covariances of the real and imaginary components of ΔM (and similarly for Δm_n):

$$\begin{aligned} \langle \mathrm{Re}\Delta M(t,\tau)\mathrm{Re}\Delta M(t',\tau')\rangle &= \frac{1}{2}\mathrm{Re}\Big(K(t,t')K(\tau',\tau) + K(t,\tau')K(t',\tau)\Big)/\mathrm{ND} \\ \langle \mathrm{Im}\Delta M(t,\tau)\mathrm{Im}\Delta M(t',\tau')\rangle &= \frac{1}{2}\mathrm{Re}\Big(K(t,t')K(\tau',\tau) - K(t,\tau')K(t',\tau)\Big)/\mathrm{ND} \\ \langle \mathrm{Re}\Delta M(t,\tau)\mathrm{Im}\Delta M(t',\tau')\rangle &= \frac{1}{2}\mathrm{Im}\Big(-K(t,t')K(\tau',\tau) + K(t,\tau')K(t',\tau)\Big)/\mathrm{ND} \\ \langle \mathrm{Im}\Delta M(t,\tau)\mathrm{Re}\Delta M(t',\tau')\rangle &= \frac{1}{2}\mathrm{Im}\Big(K(t,t')K(\tau',\tau) + K(t,\tau')K(t',\tau)\Big)/\mathrm{ND}. \end{aligned} \tag{3.18}$$

It is interesting to note that in the case where $K(t,t')$ is real, the fluctuations in the real and imaginary parts of the ACF estimates are uncorrelated. In particular, this means then that the imaginary part contains no information about the form of the ACF. If the error fluctuations of the real and imaginary part estimates were correlated, the imaginary part estimate could contain information about the form of the ACF even though the ACF were real.

The non-correlation of the real and imaginary parts of the ACF estimates $M(t,t')$ in the case of a real ACF can be extended to one important case, where the

theoretical covariance function $K(t,t')$ is complex. This can be done, if we can suppose that the theoretical covariance function $K(t,t')$ has the form

$$K(t,t') = K_0(t,t')\exp\big(iv(t-t')\big), \tag{3.19}$$

where K_0 is real and v is a real constant. This situation can arise when an otherwise symmetrical fluctuation spectrum suffers a Doppler shift due to a nonzero plasma velocity.

If the modulus and phase of a complex number z are denoted by Moz and Phz:

$$z = \mathrm{Mo}z\exp(i\mathrm{Ph}z), \tag{3.20}$$

we can derive the following approximate formulae

$$\begin{aligned}
\langle \Delta\mathrm{Mo}M(t,\tau)\Delta\mathrm{Mo}M(t',\tau')\rangle &= \frac{1}{2}\Big(K_0(t,t')K_0(\tau',\tau) + K_0(t,\tau')K_0(t',\tau)\Big)/\mathrm{ND}\\
\langle \Delta\mathrm{Ph}M(t,\tau)\Delta\mathrm{Ph}M(t',\tau')\rangle &= \frac{K_0(t,t')K_0(\tau',\tau) - K_0(t,\tau')K_0(t',\tau)}{2\mathrm{ND}K_0(t,\tau)K_0(t',\tau')}\\
\langle \Delta\mathrm{Mo}M(t,\tau)\Delta\mathrm{Ph}M(t',\tau')\rangle &= 0.
\end{aligned} \tag{3.21}$$

One can arrive at these results considering, instead of the process $z + z_n$, the sum of the complex Gaussian processes $z'(t) = z(t)\exp(-ivt)$ and $z'_n(t) = z_n(t)\exp(-ivt)$

$$z'(t) + z'_n(t) = (z(t) + z_n(t))\exp(-ivt), \tag{3.22}$$

and its covariance matrix M'. The average of this matrix is given by K_0, and the fluctuations in the real part will be approximately equal to the fluctuations in the modulus. The fluctuations of its imaginary part will be approximately equal to the fluctuations of its phase multiplied by the modulus. This is because the new matrix M' will be approximately real. Since the matrices M and M' are related by

$$M'(t,t') = \exp\big(iv(t'-t)\big)M(t,t'), \tag{3.23}$$

the modulus and phase of M and M' fluctuate in the same way. Thus, by using the formulae (3.18) for the matrix M', whose average K_0 is real, we have proved the formulae (3.21).

It should be noted that the formulae (3.21) are only true in the situation where the theoretical covariance can be expressed in the form (3.19), and where the fluctuations of the matrix M', given by (3.23) will be so small that a linear approximation can be used to relate the fluctuations of the real and imaginary parts of M' to the fluctuations of the phase and modulus of M'.

If these assumptions are valid, information about the form of the spectrum is contained only in Mo$M(t,t')$ and information about the plasma velocity is contained only in the phase of M. This fact may be useful to simplify the analysis of data.

3.4 Range-gate Estimates

Range-gating is an operation where those elements of the measured matrices $M(t,t')$, $m_n(t,t')$ or $m_c(t,t')$ are summed together which are considered to represent essentially the same kind of information. This situation arises, for example, in the long pulse measurements, where the ambiguity functions

$$W_{t,t'}(\tau, S(\vec{x})) \quad \text{and} \quad W_{t+i,t'+i}(\tau, S(\vec{x})) \tag{3.24}$$

overlap in the range direction for small i, and thus the corresponding estimated lagged products

$$M(t,t') \quad \text{and} \quad M(t+i,t'+i) \tag{3.25}$$

contain information from essentially the same plasma volume. Since the time lag is the same in both estimates in (3.25), they contain information about $\sigma_{\text{eff}}(\tau,\vec{x})$ for lags τ close to $t-t'$.

To reduce the number of data points, the correlator can be programmed to sum together sets of estimated lagged products to form *range-gate estimates.* In this case, a range-gate estimate might be formed as

$$M_{rg} = \sum_{i=0}^{N} M(t+i,t'+i). \tag{3.26}$$

In general, a range-gate estimate can be described by specifying a set of time pairs (t, t'), over which the sum is calculated. If we denote this set by

$$rg = \{(t_1, t_1'), (t_2, t_2'), ..., (t_N, t_N')\}, \tag{3.27}$$

the range-gate estimate M_{rg} for the range-gate rg can be defined in a general way by

$$M_{rg} = \sum_{(t,t')\in rg} M_{t,t'}. \tag{3.28}$$

Usually all the lags $t - t'$ for the time pairs belonging to a range-gate rg are the same. The averaging of the plasma effective correlation function over range and lag is specified by the *range-gate ambiguity functions* $W_{rg}(\tau; S)$, $W_{rg}(\tau)$ or $W_{rg}(S)$, given by

$$W_{rg} = \sum_{(t,t')\in rg} W_{t,t'}. \tag{3.29}$$

If the function $W_{rg}(S)$ has a well defined maximum or middle point r, one can say that the range-gate rg represents the range r and the lag τ that is given by the common increment of the time pairs belonging to the range-gate. Ordinarily many range-gate estimates may be formed, and the different range-gates are most easily referred to by giving their range and common lag. Thus, we can refer to different range-gates by the notation

$$rg(r, \tau), \tag{3.30}$$

where r is the centre point of range-gate ambiguity function and τ is the common lag of the range-gate.

It should be noted that information may be lost by range-gating. One loses knowledge of the individual lagged product estimates, and neither is the simple summation a statistically optimal method to combine the information in the individual lagged products.

3.5 Variances of the Range-gate Estimates

The fluctuations of the range-gate estimates are defined by

$$\Delta M_{rg} = M_{rg} - \langle M_{rg} \rangle. \tag{3.31}$$

It is then easy to write an expression for the covariance between two range-gate estimates $rg1$ and $rg2$. Using the formulae (3.28) and (3.18), we have

$$\begin{aligned}
&\langle \mathrm{Re}\Delta M_{rg1} \mathrm{Re}\Delta M_{rg2} \rangle = \\
&\quad \sum_{(t,\tau)\in rg1} \sum_{(t',\tau')\in rg2} \frac{1}{2}\mathrm{Re}\Big(K(t,t')K(\tau',\tau) + K(t,\tau')K(t',\tau) \Big)/\mathrm{ND} \\
&\langle \mathrm{Im}\Delta M_{rg1} \mathrm{Im}\Delta M_{rg2} \rangle = \\
&\quad \sum_{(t,\tau)\in rg1} \sum_{(t',\tau')\in rg2} \frac{1}{2}\mathrm{Re}\Big(K(t,t')K(\tau',\tau) - K(t,\tau')K(t',\tau) \Big)/\mathrm{ND} \\
&\langle \mathrm{Re}\Delta M_{rg1} \mathrm{Im}\Delta M_{rg2} \rangle = \\
&\quad \sum_{(t,\tau)\in rg1} \sum_{(t',\tau')\in rg2} \frac{1}{2}\mathrm{Im}\Big(-K(t,t')K(\tau',\tau) + K(t,\tau')K(t',\tau) \Big)/\mathrm{ND} \\
&\langle \mathrm{Im}\Delta M_{rg1} \mathrm{Re}\Delta M_{rg2} \rangle = \\
&\quad \sum_{(t,\tau)\in rg1} \sum_{(t',\tau')\in rg2} \frac{1}{2}\mathrm{Im}\Big(K(t,t')K(\tau',\tau) + K(t,\tau')K(t',\tau) \Big)/\mathrm{ND}.
\end{aligned} \tag{3.32}$$

If the time lag for every pair in $rg1$ is the same and similarly for $rg2$ (the lag in $rg1$ may be different from the lag in $rg2$), and if the assumptions leading to formula (3.21) are true, we can generalize (3.21) to

$$\begin{aligned}
&\langle \Delta\mathrm{Mo}M_{rg1} \Delta\mathrm{Mo}M_{rg2} \rangle = \\
&\quad \sum_{(t,\tau)\in rg1} \sum_{(t',\tau')\in rg2} \frac{1}{2}\mathrm{Re}\Big(K_0(t,t')K_0(\tau',\tau) + K_0(t,\tau')K_0(t',\tau) \Big)/\mathrm{ND} \\
&\langle \Delta\mathrm{Ph}M_{rg1} \Delta\mathrm{Ph}M_{rg2} \rangle \cdot \langle \mathrm{Mo}M_{rg1} \rangle \langle \mathrm{Mo}M_{rg2} \rangle = \\
&\quad \sum_{(t,\tau)\in rg1} \sum_{(t',\tau')\in rg2} \frac{1}{2}\mathrm{Re}\Big(K_0(t,t')K_0(\tau',\tau) - K_0(t,\tau')K_0(t',\tau) \Big)/\mathrm{ND} \\
&\langle \Delta\mathrm{Mo}M_{rg1} \Delta\mathrm{Ph}M_{rg2} \rangle = 0.
\end{aligned} \tag{3.33}$$

3.6 Practical Estimation of the Variances

The formulae (3.18) or (3.21) can be used to find the covariances of the fluctuations in the ACF estimates. For exact results, the functions $K(t,t')$ or $K_0(t,t')$ need to be known. Since the values of these theoretical expressions are not known, we must use the experimental estimates instead. Thus, when calculating experimental errors from the formulae (3.18) and (3.21), we can use the following approximations:

$$K(t,t') \approx M(t,t') \quad \text{and} \quad K_0(t,t') \approx \mathrm{Mo}M(t,t'). \tag{3.34}$$

In the design of the experiments one should take into account the fact that some parts of the $M(t,t')$ matrix may be needed in the calculation of variances that are not of interest in the estimation of the plasma parameters themselves. This is important in multiple pulse experiments, where the experimenter is perhaps not interested in the zero lag profile - that is the values of $M(t,t)$ - because of the complicated structure of the corresponding ambiguity function $W_{t,t'}$. However, if one wants to get reliable variance estimates, it is extremely important to measure the zero lag profile also.

Some points of $M(t,t')$ are never estimated, because it is known in advance from the theory that the average will be zero regardless of the plasma autocorrelation function values. This occurs for example if $t - t'$ is longer than the pulse sent plus receiver filter response time. If these time pairs occur at the right hand sides of (3.18) or (3.21), one can substitute zeroes for them.

Reliable estimation of the variances of the range-gate estimates by the formulae (3.32) and (3.33) seems to be more difficult than in the non-range-gated case, since the individual lagged product estimates $M(t,t')$ are lost in the range-gating process.

However, to justify range-gating there is an underlying assumption that the different $M(t,t')$ values summed in a range-gate are all approximately equal. If this is valid, we may get approximations for $K(t,t')$ by dividing a range-gate estimate M_{rg} for the lag $t - t'$ by the number of summands in that estimate. More specifically, let us suppose that we are interested in the covariances of the range-gate estimates

$$M_{rg(r,\tau)} \quad \text{and} \quad M_{rg(r,\tau')}, \tag{3.35}$$

both representing the same distance r, but different lags τ and τ'. In using formula (3.32) or (3.33), we could make the approximations

$$\begin{aligned} K(t,t') &\approx M_{rg(r,t-t')}/\mathrm{NOP}(rg(r,t-t')) \text{ or} \\ K_0(t,t') &\approx \mathrm{Mo}M_{rg(r,t-t')}/\mathrm{NOP}(rg(r,t-t')), \end{aligned} \tag{3.36}$$

where NOP(rg) means the number of products in a range-gate, or more exactly the number of time pairs that form the range-gate.

3.7 Signal to Noise Ratio

In the literature, variances of the lagged product estimates are often given in terms of the *signal-to-noise* ratio of the measurement. The signal-to-noise ratio SNR is defined by

$$\mathrm{SNR}(t) = k(t,t)/k_n(t,t). \tag{3.37}$$

Thus, the signal-to-noise ratio is the ratio of the signal and noise powers. As the backscattered signal is not a stationary process, the signal-to-noise ratio may be different depending on which samples one is interested in. The SNR depends both on the transmitted pulse form and the receiver filter response.

Let us now consider a situation where the sampling times t and t' are far enough apart so that the noise signals are independent of each other:

$$k_n(t,t') = 0. \tag{3.38}$$

The covariance matrix for the total signal $z + z_n$ can now be written

$$K(t,t') = k(t,t') + \mathrm{SNR}^{-1}(t)k(t,t)\delta(t,t'). \tag{3.39}$$

If SNR $\ll 1$, we can forget the first term in (3.39), and the equations (3.18) can be simplified to

$$
\begin{aligned}
\langle \mathrm{Re}\Delta M(t,\tau)\mathrm{Re}\Delta M(t,\tau)\rangle &\approx \frac{1}{2}\mathrm{SNR}^{-1}(t)\mathrm{SNR}^{-1}(\tau)k(t,t)k(\tau,\tau)/\mathrm{ND}, t \neq \tau \\
\langle \mathrm{Re}\Delta M(t,t)\mathrm{Re}\Delta M(t,t)\rangle &\approx \mathrm{SNR}^{-2}(t)k^2(t,t)/\mathrm{ND} \\
\langle \mathrm{Im}\Delta M(t,\tau)\mathrm{Im}\Delta M(t,\tau)\rangle &\approx \frac{1}{2}\mathrm{SNR}^{-1}(t)\mathrm{SNR}^{-1}(\tau)k(t,t)k(\tau,\tau)/\mathrm{ND}, t \neq \tau \\
\langle \mathrm{Im}\Delta M(t,t)\mathrm{Im}\Delta M(t,t)\rangle &= 0 \\
\langle \mathrm{Re}\Delta M(t,\tau)\mathrm{Im}\Delta M(t',\tau')\rangle &\approx \langle \mathrm{Im}\Delta M(t,\tau)\mathrm{Re}\Delta M(t',\tau')\rangle \approx 0.
\end{aligned}
\tag{3.40}
$$

The errors of the estimates belonging to different time pairs (t,τ) and (t',τ') do not correlate with each other.

Thus, we see that in this case, the variances of both real and imaginary parts are all equal and independent of each other, except for the zero lag (real) estimate, whose variance is twice that of the other estimates. This fact can sometimes (when the ACF is known to be wide) be used to reduce measurement variances, if one can measure a non-zero lag estimate instead of the zero lag estimate.

4. SPEED OF MEASUREMENTS

4.1 Introduction

In this chapter, the classical methods of coding are compared with each other. We suppose here that the freedom of choosing the codes is constrained by a maximum allowable modulation time and we suppose that this time is long enough to satisfy the lag extent requirement.

We suppose also that the resolution of the classical methods is determined by the basic pulse length and thus by the width of the ambiguity function in both directions. In other words, we suppose that no effort is made to utilize advanced techniques to improve the resolution, which conforms well with the classical way of utilizing the data. This will simplify our analysis, as we may then forget that the effective plasma cross section $\sigma_{\text{eff}}(\tau; S)$ is a function of both range and lag.

We also suppose that the SNR is poor, so that the variance of the cross product estimates is mainly determined by the white noise part in the formulae giving the variances as products of the averages of the crossed products themselves. If this is not true, the interpretations may turn out to be totally different. We restrict our study here to the monostatic case.

4.2 The Definition of the Speed of a Measurement

We have shown that the averages of the signal part of the crossed product estimates are given by

$$\langle m(t,t')\rangle = \langle z(t)\overline{z(t')}\rangle/R = P_0 V_{t,t'}^{\text{eff}} \sigma_{\text{eff}}\,(t-t';\vec{x_0})\,, \tag{4.1}$$

where the effective scattering volume is given by

$$V_{t,t'}^{\text{eff}} = A_{beam}^{\text{eff}} \frac{c}{2} l_{t,t'} = C_{beam} \frac{4\pi R_0^2}{G(0)} \frac{c}{2} l_{t,t'}\,. \tag{4.2}$$

For simplicity, we shall be interested in the variances of the real part of the crossed products only. Moreover, we suppose that the background m_n can be estimated exactly. By the results of chapter 3 we can write

$$\langle \mathrm{Re}\Delta M(t,t')\mathrm{Re}\Delta M(t,t')\rangle = \frac{1}{2}\mathrm{Re}\Big(K(t,t)K(t',t') + K(t,t')K(t,t')\Big)/\mathrm{ND}. \quad (4.3)$$

As $K(t,\tau) = k(t,\tau) + k_n(t,\tau)$, and as we have supposed that the white noise part is dominant in the error formulae, and as $k_n(t,\tau) = \kappa T A_p(t-\tau)$, we may write

$$\langle \mathrm{Re}\Delta M(t,t')\mathrm{Re}\Delta M(t,t')\rangle = \frac{1}{2}(\kappa T)^2\Big(A_p^2(0) + A_p^2(t-t')\Big)/\mathrm{ND}. \quad (4.4)$$

Let us choose a constant representative value of the electron density and denote it by n_e. It will be useful to use the dimensionless variable $\sigma_{\mathrm{eff}}/n_e$ as the unknown instead of σ_{eff}. Because we know the linear relationship of the unknown variable $\sigma_{\mathrm{eff}}/n_e$ to the measurement $\mathrm{Re}\Delta M(t,t')$ and because we know the variance of that measurement, we can write the Fisher information for (inverse of the variance of) $\sigma_{\mathrm{eff}}/n_e$ contained in that measurement as

$$\begin{aligned} Q_{t,t'} &= \frac{2\mathrm{ND}\Big(P_0 V_{t,t'}^{\mathrm{eff}} n_e\Big)^2}{(\kappa T)^2\Big(A_p^2(0) + A_p^2(t-t')\Big)} \\ &= \mathrm{ND}\frac{\Big(P_0 A_{beam}^{\mathrm{eff}} \frac{c}{2} n_e\Big)^2}{(\kappa T)^2}\,\frac{2l_{t,t'}^2}{A_p^2(0) + A_p^2(t-t')}. \end{aligned} \quad (4.5)$$

In the latter expression, only the last factor depends on the envelope of the transmission or the receiver filter impulse response p. The other factors contain all the other elements affecting the situation. Because of this, we will call the last factor the *speed* of the code for lag $t-t'$, and use the notation

$$S_{t,t'} = S_{t-t'} = \frac{2l_{t,t'}^2}{A_p^2(0) + A_p^2(t-t')}. \quad (4.6)$$

For $t = t'$ this becomes $S_{t,t'} = l_{t,t'}^2/A_p^2(0)$.

4.3 A Power Profile Measurement with a Single Pulse

Let us suppose that we send a radar pulse of constant amplitude and phase and length τ_0. Let us suppose, moreover, that the receiver impulse response is a boxcar of length τ_0, that is, let

$$\begin{aligned} \mathrm{env}_0(t) &= \begin{cases} 0, & \text{if } t < 0 \text{ or } t > \tau_0 \\ 1, & \text{if } 0 \le t \le \tau_0 \text{ and} \end{cases} \\ p(t) &= \mathrm{env}_0(t)/\tau_0. \end{aligned} \tag{4.7}$$

Then $l_{t,t} = (2/3)\tau_0$ and $A_p(0) = \tau_0^{-1}$, resulting in $S_{t,t} = (4/9)\tau_0^4$.

If the total modulation time available is τ_{tot}, it is possible to use τ_{tot}/τ_0 different frequency channels to send independent pulses of length τ_0. As the Fisher information from a set of independent measurements is the sum of the Fisher informations from each measurement separately, the total speed from this kind of arrangement will be $S = (\tau_{tot}/\tau_0)(4/9)\tau_0^4$.

4.4 A Phase Code

If env_0 is defined as a boxcar function of length τ_0 as above, an arbitrary phase code envelope of bit length n_B can then be defined by the formula

$$\mathrm{env}(t) = \sum_{i=0}^{n_B-1} a_i \mathrm{env}_0(t - i\tau_0), \tag{4.8}$$

where the a_i are complex numbers of unit modulus with the property that

$$\left| \sum_{i=0}^{n_B-j-1} a_i a_{i+j} \right| \le 1 \quad \text{for every} \quad 1 \le j < n_B. \tag{4.9}$$

That is, a phase code is a series of elementary boxcar pulses with possibly different phases. The receiver impulse responses used in conjunction with phase codes are ideally reversions of the phase code itself: $p(t) = \mathrm{env}(-t)/\tau_0$.

The simplest phase codes are Barker codes, of which we mention as examples the five bit Barker code with $(a_i) = (1, 1, 1, -1, 1)$ and the 13-bit Barker code with $(a_i) = (1, 1, 1, 1, 1, -1, -1, 1, 1, -1, 1, -1, 1)$.

It follows from the condition (4.9) and from the fact that $p(t) = \mathrm{env}(-t)/\tau_0$ that

$$(p*\mathrm{env})(t) = A_{\mathrm{env}}(t)/\tau_0 = n_B A_{\mathrm{env}_0}(t)/\tau_0 + sl(t), \tag{4.10}$$

where sl is a function called the sidelobes with a modulus not bigger than 1. The function $A_{\mathrm{env}_0}/\tau_0$ is a triangle of width $2\tau_0$ with a peak of height 1 at $t = 0$, and sloping linearly to 0 for $t = \pm\tau_0$. As $W_{t,t}(S) = (p*\mathrm{env})(t-S)(p*\mathrm{env})(t-S)$, we see that the range ambiguity function has a sharp peak of height n_B^2 and width $2\tau_0$ at t and some sidelobes (the square of sl) extending further away in the range direction. Usually the response from these sidelobes is ignored, and one is only interested in the response from the sharp centre peak.

If the sidelobes are ignored the speed of a n_B-bit phase code is given by $S_{t,t} = n_B^2(4/9)\tau_0^4$. If the total modulation time available is again τ_{tot} and if it is filled altogether with a phase code so that $n_B = \tau_{tot}/\tau_0$, the resulting speed will be $S = (\tau_{tot}/\tau_0)^2(4/9)\tau_0^4$.

4.5 Modulation Period Filled with Pulse Codes

Let us suppose that the elementary pulse length is τ_0 and the receiver impulse response is a boxcar with the same length. Then an ordinary pulse code is a sequence of elementary pulses with separations between the elementary pulses of even multiples of τ_0. The pulse sequence must satisfy the condition that the separation between any pair of the constituent elementary pulses must not be equal to the separation of any other pair.

It follows from these conditions that the length of a pulse code with n subpulses must be at least

$$1 + \sum_{i=1}^{n-1} 2i = 1 + 2(n-1)(n/2) \approx n^2 \tag{4.11}$$

units ($= \tau_0$). As the total modulation length is $N = \tau_{tot}/\tau_0$ units, the longest pulse code that can be used can have at most $\sqrt{N}$ subpulses, and consequently it will be possible to commute at most $\sqrt{N}$ of these pulse codes to the independent frequency channels in the modulation.

As a multipulse code of length $\sqrt{N}$ gives $(1/2)\sqrt{N}(\sqrt{N}-1) \approx N/2$ lagged product estimates (this is the number of different pairs of the pulses), and as the speed of each of these estimates is $S_{t,t'} = (4/9)\tau_0^4 \cdot 2$, the total speed of all the lagged product estimates available will be $S = \sqrt{N} \cdot N/2 \cdot (4/9)\tau_0^2 \cdot 2 = (\tau_{tot}/\tau_0)^{3/2}(4/9)\tau_0^4$.

4.6 Phase Coded Multipulses

The subpulses forming a multipulse group may be phase coded themselves. Let us again denote the basic pulse length by τ_0 and the total modulation time available by τ_{tot}. Moreover, let $N = \tau_{tot}/\tau_0$. Let us suppose that the phase modulation length is n_B. Then the total modulation length is N/n_B phase code sequences long, and by a reasoning similar to that above, at most $(N/n_B)^{3/2}\sqrt{2}$ lagged product estimates will be available from the multipulse groups at the possible different frequency channels. (Notice here that the separations between different phase code groups in a phase coded multipulse need not be even multiples of the pulse code length; traditionally they are integer multiples of it.)

As the speed of each of the lagged product estimates is $S_{t,t'} = n_B^2(4/9)\tau_0^4 \cdot 2$, the total speed will be $S = (\tau_{tot}/\tau_0)^{3/2}\sqrt{2n_B}(4/9)\tau_0^4$.

4.7 Comparison of the Different Methods

Let us summarize the speeds resulting from different methods of measuring σ_{eff} with a modulation whose total length is τ_{tot} and whose range resolution (=basic pulse length) is τ_0.

1) Filling the modulation time with independent single pulses results in $S = (\tau_{tot}/\tau_0)(4/9)\tau_0^4$.

2) Filling the modulation time with a full length phase code results in $S = (\tau_{tot}/\tau_0)^2(4/9)\tau_0^4$.

3) Filling the modulation time with pulse codes results in $S = (\tau_{tot}/\tau_0)^{3/2}(4/9)\tau_0^4$.

4) Filling the modulation period with phase coded multipulses with phase code length n_B will result in total speed $S = (\tau_{tot}/\tau_0)^{3/2}\sqrt{2n_B}(4/9)\tau_0^4$.

We see that the most effective measurement is achieved by phase coding. Moreover, we see here that when using this method, the speed of the measurement is proportional to the square of the basic pulse length (=spatial resolution).

When deriving these results we have supposed that the effective plasma correlation function $\sigma_{\text{eff}}(\tau, S)$ is constant in the τ direction for all lags τ not longer than the total modulation length. This assumption can of course not always be made, and in these cases, only a smaller set of lagged products with lags in a narrow enough interval may be used to form estimates for $\sigma_{\text{eff}}(\tau, S)$. Moreover, phase codes cannot be used in the situations where σ_{eff} can change significantly should the τ variable change by an amount equal to the phase code total length.

Let us suppose that the lag resolution is τ_{lag}. That is, we can suppose that $\sigma_{\text{eff}}(\tau; S)$ has a constant value σ_i for $\tau \in ((i-1/2)\tau_{lag}, (i+1/2)\tau_{lag})$ for any integer i. Phase coded multipulses with phase length $n_B = \tau_{lag}/\tau_0$ can then be used, but in this case only a fraction τ_{lag}/τ_{tot} of all the available crossed products can be used to estimate the variable σ_i. The resulting speed will thus be

$$S = \left(\frac{\tau_{lag}}{\tau_{tot}}\right)\left(\frac{\tau_{tot}}{\tau_0}\right)^{3/2}\left(\frac{\tau_{lag}}{\tau_0}\right)^{1/2}\sqrt{2}\,\frac{4}{9}\,\tau_0^4\,. \tag{4.12}$$

If the problem were solved by sending non-phase-coded pulse codes and by using all the $\tau \in ((i-1/2)\tau_{lag}, (i+1/2)\tau_{lag})$ to estimate σ_i, the speed would be

$$S = \left(\frac{\tau_{lag}}{\tau_{tot}}\right)\left(\frac{\tau_{tot}}{\tau_0}\right)^{3/2}\frac{4}{9}\,\tau_0^4\,, \tag{4.13}$$

which is a speed $\sqrt{2n_B}$ times lower than in the case where phase coded multipulses are used.

Thus, we may draw the conclusion that in the case where $\tau_0 < \tau_{lag} < \tau_{tot}$, it is best to use phase coded multipulses, where the phase code length is chosen equal to the lag resolution τ_{lag}. Moreover, if we change the spatial resolution τ_0, keeping τ_{tot} and τ_{lag} constant, the speed will in this case be proportional to τ_0^2.

4.8 Filter Effects in Multipulse Measurements

The concept of speed can be used to compare different filtering possibilities of the received signal. In this chapter, we study the case of a multipulse measurement composed of boxcar shaped elementary pulses of length τ_0 and spaced apart at intervals that are integer multiples of $2\tau_0$. By formulae (2.12) and (2.17), it is clear that ambiguities appear in the weighting functions, if the receiver filter impulse response time is longer than τ_0.

Real filters have infinite response times, so that some ambiguities are always present. To choose a suitable filter response time is thus a compromise between the magnitude of the allowable ambiguities on one hand and between the signal power compared to the resulting noise on the other hand. That is, the filter width should be chosen so that one will get the best possible measurement speed, while keeping the ambiguities within allowable limits.

Let us denote by env_0 the envelope of an elementary boxcar pulse of length τ_0 as in (4.7), so that the multipulse code for N pulses can be written as

$$\mathrm{env}(t) = \sum_{i=1}^{N} \mathrm{env}_0(t - 2a_i\tau_0)\,, \tag{4.14}$$

where the a_i are suitably chosen integers (for a four-pulse code they could be $(a_i) = (0, 1, 4, 6)$). The speed for a lag $t - t' = 2a_i\tau_0$ is given by

$$S_{t,t'} = \frac{2l_{t,t'}^2}{A_p^2(0) + A_p^2(t - t')}. \tag{4.15}$$

The effective length of the pulse code $l_{t,t'}$ is, if the ambiguous contributions are omitted, approximately given by

$$l_{t,t'} = \int_{-\tau_0}^{\tau_0} (p{*}\mathrm{env}_0)^2(t)\, dt\,. \tag{4.16}$$

The ambiguities resulting from the fact that the receiver response time exceeds τ_0 are approximately given by

$$l_{t,t'}^{amb} = \int_{-\infty}^{\infty} |(p{*}\mathrm{env}_0)(t)\,(p{*}\mathrm{env}_0)(t - 2\tau_0)|\, dt\,. \tag{4.17}$$

The relative ambiguity may be defined as the ratio of the ambiguous part of the weighting function to the main part of the weighting function. These ambiguities may appear at several different heights, and there may be approximately the same number of ambiguous heights as there are pulses in the pulse code. For example, there arise five ambiguous contributions in the first lag of a four-pulse code.

The relative ambiguities and speeds for some generally used filters are given in the following table:

$(2\tau_0 BW_p)^{-1}$	Butterworth		linear phase		boxcar	
	speed	$\frac{l^{amb}}{l}$	speed	$\frac{l^{amb}}{l}$	speed	$\frac{l^{amb}}{l}$
0.30	0.106	0.001				
0.40	0.170	0.005				
0.50	0.253	0.018				
0.60	0.347	0.033	0.195	0.000	0.230	0.000
0.80	0.539	0.065	0.294	0.000	0.344	0.000
1.00	0.696	0.075	0.380	0.000	0.444	0.000
1.20	0.814	0.082	0.446	0.002	0.521	0.002
1.40	0.899	0.096	0.498	0.012	0.580	0.010
1.60			0.541	0.033	0.627	0.028

Table 4.1 Speeds and relative ambiguities for different filter and response time combinations in a multipulse experiment.

It is clear from the table that if a maximum relative ambiguity requirement of *e.g.* 0.01 (corresponding to total relative ambiguity area of 0.05 in the four-pulse code) is posed, the bandwidth of a Butterworth filter would have to be chosen so that its speed would be less than half of the speed of a corresponding linear phase or boxcar filter. This is due to the long ringing times of the Butterworth impulse response functions.

The Butterworth and linear phase filter responses used in the calculations were obtained by digitizing the impulse responses of the filters used in the EISCAT receivers, read from an oscilloscope screen. The linear phase filters used in EISCAT are by design the linear phase with equiripple error 0.5° filters of fifth degree and the Butterworth filters are the fifth degree Butterworth filters as specified in Williams (1981).

One should note that, in the above table, the convention of bandwidth definition does not agree with the theoretical definition (2.24) for Butterworth filters. Rather, for Butterworth filters, the bandwidth is commonly defined as the half-power width of the frequency response curve. The bandwidths concerning the linear phase filters and boxcar filters are in accordance with (2.24), the bandwidth of a boxcar filter being half the inverse of its length, so that a boxcar response 20μs long has a bandwidth of 25kHz, for example. These values correspond to the value 1.00 in the first column of the table.

5. A THEORETICAL LIMIT TO THE ACCURACY OF MEASUREMENTS

5.1 Statement of the Most General Estimation Problem and its Optimization

Let us model the function $\sigma_{\text{eff}}(\tau; S)$ as a spline expansion in the $\tau - S$ plane. That is, let us suppose that σ_{eff} can be represented as

$$\sigma_{\text{eff}}(\tau; S) = \sum_{i=1}^{N} \sigma_i s_i(\tau; S), \tag{5.1}$$

where the s_i are a suitable spline basis in the $\tau - S$ plane. We may use cubic base splines in the τ-direction to take into account the fact that the plasma autocorrelation functions are smooth in the lag variable, and linear interpolation in the S direction to allow a ragged structure in the range direction, for example.

The relationship of the expectation of a measured lagged product $m(t, t')$ to the spline coefficients σ_i is given by

$$\begin{aligned}
\langle m(t,t')\rangle &= \langle z(t)\overline{z(t')}\rangle / R \\
&= P_0(0) A_{beam}^{\text{eff}} \frac{c}{2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\tau dS\, (S_0/S)^2 W_{t,t'}(\tau; S)\sigma_{\text{eff}}(\tau; S) \\
&= P_0(0) A_{beam}^{\text{eff}} n_e \frac{c}{2} \sum_{i=1}^{N} \frac{\sigma_i}{n_e} l_{t,t'}^i \quad \text{with} \\
l_{t,t'}^i &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\tau dS\, (S_0/S)^2 W_{t,t'}(\tau; S) s_i(\tau; S)
\end{aligned} \tag{5.2}$$

Here R_0 is some characteristic (fixed) range corresponding to the measurement and P_0 is the corresponding single electron scattering power. S_0 is the corresponding travel time of the scattered signal. The variation of the single electron scattering power and the effective beam cross section as a function of

range is here taken into account by the factor $(S_0/S)^2$ in the integral. This is a convenient way to handle this variation as a dimensionless form factor. The beam effective cross section is denoted by A^{eff}_{beam}. A representative constant value of electron density is denoted by n_e. It is introduced to make the spline coefficients σ_i/n_e dimensionless.

The general estimation problem for the spline coefficients σ_i/n_e is thus a linear inverse problem whose direct theory is given by equation (5.2). The measurements are sums of $m_{t,t'}$ with $(t,t')\in rg$ for all range-gates rg available. For each channel, the vector of the estimated crossed products is denoted by m_{ch}. Due to integration, the errors in each channel used are additive Gaussian errors with a covariance matrix Γ_{ch} that can be estimated by the methods given in chapter 3. The summation in (1.11) and (1.12) corresponds to summation over different frequency channels used and the matrix A_i in chapter 1 corresponds here to A_{ch}. The elements of this matrix can be indexed as $(A_{ch})_{rg,i}$ with

$$(A_{ch})_{rg,i} = P_0 A^{\text{eff}}_{beam} n_e \frac{c}{2} \sum_{(t,t')\in rg} l^i_{t,t'} = P_0(0) A^{\text{eff}}_{beam} n_e \frac{c}{2} l^i_{rg}. \tag{5.3}$$

The index i in (5.3) should not be confused with the index i appearing in (1.11) and (1.12). The spline coefficients can then be solved by the formula (1.13) with the a posteriori covariance Q^{-1} given by (1.11), if the summation index i is replaced by the channel number ch. This leads to the solution

$$\begin{aligned} Q &= \sum_{ch} A^T_{ch} \Gamma^{-1}_{ch} A_{ch} \\ (Q)_{ij} &= \sum_{ch} \sum_{rg} \sum_{rg'} (A_{ch})_{rg,i} (A_{ch})_{rg',j} (\Gamma_{ch})_{rg,rg'}. \end{aligned} \tag{5.4}$$

The best estimate for the $\overline{\sigma_i}/n_e$ is then given by

$$\overline{\sigma}/n_e = Q^{-1} \sum_{ch} A^T_{ch} \Gamma^{-1}_{ch} \overline{m}_{ch}, \tag{5.5}$$

or componentwise,

$$\overline{\sigma_i}/n_e = \sum_j (Q^{-1})_{ij} \sum_{ch} \sum_{rg} \sum_{rg'} (A_{ch})_{rg,j} (\Gamma^{-1}_{ch})_{rg,rg'} (\overline{m}_{ch})_{rg'}. \tag{5.6}$$

The optimization problem can be stated in its most general form as the problem of minimizing the posteriori covariance Q^{-1} within the system constraints.

5.2 Theoretical Accuracy Limits

It is useful to know upper limits to the Fisher information matrix Q of the spline coefficients σ_i/n_e within some basic technical constraints, such as a maximum radar power P_0 and a maximum modulation time τ_{tot}. In this way, different methods of measurement may be evaluated so that the experimenter knows whether the best possible methods are being used.

Let us suppose that we use an arbitrary modulation pattern of length τ_{tot} with a modulus not exceeding 1. Moreover, let us suppose that the signal is sampled with a normalized receiver boxcar impulse response with length $\Delta\tau$ at time intervals $\Delta\tau$, and that all crossed products are calculated. When $\Delta\tau \to 0$, this corresponds to sampling the signal at infinite resolution, so that all information is preserved, and it is guaranteed that no information is lost in the receiver filters.

Let us introduce the notations

$$m_{i,j} = m(i\Delta\tau, j\Delta\tau) \quad \text{and} \quad K_{i,j} = K(i\Delta\tau, j\Delta\tau), \tag{5.7}$$

and similarly for $k_n, \Delta m$ and other variants. The variances of the different estimates are given by (provided that the $m_n(t,t')$ values can be estimated exactly)

$$\langle \mathrm{Re}\Delta m_{i,i+\Delta i} \mathrm{Re}\Delta m_{j,j+\Delta j} \rangle = \frac{1}{2}\mathrm{Re}\Big(K_{i,j}K_{i+\Delta i,j+\Delta j} + K_{i,j+\Delta j}K_{i+\Delta i,j} \Big)/\mathrm{ND}. \tag{5.8}$$

Here $K_{i,j}$ is given by $K_{i,j} = k_{i,j} + (k_n)_{i,j}$, and if we suppose that the SNR is poor, we may use the approximation $K_{i,j} \approx (k_n)_{i,j} = \kappa T \delta_{i,j}/\Delta\tau$. It follows that

$$\langle \mathrm{Re}\Delta m_{i,i+\Delta i} \mathrm{Re}\Delta m_{j,j+\Delta j} \rangle = \frac{1}{2}(1+\delta_{\Delta i,0})\delta_{i,j}\delta_{\Delta i,\Delta j}(\kappa T)^2/(\Delta\tau^2 \mathrm{ND}). \tag{5.9}$$

The different measurements are thus noncorrelated, and we can form the Fisher information matrix Q_{ij} for the spline coefficients σ_i/n_e as the sum of the Fisher information matrices for the measurements $m_{i,i+\Delta i}$ separately.

The Fisher information matrix $Q^{i,i+\Delta i}$ for the measurement $m_{i,i+\Delta i}$ is given by

$$Q^{i,i+\Delta i}_{m\,n} = \Big(P_0(0) A^{\mathrm{eff}}_{beam} n_e \frac{c}{2} \Big)^2 \frac{2 l^m_{i,i+\Delta i} l^n_{i,i+\Delta i}}{(1+\delta_{\Delta i,0})(\kappa T)^2/(\Delta\tau^2 \mathrm{ND})} \tag{5.10}$$

The total information matrix is given by summing the above expression over i and Δi:

$$Q_{m\,n} = \left(P_0(0) A_{beam}^{\text{eff}} n_e \frac{c}{2}\right)^2 \text{ND}/(\kappa T)^2 \sum_{i,\Delta i=0}^{\infty} \frac{2 l_{i,i+\Delta i}^m l_{i,i+\Delta i}^n}{(1+\delta_{\Delta i,0})/\Delta\tau^2}. \tag{5.11}$$

To simplify the calculations, we suppose furthermore that the spline expansion used is a step function expansion of lag resolution τ_{lag} and spatial resolution τ_0. Hence we suppose that each of the spline basis functions has the form

$$s_m(\tau; S) = \begin{cases} 1, \text{ for } \tau_m \leq \tau \leq \tau_m + \tau_{lag} \text{ and } S_m \leq S \leq S_m + \tau_0 \\ 0, \text{ otherwise.} \end{cases} \tag{5.12}$$

In addition, we suppose that the range factor S_0/S is constant and equal to one over the range of interest. The spatial resolution τ_0 as defined here differs from the τ_0 in chapter 4 in the sense that in chapter 4 it was the length of an elementary pulse. Here we allow the modulation to be arbitrary, but we have fixed the details in the model fitted to be of size τ_0 in the range direction.

Now it becomes possible to estimate the diagonal components of the total Fisher information matrix in the limit $\Delta\tau \to 0$. If the modulation length is τ_{tot} and the modulus of the modulation envelope is at most 1, it is easy to see that the two-dimensional ambiguity function $W_{i\Delta\tau,(i+\Delta i)\Delta\tau}(\tau; S)$ must have its support in the vertical strip $\tau \in (\Delta i \Delta\tau - \Delta\tau, \Delta i \Delta\tau + \Delta\tau)$. Moreover, the support is limited to the area $S \in ((i+\Delta i - 1)\Delta\tau - \tau_{tot}, i\Delta\tau)$. In addition, the behaviour of the ambiguity function must be such that

$$\int_{-\infty}^{\infty} |W_{i\Delta\tau,(i+\Delta i)\Delta\tau}(\tau; S)| d\tau \leq 1. \tag{5.13}$$

Let us now estimate the sum in (5.11) for $Q_{m,m}$. Because of the condition $W_{i\Delta\tau,(i+\Delta i)\Delta\tau}(\tau; S) = 0$ for $\tau \notin (\Delta i \Delta\tau - \Delta\tau, \Delta i \Delta\tau + \Delta\tau)$, there can be at most $\tau_{lag}/\Delta\tau + 2$ terms contributing in the summation for Δi. For each fixed Δi, the sum over i can also be estimated. The length of the support of $W_{i\Delta\tau,(i+\Delta i)\Delta\tau}(\tau; S)$ in the S direction is $\tau_{tot} - (\Delta i - 1)\Delta\tau$, and as the support of s_m in the S direction is of length τ_0, these supports totally overlap for $(\tau_{tot} - (\Delta i - 1)\Delta\tau - \tau_0)/\Delta\tau$ values for i and partially overlap for $2\tau_0/\Delta\tau$ values for i. The integral $l_{t,t'}^m$ evaluates at most to τ_0 for the case of total overlap, and the average value of its square for the partially overlapping cases is at most

$\tau_0^2/3$. Summing together all the i and Δi values we see that the total Fisher information matrix diagonal components must satisfy the condition

$$\begin{aligned}\frac{Q_{m\,m}}{\left(P_0(0)A_{beam}^{\text{eff}}n_e\frac{c}{2}\right)^2\text{ND}/(\kappa T)^2} &= \sum_{i,\Delta i=0}^{\infty}\frac{2l_{i,i+\Delta i}^m l_{i,i+\Delta i}^m}{(1+\sigma_{\Delta i,0})/\Delta\tau^2}\\ &\leq\left(\frac{\tau_{lag}}{\Delta\tau}+2\right)\left(\left(\tau_{tot}-\tau_m-\tau_{lag}/2-\tau_0\right)\frac{\tau_0^2}{\Delta\tau}+\frac{2\tau_0}{\Delta\tau}\frac{1}{3}\tau_0^2\right)2\Delta\tau^2 \qquad (5.14)\\ &\to 2\tau_{lag}\left(\tau_{tot}-\tau_m-\frac{\tau_{lag}}{2}-\frac{\tau_0}{3}\right)\quad\text{as}\quad\Delta\tau\to 0.\end{aligned}$$

In the interesting cases of relatively high resolutions, $\tau_0 \ll \tau_{tot}$ and $\tau_{lag} \ll \tau_{tot}$, and we may use a simplified result

$$Q_{mm}\leq\left(P_0(0)A_{beam}^{\text{eff}}n_e\frac{c}{2}\right)^2\frac{\text{ND}}{(\kappa T)^2}\cdot 2\tau_{lag}(\tau_{tot}-\tau_m)\tau_0^2. \qquad (5.15)$$

As the variance of $\overline{\sigma_m}/n_e$ is given by $(Q^{-1})_{mm}$, and as it is a general result for positive definite symmetric matrices that $(Q^{-1})_{mm}\geq(Q_{mm})^{-1}$, we can state that for any kind of modulation used with a maximum radar power P_T and maximum modulation time τ_{tot}, the posteriori variance for $\sigma_{\text{eff}}(\tau_m, S)$ with spatial resolution τ_0 and lag resolution τ_{lag} must be larger than a limit given by

$$\text{var}\,(\sigma_{\text{eff}}/n_e)\geq Q_{mm}^{-1}. \qquad (5.16)$$

In terms of the last chapter, we may also state that the speed for a measurement with the above specified resolutions will always be limited by

$$S\leq S_{max}=2\tau_{lag}\left(\tau_{tot}-\tau_m\right)\tau_0^2. \qquad (5.17)$$

We have limited us to the case of using only one frequency channel in the discussion above. If one uses many different frequency channels commuted in the specified maximum modulation time, it can be proven that the above inequalities for the Fisher information matrix apply as well to the sum of all the Fisher information matrices for the channels separately. The reason for this is that the ambiguity functions for different channels cannot overlap, because the different frequencies cannot appear simultaneously in the transmission. The limit is thus true for frequency-commuted modulations also.

5.3 Efficiencies of Different Methods

It is interesting to compare this maximum speed with the classical methods of measurement. In the case of a phase coded multipulse with the above stated resolutions, the speed was found to be (4.12)

$$S = \left(\frac{\tau_{lag}}{\tau_{tot}}\right)\left(\frac{\tau_{tot}}{\tau_0}\right)^{3/2}\left(\frac{\tau_{lag}}{\tau_0}\right)^{1/2}\sqrt{2}\frac{4}{9}\tau_0^4 \tag{5.18}$$

If the efficiency of a measurement is defined as the ratio of its speed and the maximum speed S_{max}, we find that the efficiency of phase coded multipulses is given by

$$\text{eff} = \sqrt{2}\frac{2}{9}\left(\frac{\tau_{tot}}{\tau_{tot}-\tau_m}\right)\left(\frac{\tau_{lag}}{\tau_{tot}}\right)^{1/2}. \tag{5.19}$$

Similarly (4.13), the efficiency of a non-phase coded multipulse is at most

$$\text{eff} = \frac{2}{9}\left(\frac{\tau_{tot}}{\tau_{tot}-\tau_m}\right)\left(\frac{\tau_{lag}}{\tau_{tot}}\right)^{1/2}\left(\frac{\tau_0}{\tau_{lag}}\right)^{1/2}. \tag{5.20}$$

In the cases where one can use phase codes that are as long as the total modulation time, the lag resolution is equal to τ_{tot}. As τ_{lag} is not small, we have to use (5.14) directly instead of (5.17) leading to a maximum speed

$$S_{\max} = \tau_{lag}\tau_{tot}\tau_0^2, \tag{5.21}$$

and as the speed of a phase code was found to be $S = (\tau_{tot}/\tau_0)^2(4/9)\tau_0^4$, the efficiency of a phase code is given by

$$\text{eff} = \frac{4}{9}. \tag{5.22}$$

We see that out of the classical methods only phase coding comes close to the theoretical limits in efficiency. Both pulse codes and phase coded pulse codes fall short of it by a factor equal to a constant factor times the square root of the ratio of the lag resolution and the total modulation time.

Let us suppose that it would be possible to design a code filled altogether with phase coded groups of length τ_{lag} (with no spaces) and that it would be possible to get independent lagged product estimates corresponding to each pair with the correct lag increment τ_m. The number of these pairs would then be $(\tau_{tot} - \tau_m)/\tau_{lag}$ and consequently this would lead tho the same efficiency as in (5.22). In this case, it would be possible to get non-zero lags with the same near-maximum efficiency that is possible for zero lags with phase coded experiments. In the next chapter we will describe how this can be done in practice.

6. A NEW MODULATION PRINCIPLE FOR INCOHERENT SCATTER MEASUREMENTS

6.1 Introduction

The alternating code is a sequence of N signed elementary pulses that occupy the total available transmission period. The elementary pulses may themselves be phase coded or plain, and the receiver response is chosen to be in accordance with the nature of the elementary pulses themselves. The signs (phases) of the elementary pulses are changed from cycle to cycle in a specified manner. All lagged products corresponding to the separations of the elementary pulses have to be calculated in the signal processing hardware. It will then be possible to obtain $N(N-1)/2$ independent lagged product estimates corresponding to every pair of elementary pulses in the modulation, as if all possible combinations of a double-pulse experiment were performed with the rest of the pulses absent.

It is impossible to make the range ambiguity function of a transmission modulation of length N with no spaces in the modulation similar to the ambiguity function of a double-pulse measurement. This function is the sum of the double-pulse range ambiguity functions corresponding to all pairs with the correct or near-correct lag. However, if the signs of the elementary pulses are chosen correctly, some of the ranges in the total ambiguity function will give a positive contribution, and some other ranges will give a negative contribution. If varying signs are used from scan to scan, it will be possible that the sum of the different ambiguity functions will have a single peak only, while the contributions from the unwanted heights have cancelled each other out completely. Also, if specifically chosen signs are used instead of a direct summation in the inter-scan integration, it will be possible to exhibit this single peak corresponding to the ambiguity function of an arbitrary double-pulse pair contained in the modulation.

For the cancellation to occur, one has to be able to suppose that the plasma correlation function does not change during the time which is required to do the inter-scan summation. By this we refer to the time scale of the plasma parameter variations, not the scattering correlation time itself, which may be of the order of the inter-pulse separation in our code. This is not a serious limitation, as the number of scans required will be equal to the number of elementary pulses N in the phase coded case, or twice that number in the non-phase-coded case so that, because of the noise, the integration times will be longer anyway.

A similar coding principle has been reported by Sulzer (1986). It differs from our method in that the signs of the elementary subpulses are randomly chosen leading to a cancellation of the ambiguities in a statistical sense. In our case, the unwanted ambiguities cancel exactly after a specified number of radar scans. This is advantageous in many situations.

6.2 The Mathematical Details

The reduced ambiguity function for range S corresponding to a signal cross product $z_{\mathrm{SC}}(t)\overline{z_{\mathrm{SC}}(t')}$ for scan SC is given by

$$W_{t,t'}(S(\vec{x})) = (\mathrm{env}{*}p)(t - S(\vec{x}))(\mathrm{env}{*}p)(t' - S(\vec{x})), \tag{6.1}$$

where t and t' are the sampling times of the two terms of the crossed product estimates used, env is the transmitter modulation envelope and S is the travel time to the scattering point $\vec{x}$ and back. Let us now suppose that the basic pulse length $\tau = 1$, the receiver impulse response p is a boxcar function of length 1, and the transmitter envelope is a series of N such pulses, which may have different signs. Let us suppose that the signs of these pulses are denoted by

$$s_i, \; i = 0 \ldots N-1.$$

It is clear that the reduced ambiguity function cannot be a narrow function of S. However, if we use a different sign sequence for each scan, it will become possible to arrange the signs so that the sum of the different ambiguity functions will have a single peak only. More specifically, let us define the envelope by

$$\mathrm{env}(t) = \sum_{i=0}^{N-1} s_i(\mathrm{SC}) q_i(t), \tag{6.2}$$

where SC denotes the scan count and q_i is given by

$$q_i(t) = \begin{cases} 1, & \text{if } i < t < i+1 \text{ and} \\ 0, & \text{if } t \leq i \text{ or } t \geq i+1\,. \end{cases} \tag{6.3}$$

The ambiguity function can then be written as

$$W_{t,t'}(S) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} s_i(\mathrm{SC}) s_j(\mathrm{SC})\, (p*q_i)(t-S)(p*q_j)(t'-S) \tag{6.4}$$

Let us next choose a pair (i', j'). We wish to create an ambiguity function corresponding to the correlation of the two elementary pulses $q_{i'}$ and $q_{j'}$. For that purpose, let us form the sum

$$W_{t,t'}^{i',j'}(S) = \sum_{\mathrm{SC}=0}^{\mathrm{NP}-1} s_{i'}(\mathrm{SC}) s_{j'}(\mathrm{SC})\, W_{t,t'}(S)\,. \tag{6.5}$$

$W_{t,t'}^{i',j'}$ can then be written in the form

$$\begin{aligned} W_{t,t'}^{i',j'}(S) &= \mathrm{NP}\,(q_{i'}*p)(t-S)(q_{j'}*p)(t'-S) \\ &\quad + \sum_{i\neq i'} \sum_{j\neq j'} A_{iji'j'}\,(q_i*p)(t-S)(q_j*p)(t'-S), \end{aligned} \tag{6.6}$$

where the coefficients $A_{iji'j'}$ are given by

$$A_{iji'j'} = \sum_{\mathrm{SC}=0}^{\mathrm{NP}-1} s_{i'}(\mathrm{SC}) s_{j'}(\mathrm{SC}) s_i(\mathrm{SC}) s_j(\mathrm{SC})\,. \tag{6.7}$$

The first term in (6.6) gives the ambiguity function for the lag profile corresponding to the elementary pulse pair $q_{i'}$ and $q_{j'}$, if $t - t' = i' - j'$. Since

$$(q_i*p)(t-S) = (q_0*p)(t-S-i),$$

we may write the second term in (6.6) as

$$\sum_{i\neq i'} \sum_{j\neq j'} A_{iji'j'}\,(q_0*p)(t'+i'-j'-S-i)(q_0*p)(t'-S-j)\,. \tag{6.8}$$

The function multiplying $A_{iji'j'}$ is nonzero only if $|i'-j'-i+j| \leq 1$. Thus, we see that, provided we can choose the sign sequences $s_i(\mathrm{SC})$ so that

$$A_{iji'j'} = 0, \quad \text{if} \quad i - j = i' - j' \quad \text{or} \quad i - j = i' - j' \pm 1, \tag{6.9}$$

the double sum in (6.6) disappears, and only the first term in (6.6) remains. This is the ambiguity function corresponding to the pulse pair q_i and q_j. In the summation (6.5), all the other contributions from other pulse pairs have cancelled each other out exactly due to the alternating signs.

The problem of finding suitable sign sequences $s_i(\mathrm{SC})$ fulfilling the condition (6.9) can be solved by using the Walsh sign sequences. It is not difficult to see that it is quite impossible to satisfy the condition (6.9) if $i = j$, but if the additional condition $i \neq j$ is made, suitable sequences can be found. This means that all nonzero lags can be obtained by this method.

6.3 The Walsh Sign Sequences

The Walsh sign matrix $\mathrm{Wsh}(i,j)$ is defined by the binary representation of the positive integers i and j as follows:

$$\mathrm{Wsh}\left(\sum_{n=0}^{\infty} a_n 2^n, \sum_{n=0}^{\infty} b_n 2^n\right) = (-1)^{\sum_{n=0}^{\infty} a_n b_n}. \tag{6.10}$$

To simplify our derivations, let us define the bitwise logical operation and on integers by

$$\left(\sum_{n=0}^{\infty} a_n 2^n\right) \wedge \left(\sum_{n=0}^{\infty} b_n 2^n\right) = \sum_{n=0}^{\infty} a_n b_n 2^n, \tag{6.11}$$

and the bitwise logical exclusive or by

$$\left(\sum_{n=0}^{\infty} a_n 2^n\right) \oplus \left(\sum_{n=0}^{\infty} b_n 2^n\right) = \sum_{n=0}^{\infty} c_n 2^n, \tag{6.12}$$

with

$$c_n = \begin{cases} 0, & \text{if } a_n = b_n \\ 1, & \text{if } a_n \neq b_n \end{cases},$$

If we define the parity of a binary number as

$$\mathrm{par}\left(\sum_{n=0}^{\infty} a_n 2^n\right) = \left(-1\right)^{\sum_{n=0}^{\infty} a_n}, \tag{6.13}$$

we can rewrite the definition of the Walsh sign matrix as

$$\mathrm{Wsh}(i,j) = \mathrm{par}(i \wedge j). \tag{6.14}$$

A part of the Walsh sign matrix $\mathrm{Wsh}(i,j)$ looks like follows:

	j=0	1	2	3	4	5	6	7		8	9	A	B	C	D	E	F
i=0	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+
1	+	-	+	-	+	-	+	-		+	-	+	-	+	-	+	-
2	+	+	-	-	+	+	-	-		+	+	-	-	+	+	-	-
3	+	-	-	+	+	-	-	+		+	-	-	+	+	-	-	+
4	+	+	+	+	-	-	-	-		+	+	+	+	-	-	-	-
5	+	-	+	-	-	+	-	+		+	-	+	-	-	+	-	+
6	+	+	-	-	-	-	+	+		+	+	-	-	-	-	+	+
7	+	-	-	+	-	+	+	-		+	-	-	+	-	+	+	-
8	+	+	+	+	+	+	+	+		-	-	-	-	-	-	-	-
9	+	-	+	-	+	-	+	-		-	+	-	+	-	+	-	+
A	+	+	-	-	+	+	-	-		-	-	+	+	-	-	+	+
B	+	-	-	+	+	-	-	+		-	+	+	-	-	+	+	-
C	+	+	+	+	-	-	-	-		-	-	-	-	+	+	+	+
E	+	+	-	-	-	-	+	+		-	-	+	+	+	+	-	-
F	+	-	-	+	-	+	+	-		-	+	+	-	+	-	-	+

. (6.15)

In (6.15) the matrix has been divided into blocks in order to visualize its internal symmetries. One should note that all blocks are equal except one, which is the sign inverse of the other blocks. The Walsh matrix can be continued by an obvious method to any dimension which is a power of two.

The rows corresponding to a power of two are particularly simple, as well as the corresponding columns, since the matrix is symmetric. Let us consider row number 5, for example. It is easy to see that it is the product of rows 4 and 1, the reason being that 5=4+1 is the binary representation of 5.

It is easy to prove the following lemmas:

Lemma 6.1: For any nonnegative integers i, j and j', the following formula is true:

$$\mathrm{Wsh}(i,j)\cdot\mathrm{Wsh}(i',j) = \mathrm{Wsh}(i\oplus i',j). \tag{6.16}$$

Lemma 6.2: If N is a power of two and $0 \le i \le N-1$, then the two conditions are equivalent:

$$\text{a)} \quad \sum_{j=0}^{N-1} \mathrm{Wsh}(i,j) = 0$$

$$\text{b)} \quad i \neq 0\ .$$

The key point in finding sign sequencies satisfying condition (6.9) is that we constrain ourselves to the Walsh sequencies, *i.e.*, we start to search for sequencies which may be represented in the form

$$s_i(\mathrm{SC}) = \mathrm{Wsh}(a_i, \mathrm{SC}), \quad \mathrm{SC} = 0...\mathrm{NP}-1, \quad 0 \le a_i \le \mathrm{NP}-1. \tag{6.17}$$

We must also suppose that NP is a power of two. Using the two lemmas given, it is easy to show that condition (6.9) is equivalent to the condition

Condition 6.1: (strong condition)

$$a_i\oplus a_j\oplus a_{i'}\oplus a_{j'} \neq 0, \text{ if } i-j = i'-j' \text{ or } i-j = i'-j'\pm 1 \ (\text{ and } i \neq j\ )\ .$$

It is much easier to test condition 6.1 than the original condition (6.9) for the sign sequences. Moreover, the inversion of any bit position in all the numbers a_i or the permutation of the bits in all the numbers a_i does not affect the validity of condition 6.1. Using this, it is fairly straightforward to prove the following lemma:

Lemma 6.3: If a sequence of numbers $a_i, i = 0...N$ exists satisfying condition 6.1, it can be chosen to satisfy the additional condition

$$\log_2 a_i \le \max_{j<i}(\mathrm{int}(\log_2 a_j)) + 1. \tag{6.18}$$

This condition means that no element in the sequence can be larger than the smallest power of two which is higher than all preceding elements.

In condition (6.9), the requirement for

$$i - j = i' - j' \pm 1$$

resulted from the fact that no separation was supposed to exist between the elementary pulses. If there is a separation between the elementary pulses, longer than the receiver filter response time, or if the elementary pulses are themselves phase coded, this requirement can be relaxed, and we can write a weaker condition:

Condition 6.2: (weak condition)

$$a_i \oplus a_j \oplus a_{i'} \oplus a_{j'} \neq 0, \text{ if } i - j = i' - j' \text{ (and } i \neq j \text{)}$$

Lemma 6.3 can also be used in connection with this condition.

6.4 Practical Alternating Codes

A number of computer searches has been made to find the longest possible pulse sign alternation sequences of various cycle lengths NP satisfying the conditions 6.1 or 6.2. Lemma 6.3 was used to reduce the number of possible combinations of a_i. The results are tabulated in table 6.1. For NP = 4..32 the search was complete. For NP = 64, only the strings having the first 7 elements equal to 00,01,02,04,10,20,40 and 03 were searched through.

NP	a_0 a_1 a_2 a_3 a_4 a_5 a_6 a_7 a_8 a_9 a_{10} . . . . a_{15} . . . . . a_{23}
4	00 01 02
8	00 01 02 04
16	00 01 02 04 10 03 07 16
16	00 01 02 04 10 14 05 13
32	00 01 02 04 10 20 17 37 21 14 31 35 24 06 15 32
32	00 01 02 04 10 20 36 03 07 16 34 06 15 32 12 25
64	00 01 02 04 10 20 40 03 55 77 52 43 31 62 71 25 47 16 57 27 15 51 66 46

Table 6.1 Alternating codes satisfying the strong condition. The values of a_i are given in the octal base.

The searches were programmed in the assembly language of an Z80 microcomputer to gain speed and run in a desktop computer. The search for NP=32 lasted for some hours, while the incomplete search for NP=64 lasted for four months, and the string indicated is the longest one that could be found. It seems possible that there could exist a string of length 32 with a period NP=64, but it can be estimated that the search for it would take perhaps 20 years. This kind

of elementary bit combinatorics with small word lengths is rather fast even in a cheap microcomputer, and the use of a large machine would probably not make the search radically more effective.

Searches were made also to find strings satisfying the weak condition. Complete searches have been made for period lengths 2,4,8 and 16, and the results are given in table 6.2. While in table 6.1 only two strings of maximal length (and satisfying the condition in lemma 6.3) were found for periods NP=16 or 32, much more alternatives of maximal length were found in the case of the weak condition, and only one example is included in table 2. for each period length NP. The string with a period 32 shown is the longest one that was found in a search that lasted for about one month.

NP	a_0 a_1 a_2 a_3 a_4 a_5 a_6 a_7 a_8 a_9 a_{10} . . . . a_{15} . . . . . a_{23}
2	00 01 01
4	00 01 01 02
8	00 01 01 02 04 00 02 07
16	00 01 01 02 04 10 06 16 11 06 15 17 12 03 07 15
32	00 00 01 02 04 10 20 01 26 37 25 21 14 31 34 12 23 07 27 13 06 24 33 23

Table 6.2 Alternating codes satisfying the weak condition. The values a_i are given in the octal base.

6.5 An Example of Actual Modulation

A practical use of the alternating codes might be the following. Let us suppose that the experimenter wishes to measure the lower ionosphere with a spatial resolution requiring a 40μs pulse length, and let us suppose that a 40μs lag resolution is sufficient for his purposes. Then an alternating code of eight subpulses giving a total modulation length of 320μs could be used.

In this case, the strong condition has to be used, and so the period length of the alternating code would be 16. The signs of the subpulses q_i in each scan SC are given by

$$s_i(\mathrm{SC}) = \mathrm{Wsh}(a_i, \mathrm{SC}),$$

where the string a_i is given in table 6.1 :

$$a_0 \ldots a_7 = \ 00\ 01\ 02\ 04\ 10\ 03\ 07\ 16\ .$$

The pulse sequences for each scan in the period are shown in figure 6.1.

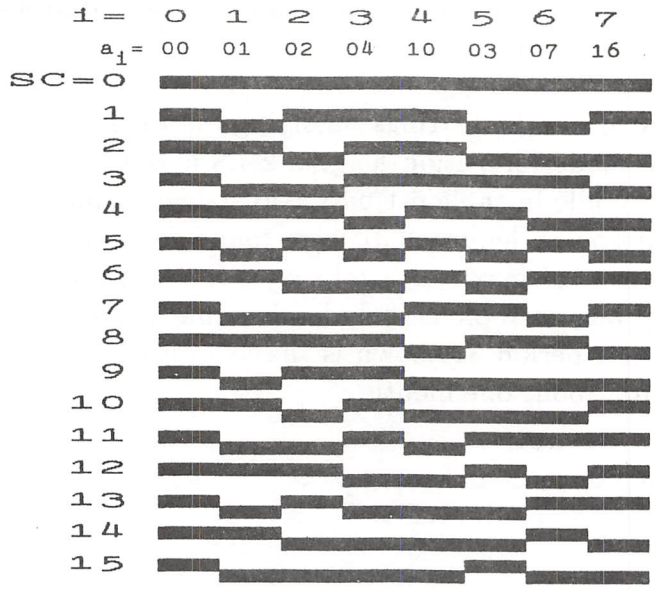


Figure 6.1 The actual alternating codes of 8 elementary pulses satisfying the strong condition.

If the experimenter is interested in a high-resolution experiment with $3\mu s$ basic pulse length and $40\mu s$ lag resolution, he could program this by using an alternating code built of 13-bit Barker codes with $3\mu s$ baud length. The lag resolution would become $39\mu s$ and if an alternating code of 8 subpulses is used, the total modulation length would be $312\mu s$. Because of the phase coding, only the weak condition has to be satisfied, and thus the period would be only 8 scans long. Using the 8 pulses long string from table 6.2, the actual scans would look like in figure 6.2.

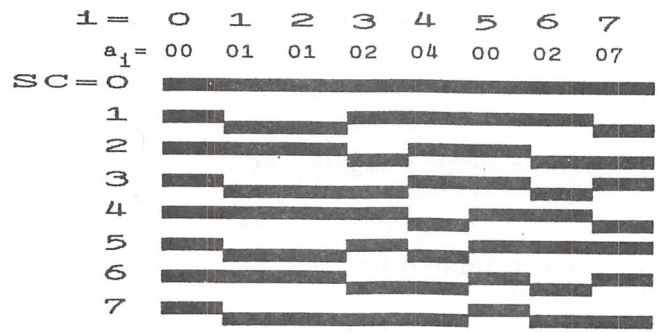


Figure 6.2 The actual alternating code sequences for a code of 8 elementary pulses satisfying the weak condition. When using the weak condition, each of the subpulses shown should itself be phase coded.

6.6 Independence of the Errors

We have stated that it is possible to obtain independent lagged product estimates corresponding to every pair of subpulses by using alternating codes. However, all the estimates corresponding to subpulse pairs

$$(q_i, q_j) \text{ and } (q_k, q_l), \quad i - j = k - l = t - t' \tag{6.19}$$

are formed from the same set of data

$$z_{\mathrm{SC}}(t)\overline{z_{\mathrm{SC}}(t')}/R, \quad \text{with SC} = 0 \ldots \mathrm{NP} - 1. \tag{6.20}$$

If we suppose that the main noise contribution is white noise, it is clear that the sets of data corresponding to different time pairs t and t' are independent. The independence of the lagged product estimates formed from the set of data (6.20) can be proven by the following calculation.

The final lagged product estimate for the elementary pulse pair (q_i, q_j) is formed as

$$M_{t,t'}^{i,j} = \sum_{\mathrm{SC}=0}^{\mathrm{NP}-1} s_i(\mathrm{SC}) s_j(\mathrm{SC})\, z_{\mathrm{SC}}(t)\overline{z_{\mathrm{SC}}(t')}\,. \tag{6.21}$$

and the lagged product estimate for the elementary pulse pair (q_k, q_l) is given by a similar formula with i replaced by k and j replaced by l.

The covariances of the errors of the two lagged product estimates are then given by expressions like

$$\begin{aligned}
\langle \Delta M_{t,t'}^{i,j}, \Delta M_{t,t'}^{k,l} \rangle &= \sum_{\mathrm{SC}=0}^{\mathrm{NP}-1} s_i(\mathrm{SC}) s_j(\mathrm{SC}) s_k(\mathrm{SC}) s_l(\mathrm{SC})\, \langle z_{\mathrm{SC}}(t)\overline{z_{\mathrm{SC}}(t')}\rangle^2 \\
\langle \Delta M_{t,t'}^{i,j}, \overline{\Delta M_{t,t'}^{k,l}} \rangle &= \sum_{\mathrm{SC}=0}^{\mathrm{NP}-1} s_i(\mathrm{SC}) s_j(\mathrm{SC}) s_k(\mathrm{SC}) s_l(\mathrm{SC}) \\
&\quad \cdot \langle z_{\mathrm{SC}}(t)\overline{z_{\mathrm{SC}}(t)}\rangle\, \langle z_{\mathrm{SC}}(t')\overline{z_{\mathrm{SC}}(t')}\rangle
\end{aligned} \tag{6.22}$$

Here we have supposed that the signals from different scans do not correlate with each other.

If the main contribution to the signal is white noise, the expectations in the sums of (6.22) do not depend on the scan count SC, and by the conditions 6.1 and 6.2 and the lemmas 6.1 and 6.2, the sums in (6.22) equal zero if $(i,j) \neq (k,l)$. We have thus proven the independence of the errors of the final lagged product estimates belonging to the different subpulse pairs.

6.7 Signal Processing Hardware

Though the theory seems rather complicated, the signs of the partial pulses are easy to determine in practice: each subpulse q_i has a certain Walsh index a_i, which is a small integer. Then the sign of the transmitted envelope for that pulse can be determined by the formula

$$\text{sign of } q_i = \text{par}(\text{SC} \wedge a_i) \tag{6.23}$$

In calculating the lag profile corresponding to the correlation of the pulse pair (q_i, q_j), the sign in the summation to the correlator result memory is given by

$$\text{sign in the summation for pair } q_i, q_j = \text{par}(\text{SC} \wedge (a_i \oplus a_j)). \tag{6.24}$$

Both of these signs are trivial to generate by a couple of logical ports. Thus, to use this method one needs a lag profile correlator with a programmable Walsh index $a_i \oplus a_j$ which determines the sign of result memory summation by the above formula depending on scan count for each lag profile. The lag profiles for all pairs $(q_i, q_j), (q_{i+1}, q_{j+1}), ..$ can be summed on top of each other, since they just give independent measurements of the same lag profile. To do this, the starting location of each profile in the result memory has to be programmable.

In the above solution, the lag profiles corresponding to every pair of subpulses in the modulation have to be calculated. This means that N(N-1)/2 lagged product profiles have to be calculated for each scan. The result memory will then contain N-1 decoded lag profiles. If there is no shortage of result memory, the number of calculations can be diminished by calculating N-1 lagged product profiles in independent areas of the result memory for each scan, and by doing the decoding later from the data.

In order to save calculation time, a third possibility is to change the code at longer intervals and to dump the result memory data on tape between these intervals. All side diagonals of the signal autocorrelation matrix with lag of an integer multiple of the pulse separation up to some maximum value should be calculated. The decoding can then be performed afterwards, and as the correlator operations are the same as are involved in calculating multipulse data, it should be possible to implement this variant in all existing installations capable

of multipulse measurements. The cycle time will of course be longer than in the first two variants depending on how rapidly the codes can be changed.

Due to the alternating signs no background subtraction is necessary. Possible DC offsets cancel out, too.

A completely general correlator would be a device capable of calculating a large number of lag profiles with the following parameters separately programmable for each profile:

- starting location of the first member of the product in data memory
- starting location of the second member of the product in data memory
- location of the first product in result memory
- data memory step size
- result memory step size
- length of the profile
- Walsh index for the profile

When calculating each profile, the correlator would run in a loop containing the following operations:

1) take the two data values and add their product in the result memory,
2) add the data memory step size to the addresses of the two data points,
3) add the result memory step size to the result memory address
4) repeat the loop until the whole profile has been calculated.

When the whole profile has been calculated, the correlator should load the parameters for the next profile and repeat until all profiles are calculated.

It is possible to implement all existing radar algorithms, including long pulse range-gated algorithms and pulse-to-pulse correlations with a correlator device of this kind.

7. SOME OTHER APPLICATIONS OF INVERSION METHODS

7.1 The Use of Multipulse Zero Lag Data to Improve Power Profile Accuracy

The zero lag response from a multipulse measurement is ambiguous in that its ambiguity function contains several peaks corresponding to each of the separate pulses in the sequence. Usually the multipulse zero lag data is discarded because of this mixed response from several ranges simultaneously. It is possible, however, to understand this situation as a linear inverse problem where the expectation of the data measured consists of sums of the wanted variables plus additional noise. It turns out that the posteriori covariances for the wanted variables can become smaller by a factor of more than 5 in practical situations when the multipulse zero lag data is also used, alone or together with single pulse data in the analysis.

Let us denote the squares of the signals for multipulse and single-pulse channels by m_t^m and m_t^s, respectively. Thus

$$\begin{aligned} \langle m_t^m \rangle &= \langle z^m(t)\overline{z^m(t)} \rangle / R \quad \text{and} \\ \langle m_t^s \rangle &= \langle z^s(t)\overline{z^s(t)} \rangle / R \end{aligned} \tag{7.1}$$

where z^m is the filtered signal for the multipulse channels and z^s is the signal for the single pulse channels. The envelopes of the transmitted signals for the different channels are denoted by

$$\text{env}^m(t) \quad \text{and} \quad \text{env}^s(t), \tag{7.2}$$

respectively.

The time is defined for each channel separately so that each transmitted envelope starts at time 0 and the time scale is chosen so that

$$\text{env}^s(t) = \begin{cases} 1, & \text{when } 0 < t < 1 \text{ and} \\ 0, & \text{otherwise.} \end{cases} \tag{7.3}$$

The pulse code is composed of elementary pulses positioned at suitable multiples of 2 apart from each other. In addition, we suppose that the envelope of each pulse in the multipulse group is similar to that of the single pulse. Hence

$$\mathrm{env}^m(t) = \sum_{i=1}^{N} \mathrm{env}^s(t - 2a_i). \tag{7.4}$$

For example, for a three-pulse code N=3 and the a_i might be chosen as $(a_i) = (0,1,3)$. For a four-pulse group N=4 and the a_i could be chosen as $(a_i) = (0,1,4,6)$. The five-pulse group used is defined by $a_i = (0,3,4,9,11)$.

Since we only discuss lagged products with equal sample times, we drop one of the indices t in the spatial ambiguity functions $W_{t,t}$ and use the notations

$$\begin{aligned} W_t^m(\vec{x}) &= |(p{*}\mathrm{env}^m)(t - S(\vec{x}))|^2 \quad \text{and} \\ W_t^s(\vec{x}) &= |(p{*}\mathrm{env}^s)(t - S(\vec{x}))|^2 \end{aligned} \tag{7.5}$$

It follows that

$$W_t^m(\vec{x}) = \sum_{i=1}^{N} W_{t-2a_i}^s(\vec{x}). \tag{7.6}$$

The zero lag products have averages

$$\langle m_t^s \rangle = \int d^3\vec{x} W_t^s(\vec{x}) P_0(\vec{x}) \mathrm{SV}(\vec{x}) \sigma_{eff}(0;\vec{x}) = x_t, \tag{7.7}$$

and, due to the formula for W_t^m, we can write

$$\langle m_t^m \rangle = \int d^3\vec{x} W_t^m(\vec{x}) P_0(\vec{x}) \mathrm{SV}(\vec{x}) \sigma_{eff}(0;\vec{x}) = \sum_{i=1}^{N} x_{t-2a_i}. \tag{7.8}$$

Here we have denoted the single-pulse averages as x_t which can be understood as a raw uncorrected power profile with different values of t corresponding to different ranges. Suppose that we are interested in the x_t values for $t = t_0 ... t_n$, and have for that purpose measured the values

$$\begin{aligned} &m_t^s, t = t_0^s \ldots t_n^s \quad \text{and} \\ &m_t^m, t = t_{2a_N}^m \ldots t_n^m . \end{aligned} \tag{7.9}$$

From eq. (7.8) it is clear that the measurements of m_t^m contain information about x_t also, but in such a way that the average of a m_t^m represents a sum of several different x_t values.

We may collect the m_t^s and m_t^m values to single vectors m^s and m^m, respectively. We can then describe the measurement as two independent vector-valued measurements m^s and m^m of x by

$$\begin{aligned} m^s &= \mathbf{1}x + \varepsilon^s \quad \text{and} \\ m^m &= Ax + \varepsilon^m \end{aligned} \tag{7.10}$$

where ε^s and ε^m are two independent Gaussian vector random variables with zero mean and with covariances determined by the noise level of the system and by the plasma properties. A good approximation for the covariances in the situation with a poor signal-to-noise ratio is

$$\begin{aligned} \Gamma_s &= \langle \varepsilon^{s^T} \varepsilon^s \rangle = \xi^2 \mathbf{1} \\ \Gamma_m &= \langle \varepsilon^{m^T} \varepsilon^m \rangle = \xi^2 \mathbf{1}. \end{aligned} \tag{7.11}$$

The matrix A is determined by the summation in (7.8):

$$A_{kl} = \begin{cases} 1 & \text{if } k - l = a_i \text{ for some } i \text{ and} \\ 0 & \text{otherwise.} \end{cases} \tag{7.12}$$

For example, in the above four-pulse experiment,

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} . \tag{7.13}$$

By using (1.13), the a posteriori distribution for x, given the independent measurements m^m and m^s, is Gaussian with centre point

$$\overline{x} = \left(\Gamma_s^{-1} + A^{\mathrm{T}}\Gamma_m^{-1}A\right)^{-1}\left(\Gamma_s^{-1}m^s + A^{\mathrm{T}}\Gamma_m^{-1}m^m A\right) \tag{7.14}$$

and a posteriori covariance matrix

$$Q^{-1} = \left(\Gamma_s^{-1} + A^{\mathrm{T}}\Gamma_m^{-1}A\right)^{-1} . \tag{7.15}$$

The a posteriori covariance matrix is also the covariance matrix for $\overline{x}$, when $\overline{x}$ is used as an estimator of the raw power profile. This estimator combines multipulse zero lag data with single-pulse data, and it has been shown in theorem 1.1 that it has smaller variance than any other estimator.

In the case of more than one different multipulse or single pulse groups, the formulae above can be readily generalized to

$$\overline{x} = \left(\sum_s \Gamma_s^{-1} + \sum_m A_m^{\mathrm{T}}\Gamma_m^{-1}A_m\right)^{-1}\left(\sum_s \Gamma_s^{-1}m^s + \sum_m A_m^{\mathrm{T}}\Gamma_m^{-1}m^m\right) \tag{7.16}$$

and

$$Q^{-1} = \left(\sum_s \Gamma_s^{-1} + \sum_m A_m^{\mathrm{T}}\Gamma_m^{-1}A_m\right)^{-1} , \tag{7.17}$$

where the summations are over all single or multipulse channels.

A technical simplification results if one divides the data sets m_t^m and m_t^s in two parts, one containing all even sample times t and the other one all odd sample times. This is due to the fact that the coupling expressed by (7.8) exists only between sample times whose separation is a multiple of two. One can do the analysis with the two data sets separately, resulting in smaller matrices in the computations. The matrix A in this case becomes

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & \dots \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{7.18}$$

The methods were applied to the EFOR-V4 code designed by Tauno Turunen. The code is similar to the code B by Turunen and Silén (1984) and consists of

four four-pulse groups and two power profile pulses, making 6 channels total. The modulation is shown in figure 7.1.

Let us denote by A' the matrix corresponding to (7.18) for the inverted four-bit codes. Because the error covariances are diagonal, $\Gamma_s = \Gamma_m = \xi^2\mathbf{1}$, the matrix Q in (7.17) becomes

$$Q = \xi^{-2}(2\cdot\mathbf{1} + 2\cdot A^{\mathrm{T}}A + 2\cdot A'^{\mathrm{T}}A')$$

$$= \frac{2}{\xi^2}\begin{pmatrix} 3 & 1 & 1 & 0 & 1 & 1 & 2 & 0 & 0 & \dots \\ 1 & 4 & 1 & 1 & 1 & 1 & 2 & 2 & 0 & \dots \\ 1 & 1 & 5 & 1 & 1 & 2 & 2 & 2 & 2 & \dots \\ 0 & 1 & 1 & 5 & 1 & 1 & 2 & 2 & 2 & \dots \\ 1 & 1 & 1 & 1 & 6 & 1 & 2 & 2 & 2 & \dots \\ 1 & 1 & 2 & 1 & 1 & 7 & 2 & 2 & 2 & \dots \\ 2 & 2 & 2 & 2 & 2 & 2 & 9 & 2 & 2 & \dots \\ 0 & 2 & 2 & 2 & 2 & 2 & 2 & 9 & 2 & \dots \\ 0 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 9 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{7.19}$$

if the data is divided into the odd and even sets as described above. The a posteriori covariance matrix is then given by the inverse of (7.19)

$$Q^{-1} = \frac{\xi^2}{2\cdot 100}\begin{pmatrix} 49 & -6 & -5 & 4 & -8 & -7 & -10 & 5 & 4 & 4 & 5 & 2 & \dots \\ -6 & 36 & 1 & -3 & -3 & -4 & -7 & -8 & 3 & 5 & 4 & 3 & \dots \\ -5 & 1 & 31 & -1 & 0 & -7 & -6 & -8 & -6 & 4 & 5 & 6 & \dots \\ 4 & -3 & 1 & 31 & 3 & 0 & -6 & -6 & -8 & -7 & 3 & 5 & \dots \\ -8 & -3 & 0 & 3 & 26 & 4 & -1 & -5 & -7 & -7 & -6 & 3 & \dots \\ -7 & -4 & -7 & 0 & 4 & 23 & 2 & -1 & -3 & -6 & -7 & -5 & \dots \\ -10 & -7 & -6 & -6 & -1 & 2 & 21 & 3 & 0 & -3 & -6 & -6 & \dots \\ 5 & -8 & -7 & -6 & -5 & -1 & 3 & 20 & 3 & 0 & -3 & -6 & \dots \\ 4 & 3 & -7 & -8 & -7 & -3 & 0 & 3 & 20 & 3 & 0 & -4 & \dots \\ 4 & 5 & 4 & -7 & -7 & -6 & -3 & 0 & 3 & 19 & 3 & -1 & \dots \\ 5 & 4 & 5 & 3 & -6 & -7 & -6 & -3 & -1 & 3 & 19 & 3 & \dots \\ 2 & 3 & 6 & 5 & 3 & -5 & -6 & -6 & -4 & -1 & 3 & 19 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{7.20}$$

If only single pulse data were used in the analysis, the a posteriori covariance matrix would be diagonal with all diagonal elements equal to 100 instead of 49...19 as in (7.20). Thus one can see that the method will improve the

variances with a factor even greater than 5. In the first and last data points the improvement is not as great, due to edge effects which rapidly disappear, however.

Figure 7.2 shows electron density profiles calculated with the inversion and direct methods. The data was measured during a special Finnish EISCAT campaign on 17 Nov, 1983. Before applying the different methods, the channels were balanced by the background and calibration measurements. Examples of electron density profiles in cases of a low and a high SNR are given in panels a and b, respectively. In the first case, the estimate calculated by equation (7.16) shows an overall smoother behaviour and in the lower F-layer, where the single pulse measurements totally collapse, the inversion method still produces a readable profile. In the high SNR case both methods give good results in the E-layer peak, but the inversion estimate is better. In the lower F-layer the improvement is evident.

The posteriori covariance matrices were calculated for a number of other combinations of different multipulse/single-pulse schemes. The posteriori variances for different combinations are given in table 7.1. The values give the variances from the middle point of the profiles whose total length is 2x25, so that the end effects cannot be seen. The values for resolution 10 are obtained by inserting the additional priori knowledge to the analysis that 2x5 adjacent gates have the same density. The column labelled "largest eigenv." gives the largest eigenvalue of a 10x10 submatrix Γ_p of the posteriori covariance matrix Q^{-1}. One should note that, with some combinations, the single pulse channels can be dropped out altogether.

code group	variance for res=1	variance for res=10	largest eigenval.	code group	variance for res=1
1	1.000	0.100	1.000		
1 + 3	0.340	0.016	0.479	3	∞
1 + 3 + 3'	0.220	0.009	0.330	3 + 3'	0.344
1 + 4	0.286	0.024	0.386	4	∞
1 + 4 + 3	0.159	0.009	0.194	4 + 3	0.196
1 + 4 + 4'	0.191	0.017	0.284	4 + 4'	0.607
1 + 5	0.305	0.039	0.441	5	∞
1 + 5 + 3	0.154	0.009	0.221	5 + 3	0.364
1 + 5 + 4	0.140	0.011	0.186	5 + 4	0.221
1 + 5 + 5'	0.197	0.027	0.284	5 + 5'	0.512
1 + 5 + 4 + 3	0.100	0.006	0.124	5 + 4 + 3	0.119

Table 7.1 The estimation variances for different combinations of multipulse and single pulse groups. The numbers in the column "code group" denote the number of pulses in the pulse code groups used. Primes mean that the groups have been inverted.

The analysis of the error covariances as well as the practical calculations are discussed in some more detail in Lehtinen and Huuskonen (1986).

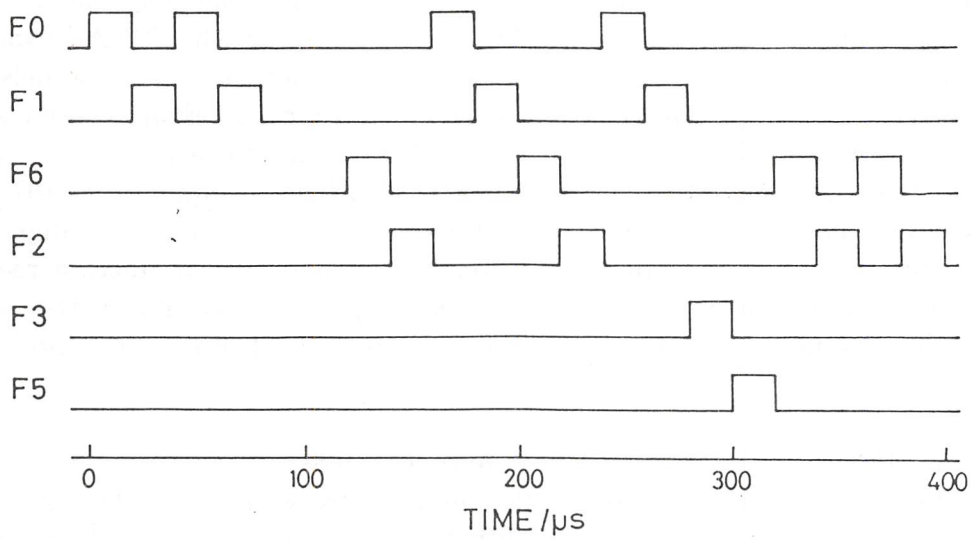


Figure 7.1 The monostatic part of the example modulation.

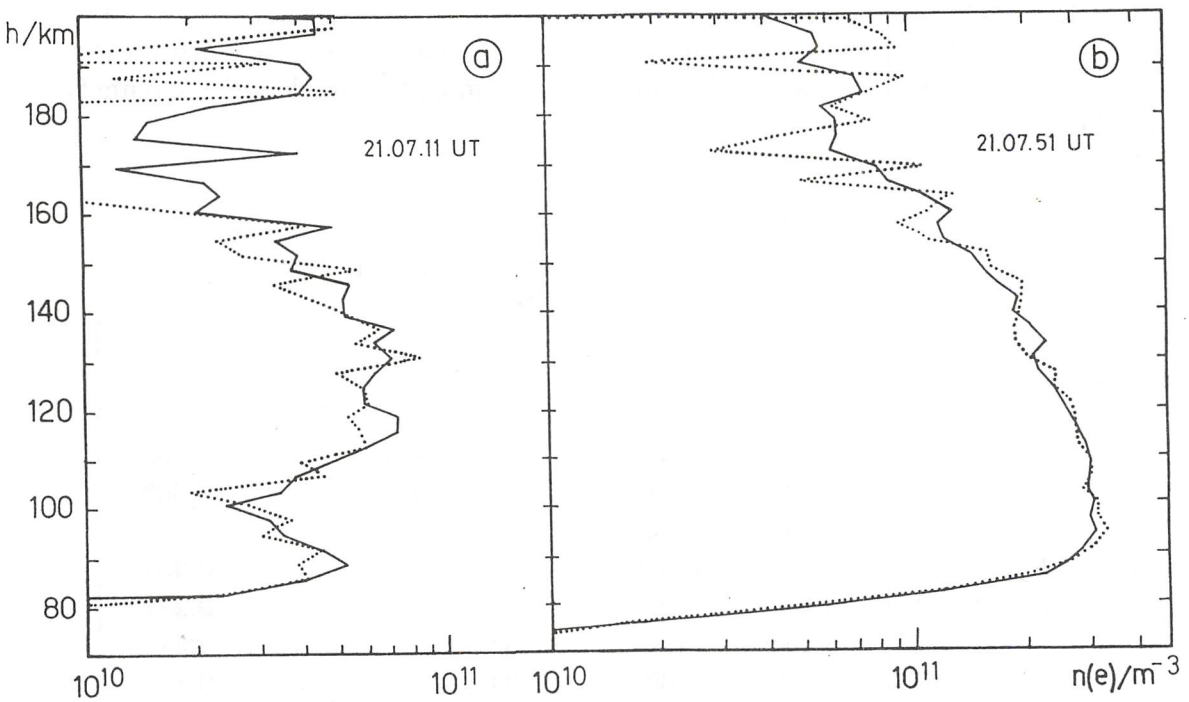


Figure 7.2 Electron density profiles in cases of a low and a high SNR are displayed in panels a and b, respectively. The solid line gives the result of the inversion method, and the dotted profile was obtained with the single-pulse data only. The data was measured on 17 Nov 1983 at times displayed in panels. The integration time is 10 seconds.

7.2 Variances of the Fitted Plasma Parameters

Finding the plasma parameters is an inversion problem where the unknowns are the physical properties of the plasma and where the theory mapping is a combination of the physical theory describing plasma scattering spectra and the theory developed in chapters 2 and 3 relating the plasma scattering cross section to the measured cross products of the signal and their variances. This inversion problem may be understood in many ways, depending on how one chooses the space of unknowns. The most usual approach can be described in our terms as follows.

A range interval of interest is fixed and a group of measured lagged products is chosen so that their ambiguity functions fit in that interval. This group is used as the set of measurements to which the theoretically calculated autocorrelation values are fitted. The plasma is supposed to be constant with respect to the range variable in the range interval of interest, and the constant values of its physical properties are the unknowns in the analysis. The space of unknowns is thus finite-dimensional. The variation of $\sigma_{\text{eff}}(\tau; S)$ with respect to τ is then taken care of through filter corrections. Correctly performed, the filter corrections are calculated by equations (2.18) and (2.35) giving the expectations of the crossed products by the reduced lag ambiguity functions. The possible variation of the plasma with respect to range is not taken care of in these kinds of methods. The direct theory is thus given by the formulae (0.2) and (2.18) and the variances by (3.18).

The a posteriori variances (=the second moments of the a posteriori distribution) of the plasma parameters can be evaluated for this kind of estimation problem. In (0.2), let us suppose that the parameters N_e, T_i and T_e/T_i are constant in space and let us consider them to be the unknown parameters of an inversion problem. The measurements of the inversion problem are crossed products of the signal as given by (2.18) and the measurement variances as given in chapter 3. We suppose that a set of codes as in figure 7.1 is used. Six lag profiles are calculated for each of the four-pulse channels and a zero lag profile for both of the single-pulse channels. For each height two independent power values and four independent estimates of each of the six lags are used. The inversion method of chapter 7.1 for the zero lag profiles of the multipulse channels is not used. The receiver impulse response is a boxcar of the same length (=20μs) as the pulses in the transmitted codes. The radar frequency is 933 MHz. The lag resolution is thus 40μs and the longest lag calculated is 240μs. Moreover, we suppose that the signal strength and integration time is such that each of the power profiles can be estimated with 1% accuracy. (For example, if SNR = 0.1 and the integration time is 10^6 cycles, this is true).

We denote the square roots of the a posteriori variances (a posteriori error bars) of the parameters by ΔN_e, ΔT_i and $\Delta(T_e/T_i)$. The values $\Delta N_e/N_e$, $\Delta T_i/T_i$ and $\Delta(T_e/T_i)$ depend only on the width of the scattering spectrum and on T_e/T_i.

The width of the scattering spectrum depends on T_i/m_i. In figure 7.3, we have chosen $T_e/T_i = 1.5$ and plotted the quantities $100 \cdot \Delta N_e/N_e$, $100 \cdot \Delta T_i/T_i$ and $100 \cdot \Delta(T_e/T_i)$ as a function of T_i/m_i in the left hand panels. For example, the point 100 in the x-axis corresponds to $T_i = 1600°\mathrm{K}$ if the ion is O^+ with $m_i = 16$. The y-axis can be understood to give relative percentage error bars.

The analysis of figure 7.3 gives us some information about how the measured lags should be distributed if these kinds of parameters are going to be estimated. As the width of the plasma autocorrelation function is inversely proportional to the square root of T_i/m_i, it follows that the lag resolution becomes the critical factor when the errors start to increase at the right hand side of figure 7.3 and correspondingly, a too short lag extent causes the increase of the errors at the left in figure 7.3. It can be seen that a lag resolution of 40μs is sufficient up to $T_i/m_i = 60$ corresponding to 2000°K for NO^+, 1000°K for O^+ or 60°K for H^+. Correspondingly, a lag resolution of 20μs would be sufficient up to 8000°K/NO^+, 4000°K/O^+ or 240°K/H^+. The lag extent requirements are as follows: A total code length of 240μs is good from $T_i/m_i = 5$ upwards, corresponding to 160°K/NO^+ or 80°K/O^+. What has here been said for O^+ and a 933 MHz radar corresponds to H^+ and a 933 MHz/4 radar because of the way the ACF widths depend on the ion mass. Thus, the lag resolution and extent requirements for a VHF radar of 233 MHz when measuring H^+ are the same as those for an UHF radar measuring O^+ with 933 MHz.

A peak appears in the error bar curves in figure 7.3. The width of the ACF corresponding to the peak is such that the first lag measured is 23% longer than the first zero crossing of the ACF. The peak is caused by an ill-conditioned behaviour of the set of the partial derivatives of the direct theory with respect to the fitted parameters. At the peak, these vectors are almost linearly dependent. If more than three parameters are fitted, the partial derivatives might get ill-conditioned also earlier.

We have plotted the correlation coefficients of the fluctuations of the various parameters in the right hand panels of figure 7.3 and also the eigenvalues of the a posteriori covariance matrix. It can be seen that the errors of the different parameters are highly correlated and the posteriori error distribution is much wider in some directions than in others.

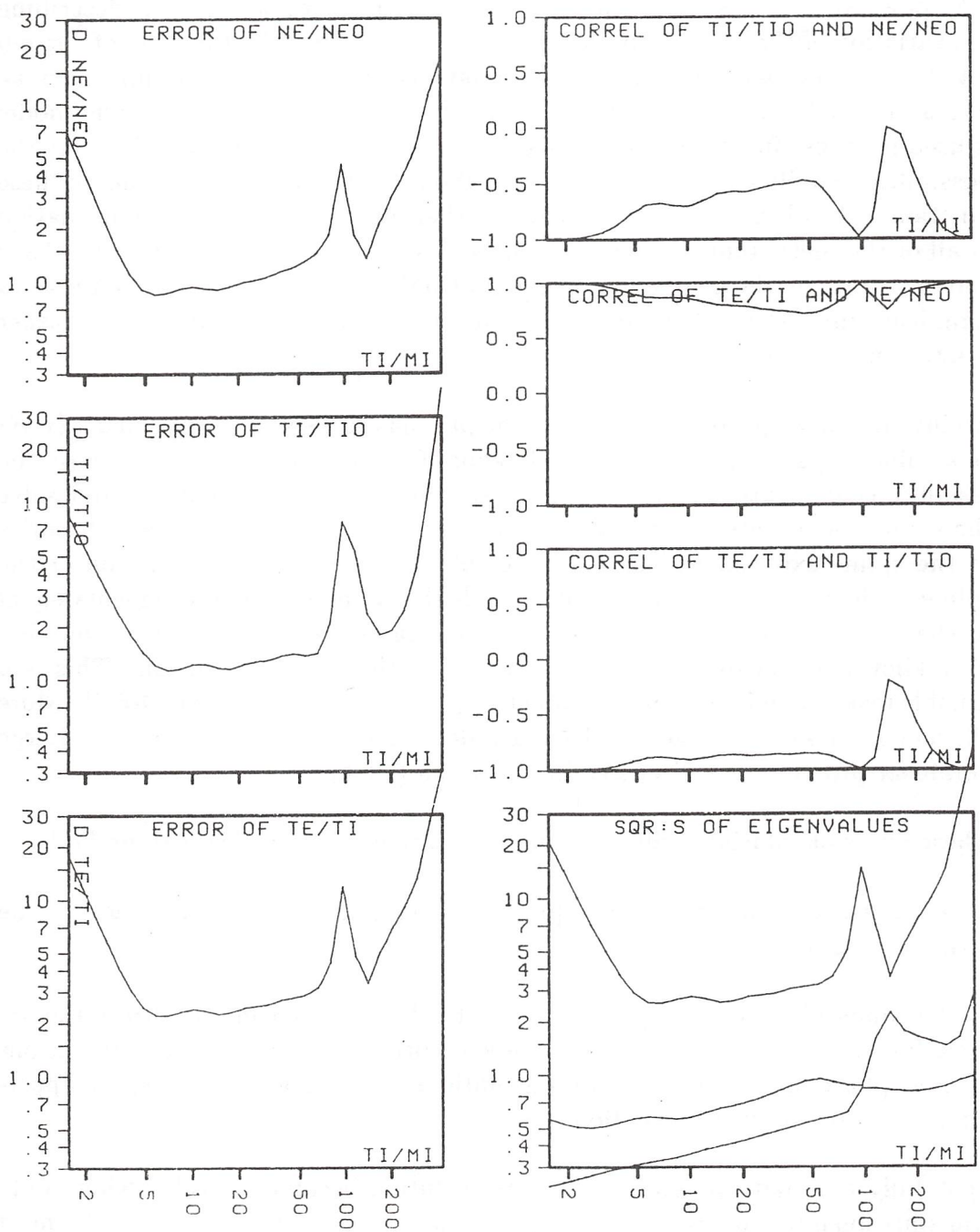


Figure 7.3 The a posteriori errors and correlation coefficients of the errors of the fitted plasma parameters as a function of ACF width.

7.3 A General Inversion Approach to the Analysis of Plasma Parameters

A more general approach to the inversion problem of estimating the plasma parameters would also take into account possible variations of the plasma in the range direction. The unknowns in this approach would then be plasma parameter profiles over the whole ionosphere. It is more difficult to realize this kind of an analysis, as the unknowns would now be a set of functions describing the variation of the physical parameters with respect to height instead of a small set of physical constants. The space of unknowns would be infinite-dimensional. A practical solution can be found, if some finite-dimensional parametric model is used for these functions. A spline expansion would give the experimenter the possibility to build in a priori information about the possible height scales of these functions. A full inversion approach can then make use of all the information in all of the measurements, including also classically too broad responses. As is demonstrated by the example of using multipulse zero lag data, this can lead to significant improvements in accuracy. This approach will be numerically rather heavy, because the problem is non-linear.

An intermediate approach is to model the plasma cross section function $\sigma_{\text{eff}}(\tau; S)$ as a spline expansion (5.1) with full freedom for the experimenter to choose the resolutions as seems sensible. The solution will then be in two steps: first solve the spline coefficients σ_i by (5.4-5.6) and then fit the nonlinear plasma theories to the spline expansion at as many heights as necessary. The solution of the spline coefficients is a linear problem with the number of measurements equal to the number of crossed products (or range-gates) calculated and the number of unknowns equal to the number of terms in the spline expansion. This is a feasible task, though perhaps not directly by the matrix formulae (5.4-5.6). More efficient methods for large problems are described *e.g.* in Nash (1979). Large nonlinear problems are discussed in Tarantola (1984).

These kinds of solutions will have many advantages to the classical methods:

1) They use the information completely, leading to improved accuracy in the estimates, as demonstrated in 7.1.

2) All kinds of "corrections" are automatically included in the direct theory, including filter corrections, plasma gradient corrections, sidelobe deconvolutions of short phase codes, ambiguity deconvolutions of ambiguous codes, spatial post-integration to specified resolutions etc.

3) A fully automatical general computer solution becomes possible where only transmission pulse forms, receiver responses and correlator summation rules need be specified for each different experiment and the computer can handle the analysis using specified resolutions without the need for reprogramming when the codes are changed.

7.4 Sufficiency of the Crossed Product Matrix

Let us denote the unknown variable in an inversion problem by x and a set of independent measurements by m_i, $i = 1 \ldots N$. We suppose that the transition densities are of the form

$$D(m_i|x) = \frac{1}{(2\pi)^{n_{M_i}/2}|\Gamma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} m_i^{\mathrm{T}} \Gamma^{-1}(x) m_i\right), \tag{7.21}$$

where $\Gamma(x)$ is a non-negative matrix function of the unknowns x. The a posteriori density for x, given measurements $\overline{m_i}$, is the product of the a priori density and the transition densities. It can be written in the form

$$D_p(x|\overline{m_i}) \sim D_{pr}(x) \cdot |\Gamma(x)|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathrm{Tr}\left(\Gamma^{-1}(x) \sum_{i=1}^{N} \overline{m}_i \overline{m}_i^{\mathrm{T}}\right)\right). \tag{7.22}$$

The trace of a matrix has been denoted by Tr. It is necessary to include $|\Gamma(x)|$ in the formula, because it is not a constant due to the dependence on x. We see that the a posteriori density depends on the measurements only through the matrix

$$M = \sum_{i=1}^{N} \overline{m}_i \overline{m}_i^{\mathrm{T}}. \tag{7.23}$$

In this case, the matrix M is called a sufficient statistic for the unknowns x. All information concerning x is contained in the sum of crossed products M and there is no loss of information if the original data m_i is lost and only M is used.

The standard way of handling sampled incoherent scatter data is similar to the situation above. The sampled signal is a Gaussian random vector with zero mean. The theories of plasma density fluctuations and ambiguity functions give the second moments of the sampled signal, corresponding to $\Gamma(x)$ above. The summation over i corresponds to the summation over different radar scans and frequency channels. It can be understood that the unknown x consists of the physical parameters of the plasma as well as the unknown system parameters, like background noise T. If the signal is sampled at time intervals $\Delta\tau$ and if the corresponding crossed product matrix $M_{ij} = M(i\Delta\tau, j\Delta\tau)$ as defined in (3.10) is calculated, it is thus guaranteed that no information is lost in the transition from the sampled signal itself to the crossed product matrix. In particular, no

other methods of signal estimation, including maximum entropy methods, can provide any statistical information about the plasma parameters that could not be found by studying the matrix $M(i\Delta\tau, j\Delta\tau)$.

If some system parameters (like DC offsets) could cause the signal to have a non-zero mean, the signal averages should also be calculated. In this case it can be shown that the averages together with the crossed product matrix are a sufficient statistic for the unknown parameters.

Some information is lost in the sampling of the signal. The sampling and the corresponding filtering loses small details of the signal. Calculations have shown that a dense sampling of the signal will lead to a gain in speed of the order of 1.5...2 in usual multiple pulse situations. Most of this gain is achieved already when the sampling interval and the receiver response time are divided by two.

The conclusions of the chapters 4 and 5 were based on the assumption that the measurements are the crossed product estimates $M(i\Delta\tau, j\Delta\tau)$ and that the variances of these estimates can be derived as in chapter 3. Thus, we see that it was a sound approach and nothing better could be said by using the signal values $Z(i\Delta\tau)$ directly. Strictly speaking, it is true that the discussion of a posteriori variances should be based on formula (7.22) above, but as it is rather difficult to handle due to the matrix formula in the exponential, the variances as given in chapter 3 give a good and an accurate enough way to handle the a posteriori distribution. This is true because a large number of independent measurements is made during the integration so that the sums in (3.10) will have almost Gaussian distribution and the K values appearing in (3.18) can be accurately estimated from the data.

7.5 The Lag Profile Algorithms

In a correlator device the crossed product matrix $M_{ij} = M(i\Delta\tau, j\Delta\tau)$ (see 3.10) is calculated for each channel used. From a theoretical point of view, it is completely uninteresting, whether the matrix is calculated row by row, column by column or in the order of ascending side diagonals. It does not affect the matrix element values or the information content of the result.

In practise, the algorithms may look completely different depending on the order with which the different elements of M_{ij} are handled. The most important reason for this is that some elements of M_{ij} may contain no or very little information about σ_{eff} (*i.e.* the a posteriori distribution does not depend on them). This is true if the ambiguity function for M_{ij} is zero or near to zero and if the error fluctuation corresponding to that term is independent or almost independent of the error fluctuations of the relevant terms. To save correlator memory and calculation time, these elements are not calculated at all.

The forms of the ambiguity functions $W_{t,t'}$ depend only on the time lag $t - t'$, but not otherwise on the values of t or t'. If t and t' change so that $\tau = t - t'$ stays constant, the ambiguity function $W_{t,t'}(\tau; S)$ moves in the range direction without changing its shape or its position in the lag direction. This is the reason why the side diagonals of M or parts of them $M_{i+k,j+k}$, $k = 0...N$ are called *lag profiles*. The ambiguity functions corresponding to the products in a lag profile form a series of similar shapes moving higher in range as k grows. Their position with respect to the lag variable stays constant.

If a point M_{ij} does not contain any information about σ_{eff}, this is true for the whole lag profile $M_{i+k,j+k}$. All the points in a lag profile represent the same kind of information about σ_{eff},only the height of the response changes. Thus it is most natural to arrange the calculation of the matrix M so that a whole lag profile is calculated in the innermost loop. The lag profiles can be specified by the locations of the two factors of the first product in the data memory, the location of the first result in the result memory and by the length of the profile.

If the matrix M_{ij} is handled row by row, the consecutive crossed products M_{ii}, $M_{i,i+1}$, ..., $M_{i,i+k}$,... will generally have quite different ambiguity functions, where the response height may change up and down in an irregular manner (see fig 3.3). Moreover, it may be necessary to increment k irregularly to avoid the calculation of non-informative products. A lag profile approach leads to simpler and more efficient correlator software and hardware (see Ho et al., 1983).

Another much used technique is the fast Fourier transform. The square of the Fourier transform of a finite length sampled signal $Z(i\Delta\tau), ..., Z((i+k)\Delta\tau)$ is the Fourier transform of the autocorrelation function $A_Z(l)$ of the finite-length sampled signal. The signal is usually continued with a proper number of zeros to make this strictly true. The autocorrelation function $A_Z(l)$ in turn is a function of lag l and it is then related to the matrix M by

$$A_Z(l) = \sum_{j=0}^{k-l} M_{i+j,i+j+l} \,. \tag{7.24}$$

The Fourier transform method thus gives sums a of partial lag profiles. It follows that it is not so useful in cases where individual lag profile points are necessary, but it can be used to save a lot of computation time in range-gated situations.

LITERATURE

Buchert,S.: Das Studium der polaren Ionosphäre mit der Thomsonstreutechnik, Diplomarbeit, Technische Universität Munchen (1983)

Buneman,O.: Scattering of radiation by the fluctuations in a nonequilibrium plasma, J. Geophys. Res., **67**, 2050-2053 (1962)

Cohen,M.H. : Hydrodynamic theory of plasma density fluctuations, J. Geophys. Res., **68**, 5675-5679 (1963)

Dougherty,J.P. and Farley,D.T.: A theory of incoherent scattering of radio waves by a plasma, Proc.R.Soc., **A 259**, 79-99, (1960)

Dougherty,J.P. and Farley,D.T.: A theory of incoherent scattering of radio waves by a plasma II: Scattering in an magnetic field, Proc.R.Soc., **A 263**, 238-258 (1961)

Dougherty,J.P. and Farley,D.T.: A theory of incoherent scattering of radio waves by a plasma III: Scattering in a partly ionized gas, J. Geophys. Res, **68**, 5473-5487, (1963)

Farley,D.T.: Incoherent scatter correlation function measurements, Radio Sci., **4**, 935-953 (1969)

Farley,D.T.: Multiple-pulse incoherent-scatter correlation function measurements, Radio Sci., **7**,661-666 (1972)

Fejer,J.A.: Radio-wave scattering by an ionized gas in thermal equilibrium, J. Geophys. Res., **81**, 3441-3443 (1960a)

Fejer,J.A.: Scattering of radio waves by an ionized gas in thermal equilibrium, Can. J. Phys., **38**, 1114-1133 (1960b)

Fejer,J.A.: Scattering of radio waves by an ionized gas in thermal equilibrium in the presence of a uniform magnetic field, Can. J. Phys., **39**, 716-740 (1961)

Gihman,I.I. and Skorohod,A.V.: The theory of stochastic processes I, Springer Verlag, Berlin (1980)

Gray,R.W. and Farley,D.T.: Theory of incoherent-scatter measurements using compressed pulses, Radio Sci., **8**,123-131 (1973)

Hagfors,T.: Density fluctuations in a plasma in a magnetic field, with applications to the ionosphere, J. Geophys. Res., **66**, 1699-1712 (1961)

Hagfors,T and Brockelman: A theory of collision dominated electron density fluctuations in a plasma with applications to incoherent scattering. Phys. Fluids, **14**, 1143-1151 (1971)

Ho,T., Turunen,T., Silén,J., Lehtinen,M.: The lag profile routine and the universal program for the EISCAT digital correlators, EISCAT Technical note 83/37, EISCAT Scientific Association, Kiruna, Sweden (1983)

Lehtinen,M.S., Päivärinta,L., Somersalo,E.: Linear inverse problems for generalized random variables, (to be published)

Lehtinen,M.S. and Huuskonen,A.: The use of multipulse zero lag data to improve incoherent scatter radar power profile accuracy, J. atmos. terr. Phys., (in press, 1986)

Moorcroft,D.R.: On the determination of Temperature and ionic composition by electron backscattering from the ionosphere and magnetosphere, J. Geophys. Res., **69**, 955-970 (1964)

Murdin,J.: SNR for the EISCAT UHF system, KGI report no.78.1, Kiruna Geophysical Institute, Kiruna Sweden (1978)

Priestley,M.B.: Spectral analysis and time series, Academic Press, London (1981)

Rino,C.L.: Incoherent scatter radar signal processing techniques, Technical memo, SRI international, 333 Ravenswood Avenue, Menlo Park, California 94025 (1978)

Salpeter,E.E.: Electron density fluctuations in a plasma, Phys. Rev., **120**, 1528-1535 (1960)

Salpeter,E.E.: Plasma density fluctuations in a magnetic field, Phys. Rev., **122**, 1663-1674 (1961)

Sheffield,J.: Plasma scattering of electromagnetic radiation, Academic Press, New York (1975)

Sulzer,M.P.: A radar technique for high range resolution incoherent scatter autocorrelation function measurements utilizing the full average power of klystron radars, Radio Sci. (to be published 1986)

Tarantola,A. and Valette,B.: Inverse problems = quest for information, J. Geophys., **50**, 159-170 (1982a)

Tarantola,A. and Valette,B.: Generalized nonlinear inverse problems solved using the least squares criterion, Rev. Geoph. and Space Phys., **20**, 219-232 (1982b)

Tarantola,A. and Nercessian,A.: Three-dimensional inversion without blocks, Geophys. J. R. astr. Soc., **76**, 299-306 (1984)

Tarantola,A.: The seismic reflection inverse problem, in Inverse problems for acoustic and elastic waves, Santosa, F. et al, eds, SIAM, Philadelphia (1984)

Turunen,T. and Silén,J.: Modulation patterns for the EISCAT incoherent scatter radar, J. atmos. terr. Phys., **46**, 593 (1984)

Williams,A.: Electronic filter design handbook, McGraw-Hill Book Company, New york (1981)

Woodward,P.M.: Probability and information theory with applications to radar, Pergamon Press Ltd, London (1953)

Woodman,R.F. and Hagfors,T.: Methods for the measurement of vertical ionospheric motions near the magnetic equator by incoherent scattering, J.Geophys.Res., **74**, 1205-1212 (1969)

Zamlutti,C.J.: Design of Barker coded multiple pulse experiments, J. atmos. terr. Phys., **42**, 975-982 (1980)

LIST OF NOTATIONS

0. Introduction:

$\sim$	equal, except for normalization ($D_1(m) \sim D_2(m) \Leftrightarrow D_1(m)/D_2(m) = \text{constant}$)
$\approx$	approximately equal sign
$e(t; d^3x)$	complex envelope of signal scattered from d^3x
R	receiver input impedance
$\vec{x}$	a point in space where scattering occurs
$\vec{x}_0$	a fixed point in scattering volume
$P_0(\vec{x})$	the single electron scattering power at $\vec{x}$
ω	angular frequency of a signal
ω_{th}	Doppler shift angular frequency corresponding to ion thermal velocity
ω_0	radar transmission angular frequency
D	Debye length
k	length of scattering wave vector
$W(\omega/\omega_{th})$	the Fried–Conte function
N_e	electron density
T_e	electron temperature
T_i	ion temperature
θ	angle between incident and scattered waves
c	speed of light in vacuum
ε_0	dielectric constant of the vacuum
σ_{eff}	effective plasma cross section
S_{eff}	plasma scattering spectrum

1. Elementary Inversion Theory:

Ω	a measurable space, the ensemble space for random variables
$\mathbf{R}^n$	the euclidean space of n dimensions
M_0, X	an euclidean space, usually the space of unknowns in an inverse problem
M_i	an euclidean space, usually the space of measurements in an inverse problem
m_0, x	a random variable in M_0 or X, usually the unknown in an inverse problem
m_i	a random variable in M_i, usually a measurement in an inverse problem
exp	the exponential function

$D(m_0, ..., m_N)$	probability density of the distribution of the variables $m_0 ... m_N$
$D(m_i \mid m_0)$	conditional density of m_i, given m_0
$D_{pr}(m_0)$	the a priori density for m_0, marginal density of m_0
$D_p(m_0)$	the a posteriori density for m_0, conditional density of m_0, given measurement values $\overline{m}_i$
ε	a random vector, usually noise
Γ, $\lvert\Gamma\rvert$	a covariance matrix and its determinant
A_i	linear operators from the space of unknowns to the space of measurements
Q	the Fisher information matrix
T in A^{T}	the transpose of a matrix. The vectors are column vectors so that $m^{\mathrm{T}}m$ is a scalar, while mm^{T} is a matrix
error(f)	the error covariance matrix of an estimator f
$\mathbf{E}x$	expected value of x, statistical average (same as $\langle x \rangle$
$\mathbf{E}(x \mid m)$	conditional expectation of x, given m

2. Direct Theory of Pulsed Radar Measurements:

$G_0(\vec{x})$	transmitter antenna gain in the direction of $\vec{x}$
$G_1(\vec{x})$	receiver antenna gain in the direction of $\vec{x}$
$G_0(0)$	transmitter antenna gain in the direction of the beam axis
$G_1(0)$	receiver antenna gain in the direction of the beam axis
$R_0(\vec{x})$	distance from transmitter antenna to scattering point $\vec{x}$
$R_1(\vec{x})$	distance from receiver antenna to scattering point $\vec{x}$
$S(\vec{x})$	total travel time from transmitter through $\vec{x}$ to the receiver
R	receiver impedance
λ	radar wavelength
A_e	antenna effective area
χ	polarization angle
r_0	classical electron radius
P_t	transmitter power
$e(t; d^3\vec{x})$	complex envelope of signal scattered from $d^3\vec{x}$ by a continuous wave incident signal
$\delta(t)$	the Dirac delta function
$e(t)$	complex envelope of the received scattering signal
$e_n(t)$	complex envelope of the received noise signal
$\int_{\vec{x}}$	a space integral with respect to the position vector $\vec{x}$
env(t)	the complex envelope of the transmission modulation
$p(t)$	receiver impulse response function
$z(t)$	the filtered signal
$z_n(t)$	the filtered noise signal
$Z(t)$	the total filtered signal, $Z = z + z_n$
$W_t^A(\tau; \vec{x})$	the amplitude ambiguity function for signal sampled at t
$A_p(t)$	the autocorrelation function of the receiver impulse response p

$A_{\text{env}}(t)$	the autocorrelation function of the transmitter modulation envelope p
$W_{t,t'}(\tau;\vec{x})$	the two-dimensional ambiguity function, understood as a function of lag τ and position $\vec{x}$. Corresponds to signal pair sampled at t and t'
$W_{t,t'}(\tau;S(\vec{x}))$	same as above
$W_{t,t'}(\tau;S)$	the two-dimensional ambiguity function, understood as a function of lag τ and range S
$W_{t,t'}(\tau)$	the reduced lag ambiguity function
$W_{t,t'}(\vec{x})$	the reduced range ambiguity function, understood as a function of position $\vec{x}$
$W_{t,t'}(S(\vec{x}))$	same as above
$W_{t,t'}(S)$	the reduced range ambiguity function, understood as a function of range S
$l_{t,t'}$	the effective pulse length
$\text{SV}(\vec{x})$	form of scattering volume, as determined by the radar beam geometries
BW_p	the bandwidth of the receiver filter
$V^{\text{eff}}_{t,t'}$	the effective scattering volume for a sampled signal pair
C_{beam}	a radar beam geometrical constant (≈ 0.460)
$G(\theta,\phi)$	antenna gain in polar coordinates
A^{eff}_{beam}	the beam effective cross section
κ	the Boltzmann constant
T	system noise temperature

3. Variances of the ACF Estimates:

$k(t,t')$	average of the scattering signal crossed product
$k_n(t,t')$	average of the noise signal crossed product
$K(t,t')$	average of the total signal crossed product
$M(t,t')$	an element of the measured crossed product matrix
ND	integration scan count for the crossed product estimates
$m_n(t,t')$	an element of the crossed product matrix for background estimates
NB	integration scan count for the background estimates
$m_c(t,t')$	an element of the crossed product matrix for the calibration estimates
NC	integration scan count for the calibration estimates
$m(t,t')$	the crossed product estimates, with background subtracted; $m(t,t') = M(t,t') - m_n(t,t')$
$\Delta M(t,t')$	the fluctuation in $M(t,t')$
$\Delta m_n(t,t')$	the fluctuation in $m_n(t,t')$
$\Delta m(t,t')$	the fluctuation in $m(t,t')$
Re	the real part of a complex number
Im	the imaginary part of a complex number
Mo	the modulus of a complex number

Ph	the phase of a complex number
rg	a range-gate
M_{rg}	a range-gated crossed product estimate
ΔM_{rg}	the fluctuation of a range-gated crossed product estimate
SNR	the signal-to-noise ratio

4. Speed of Measurements:

n_e	a constant, representative value of electron density
$Q_{t,t'}$	the Fisher information for σ_{eff}/n_e contained in the crossed product estimate $M(t,t')$
$S_{t,t'}$	the speed contribution of the measurement $M(t,t')$
S	the total speed of a measurement
env_0	the modulation envelope of a simple boxcar pulse
τ_0	the elementary pulse length, spatial resolution
τ_{tot}	total modulation time
τ_{lag}	lag resolution
N_B	number of elementary pulses in a phase code
N	number of elementary pulses in the total modulation time
l^{amb}	the effective length of the ambiguous parts of the ambiguity function for a multipulse code

5. A Theoretical Limit to the Accuracy of Measurements:

$s_i(\tau;S)$	a two-dimensional base spline
σ_i	coefficients in the spline expansion of σ_{eff}
A_{ch}	a matrix that relates the crossed products (or range-gates) measured with frequency channel ch to the spline coefficients
Γ_{ch}	the covariance matrix of all the lagged product estimates for channel ch
$\overline{m}_{ch}$	the vector of measured crossed product estimates for each channel
$\overline{\sigma}_i$	the inversion solution to the spline coefficients using measured values $\overline{m}$
$\overline{\sigma}$	a vector with components σ_i
$l^i_{t,t'}$	contribution of a base spline s_i to the crossed product estimate for (t,t')
$\Delta\tau$	sampling interval of the signal
$m_{i,j}$	same as $m(i\Delta\tau, j\Delta\tau)$
$K_{i,j}$	same as $K(i\Delta\tau, j\Delta\tau)$
$\delta_{i,j}$	the Dirac delta matrix
$Q^{i,i+\Delta i}_{m,n}$	the Fisher information matrix components corresponding to the unknowns σ_m and σ_n and the measurement $m(i\Delta\tau,(i+\Delta i)\Delta\tau)$
τ_m	the lag value studied
S_m	the range value studied

τ_{lag}	the lag resolution
τ_0	the spatial resolution
$\Delta\tau$	sampling interval of the signal
S_{max}	the maximum speed possible
eff	efficiency of a method

6. A New Modulation Principle for Incoherent Scatter Measurements:

N	number of elementary pulses in the total modulation time
SC	scan count
NP	number of different periods in the alternating code sequence
$z_{\mathrm{SC}}(t)$	the signal sampled at scan count=SC
$s_i(\mathrm{SC})$	the sign of the i'th pulse in an alternating code sequence
$q_i(t)$	an elementary boxcar pulse starting at time i
$W_{t,t'}^{i',j'}$	the ambiguity function corresponding to the correlation of the pulses q_i' and q_j'
$A_{iji'j'}$	coefficients of the unwanted contributions in $W_{t,t'}^{i',j'}$
Wsh(i,j)	the Walsh sign matrix
a_n, b_n, c_n	bits (0 or 1) in the binary representation of integers i or j
$\wedge$	logical bitwise and of integers
$\oplus$	logical bitwise exclusive or of integers
par(i)	parity of the binary representation of i
int(x)	the largest integer not greater than x
$\log_2 x$	logarithm of x in base 2
$M_{t,t'}^{i,j}$	the final (decoded) lagged product estimate corresponding to sample times t and t' and pulse pair (q_i, q_j)

7. Some Other Applications of Inversion Methods:

m_t^m	square of the signal for a multipulse channel m
m_t^s	square of the signal for a single-pulse channel s
$\mathrm{env}^m, \mathrm{env}^s$	the modulation envelopes for multipulse and single-pulse channels
W_t^m, W_t^s	same as $W_{t,t}$ for the different channels
$\mathbf{1}$	the identical matrix
$\varepsilon^m, \varepsilon^s$	the noise in the estimates in different channels
Γ_m, Γ_s	covariance matrix of the noise
ξ^2	variance of the different lagged product estimates
A	a matrix relating the single pulse data averages to the multipulse data averages
Tr	the trace of a matrix
$\Gamma(x)$	a covariance matrix that depends on the unknowns x
$A_Z(l)$	discrete autocorrelation function of a sampled signal Z